

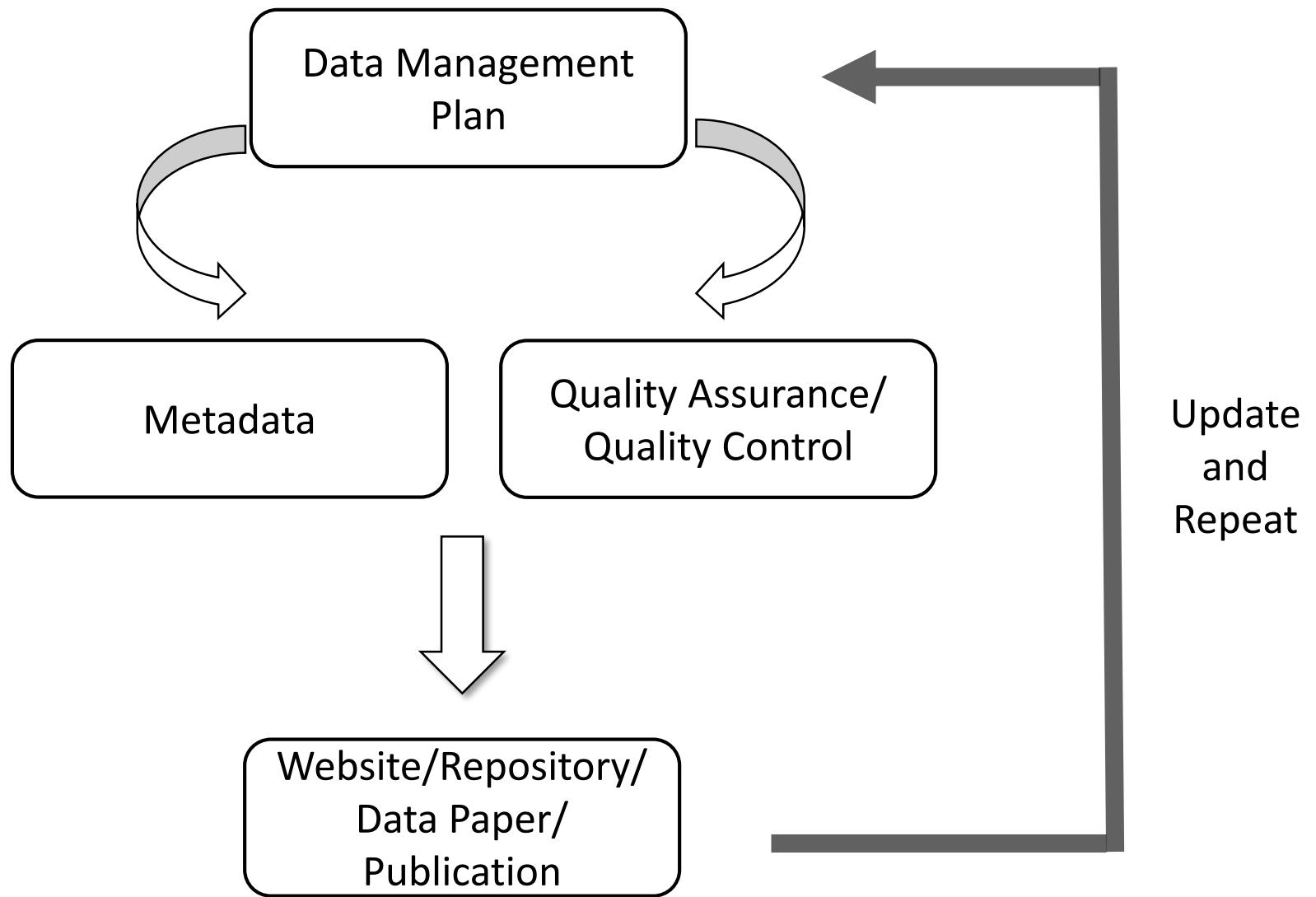


Interagency Ecological Program Data Management Showcase

Presented by The IEP Data Utilization Workgroup (DUWG)

August 13, 2019, 9:00-3:00

California Natural Resources Building Auditorium
1416 9th St Sacramento, CA



Update
and
Repeat

Opening the Door to Data FAIRness

Vanessa Tobias
vanessa_tobias@fws.gov

Characteristics of Open Data

Findable

Accessible

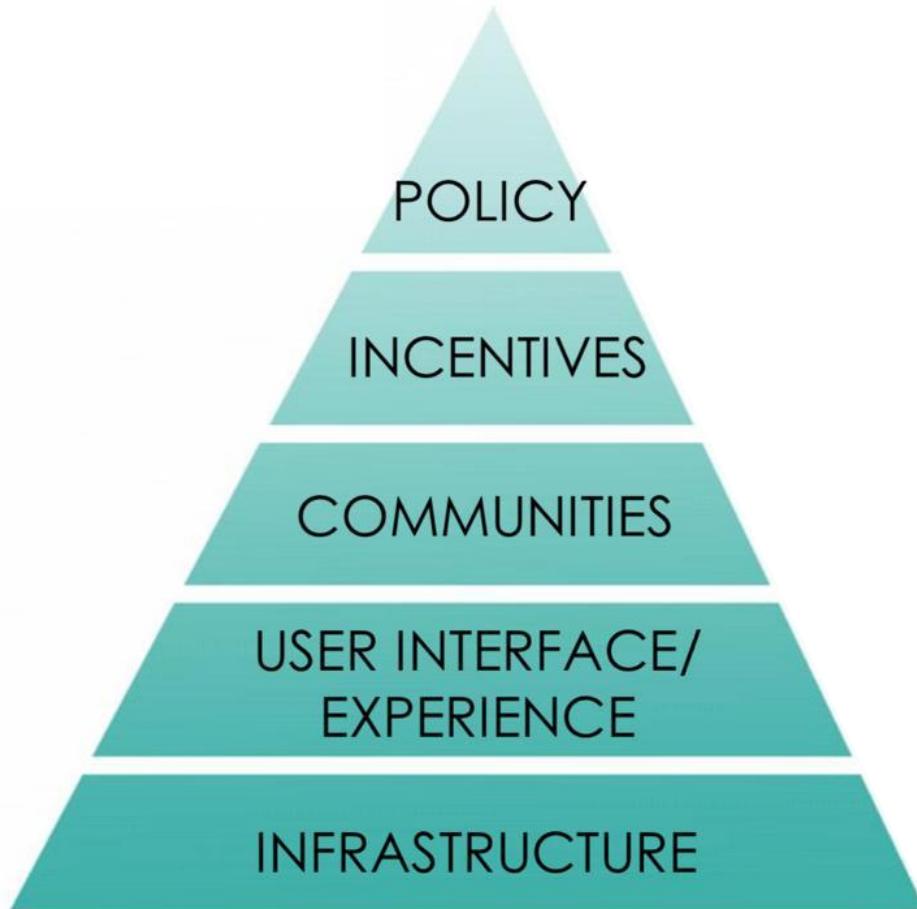
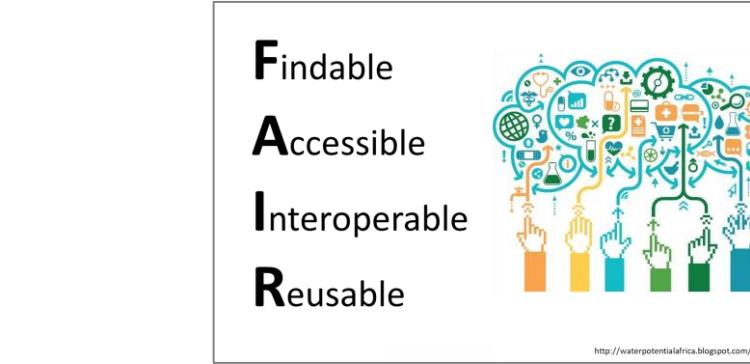
Interoperable

