Data Management Lessons Learned from Environmental Cleanup Projects

Presented by:
Brian Jordan, USACE
and
Nazy Abousaeedi, CSC

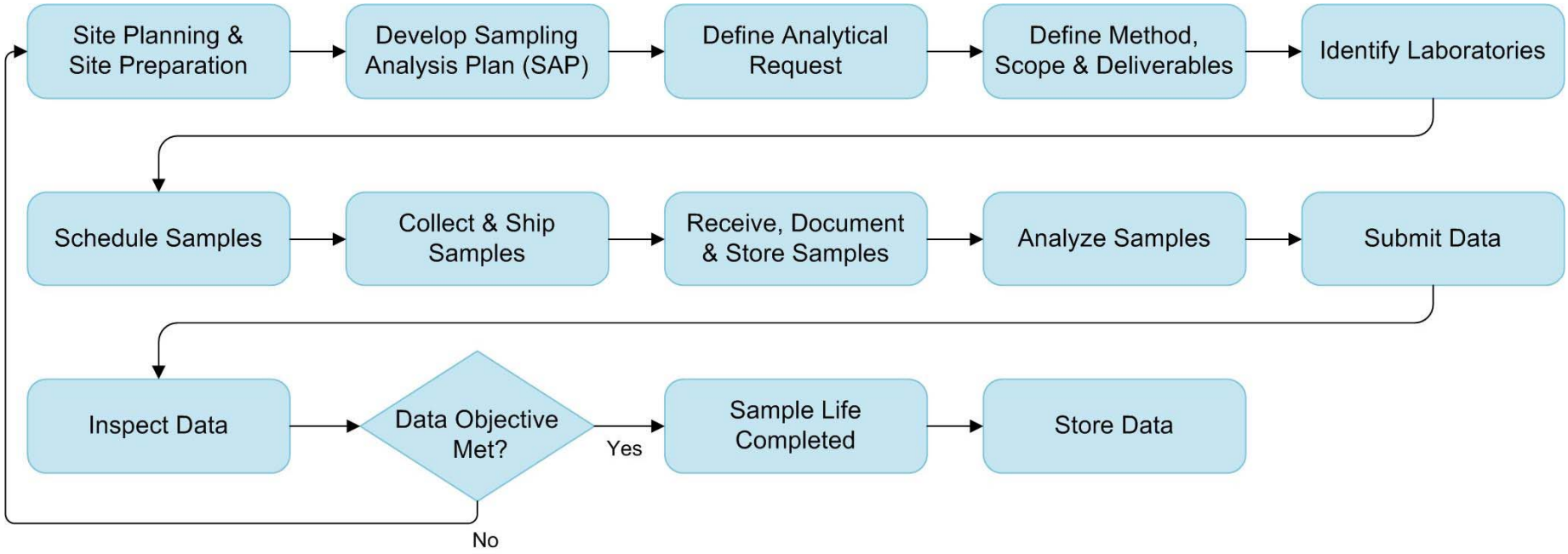
What all Cleanup Projects have in Common?

- Data is collected at multiple levels:
  - Program and/or Project
  - Field and Chain of Custody (COC)
  - Laboratory analysis
  - Analytical data assessment results
  - Data usability status
  - Financial data (if applicable)
- Collected data is used by multiple parties to make decisions about the site and/or public health.
- **Data is stored for long-term usage and retrieval.**

Common Lessons Learned

- What have we learned after 30 years of managing environmental data from cleanup projects:
  - The QAPP and/or sampling plan are not well defined and/or lack the necessary level of detail.
  - The appropriate method/process for examining analytical data are not used to ensure that the collected data is precise, accurate, and adequate for the intended use.
  - The appropriate amount of data points are not collected to support the decision-making process.
  - Data are stored in a format that cannot easily be used and shared by all parties.
  - The level of validation and assessment performed cannot be deduced or understood by looking at the project outputs and reports.
- After all the time, effort, and money spent we cannot really use the data for its full intended purposes.
- Program and Project Mangers are required to answer the same pressing questions with shrinking resources and funding pools.
- Environmental and human health decisions must be made based on credible data.

Life of a Sample



Note: The sample can be an environmental sample collected from a Superfund site or a sample collected in support of the FDA, NIH, DOD, etc.

