ITRC has developed a series of fact sheets that summarizes the latest science, engineering, and technologies regarding environmental data management (EDM) best practices. This fact sheet describes:

- defining categories of data collected in the field
- low volume data
- high volume data
- unstructured data
- metadata

Additional information related to field data collection is provided in the Introduction to and Overview of Field Data Collection Best Practices fact sheet, as well as fact sheets on Field Data Collection Process Development Considerations, Field Data Quality Assurance and Quality Control (QA/QC), Field Data Collection Training Best Management Practices, and Other Considerations for Field Data Collection.

1 OVERVIEW

Data collected in the field include a vast range of potential observations, measurements, and sample types depending upon the data and information needs of the project. Matrixes observed, measured, or sampled may include air, water, soil, sediment, tissue, etc. Data may be collected over wide-ranging space and time scales and may be recorded in hard copy, digital, or both formats.

Before an organization can develop processes and implement field data collection programs, it is important to understand the basics of what categories of data you are collecting, what form is it in, and data quality objectives. The following sections will elaborate on the various categories of data and provide a framework for understanding your project before you begin the process of developing a field data collection program.

2 SAMPLE AND LOCATION IDENTIFIERS

Sample identification is a key element in the field data capture process and is fundamental to managing data. Proper sample identifier assignment ensures that laboratory results from each sample can be traced to their point of origin and throughout the data lifecycle and can be used to interpret the results correctly. Data management plans should have defined sample-naming procedures. Essentially, every sample collected or generated in the field must be assigned a unique sample identifier that is traceable to the field collection records. This applies to normal field samples, different fractions created in the field (for example, filters, filtrates), and quality control (QC) samples. Location identifier naming procedures should also be prescribed in data management plans when locations are revisited or if there is the possibility that sampling locations will be revisited. For more information on developing guidance on these naming conventions in a data management plan or governance document, see the Data Governance and Data Lifecycle fact sheets. Best practices for the collection and management of geospatial data can be found in the Geospatial Data Collection Consistency subtopic sheet.

In addition to sample naming, the relationships between samples need to be documented. Data users will need to know, for example, that a sample was split in the field into two subsamples, or that multiple individual samples were composited in the field, or that two soil samples are from the same boring, and so forth. This is often done by using fields within sample identifiers to contain linking information, but in addition to this it is preferable to also record the relationship information in metadata collected in the field and with additional identifiers in the database. A soil boring, for example, would be assigned a unique boring identifier, and every soil horizon from that boring would have not only its own unique sample identifier but also (as metadata) the associated soil boring identifier. Additional information on metadata is included below in Section 3.4.

Location identification is essential in the field data capture process for sensor data as well. When a handheld sensor internally records the location (using GPS) and date/time of a measurement, then location identifiers are of secondary importance. However, proper location identifier assignment ensures that field measurements made with manually recorded data and metadata can be traced to their point of origin.

3 FIELD DATA CATEGORIES

The text and tables below describe a wide range of field data categories. They are not intended to be all inclusive but do provide a range of categories and examples. The data categories are grouped by the volume and kind of data they generate,

that is, low-volume, high-volume, and unstructured data. Along with any measurement in the field, the metadata should also be recorded.

- **Low-volume data** are discrete field or lab data collected at a single point in time during a field sampling event, or analysis of a field sample in the lab. The amount of data is low compared to ongoing monitoring done by sensors collecting data every 5–10 minutes.
- **High-volume data** are those data generated in short time intervals for continuous monitoring, thus producing a large data set of frequent measurements. High-volume data is stored on the sensor device and may be transmitted in real time to an environmental data management system (EDMS) using telemetry to provide database users near real-time data results and alerts. Otherwise, sensor data can be downloaded from sensor devices periodically and incorporated into an EDMS.
- **Unstructured data** are not organized in a structured database schema table. Examples include files that can be organized in a file folder structure, such as video or audio recording files, photograph files, and handwritten scanned documents or PDF files.
- **Metadata** are the who, what, where, when, and how low-volume, high-volume, or unstructured data are collected and recorded.

