

The Data Management Lifecycle (in 5 minutes)

Catarina Pien

Bureau of Reclamation

IEP Workshop | April 18, 2024

Data Management Lifecycle



Plan

Data Management Planning

How data are collected and administered over its lifecycle

Accompany other types of documentation (SOP, metadata, QAPP) that may include more details

Living document (continuously updated)

Applicable to: Monitoring study, short field study, analysis, others



See [IEP Data Management Plan Template \(ca.gov\)](#)
See example [IEP Data Management Plans \(ca.gov\)](#)
See additional [Data Management Planning](#)

Components of Data Management Plan

- Principal Investigator and Point of Contact
- Description of data [source of data, how it is being acquired, collection methods, scale of data]
- Data processing [how data are processed, relevant workflows]
- Quality assurance and control [brief procedures for ensuring data quality]
- Formats [data format, necessary software]
- Metadata [how to access metadata, metadata standards]
- Data storage and backup, security [short-term storage and backup procedures, software, security]
- Data preservation/archiving [long-term archiving, preservation, succession plans]
- Access and sharing [How and where data will be shared, timeframe/frequency for sharing and publishing]
- Rights and requirements [Instructions or link for rights, requirements for data use]

Acquire

Data Management Planning

Data collection

Collect new data

Use, transform existing data



Maintain

Data Management Planning

Data collection

Process data

Database management

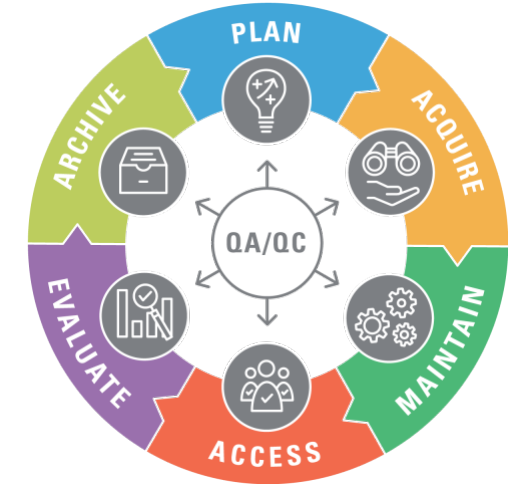
Check data (QA/QC)

Record/ load data

Rename data

Combine data

Convert data into useable format



	A	B	C	D	E	F	G
1	Station	StationName	StartDateDataset	EndDateDataset	Agency	Latitude	Longitude
2	ANC	ANTIOCH (USBR)	2/28/1999	6/9/2016	USBR	38.01782	-121.802965
3	ANH	SAN JOAQUIN RIVER AT ANTIOCH	12/31/1994	7/20/2017	DWR	38.01783	-121.8029633
4	BAC	BACON ISLAND AT OLD RIVER	4/17/2008	#####	DWR	37.96935	-121.5721741
5	BDL	BELDON LANDING	2/27/2008	#####	DWR	38.1869	-121.9708
6	BDT	SAN JOAQUIN R AT BRANDT BRIDGE	4/6/2005	1/3/2008	DWR	37.865	-121.3231
7	BET	BETHEL ISLAND AT PIPER SLOUGH	3/30/2006	9/12/2019	DWR	38.04109	-121.6321869
8	BIR	MIDDLE R AT BACON ISLAND RD	3/6/2018	#####	DWR	37.94306	-121.53389
9	BKS	BARKER SLOUGH PUMPING PLANT (KG000000)	7/1/2005	#####	DWR	38.2759	-121.7965
10	BLL	BLACKLOCK (NE1)	9/16/2001	7/1/2019	DWR	38.1801	-121.9069
11	BLP	SAN JOAQUIN RIVER AT BLIND POINT	2/17/2010	5/10/2018	DWR	38.03245	-121.7190399
12	CCS	CACHE SLOUGH	2/22/1999	#####	USBR	38.2918	-121.742
13	CLC	CLIFTON COURT	7/7/2004	#####	DWR	37.8298	-121.5574
14	CLL	COLLINSVILLE	2/23/1999	3/29/2019	USBR	38.07398	-121.850123
15	CNT	CONTRA COSTA	3/1/1999	12/9/2009	USBR	37.99495	-121.702809
16	CPP	CORDELIA PUMPING PLANT (KG002111)	12/31/2000	6/30/2009	DWR	38.2276	-122.1347
17	CSE	COLLINSVILLE ON SACRAMENTO RIVER	5/19/2009	5/15/2018	DWR	38.074	-121.8501
18	CYG	CYGNUS-CORDELIA SLOUGH	4/29/2010	5/31/2019	DWR	38.1527	-122.0908
19	DAR	DRAINAGE AT ARBOR ROAD	1/22/2016	11/6/2019	DWR	37.76896	-121.419414

Maintain

Data Management Planning

Data collection

Process data

Document data (metadata)



Description of data (equipment, field and lab procedures, parameters)

Document historical changes and each version's changes

i annotated ur document...