RReusable





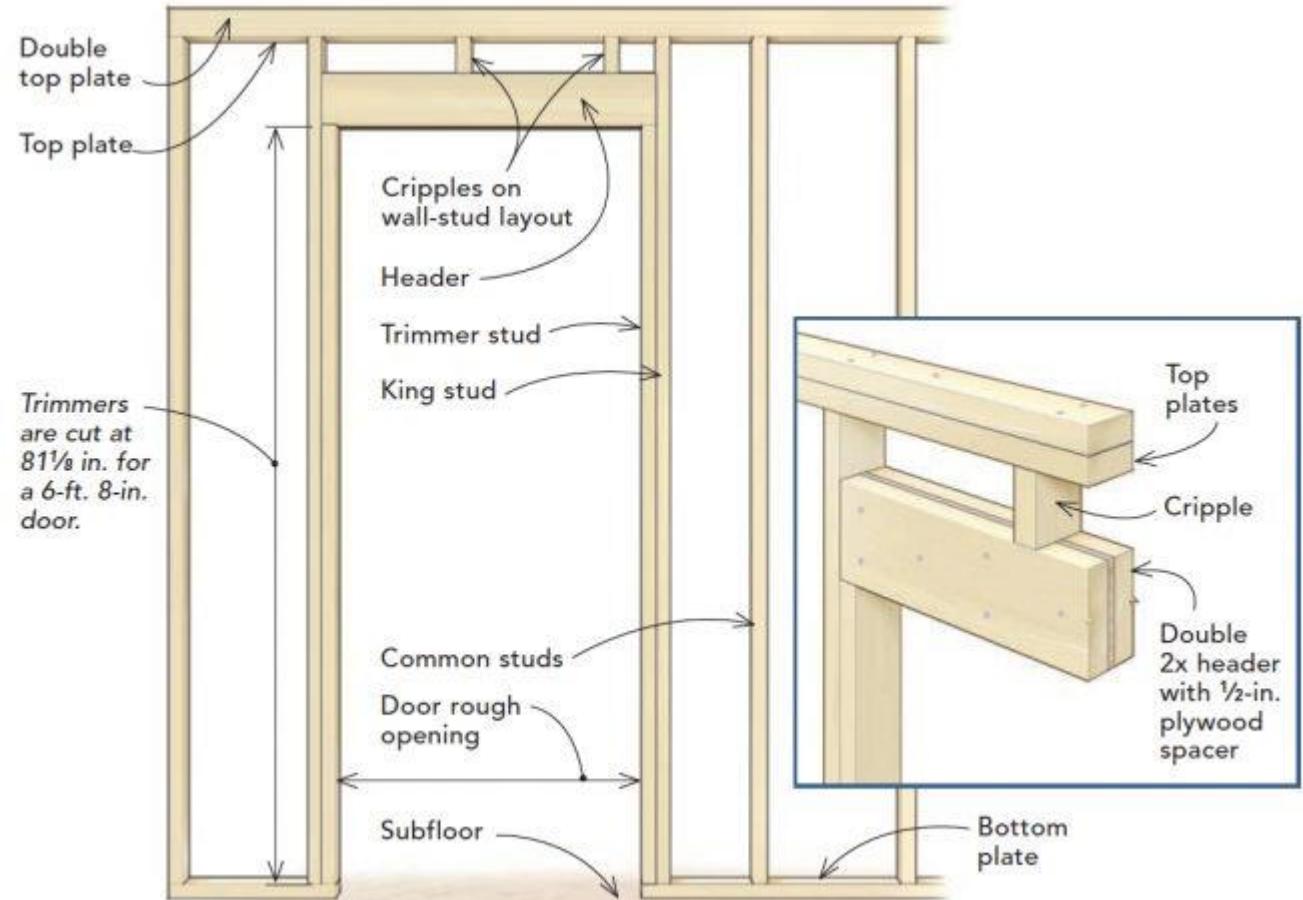
Photo: etsy



Creating a Culture of Openness

Image: Brian Nosek, Center for Open Science

Infrastructure make it possible to share data



User support

templates, help, user groups make sharing easier



Community

helps build trust and credibility



Incentives

make sharing rewarding



Policy

make data sharing required

AB 1755: The open and transparent water data act



SCIENTIFIC REPORTS

Availability of materials and data

An inherent principle of publication is that others should be able to replicate and build upon the authors' published claims. Therefore, a condition of publication in Scientific Reports is that authors are required to make materials, data and associated protocols promptly available to readers without undue qualifications in material transfer agreements. Any restrictions on the availability of materials or information must be disclosed to the publishing team at the time of submission. Any restrictions must also be disclosed in the submitted manuscript, including details of how readers can obtain materials and information. If materials are to be distributed by a for-profit company, this must be stated in the paper.

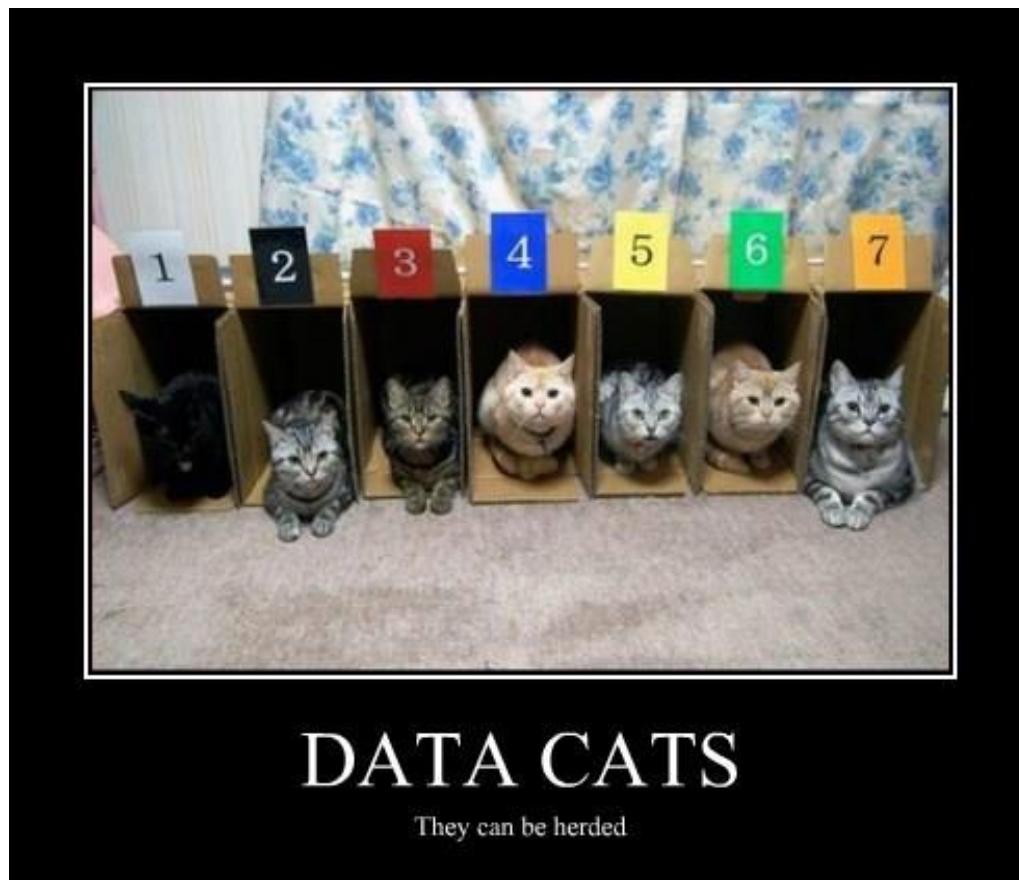
A screenshot of a digital document interface, likely a PDF viewer. At the top right are icons for search (magnifying glass), email (envelope), and file (document). The main content area contains text about the availability of materials and data, as described above.

Join us! We're holding the door for you.



Photo: shutterstock

Data Management Plans



Rosemary Hartman
Department of Water Recourses
Rosemary.Hartman@water.ca.gov

IEP Data Management Showcase
August 13, 2019

Why do we make you do DMPs?



WHERE THE HELL DID
ALL MY DAMN DATA GO
!?!?!?!?!

What happens when you leave the room before saving!

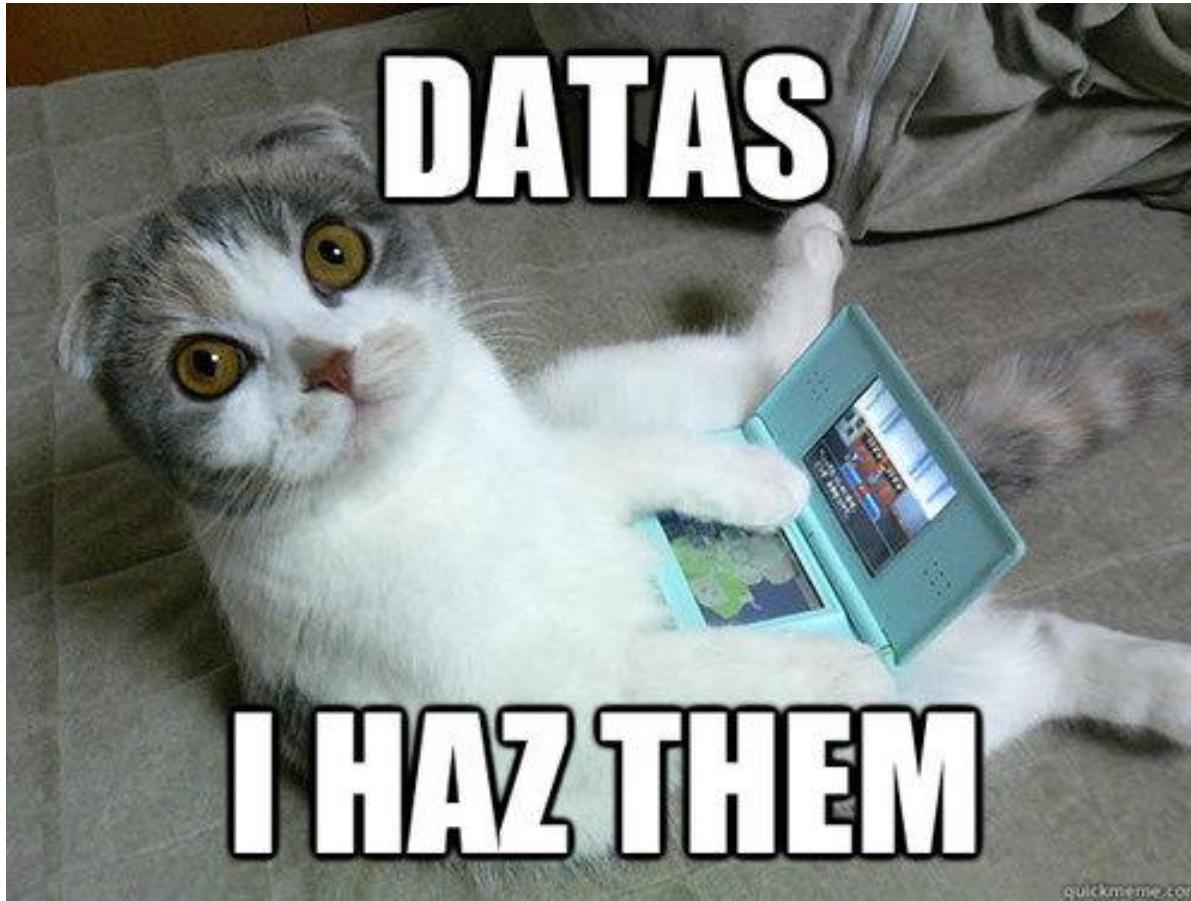
Why do we make you do DMPs?



Managing data is easier if you plan ahead



This DMP is not up to IEP standards



But we do want to keep it simple

- Contact info
- Data description
- Metadata description
- Backups
- Archiving
- Sharing
- Formats
- Quality assurance
- Rights and requirements



<https://github.com/InteragencyEcologicalProgram/Open-Data-Workshop/tree/master/resources>

PI and Point of Contact



(May or may not be the same cat)

PI and Point of Contact



Example: Fish Restoration Program

PRINCIPAL INVESTIGATOR:

Dave Contreras/CDFW/dave.contreras@wildlife.ca.gov/209-234-3459

Daniel Ellis CDFW/daniel.ellis@wildlife.ca.gov/209-234-3680

POINT OF CONTACT:

Dave Contreras/CDFW/dave.contreras@wildlife.ca.gov/209-234-3459

Data description

- What was measured?
- What methods did you use?
- What time period does it cover?
- How big is the data set?



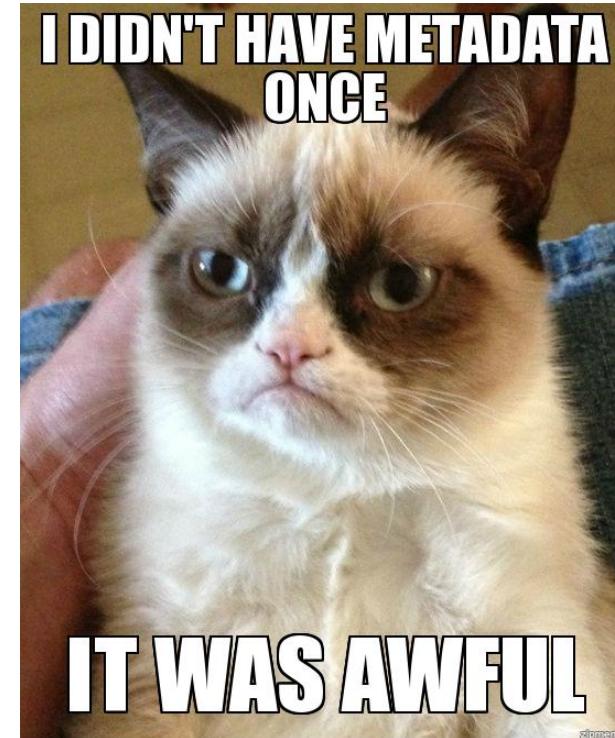
Data description

Example: FRP

The Fish Restoration Program (FRP) samples fish and invertebrates in tidal wetlands and their adjacent channels in order to inform monitoring of restoration sites. In 2020 sampling will occur around the current and future tidal wetland locations. Data includes fish and invertebrate catch, water quality, submerged aquatic vegetation biomass, chlorophyll-a concentrations, associated sampling information (effort), and associated environmental variables (weather, tides, etc). Data from the 2020 workplan will be added to the existing FRP database, and is estimated to be 30 MB in size.