Some unstructured data can be transcribed in various ways into structured data; for example, information recorded on a hard copy field data sheet form is transcribed into a spreadsheet, or numbers and species of trees are counted on aerial photographs and transcribed into a data table. Considerations for migration of unstructured data to a structured EDMS are discussed in the Data Migration Best Practices Fact Sheet.

Tables 1, 2, and 3 provide examples of low-volume, high-volume, and unstructured data, respectively. Due to the many potential sources of data, a discussion regarding data sources is beyond the scope of this document. Figure 1 shows the general flow of field data.

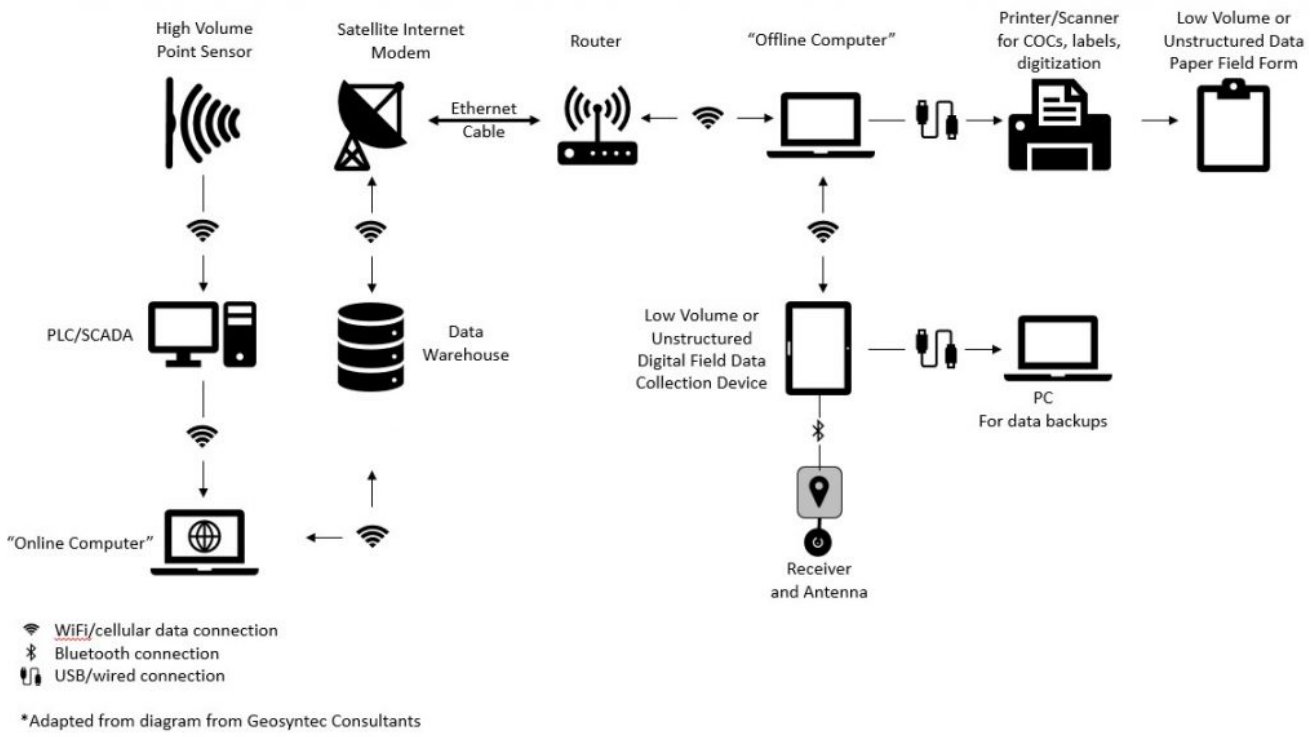


Figure 1. Example flow of field data.
Source: Adapted from Geosyntec Consultants, Undated. Used with permission.