...but i eated it!



<https://geoukda.wordpress.com/2011/05/27/progress-update-2/>

See [Metadata template](#)

See [Historical metadata template](#)

More later in the workshop!

Access

Data Management Planning

Data collection

Process data

Document data (metadata)

Publication of data

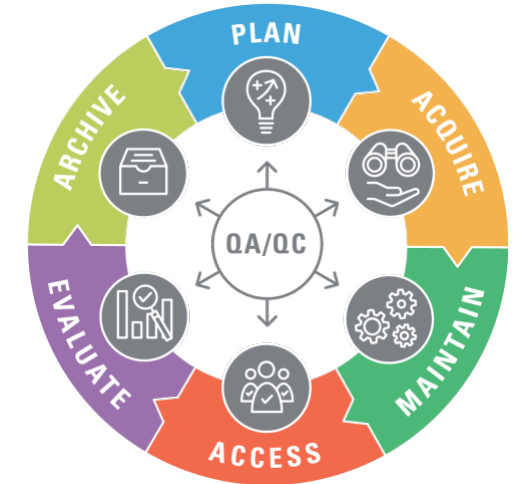
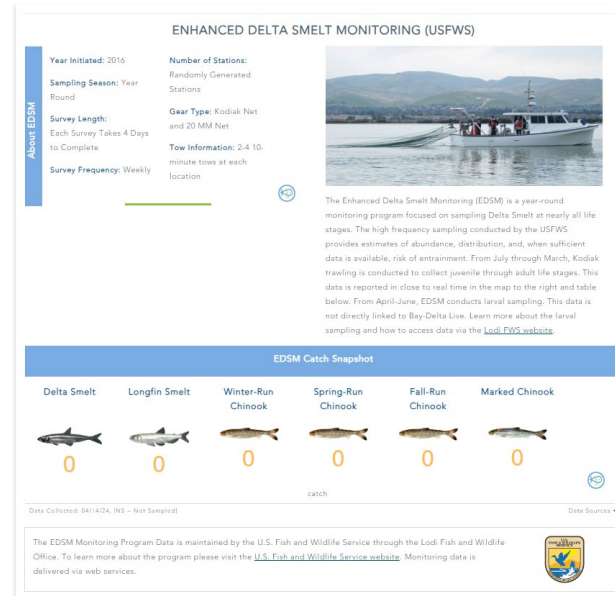
Web access

Data portal

Data paper

Reports, Manuscripts

More later in the workshop!



See [Guidelines for Data Publishing and Formatting](#)

See [Instructions for creating metadata for EDI](#)

See [EDI Resources for Data Authors](#)

See [EDI Resources for Information Managers](#)