Description of metadata

- Format – (EML? Word Document? Machine readable?)
- Location/access
- Update frequency



Data about metadata – that's “METAMETADATA!”

Description of metadata



Example: FRP

Metadata, including citation information, geographic scope, standard operating procedures, and QA/QC procedures is available online in

Ecological Metadata Language

(<https://portal.edirepository.org/nis/mapbrowse?packageid=edi.269.2>)

, and updated annually. Tables within the MS Access Database contain detailed descriptions of all the metrics collected.

Backups – data recovery

- How will you recover your data if your computer crashes?
- Usually the same format that you enter data
- Your server probably has backups you don't even know about.



Quick Poll! – Where do you back up your data?

- Go to www.menti.com and use code 64 14 16

Backups – data recovery

- Example: FRP

The FRP database is located on a shared drive housed by DFW. This server is backed up in its entirety once per week, with updates backed up mid week. An additional back up version of the database are stored on the environmental scientist's computer every week.



Archiving – saving it for later

- Should be QC'd, finalized data
- Accompanied by good metadata
- Stable formats



Data is where?

Cannot find.

Follow the 3-2-1 rule:

- 3 copies of the data,
- stored on 2 different media,
- with at least 1 **copy** stored off-site or in the "cloud".



Archiving – saving it for later

- Example – FRP

Data is currently preserved on paper and digitally. Paper data sheets will be stored between both of the Environmental Scientist's cubicles in binders. Digital data will be stored temporarily on the local CDFW server, to be backed up on CDFW's Tier 3 server. Yearly updates are also posted to the Environmental Data Initiative data repository in .csv format. An archiving system is currently in development with the Department of Water Resources Data Management Branch.

Quick Poll! – How often do people contact you about your data?

- Go to www.menti.com and use code 64 14 16

Sharing



- Be Open!
- EDI is IEP's recommendation
- Many other online repositories
- Share metadata, QA info with data
- Short embargos OK

Sharing

- Example: FRP

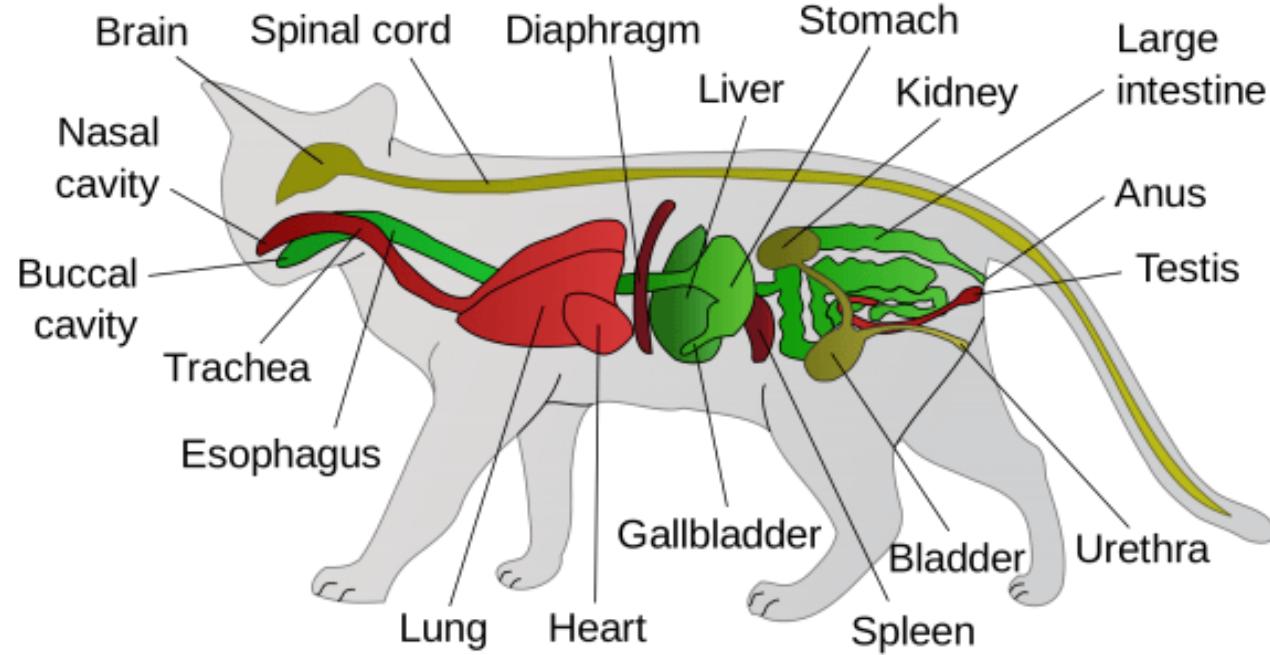
Data will be shared yearly on the Environmental Data Initiative data repository as well as CDFW's Biogeographic Information and Observation System (BIOS) as soon as reasonably possible, no more than one year after collection.

<https://portal.edirepository.org/nis/mapbrowse?packageid=edi.269.2>

Annual reports will be distributed on the FRP website:

<http://wdl.water.ca.gov/environmentalservices/frpa.cfm>

Format



- Include general format, and file extension
- Software requirements
- List formats for both storage and sharing

Format

- Example: FRP



Field data will be collected on paper datasheets or Ipads with Pendragon Forms Software. Data collected in lab will be entered directly into computer database when possible, or recorded on paper and transferred to the computer database. Data is currently stored in an Microsoft Access 2016 (.accdb) database, and published on EDI as a series of comma-separated (.csv) flat files. Data is shared on BIOS as an ESRI shapefile.

QA/QC

- Do you have a QAPP? How can your data users access it?
- SOPs and other quality documentation?
- Can a data user trust your data?



QA/QC



- Example: FRP

All data collection SOPs have quality assurance and quality control methods, including regular calibration of instruments, checks of identifications, and checks of data entry. See SOPs described in metadata for details. Contact D. Contreras for documents.

Rights and Requirements

- IEP recommends “cc by”
- Department/Agency requirements
- Disclaimers



Example - FRP

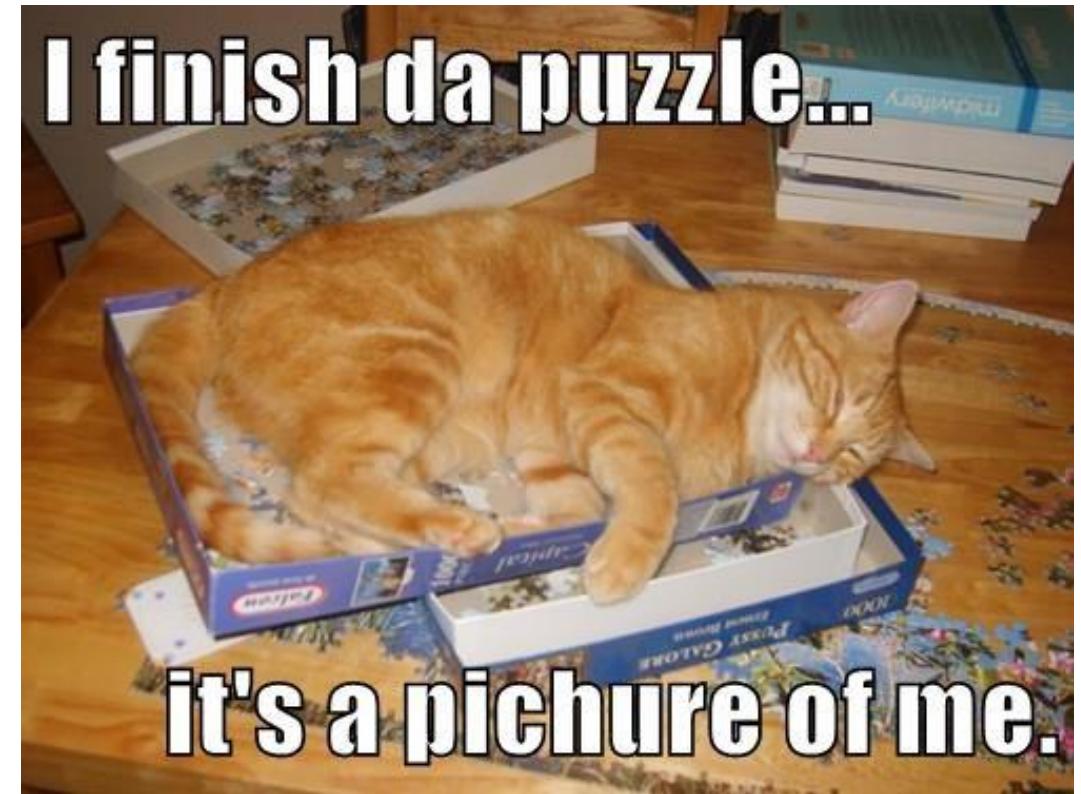
We adhere to all CDFW and CDWR policies on data quality.

<https://www.wildlife.ca.gov/Science-Institute/Policies-and-Guidelines>

Data is published online using the “CC by” license

Synthesis DMPs

- Documentation of data sources
- Integrated datasets
- Code



I finish da puzzle...

it's a pichure of me.

Data management plans rock

- Help you get organized
- Let your funders know what you are doing
- Help your users understand your data



Questions?



memecrunch

It's All About Quality!

Rachel Pisor, DWR, Senior Environmental Scientist (supervisor)

JohnFranco Saraceno, DWR, Senior Environmental Scientist (specialist)

Outline

- What is QA/QC?
- QA Connection to Open Data
- What is the IEP QA Survey?
- Key QA planning components
- Key QC components
- Resources



Quality Assurance vs. Quality Control

- What is quality assurance?
 - An integrated system of management activities
 - Prevention based
- What is quality control?
 - An overall system of technical activities
 - Detection based



What are some benefits to doing QA?



What are the benefits to QA?

- Consistency!
- Confidence!
- Avoid most of the guesswork!