3.1 Low-Volume Data

3.1.1 Field Samples and Field Observations

When field samples are collected for later analysis (for example, laboratory samples), the field team members must capture information about the sample collection process. This includes a unique sample identifier (see Section 2 above), date/time,

location, collection methods, collector, and preservation method on the sample chain of custody for each sample collected. Field observations can vary widely and may take the form of soil boing logs (lithology, color, odor, moisture), well construction logs (screen interval, filter pack, construction materials), water quality logs (temperature, pH, dissolved oxygen, turbidity), etc. Sample information and field observations may be recorded data by hand or digitally but should be sufficient to ensure that field data capture provides complete and consistent information. The importance of a well-designed field form cannot be overstressed in this context, and this topic is discussed at length in Field Data Collection Process Development Considerations.

3.1.1.1 Chain of Custody (COC) Data for Discrete Samples

COC procedures provide a record of physical sample custody to track samples from the time of collection to delivery at a destination laboratory. COC documentation is the record that a sample was collected, handled according to method requirements (for example, preservation), shipped to a laboratory, and directly linked to the analytical results. Sample tracking can be simplified and expedited using bar-coded labels. A separate COC form of one or more pages may be prepared for each shipping container. When multiple shipping containers are used, it can be helpful to have the shipping containers numbered (for example, Cooler #1) and the number of bottles in each container listed on the COC. COCs are often designed and controlled by the laboratory. However, not all laboratories require COCs, and different laboratories may have different nomenclature for submittal. Communicate with laboratories to ensure that the data captured are sufficient. COCs can be completed digitally or as paper documents as shown in Figure 2. The following data may be included for each sample in the shipping container:



Figure 2. Field technician completing a paper chain of custody form.

- sample identifier
- date/time of collection
- number and type of containers
- preservation
- requested analyses, which should be specified using controlled vocabulary valid values to avoid miscommunication
- sample turn-around time requirements
- signatures/dates of sample collectors and individuals involved in sample transfer

3.1.2 Point Sensors

Point sensors from various manufacturers measure a plethora of parameters at a single point in space and in media such as air, water, pore water, confined spaces, and process liquid streams. One example is shown in Figure 3. Low-volume data sensor configurations are typically handheld devices measuring a single or limited number of parameters at one moment in time. With handheld devices the sensor readings are digitally displayed for the user to record on paper or to record in a mobile digital electronic form, or they are recorded in the device memory for later download into a structured electronic table or file.



Figure 3. Field technician using low-volume point sensor.

3.1.2.1 Point Sensor Calibration and Calibration Data Handling

The instrument sensors should be maintained and calibrated as specified by the manufacturer. Best practice suggests that sensor calibration also be checked prior to going into the field using manufacturer recommended calibration procedures to verify accuracy. It is also recommended that the calibration fluids be taken to the field as well in case the instrument falls out of calibration during the course of the day. The calibration data for each instrument/sensor should be maintained in the project data archive along with the collected sensor data and metadata. During post-field data analysis, it is ordinarily necessary to use instrument calibration data to convert from instrument values/units to final engineering values/units.

3.1.2.2 Data Unpacking, Processing, and Archiving

When sensor data from low-volume sensors are recorded manually on log forms or in a field notebook, the data are often transcribed into electronic form. After transcription, QA/QC procedures should be applied to ensure no transcription errors occur. For more guidance on QA/QC of field data, see the Field Data Collection Quality Assurance and Quality Control (QA/QC) Fact Sheet.

When sensor data from low-volume sensors are recorded electronically in instrument memory, data are downloaded from the instrument(s) using communications software provided by or recommended by the manufacturer. These sensor data may include metadata, measurements, and calibration data. Even if the data are copied into a database, the original files should be preserved in their native electronic format and should never be renamed to serve as a record of where the data originated. Files transferred from field recording devices should be copied to the directory where project files are maintained, typically a network server folder(s). Here the data files, metadata files, and sensor calibration data files will be used for data processing and will remain for archiving. These folders should be organized with future discovery and access in mind, in a structure that maintains context in the folder path names. (See Figure 4 for an example folder structure.) Complete records of the provenance of these files should be maintained.

All these practices are suggestions, and your individual process should apply these suggestions as applicable to meet your organization's existing document and data archival policies and regulations. Additional information on this topic is available in the Data Storage, Documentation, and Discovery Fact Sheet.