Inconsistency in Data Reporting

- One of the challenges faced by environmental laboratories and data users is the inconsistency in data reporting.
- Inconsistencies in data reporting requirements:
  - Create unnecessary confusion and inefficiencies for laboratories generating different Electronic Data Deliverable (EDD) packages while serving multiple clients.
  - Utilize multiple operating procedures to generate different deliverables based on client and QAPP specifications.
  - Increase data reporting turnaround time resulting in delay to data users.
  - Increase a laboratory's overhead that will be passed onto the data user or the laboratory will try to cut corners to stay within cost.
  - **Discourage laboratories from participating in projects?**

Failure to Establish Data Reporting Standards

- Confusion in data reporting standards:
  - Can be caused by failure to identify and establish valid values at the start of the project and prior to the generation of analytical data.
  - Causes delays and/or errors as result of poor communication.
  - Introduces reporting variances by laboratories due to lack of guidance.
    - What are the required data elements?
    - What valid values should I use?
      - A simple example can be “W” or “Water”, “L” or “Low”
      - More complex ones – Do I report the prep method or leave the field blank/null?
      - How should I report date and time?
      - How many significant figures, etc.?
  - Introduces reporting variances by allowing subcontractors to provide different data reporting instructions to different laboratories.
- **Data reporting standards are as crucial as the method itself.**

Providing Access to Appropriate Automated Tools

- Method Compliance Assessment:
  - Self-inspection – performed by a laboratory prior to delivery.
  - Performed by a third party to assess compliance with the requested method and data reporting procedures.
  - Used by an organization to assess the compliance of their data (e.g., State, EPA, USACE, Abs, etc.).
- Data Usability Assessment:
  - Used by a third party to assess and establish a data usability determination.
  - Used by the organization (e.g., State, EPA, USACE, Abs, etc.).
  - Makes an educated decision (confidence) on the site or at a sampling event (e.g., OSCs, RPMs, Health, or other agencies).
- Accreditation Bodies (AB) can use these tools to assess the laboratory's capability to provide the requested data in an acceptable data reporting format and to ensure that the data is precise, accurate, and adequate.
- Expedites the data assessment process resulting is cost and time savings for users.

Identifying and Using the Appropriate Data Validation Process and Tools

- Automated/electronic data review tools currently on the market:
  - Accept a limited set of EDD formats (one or two).
  - Perform data review based only on existing methods and/or specifications – can pick only one of the listed methods.
  - Tests/checks performed are hardcoded/set – making changes costly and time-consuming.
  - Have limited scalability.
  - If any of the above requirements change, the tool will need major updates or a completely new version release.
- What is missing/needed:
  - Format neutral data review tool.
  - Provide data users the ability to modify review specifications based on their needs.
  - User friendly web User Interfaces (UI) to set-up requirements and allow customization on an as needed basis – in real-time – with no development delay.
  - Ability to recalculate/verify results.
  - Scalable and configurable tool - accommodating change if details emerge leading to its expansion from screening level to definitive level testing.

Store Data in Proper Format

- Data and results were scattered in multiple formats and in multiple databases.
- Data are stored in a hardcopy report format, making it nearly impossible to gain access to the complete set of results.
- Stakeholders do not have access to the field and analytical data, as well as the validation results, in one centralized database.
- Results are stored in a format that cannot be easily understood and used for decision-making and trend analysis purposes.

Possible Solutions

- A standardized electronic reporting format that outlines the EDD structure as well as acceptable valid values expected by the user - eliminating the need to normalize the data after the assessment has been completed.
- Develop and conduct training programs for **the environmental community** to better understand EDD format and why data is reported in a specific order.
- Make use of automated data assessment part of our culture and day to day function.

Possible Solutions (Cont.)

- Use of automated data assessment tools that are configurable and scalable:
  - Can grow with your project needs – can add methods and projects on the fly.
  - Are format neutral - can accept multiple EDD formats (e.g., SEDD, NEDD, ERPIMS, EXCEL, etc.) and generate standard reports for all data delivery types.
  - Stores data in a secure centralized database that can be accessed by stakeholders with different access rights.
  - Allows users to download reports and has an audit track record.
  - Reproducible and independent electronic review improves overall project transparency.

Inspiring Quotes

- “Our past experiences are our roadmaps in making a better decision for the future.” *Anonymous*
- “The first step toward getting somewhere is to decide that you are not going to stay where you are.” *Anonymous*
- “There is a better way to do it – Let’s find it.” *Thomas Edison*


QUESTIONS ?

Contact Information:

- Brian Jordan, USACE
  - Phone: 505-506-2189
  - Email: Brian.D.Jordan@usace.army.mil
- Nazy Abousaeedi, CSC
  - Phone: 703-461-2328
  - Email: fabousaeedi@csc.com