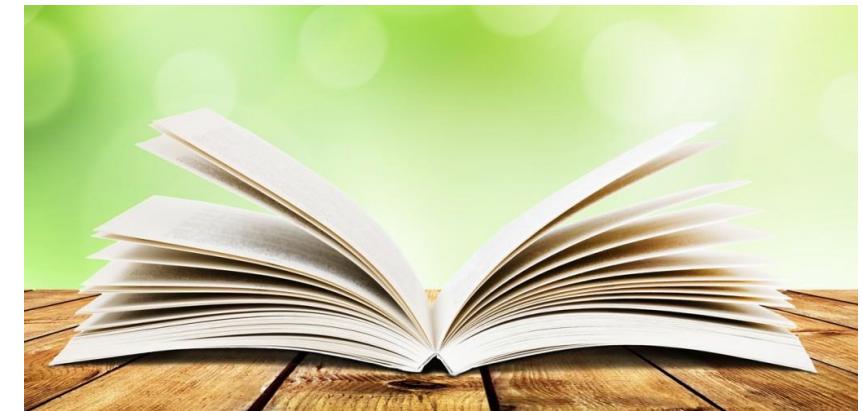
What are some benefits of QA?

- Promotes consistency across time, people, and space
- Increases transparency
 - Establish/further support confidence in your data by its users
 - Provide details and context for your data
- Legal defensibility
- ❖ **Critical for synthesis work**



QA in open data

- State water-related data is mandated to be open
- QA is a foundational component of open data and fosters:
 - Transparency of methods and data
 - Comparability which facilitates data synthesis efforts
 - Interoperability that makes it simpler and easier to integrate disparate datasets



Why QA matters to you...

- You may already be practicing QA in your study, however...
- Formally document methodology
- Establish known and documented quality to your data
- Promote the appropriate use of your data
- Reproduceable data that improves trust and credibility



IEP QA Survey

- Why was there a QA Survey?
- What is the next step with the survey?
- What is the future of QA in IEP?

IEP Quality Assurance Survey

The purpose of this survey is to gain a better understanding of the quality assurance documentation used throughout projects and programs in the IEP. This survey is intended as a 1-time effort. The results of this survey will guide future quality assurance efforts in IEP. Your honesty is key and only aggregated results will be shared.

Instructions: The answers you provide in this survey should reflect the 2020 Work Planning cycle, and not the overall history of the project (if there is a discrepancy between the two). Please complete 1 survey per workplan element, filling out the survey completely, and leaving no questions blank. Multiple boxes may be checked for any question. For new proposals, complete the survey with what data you are planning to collect and what documentation you are planning to develop. If you have any questions, please contact Rachel Pisor at Rachel.Pisor@water.ca.gov or 916-376-9711. Thank you for participating!

Information

IEP PEN:

PI Name (First, Last):

PI e-mail address:

Name of Study:

Data

Primary types of data collected for your study (if it is a synthesis study with no new data collection, please check "synthesis project"):

- Water quality
- Fish
- Phytoplankton
- Zooplankton
- Vegetation
- Benthic
- Synthesis Project
- Other: _____

Duration of your study or monitoring program:

- <1 year/Pilot project
- 1-5 years
- >5 years
- Other: _____

Quality Assurance Planning Components

- Project Plans
- QAPP's
 - Data Quality Objectives
 - Standard Operating Procedures
 - Data Management and Metadata
 - Quality Control Requirements
 - Data Review and Validation Procedures
 - Data Analysis Techniques

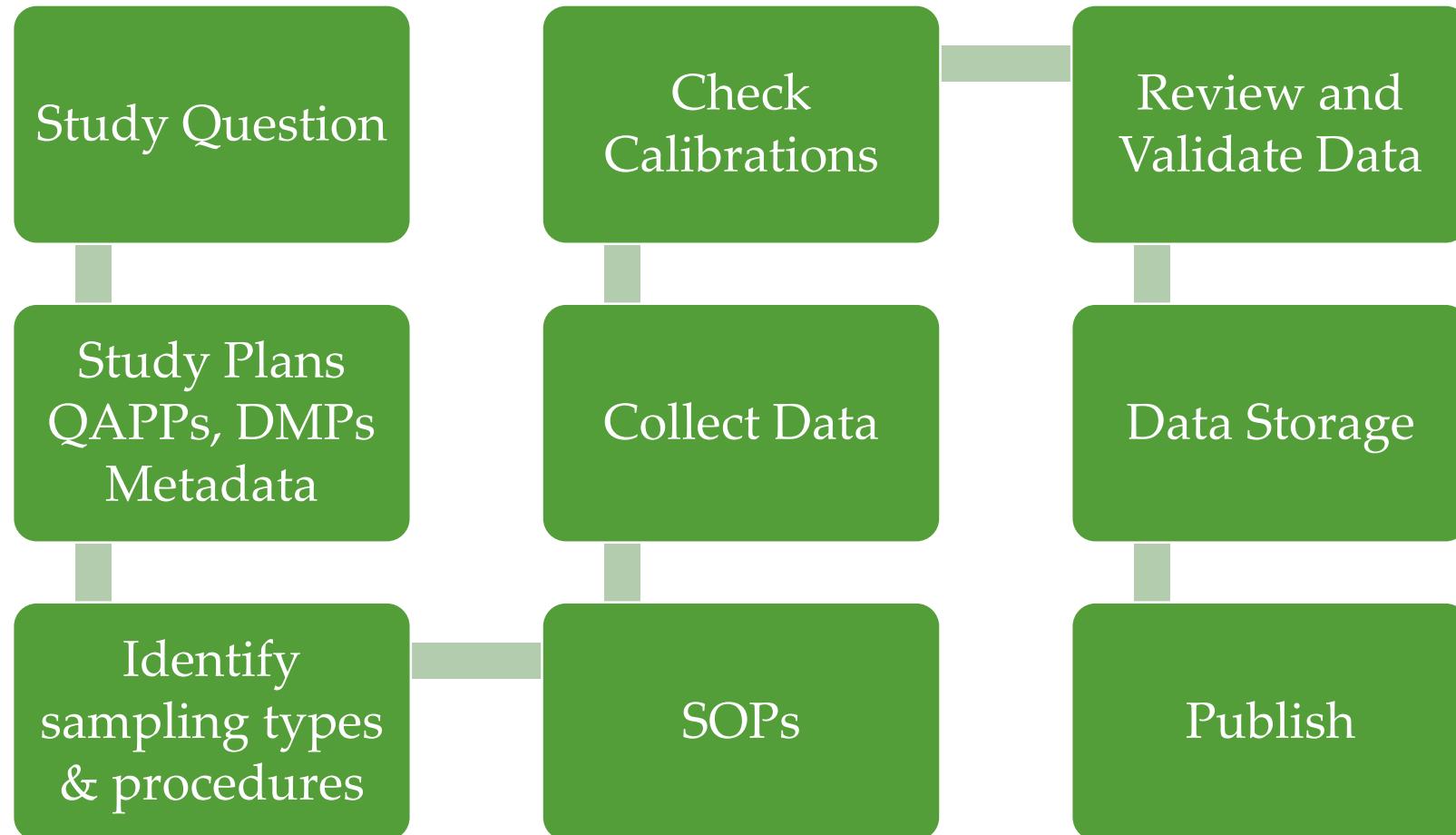


Key Quality Control Components

- Checklists
- Calibration and maintenance records
- Training records
- Consumables records
- Issue logs
- Data flagging



QA in practice



Resources

- EPA Quality
www.epa.gov/quality
- State Board Quality Assurance
www.waterboards.ca.gov/water_issues/programs/quality_assurance/
- Chesapeake Bay Program
https://www.chesapeakebay.net/what/programs/chesapeake_bay_quality_assurance_program
- USGS Quality Systems Branch
https://www.usgs.gov/mission-areas/water-resources/science/quality-systems-branch-qsb?qt-science_center_objects=0#qt-science_center_objects



To Contact Us:

- Rachel Pisor (DWR QA Program Section chief)

Rachel.Pisor@water.ca.gov

- JohnFranco Saraceno (DWR QA Officer)

JohnFranco.Saraceno@water.ca.gov



KEEP
CALM
AND
ASK
QUESTIONS

Metadata Magic

Catarina Pien & Jenna Rinde



Magic Trick







What is metadata?

Data about data

Essential to understanding the dataset



<https://www.ontotext.com/knowledgehub/fundamentals/metadata-fundamental/>



who:

- Contact information for data manager/program manager
- Funding sources

what:

- Data types (WQ, Fish, Zooplankton)

where:

- Study region, site coordinates

when:

- Sampling dates

why:

- Project background, why data are collected

how:

- Gear and instrument specifications
- Sampling procedures
- Lab procedures
- QAQC Procedures
- Methods for calculations (e.g. CPUE)
- Issues affecting samples quality
- Changes to procedures over time



Mentimeter questions

What types of metadata do you
find most useful as a data user?
(Pick up to 3)

How are your metadata currently
documented?

What are your greatest challenges
with metadata?

BENEFITS OF METADATA

- Cheat sheet for data producer and users
- Saves time long-term
- Data properly used
- Synthesis efforts



**"He's using a cheat sheet, ump!
Can he do that?"**

Human Readable

Metadata Formats

EDI Metadata Template (2018)¹

Dataset Title

Interagency Ecological Program: Discrete dissolved oxygen monitoring in the Stockton Deep Water Ship Channel, collected by the Environmental Monitoring Program, 1997-2017.

Short name or nickname you use to refer to this dataset:

IEP-DOSDWSC

Abstract

Dissolved oxygen levels in the Stockton Deep Water Ship Channel (SDWSC) have been monitored since 1968 by the Interagency Ecological Program's (IEP) Environmental Monitoring Program (EMP). The SDWSC is located on the San Joaquin River near Stockton, California. Beginning in 1997, 14 stations were routinely monitored typically in summer and fall. Dissolved oxygen impairment can occur in the SDWSC; therefore, two water quality objectives were established. The objectives of the dissolved oxygen monitoring study in the SDWSC are to: (1) determine if dissolved oxygen levels comply with the water quality objectives, (2) monitor long term trends, and (3) detect and document changes along the SDWSC. The EMP collects discrete dissolved oxygen readings near the surface and bottom of the water column during ebb slack tide. The 14 stations are located between Prisoner's Point on the San Joaquin River and ends at the terminus of the channel called Turning Basin. The site locations were selected at the channel markers on the San Joaquin River; therefore, may be referred as station number or channel marker they are located at. Dissolved oxygen and water temperature were recorded 1-meter below surface and 1-meter above the bottom of the channel. Over the period of record the following water quality parameters have been added: water temperature, specific conductance, pH, fluorescence, turbidity, secchi disk and a rating score for the blue-green algae, *Microcystis aeruginosa*.