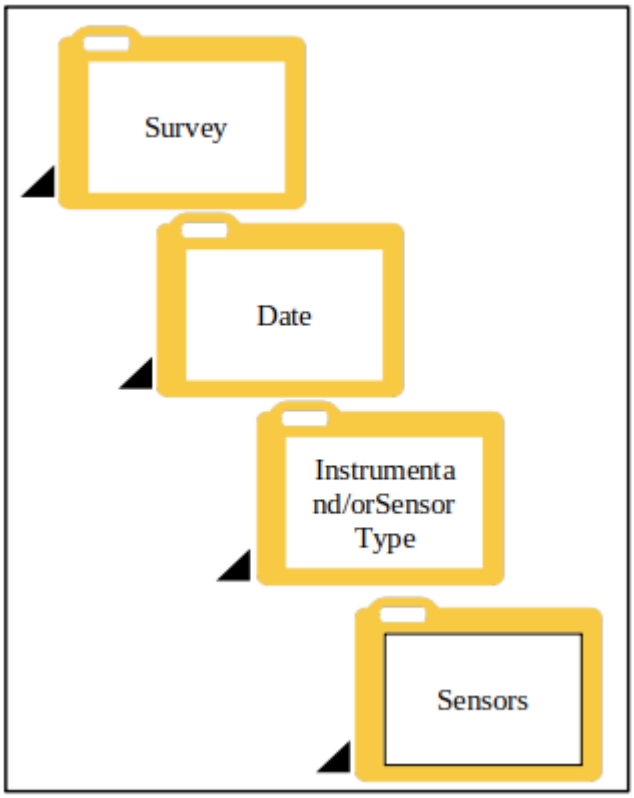


Figure 4. Example folder structure that maintains context.

Table 1. Examples of low-volume data

Category	Description	Spatial/ Temporal Extent	Method	Data	QA/QC
Field Samples & Field Observations					
Field Samples Collected for Further Analysis (For Example, Laboratory Samples) as Recorded on Sample Chain of Custody Forms	Air, water, sediment, soil, rock, tissue collected for chemical analysis, for sample dating, for species abundance, toxicity studies, etc. Examples: Air samples for polychlorinated biphenyls (PCB) analysis; tree ring cores for dating; benthic; sediment for invertebrate abundance; lithology	Single point space and time	Physical collection (for example, box core, slurp gun, air puff)	The data will be generated in the lab. Whether that is 1,000 miles away or in a trailer at the job site, the data are not generated at the point of sampling.	Chain of custody, QC sampling (including field replicates, field blanks, trip blanks, etc., as appropriate)

Subjective Observations	Biological, habitat Examples: vegetation counts, bio index	A sample area, for example, a river reach or habitat grid	Observation in the field	Recorded observations	Replicated observations and double-entry or read-back of recorded (audio, video, or other) observationsVerification against maps or other information sources
Operation, Maintenance, and Monitoring	Data relating to equipment used for environmental remediation, which may include equipment temperature, pressure, flow, power consumption, discharges, etc.	Single point measurement(s) over time	Programmable logic controllers (PLC), supervisory control and data acquisition (SCADA), or manual readings	Recorded observations, parameter measurements	Equipment calibrations
Point Sensors					
Water Sensors	Sensors that measure physical, chemical, or biological parameters in water. Examples: temperature, turbidity, conductivity, biological oxygen demand (BOD), pH, dissolved oxygen, chlorophyll	Single measurement at a single point in space and time	In situ point sensor measurement or measurement of collected water samples; typically handheld devices	Parameter measurements	Sensor calibration
Air Sensors	Sensors that measure chemical parameters in air or other gas environments Examples: O_2, CO, H_2S, VOCs, etc.	Single measurement at a single point in space and time	In situ point sensor measurement; typically handheld devices	Parameter measurements	Sensor calibration
Other In Situ Sensors	Examples: membrane interface probe, laser-induced fluorescence probe, etc.	Single measurement at a single point in space and time	In situ point sensor measurement	Parameter measurements	Sensor calibration