Evaluate

Data Management Planning

Data collection

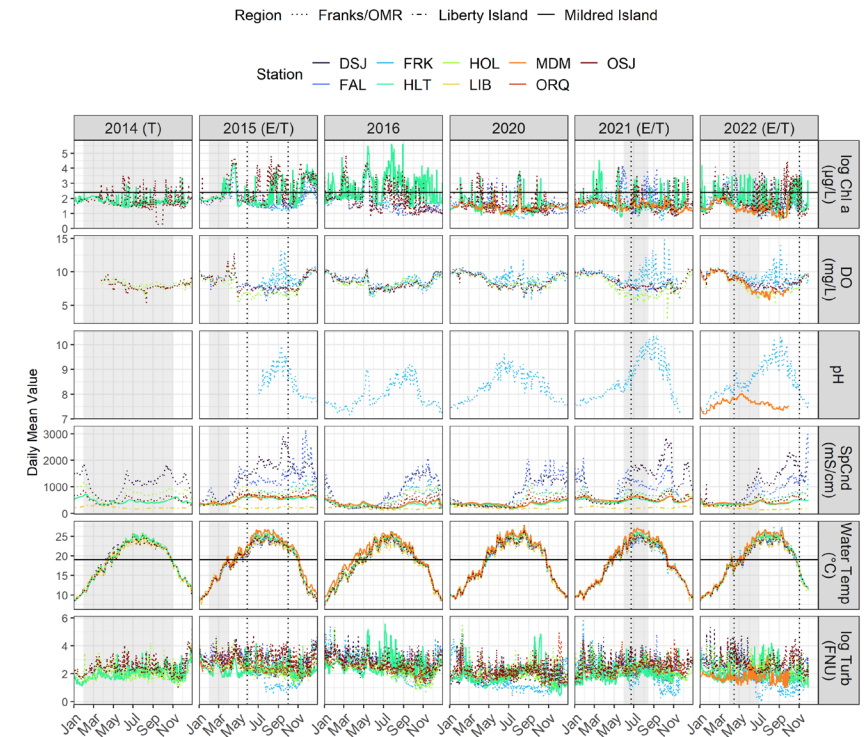
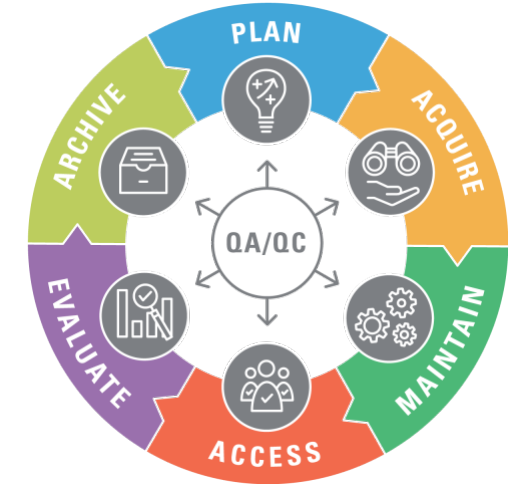
Process data

Document data (metadata)

Publication of data

Evaluate data

Analyze and use data
Develop reproducible
workflows



Archive

Data Management Planning

Data collection

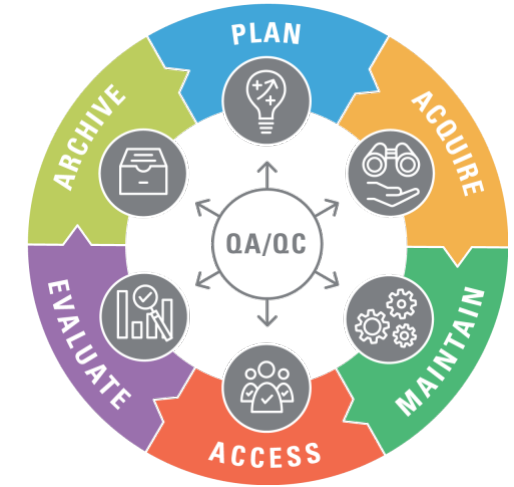
Process data

Document data (metadata)

Publication of data

Evaluate data

Archive and secure data



Multiple copies of data/database
Scanned datasheets (if physical)
Storage in a shared and backed up location
Secure confidential data

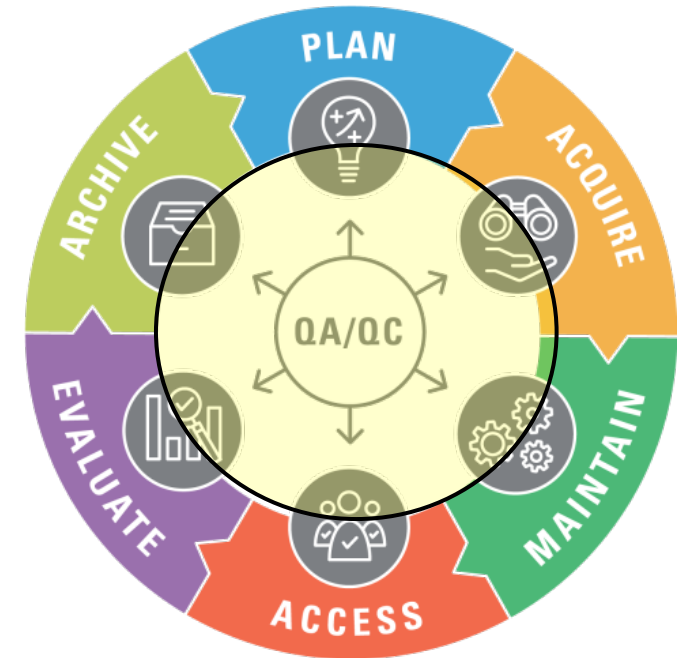
Quality Assurance/ Quality Control

Relevant along the entire process

Quality Assurance: Ensuring data quality on the front-end of data collection

Quality Control: Ensuring data quality post-collection

Coming up next!



See example [SOP template](#)

See [Digital Datasheet Best Practices](#)

See example [Fish QC Best Practices](#)

See example [Code for QC](#)