Investigators

First Name	Middle Initial	Last Name	Organization	e-mail address	ORCID ID (optional)
		Interagency Ecological Program (IEP)			creator
Sarah Lesmeister		Dissolved Oxygen Monitoring in the Stockton Deep Water Ship Channel	California Department of Water Resources		creator

Machine Readable

```
<?xml version="1.0" encoding="UTF-8"?><eml:eml xmlns:eml="eml://ecoinformatics.org/eml-2.1.1" xmln
r>    <creator>        <individualName>            <givenName>Rosemary</givenName>            <surName>Hartm
e</organizationName>        <electronicMailAddress>Dave.Contreras@wildlife.ca.gov</electronicMailAdd
Name>California Department of Fish and Wildlife</organizationName>        <electronicMailAddress>Rya
ts, and zooplankton trawls. Fish data was collected using a lampara net and beach seine. For eac
eyword>        <keyword>California Department of Fish and Wildlife</keyword>        <keyword>CDFW</key
a and associated documentation, complete accuracy of data sets cannot be guaranteed. All data are
</geographicCoverage>        <temporalCoverage>            <rangeOfDates>                <beginDate>
etlands Monitoring Project Workteam. The SOPs are available online: http://nrm.dfg.ca.gov/FileHa
alifornia Department of Fish and Wildlife, Stockton, CA.</para>        <para> Environmental Pa
ecember. Sampling sites were typically void of vegetation and composed of sand and mud substrate
shallow water gear type that is deployed from shore by crewmembers. It measures 15 m long x 1.2 m
9.5 mm stretch mesh and connects to two wings composed of variable stretched mesh (69.9, 146.1,
30 individuals of each species. All additional individuals are counted, but not measured.</para>
nels or accessible near-shore habitat were too short to take a full five-minute tow, the tow time
was less dense and easier to sample. This is also the area where salmonids have been shown to fo
7). In the laboratory, the ropts were dried to a constant weight. Samples were taken in whichever
n) tow, neuston tow, and benthic cores/ponar grabs.</para>        <para> Zooplankton and Mysid
ted in a cod end was preserved in 70% ethanol for later identification.</para>        <para> B
values. Effort was calculated as catch per surface area of substrate sampled. Neuston tow: Emergi
ll invertebrates are extracted from detritus. One in five samples will be checked for identifica
dealt with on a case-by-case basis.</para>        <para> Methods references: IEP Tidal Wetland
    <para> Howe, E. R., C. A. Simenstad, J. D. Toft, J. R. Cordell, and S. M. Bollens. 2014. Mac
udyExtent>        <coverage>            <geographicCoverage>                <geographicDescription
>            </boundingCoordinates>        </geographicCoverage>                </coverage>
/northBoundingCoordinate>        <southBoundingCoordinate>38.18178043</southBoundingCoordi
gCoordinate>121.7090585</eastBoundingCoordinate>        <northBoundingCoordinate>38.089885
            <westBoundingCoordinate>121.7926</westBoundingCoordinate>                <eastBoundingCoord
n>724 FRP</geographicDescription>        <boundingCoordinates>                <westBoundingC
</coverage>        <coverage>            <geographicCoverage>                <geographicDescri
BoundingCoordinate>        </boundingCoordinates>                </geographicCoverage>
8.26321378</northBoundingCoordinate>        <southBoundingCoordinate>38.26321378</southBou
ndingCoordinate>121.92172</eastBoundingCoordinate>        <northBoundingCoordinate>38.1824
            <westBoundingCoordinate>121.8675863</westBoundingCoordinate>                <eastBoundingCoord
geographicDescription>CA3</geographicDescription>        <boundingCoordinates>
            </coverage>        <coverage>            <geographicCoverage>                <geographicDescri
```

Interoperability!

IEP-EDI Metadata Template (2018)¹

Dataset Title

Short name or nickname you use to refer to this dataset:

Abstract

Investigators

First Name	Middle Initial	Last Name	Organization	e-mail address	ORCID ID (optional)
		Interagency Ecological Program (IEP)			creator

Other personnel names and roles

First Name	Middle Initial	Last Name	Organization	e-mail address	ORCID ID (optional)	Role in project

License

CCBY

Keywords

Funding of this work:

Add rows to table if several grants were involved, list only the main PI, start with main grant first:

Metadata template

- Modified from the Environmental Data Initiative's Word document
- Easily convertible into Ecological Metadata Language (EML)



- **Dataset Title:**

- Interagency Ecological Program: Title

- **Short Name:**

- IEP-DOSDWSC

- **Abstract:**

- Summary of 5 W's and How

- **Investigators:**

- IEP (if applicable)
 - Principal Investigator(s)

- **Personnel:**

- Associated staff with study

Investigators

Sign Up at
<http://orcid.org/>

First Name	Middle Initial	Last Name	Organization	e-mail address	ORCID ID (optional)	Role in project
		Interagency Ecological Program (IEP)				creator
Sarah		Lesmeister	California Department of Water Resources	sarah.lesmeister@water.ca.gov	0000-0001-5492-1188	Principal Investigator
Jenna		Rinde	California Department of Water Resources	jenna.rinde@water.ca.gov	0000-0001-6677-4701	Data Manager

Personnel

First Name	Middle Initial	Last Name	Organization	e-mail address	ORCID ID (optional)	Role in project
Sarah		Lesmeister	California Department of Water Resources	sarah.lesmeister@water.ca.gov	0000-0001-5492-1188	supervisor
Jenna		Rinde	California Department of Water Resources	jenna.rinde@water.ca.gov	0000-0001-6677-4701	lead
Karen		Gehrts	California Department of Water Resources			supervisor
Shaun		Philippart	California Department of			supervisor

License

- **Creative Commons (CC BY)**
 - Open access but must give appropriate credit

Keywords

- *NSF's Long Term Ecological Research (LTER) vocabulary:*
 - <http://vocab.lternet.edu/vocab/vocab/index.php>
 - water quality, dissolved oxygen, monitoring
- *Custom:*
 - Interagency Ecological Program for the San Francisco Bay Delta Estuary, California Department of Water Resources

Funding

- Not applicable is ok

PI First Name	PI Middle Initial	PI Last Name	PI ORCID ID (optional)	Title of Grant	Funding Agency	Funding Identification Number

Timeframe

- Begin date: 1997-08-04
- End date: 2017-11-6
- Data collection: Ongoing

Geographic Location

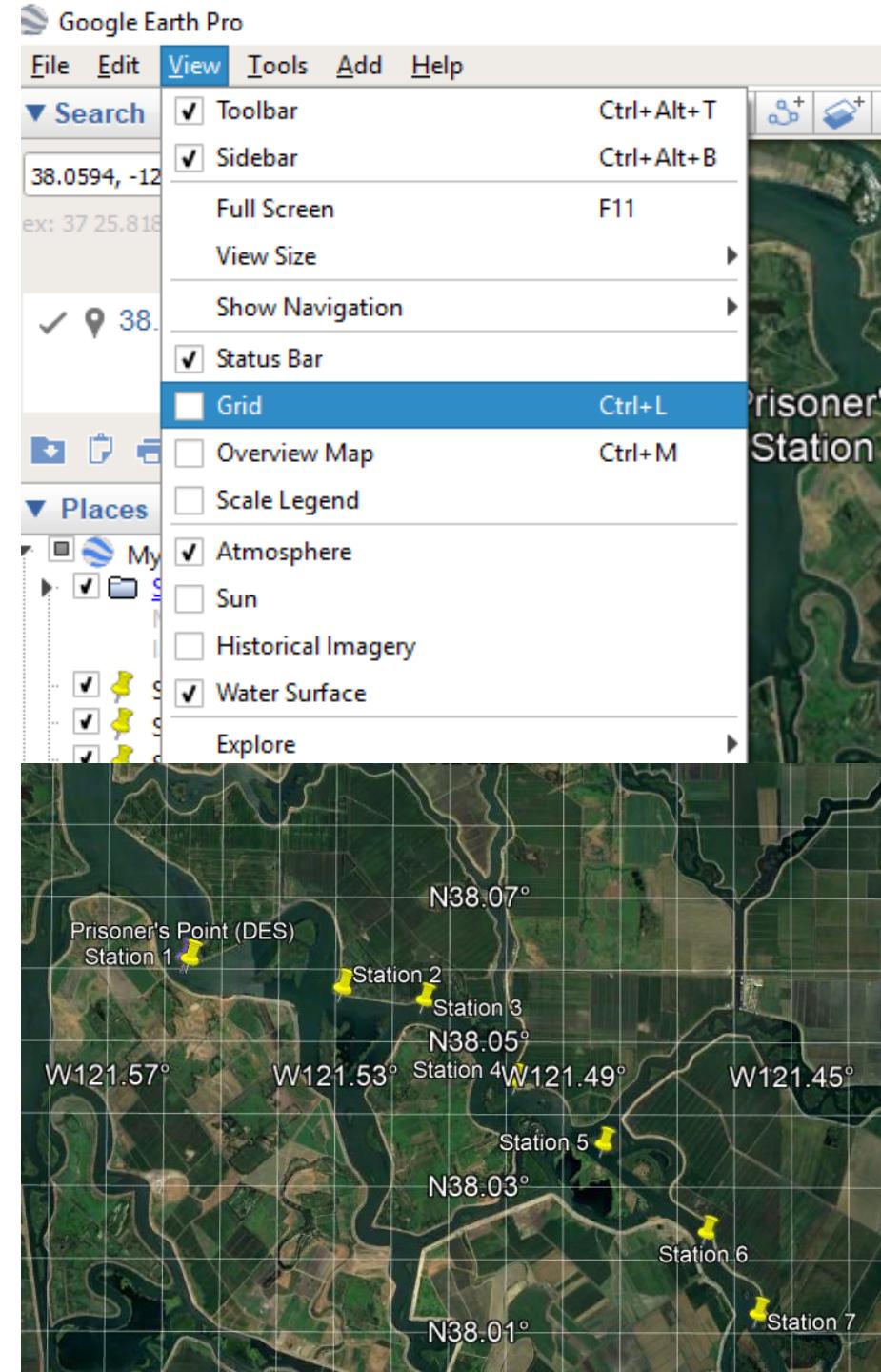
- Verbal description:
- North bounding coordinates (decimals):
- South bounding coordinates (decimals):
- East bounding coordinates (decimals):
- West bounding coordinates (decimals):

Google Earth

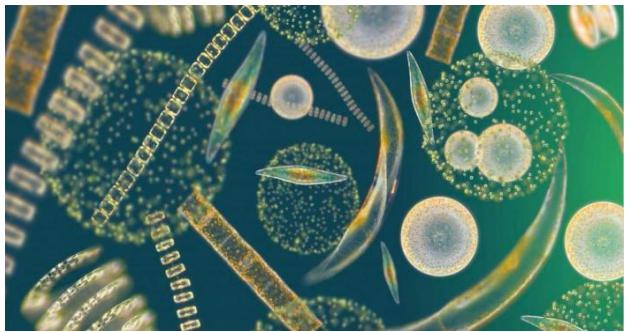
- Plot stations in Google Earth
- View (tab)
 - Check grid