3.2 High-Volume Data

High-volume sensor data require the same information previously described for sample and location identifiers, low-volume point sensor calibration data, and data unpacking/storage. Data from high-volume sensors is often handled differently, however, due to its nature. High-volume sensor data may be recorded in instrument memory, as with low-volume sensors, or it may be transmitted in real time or near real-time to servers elsewhere using telemetry. In this case, data unpacking may be done manually as described previously for low-volume data, or software may be created/used to process the data in real or near real-time. In the case of real or near real-time data processing, software quality control systems should be in place to ensure software created to unpack and process transmitted data in real or near real-time are faultless and fault-tolerant. This includes software for unpacking data, deriving final measurement values by applying instrument calibrations or corrections for instrument drift, or any other real-time data processing needed. The final archive of data files should follow a similar project folder structure described above to maintain the data files in their original context.

Table 2. Examples of high-volume data

Category	Description	Spatial/ Temporal Extent	Method	Data	QA/QC
Point Sensor Measurements					
Water Sensors	Sensors that measure physical, chemical, or biological parameters in water Examples: temperature, turbidity, conductivity, BOD, pH, dissolved oxygen, chlorophyll	Repeating measurements at a fixed location over time or repeating measurements while the sensor is moved in space	In situ point sensor measurement. Examples: sensor fixed on a mooring; conductivity, temperature, and depth (CTD) probe lowered thru the water column	Parameter measurements	Sensor calibration and drift corrections
Air Sensors	Sensors that measure chemical parameters in air or other gas environments Examples: O_2, CO, H_2S, VOCs, etc.	Repeating measurements at a fixed location over time	In situ point sensor measurement. Examples: Fixed air sampling station	Parameter measurements	Sensor calibration and drift corrections
Meteorological Sensors	Sensors that measure physical meteorological parameters Examples: wind direction, wind speed, VOC atmospheric pressure, air temperature	Repeating measurements at a fixed location over time or repeating measurements while the sensor is moved in space	In situ point sensor measurement. Examples: Fixed weather station, weather balloon	Parameter measurements	Sensor calibration
Remote Sensing					

Air and Water Movement Sensors	Acoustic sensors used to measure water current velocities or wind speed profiling over a depth or height range. Example: acoustic Doppler current profiler (ADCP), sonic detection and ranging (SODAR)	Repeating measurements of a profile at a fixed location over time or repeating measurements of a profile while the sensor is moved in space	Acoustic remote sensing measurement	Parameter measurements	Sensor calibration
Bathymetric Sensors	Acoustic sensors used to measure distance to the seafloor, lake, or riverbed Examples: Echo sounder, multibeam, side scan sonar	Repeating profile acoustic measurements while the sensor is moved in space	Single acoustic beam or multi acoustic beam measurements of distance to seafloor	Time, location, elevation, distance to seafloor	Sensor calibration
Water Level Sensors	Includes transducers that measure pressure in groundwater monitoring wells, corrected for barometric pressure, to determine groundwater depths used to calculate groundwater elevations	Repeating measurements at a fixed location over time or repeating measurements while the sensor is moved in space	In situ point sensor measurement	Parameter measurements	Sensor calibration

3.3 Unstructured Data

Any relevant unstructured field data should be captured appropriately into the study data management system. Field team members should be prepared to manage the following unstructured data and their metadata and to archive them as appropriate:

- photo documentation
- video documentation
- sound recordings
- handwritten documentation

Hard copy handwritten documentation should be maintained with project documents in the organization’s project file folder structure, but these records contain information that is less usable to the project team and may become more and more difficult to find over time. Good practice is to uniquely number each loose form sheet and each field notebook or logbook. Scanning hard copy documentation and archiving the files in network server folders with chronological file names is recommended to help protect against loss of project data. Data files (for example, spreadsheets) created in the field should be named in such a way as to ensure their uniqueness. When scanning/photographing any field form, ensure that the digital version is readable.

The following unstructured data should be captured and maintained electronically in the project folder file structure defined above.

- field logs including observations (that is, PDF copies of handwritten field logs)
- field measurement data sheets (that is, PDF copies of handwritten data sheets)
- instrument calibration sheets (that is, PDF copies of handwritten calibration sheets)
- COC forms (that is, PDF copies of handwritten COC sheets)

The following information describes processing and archival considerations for unstructured data. As with sensor data files, unstructured data files (including photographs, videos, PDF log forms, and PDF COC forms) should never be renamed and should be preserved in their native electronic format. Files are typically transferred from field recording devices or hard copy scanners to network server folders where they should be organized in file folders whose structure maintains type and context in the folder path names. Complete records of the provenance of these files should be maintained.

Table 3. Examples of unstructured data

Category	Description	Spatial/ Temporal Extent	Method	Data	QA/QC
Photographs	Photographs taken in the field. Examples: field operations documentation, field samples, habitat surveys	Single point in space and time	Photograph	Photographic image files	Verify accuracy of captions/metadata
Video	Video recorded in the field. Examples: habitat surveys	Single point in space over time	Video	Video files	Verify accuracy of captions/metadata
Scanned Field Notes (PDF or JPG)	Notes that are written in field logbooks or completed data sheet field forms that are scanned with a camera, creating a JPG, or using a tool that results in a PDF	Single point in space over time	Scanning	PDF Files	Verify images of scanned field notes are legible and complete

3.4 Metadata

Metadata are the additional pieces of information that are critical and should be collected to make the data usable now and in the future as it pertains to the larger data set. Often these metadata, such as location, are recorded automatically by instruments, sensors, or photography/videography equipment. Where these metadata are not recorded automatically, simple methods such as writing the location identifier on a small whiteboard and including it in the photo frame can help ensure that the correct metadata are captured and remain associated with the data. It is also important to establish valid values for metadata where possible.

Table 4 provides examples of the metadata associated with low-volume, high-volume, and unstructured data. For information specific to geographic metadata, see the Geospatial Metadata subtopic sheet.

Table 4. Examples of metadata

Category	Metadata
Low-Volume Data	
Field Samples & Field Observations	

Field Samples Collected for Further Analysis (For Example Laboratory Samples) as Recorded on Sample Chain of Custody forms	Date/time (including time zone, if appropriate) Location Collector Collection equipment Collection method Preservation method Unique sample identifier Sub-sample identifier (where splits or aliquots are created in the field) Composite type and scope (where composite samples are created in the field) Sample material (matrix) Sample type (for example, normal field sample or quality control sample) Number of sample containers
Subjective Observations and Operation, Maintenance, and Monitoring	Date/Time (including time zone, if appropriate) Location Collector Collection equipment Collection method
Point Sensors	
Water, Air, and Other In Situ Sensors	Date/time (including time zone, if appropriate) Location Collector Sensor/instrument serial number Sensor/instrument settings Sensor/instrument serial number
High-Volume Data	
Point Sensors	
Water and Air Sensors	Date/time (deployed and recovered) (including time zone, if appropriate) Location (fixed or moving) Collector Sensor/instrument model Sensor/instrument serial number Sensor/instrument settings
Meteorological Sensors	Date/time (deployed and recovered) (including time zone, if appropriate) Location (fixed or moving) Collector Sensor/instrument model Sensor/instrument serial number Sensor/instrument settings
Remote Sensors	
Air and Water Movement Sensors	Date/Time (deployed and recovered) (including time zone, if appropriate) Location (fixed or moving) Collector Sensor/instrument model Sensor/instrument serial number Sensor/instrument settings
Bathymetric Sensors	Date/Time (deployed and recovered) (including time zone, if appropriate) Collector Instrument serial number

Water Level Sensors	Date/time (including time zone, if appropriate) Location Collector Sensor/instrument model Sensor/instrument serial number Sensor/instrument settings
Unstructured Data	
Photographs	Asset file name File type Date/time (including time zone, if appropriate) Photographer Location Direction Photographer/videographer
Video	
Scanned Field Notes (PDF or JPG)	Asset file name File type

4 REFERENCES AND ACRONYMS

The references cited in this fact sheet, and the other ITRC EDM Best Practices fact sheets, are included in one combined list that is available on the ITRC web site. The combined acronyms list is also available on the ITRC web site.