Geographic Location

- Location
- North Bounding:
- South Bounding:
- East Bounding:
- West Bounding:



Taxonomic species or groups



- Common/relevant groups (e.g. top 10)
- Main taxa groups – phylum, class, order, family
- Full list of taxa

[Order/Suborder]: Amphipoda, Calanoida, Cladocera, Cyclopoida, Diptera, Harpacticoida

[Family]: Acartiidae, Bosminidae, Centropagidae, Chironomidae, Chydoridae, Corophiidae, Cyclopettidae, Cyclopidae, Cyclopodidae, Daphniidae, Diaptomidae, Ergasilidae, Eury cercidae, Ilyocryptidae, Macrothricidae, Moinidae, Pseudodiaptomidae, Sididae, Temoridae, Tortanidae



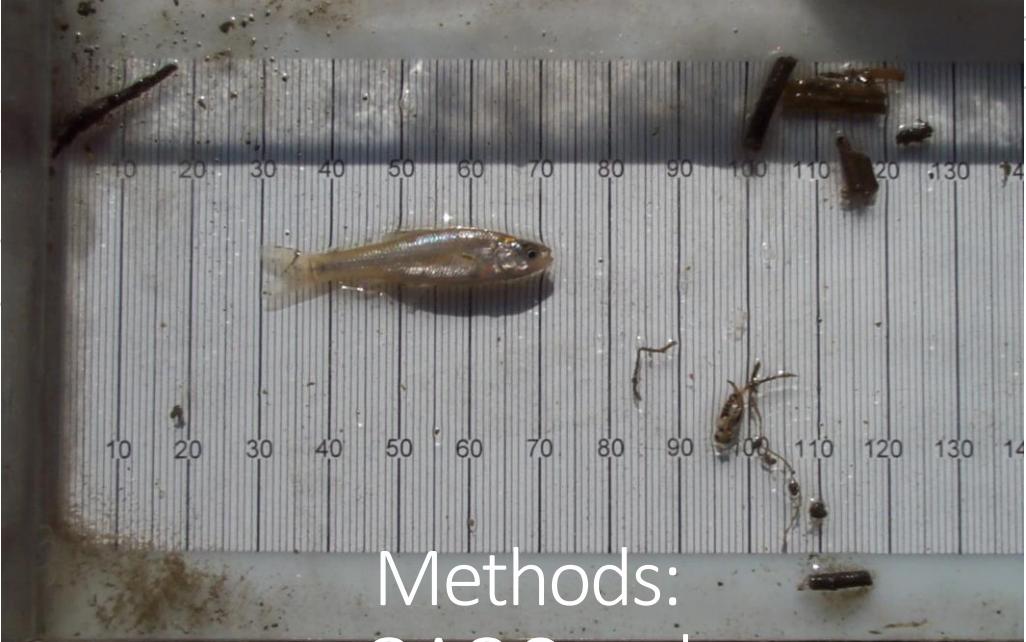
Methods: Field

- **WATER QUALITY:** Water temperature (degrees Celsius), electrical conductivity ([microSiemens/cm](#)), specific conductance (microSiemens/cm), pH, dissolved oxygen (mg/L), and turbidity (NTU) are sampled with a YSI ProDSS handheld meter.
- **ZOOPLANKTON:** The 150 micron net is fished for [5 minutes](#), and the 50 micron net is fished for 3 minutes, though [sampling times may be shortened when high levels of debris interfere](#) with sample collection. Samples are washed down with deionized water and preserved in [10% Formalin with Rose Bengal dye](#) to aid in separating organisms from detritus and algae.
- **INSTRUMENTS/EQUIPMENT:** Dimensions of the 150 micron plankton net are: 0.50 m diameter mouth, 2 m in length, polyethylene cod-end jar screened with 150 micron mesh. Water volumes are recorded using [General Oceanics flow meters \(Model 2030R\)](#) that are mounted inside each plankton net.



Methods: Lab

- **SAMPLE TRANSFER:** Zooplankton samples are transferred **8% Lugol's Iodine** solution after being stored for a minimum of **two weeks in formalin**. Samples are then shipped to contracted taxonomists.
- **SUBSAMPLING:** Contractors rinse and filter each sample through appropriately-sized sieves. Samples filtered through a 43 micron sieve are retained for analysis. For mesozooplankton, **200-250 organisms are taken per sample**.
- **IDENTIFICATION AND ENUMERATION:** Subsamples are extracted with a Hensen Stempel subsampling pipette, dispensed into a Ward zooplankton counting wheel, and examined under a compound microscope at a minimum of 100x magnification. **Zooplankton are identified to the lowest taxon possible.**



Methods: QAQC and Archiving



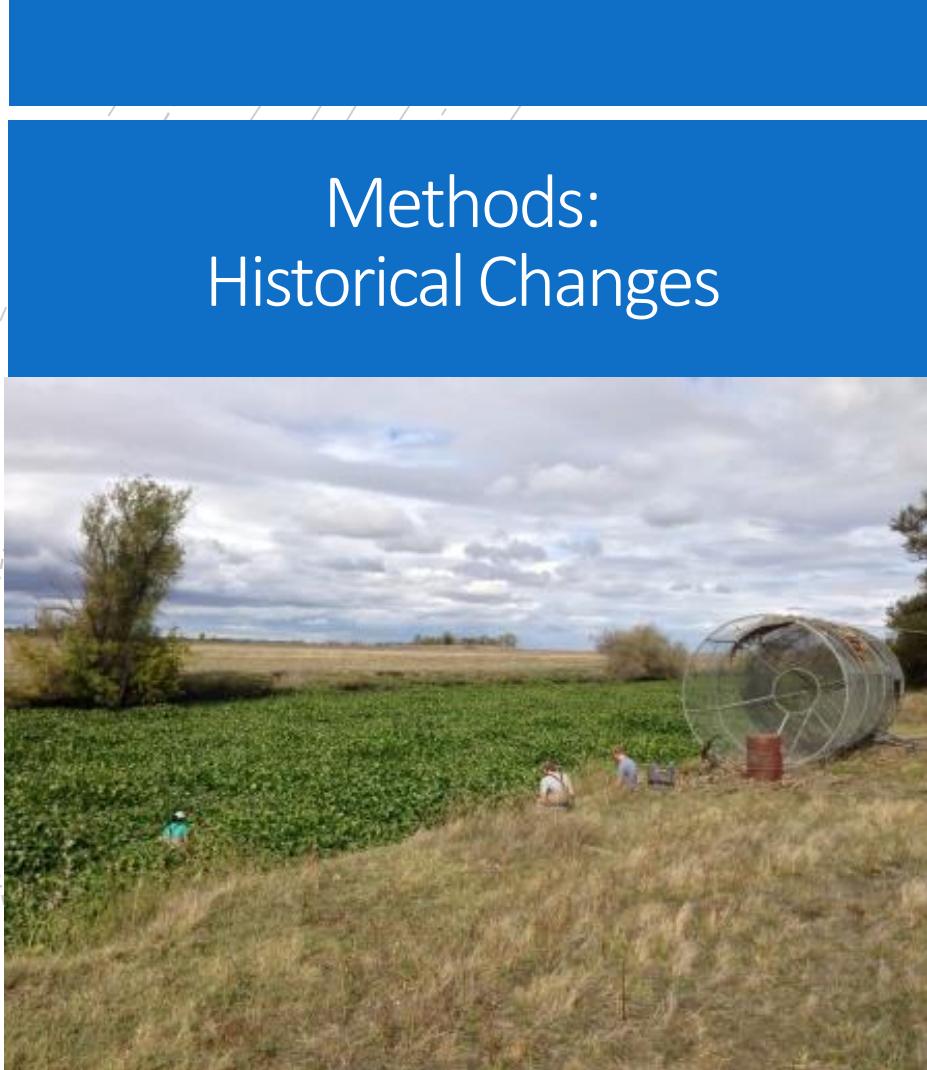
12.18.2

CALIBRATIONS: YSIs are calibrated for pH, turbidity, dissolved oxygen (DO), and electrical conductivity (EC) monthly.

SAMPLE QC: Contractors re-identify 10% of samples to ensure 90% similarity.

DATA QUALITY CONTROL: Datasheets are checked while being entered into the Microsoft Access database, which has customized error-checking and data validation checks, and historical water quality data are visualized and analyzed in R to look for outliers by station, year, and month.

ARCHIVING: Paper datasheets are archived in binders that are stored at the West Sacramento DWR office, and electronic copies are archived on DWR/AES Network drives.



Methods: Historical Changes

CHANGE IN GEAR

In 2002 there was a design change in the fyke trap and three fyke traps were deployed

CHANGE IN SITES

2010 - fall, the additional above Lisbon Weir (AL 1-4) and below Lisbon (BL 1-5) were added as year-round sampling to provide better spatial and temporal data on fish assemblages within the Yolo Bypass Toe Drain.

ENVIRONMENTAL OBSTACLES

In 2014 and 2015 clogging by water hyacinth rendered the trap likely ineffective for some period and the trap was out of operation and then relocated (from October to December).

CHANGE IN FREQUENCY OF SAMPLING

In 1998 and 1999 there were two traps fishing four months a year and seven days a week.

Methods: Calculations



Roland Birke/Getty Images)



DR RICHARD KIRBY

Calculations for CPUE are conducted in R. Separate calculations are conducted by net type, flowmeter type, and mesozooplankton & macrozooplankton versus microzooplankton, as these categories have different associated values for the equation below:

$$CPUE = (C/(V_{sub} \times V_{samp})) / V$$

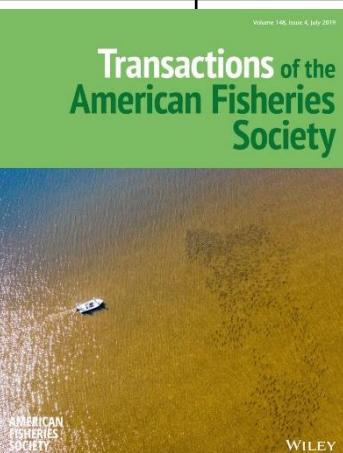
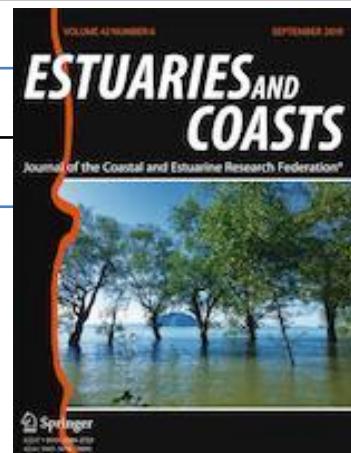
- **C:** Organisms counted in sample
- **V_{sub} :** Total subsample volume = *Subsample volume x Number of subsamples*
- **V_{samp} :** Total sample volume
- **V:** Volume of water sampled through net =
$$\frac{(Flowmeter\ count\ end - Flowmeter\ count\ start) \times R}{999999} * \frac{NV}{4}$$
- **R:** Rotor constant (Regular flow = 26873, Low flow = 57560)
- **NV:** Net volume = $3.14 * 0.0625$

Data Table

	A	B	C	D	E
1	Date	Time	StationID	SpCndSurface	SpCndBottom
2	8/4/1997	1040	PP	210	NA
3	8/4/1997	1055	lt4		
4	8/4/1997	1100	lt6		
5	8/4/1997	1115	lt12		
6	8/4/1997	1125	lt14		
7	8/4/1997	1135	lt18		
8	8/4/1997	1145	lt19		
9	8/4/1997	1200	lt28		

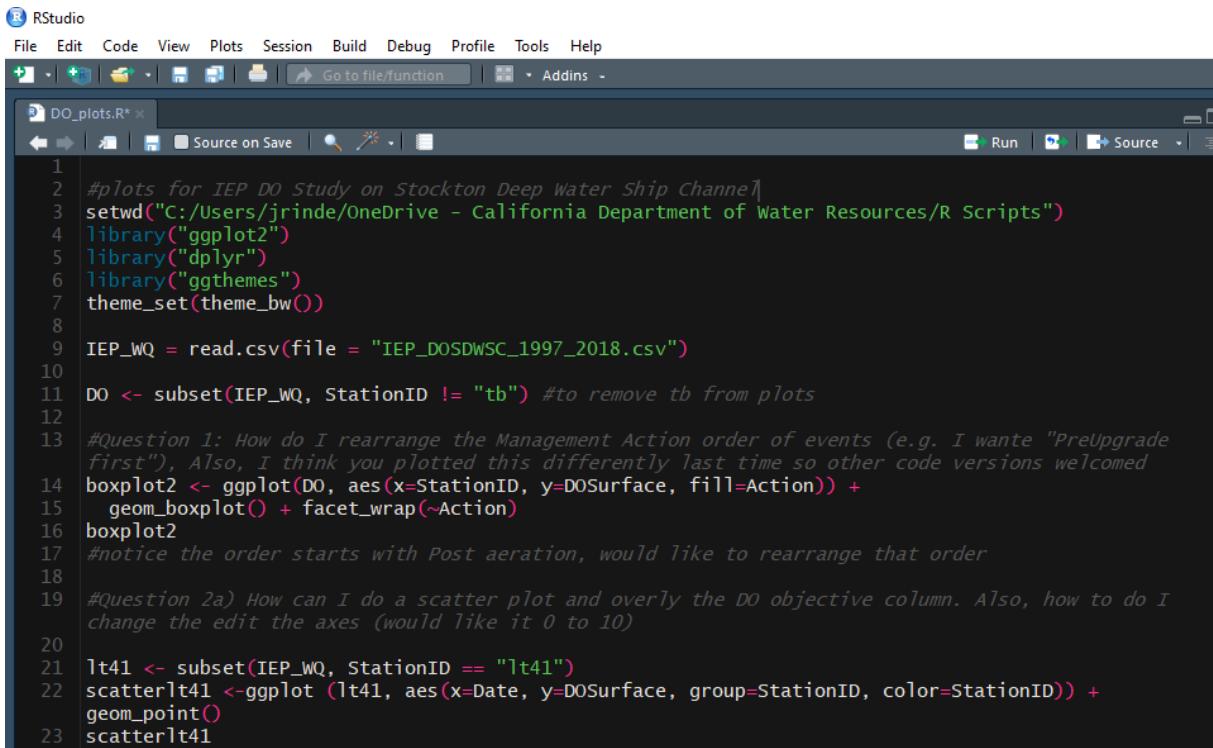
Column name	Description	Unit	Empty value code
Date	Date of sample	MM/DD/YYYY	
Time	Time of sample in Pacific Standard Time (PST)		
StationID	Station Code		
SpCndBottom	Specific conductivity 1-meter above bottom of channel	microSeimenPerCentimeter	-999999
pHBottom	Total pH 1-meter above bottom of channel	dimensionless	-999999
Fluorescence_Turner	Absorbed light in water, related to chlorophyll <i>a</i> concentrations	microgramsChlorophyllaPerLiter	-999999

Articles

Article DOI or URL (DOI is preferred)	Article title	Journal title
http://escholarship.org/uc/item/0tb0f19p	Low dissolved oxygen in an estuarine channel (San Joaquin River, California): mechanisms and models based on long-term time series	San Francisco Estuary and Watershed Science
https://link.springer.com/article/10.1007/BF02803533	Sources of oxygen demand in the lower San Joaquin River, California	Estuaries and Coasts
		



Scripts/code (software)



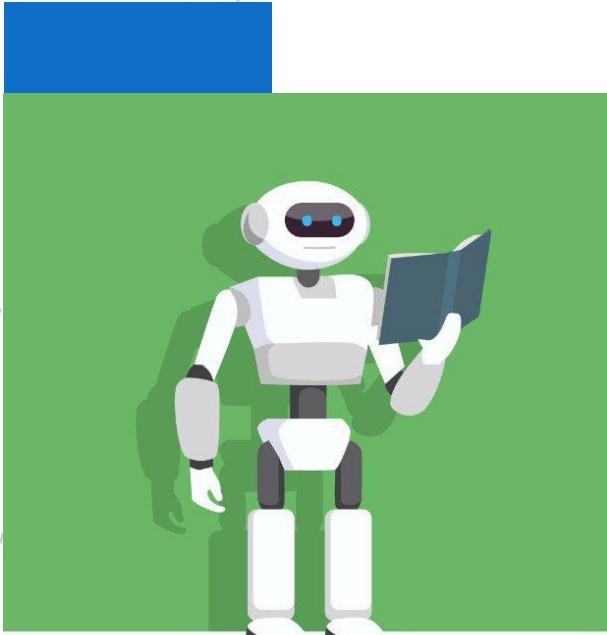
The screenshot shows the RStudio interface with the title bar "RStudio" and menu bar "File Edit Code View Plots Session Build Debug Profile Tools Help". Below the menu is a toolbar with various icons. The main area is a code editor titled "DO_plots.R*". The code is written in R and performs the following tasks:

- Changes the working directory to "C:/Users/jrinde/OneDrive - California Department of Water Resources/R Scripts".
- Imports libraries: ggplot2, dplyr, and ggthemes.
- Sets a theme using theme_set(theme_bw()).
- Reads a CSV file named "IEP_DOSDWSC_1997_2018.csv" into a data frame named IEP_WQ.
- Creates a subset of the data frame IEP_WQ, excluding rows where StationID is "tb".
- Creates a boxplot of DO Surface by StationID, colored by Action. A note in the code asks how to rearrange the Management Action order of events.
- A note in the code asks how to do a scatter plot of DO objective over time.
- Creates a subset of the data for StationID "lt41".
- Creates a scatter plot of DO Surface over Date for StationID "lt41".

Notes and Comments

Station 10 (stationid lt40) is the same as station P8 from IEP's Environmental Monitoring Program.

Machine Readable



Title: Interagency Ecological Program: Discrete dissolved oxygen monitoring in the Stockton Deep Water Ship Channel, collected by the Environmental Monitoring Program, 1997-2018

Creators: Interagency Ecological Program (IEP)
Lesmeister, Sarah; California Department of Water Resources
Rinde, Jenna; California Department of Water Resources

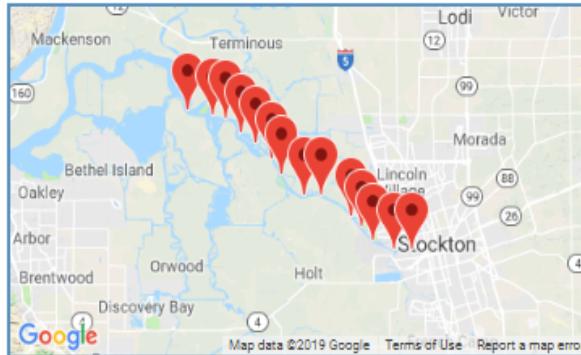
Publication Date: 2019-01-25

Citation: Interagency Ecological Program (IEP), S. Lesmeister, J. Rinde. 2019. Interagency Ecological Program: Discrete dissolved oxygen monitoring in the Stockton Deep Water Ship Channel, collected by the Environmental Monitoring Program, 1997-2018. Environmental Data Initiative. <https://doi.org/10.6073/pasta/4f254205afb605220a6453d933fb5d47>. Dataset accessed 8/05/2019.

Abstract: Dissolved oxygen levels in the Stockton Deep Water Ship Chanel (SDWSC) have been monitored since 1968 by the Interagency Ecological Program's (IEP) Environmental Monitoring Program (EMP). The SDWSC is located on the San ...

[Show more >](#)

Spatial Coverage:



[Geographic Coordinates](#)

Package ID: edi.276.1 (*Uploaded 2019-01-25*)

Resources: [View Full Metadata \(506 views\)](#)

[View Data Package Report](#)

[Download Data](#)

1. *Name:* IEP_DOSDWSC_1997_2018.csv

File: IEP_DOSDWSC_1997_2018.csv (243K; 34 downloads)

2. *Name:* IEP_DOSDWSC_site_locations_latitude_and_longitude.csv

File: IEP_DOSDWSC_site_locations_latitude_and_longitude.csv (398B; 25 downloads)

[Download Zip Archive](#)

Contact Information

Catarina Pien: Catarina.Pien@water.ca.gov

Jenna Rinde: Jenna.Rinde@water.ca.gov





Up Next – Data Publishing

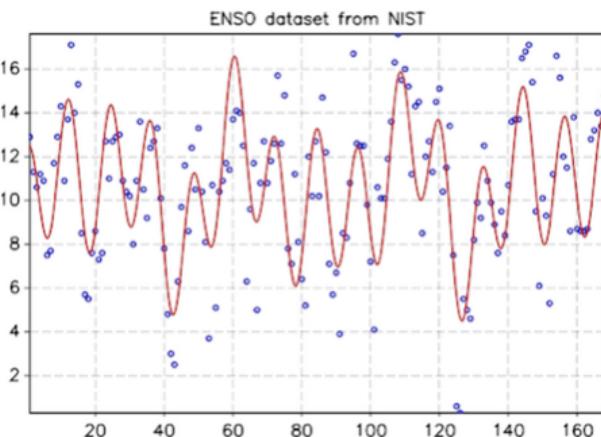


Making It Big with Data Publishing

Melinda Baerwald

Data Publishing in a Nutshell

1. Take a dataset



2. Describe it

Title
Authors
Year
Description
And others...

3. Assign a DOI



10.1234/exampledata



4. Reuse and reference!

ATLAS Collaboration, "Data from Figure 7 from: Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC: $H \rightarrow \gamma\gamma$,"
<http://doi.org/10.7484/INSPIREHEP.DATA.A78C.HK44>



Unique



Persistent

5. Enjoy the benefits!

Findability

Track citations

Reusability

Measure impact

What Is Data Publishing?

Releasing research datasets & associated information
in published form for re(use) by others

Why Do It?

Stop time-consuming data requests & start getting the praise you deserve

Turn this



Into this



Citations with



What Is a DOI?

Unique string assigned by a registration agency to identify content & provide a persistent link to its location on the internet.

Example: Delta Juvenile Fish Monitoring Program's DOI
<https://doi.org/10.6073/pasta/ea00fc37f0658dae21b817b1f93911cf>

Why Should You Want One?

Enables Preservation, Re-Use, & Exchange of Data

- Universally recognized
- Facilitates data discovery, transparency, & usage tracking
- Document dataset versions
- Retain relationship between dataset & metadata

Some Publishing Options

1. Hosting data on publicly available website with files available for download

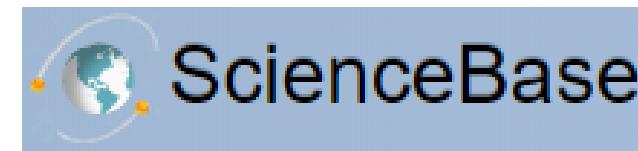
The screenshot shows a website for the California Department of Water Resources. At the top, there's a navigation bar with links for About, Contact, Current Conditions, Settings, and a search bar. Below the navigation is a main menu with Water Basics, What We Do, Programs, Work with Us, News, Library, and Search. The main content area displays four data series in a table:

			the estuary					
Fall Midwater Trawl: Zooplankton	CDFW	Randy Baxter, (209) 234-3483	A subset of stations are sampled for zooplankton: Clarke-Bumpus (CB) and mysid nets	2005 Sep only, 2011 Sep-Dec	Map	Metadata	Data Data Management Plan	
Summer Trawl: Zooplankton	CDFW	Randy Baxter, (209) 234-3483	Zooplankton sampled during first tow (CB sample).	2005	Map	Metadata	Data Data Management Plan	
20-mm Survey: Zooplankton	CDFW	Bob Fujimura (209) 234-3485	Zooplankton; CB samples, March-July	1995	Map	Metadata	Data	
Bay Study: Benthics	CDFW	Kathy Hieb (209) 234-3484	Decapods: shrimp, and crabs using otter trawls and midwater trawls throughout the Bay and upper estuary.	1980		Metadata included in zipped folders with Access databases	Data Data Management Plan	

<https://water.ca.gov/Programs/Environmental-Services/Interagency-Ecological-Program/Data-Portal>

Some Publishing Options

1. Hosting data on publicly available website with files available for download
2. Data repository
 - collect, manage, and store data sets for data analysis, sharing and reporting



Some Publishing Options

1. Hosting data on publicly available website with files available for download
2. Data repository
3. Data paper



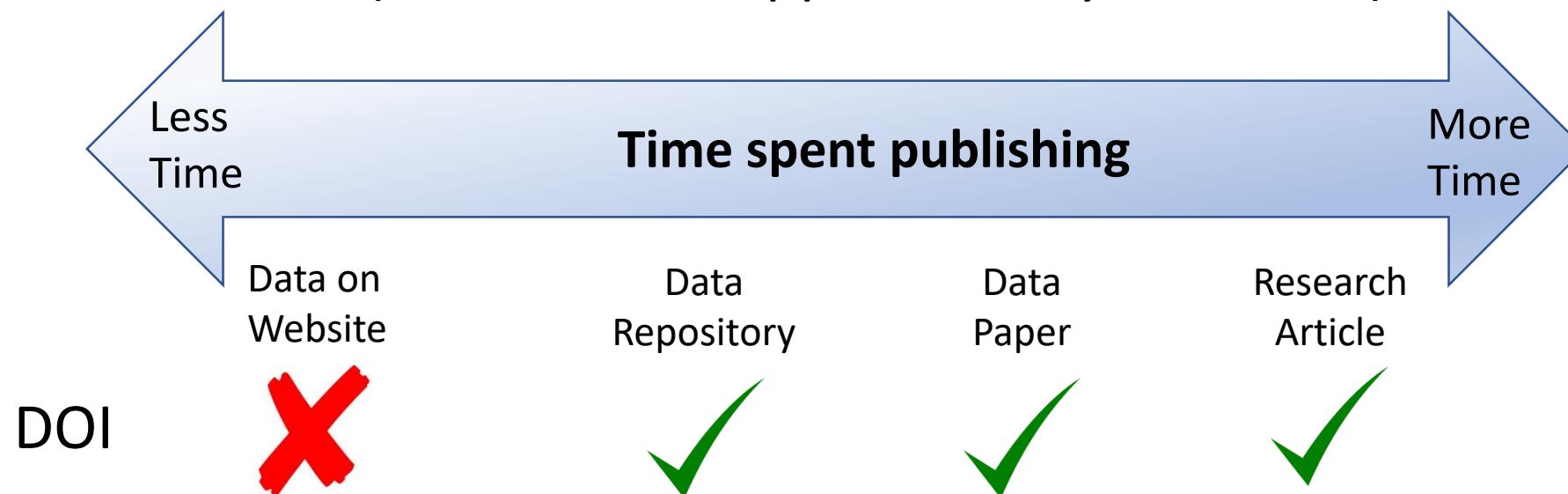
**BMC
Ecology**



**BioMed
Central**
The Open Access Publisher

Some Publishing Options

1. Hosting data on publicly available website with files available for download
2. Data repository
3. Data paper
4. Research article (data files in supplementary materials)

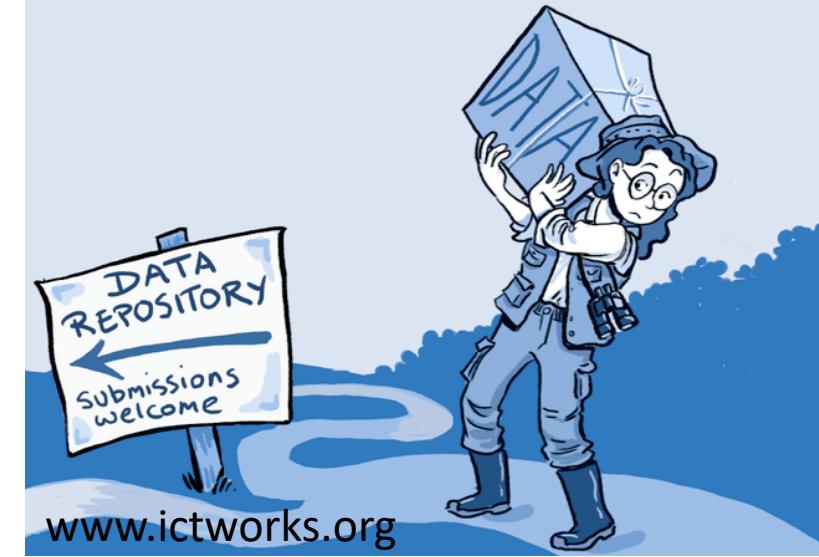


Mentimeter

- What data publishing options have you previously used?
- What data publishing options that you haven't previously used do you plan or hope to use in the future?

Data Publishing Best Practices

- Include metadata
- QA/QC & clean data (e.g. check for errors, missing values)
- Structure data in way that makes reuse easier
- Controlled vocabulary
- Obtain a persistent identifier (e.g. DOI)
- Update with a frequency that makes sense



Data Repository Recommended by IEP

Environmental Data Initiative

- Create . Package . Archive . Discover . Reuse -

<https://environmentaldatainitiative.org>



- Generates DOIs
- Focused on curation & archiving of ecological data
- Rigorous metadata standards (EML) & QC checks
- Version control
- Fast data exports & code generation
- Free (NSF-funded), yet has considerable resources & personal support
- Support for data synthesis
- Interoperable; Creative Commons attribution



Citations for IEP Datasets in EDI

- Interagency Ecological Program (IEP), B. Mahardja, J. Speegle. 2019. Interagency Ecological Program: Over four decades of juvenile fish monitoring data from the San Francisco Estuary, collected by the Delta Juvenile Fish Monitoring Program, 1976-2018. Environmental Data Initiative. <https://doi.org/10.6073/pasta/87dda12bed2271ce3d91abdb7864c50c>. 
- Interagency Ecological Program (IEP), B. Schreier, B. Davis, N. Ikemiyagi. 2019. Interagency Ecological Program: Fish catch and water quality data from the Sacramento River floodplain and tidal slough, collected by the Yolo Bypass Fish Monitoring Program, 1998-2018. Environmental Data Initiative. <https://doi.org/10.6073/pasta/b0b15aef7f3b52d2c5adc10004c05a6f>. 
- Interagency Ecological Program (IEP), S. Lesmeister, J. Rinde. 2019. Interagency Ecological Program: Discrete dissolved oxygen monitoring in the Stockton Deep Water Ship Channel, collected by the Environmental Monitoring Program, 1997-2018. Environmental Data Initiative. <https://doi.org/10.6073/pasta/4f254205afb605220a6453d933fb5d47>. 
- California Department of Fish and Wildlife, R. Hartman, S. Sherman, D. Contreras, D. Ellis. 2018. Fish catch, invertebrate catch, and water quality data from the Sacramento-San Joaquin Delta collected by the Fish Restoration Monitoring Program, 2015-2017. Environmental Data Initiative. <https://doi.org/10.6073/pasta/86810e72766ad19fccb1b9dd3955bdf8>. 

Getting Started with EDI

Options for uploading data
& metadata

1. Send EDI clean data and Word doc with metadata



Good for

Fewer datasets;
infrequent updates;
inexperienced R user

2. Do-It-Yourself with R code



More datasets;
More frequent updates;
Experienced R user
(or willing to learn)

Help is Available!

- Instructions on Github
 - <https://github.com/InteragencyEcologicalProgram/IEP-to-EDI-Publishing>
 - <https://github.com/InteragencyEcologicalProgram/Getting Started>
- Environmental Data Initiative support staff
 - Corinna Gries; cgries@wisc.edu
 - Mark Servilla; mark.servilla@gmail.com
 - Colin Smith; colin.smith@wisc.edu
- DUWG members
 - Co-chairs: Rosemary Hartman (rosemary.hartman@water.ca.gov)
Steve Culberson (steve.culberson@deltacouncil.ca.gov)

A Tour of EDI

- Online trainings & webinars
 - <https://environmentaldatainitiative.org/events/training-webinars-workshops/>
- Example IEP datasets
 - How to access: <https://portal.edirepository.org/nis/home.jsp>
 - How to download
- Featured Data Contributions
 - <https://environmentaldatainitiative.org/data/edis-featured-data-contributions/>