An Integrated List of Data Elements: Unifying Concepts in Action

Revital Katznelson
University of California Extension, Berkeley (Instructor)
revitalk@sbcglobal.net

Katznelson, R. 2010. An Integrated List of Data Elements: Unifying Concepts in Action. In: Proceedings of the 7th National Water Quality Monitoring Conference, April 26-29, 2010, Denver, CO.

ABSTRACT

Data Elements are placeholders for bits of information that describe environmental monitoring results. In many ways, data elements are equivalent to the “data fields”, or column headers, in a database table. The people who generate, verify, validate, and manage monitoring data (at the Project level) need a comprehensive array of data elements (i.e., a “long list”) to organize and document their datasets. This paper presents a newly revised list of data elements which addresses that need. The author’s original comprehensive list has been molded into the NWQMC's Water Quality Data Elements (NWQMC Tech. Report no. 3) to create this new list and to address the following challenges:

(a) Integration of data elements from different areas of inquiry: The content and types of information needed for environmental monitoring projects vary widely, and many data elements are specific to a certain area of inquiry. However, monitoring activities in all major areas of inquiry share several groups of common elements, or “unifying concepts”. These unifying concepts are the foundation of the revised list, and will facilitate integration of monitoring and assessment data from various areas of inquiry, including chemical analyses, discrete and continuous field measurements, bacterial counts, toxicity testing, biological assessments such as fish, benthic macroinvertebrate, or periphyton assemblages, and physical habitat assessments.

(b) Independence of database structure and data models: the list is built by subject matter, i.e., by “who, what, why, when, where, and how”. In other words, data elements are categorized, grouped, and placed in the list based on what each of them describes.

(c) Comprehensiveness with flexibility: this improved list provides an extremely detailed foundation that can be easily tailored to any Project by culling irrelevant elements or by adding elements into the right place in the list if additional detail is needed.

(d) Standardization of contents: for consistency and comparability among projects, the revised list of data elements is already associated with an array of pick-lists (a.k.a. Lookup tables), that are also expandable to maintain inclusiveness and flexibility.

(e) Data sharing and comparability: Only a sub-set of the data elements in this list are necessary for data sharing beyond the Project, and these can be easily extracted to form a “short list” of core data elements that accompany the monitoring results as they are uploaded into central data management systems.

Key words: Data elements, areas of inquiry, data management, unifying concepts

1. INTRODUCTION

This list of data elements began its life in the late 1990s in response to a need to communicate. While developing a basic spreadsheet for field measurement results, the author kept adding placeholders for more and more 'bits of information' that she – and her colleagues who might use these data - would want to know about the results. Organizing the information bits by subject-matter made it easier to perceive the contents without the complications of data structures. The list, which was focused on field measurements at first, grew and expanded to include water chemistry and toxicity. By 2001, it contained the majority of core Water Quality Data Elements (WQDE) recommended by the Methods and Data Comparability Board (Methods Board).

The first installment of WQDE, for chemistry and microbiology, was assembled by the WQDE workgroup of the Methods Board and approved in 2001. The WQDE workgroup then created two supplemental lists of data elements, one for toxicity and the other for population and community level analyses (NWQMC 2006). The author joined the WQDE workgroup in 2004 and participated in the efforts to develop the third supplement, a list of elements for physical habitat assessments. However, the separate lists created for various areas of inquiry have not been integrated.

Integration of data elements can facilitate integration of monitoring and assessment data from various areas of inquiry, including chemical analyses, discrete or continuous field measurements, bacterial counts, toxicity testing, biological assessments such as fish, benthic macroinvertebrate, or periphyton assemblages, and physical habitat assessments. One way to integrate the lists is based on the concepts that unify various areas of inquiry, i.e., entities such as a measurement system that are common to all areas of inquiry.

This paper describes a new, integrated list of data elements that makes use of many of these unifying concepts. It is hoped that the integrated data element list will facilitate the integration of data from diverse areas of inquiry, enhance data comparability, and enable streamlined data exchange.

2. PURPOSE OF THE INTEGRATED LIST OF DATA ELEMENTS

The integrated data elements list is designed for use at the monitoring Project level, by Project operators (a monitoring Project, as referred to in this paper, is a "data collection effort performed by one or more Organizations and limited to defined space and time"). The list is meant to be tailored to the needs of each Project; it provides the initial 'catch-all' basis from which the irrelevant data elements should be mercilessly deleted, and can be extended to accommodate Project-specific or missing elements as needed. It must be emphasized that this list is not a Standard, and that it needs to be augmented by the Data Standards that are appropriate for the Project. The reader is referred to the latest publications of the Environmental Data Standards Council (EDSC) for details (EDSC 2006).

The integrated data elements list, as augmented and tailored for a specific Project, can be used in a number of ways. In the absence of information technology (IT) support, this list can be easily converted to a simple data quality management system, e.g., an array of spreadsheets, in which all the information relevant to the monitoring Project can be captured, processed, and stored. If IT support is available, this list can provide a basis for a variety of automated data transfer protocols (e.g., for sensors' data-logger files with their associated calibration records and quality checks). The integrated data element list, as organized by subject-matter, can also provide an understanding of the information function of each element; this can be very instructive to the data structure architects involved in the creation of data models.

3. CONTENTS OF THE INTEGRATED DATA ELEMENTS LIST

The original list created by the author began with the data itself (i.e., the "What", or the results) and then went on to their descriptors. However, this order changed when the list was merged with WQDE and the current order conforms to the "who, what, why, when, where, and how" format. Thus, the list starts with WQDE "Contacts" module which describes the Project and its organizations, continues with the "Result" module, and then proceeds to the WQDE "Reason" module which describes the intent and the design of each study-dataset. The fourth module (WQDE "Date/Time") provides the Site-Visit information, and this flows naturally into the fifth module, "Location". The sixth WQDE module "Sample Collection" has been expanded to describe all manner of Field Activities (where the word "activity" is used in the STORET sense), and the seventh WQDE module "Sample Analysis" is now the group of elements describing the Measurement System with all its actions to affect and to check the quality of the Results.

Table 1 shows a 'skeleton' outline of the integrated data elements list. The seven new Categories essentially correspond to the WQDE Modules (shown in parentheses). Each Category has a number of data element Groups. The following paragraphs provide further information about each Category.

3.1 The Monitoring Project and its Organizations

As mentioned above, a monitoring Project is a data collection effort that is limited in space and time (e.g., routine monitoring of one watershed over one year, or a special study to identify the source of a particular constituent, or a probability-based survey of stream health in a given area over one season). The "Who", i.e., each of the folks involved in the data collection effort, belongs to one of the Project's Organizations. A Project can be done by several Organizations (e.g., the Agency, the Laboratory, the citizen monitoring group, the Contractor(s), or the academic institution). Each Organization, no matter what its function is within the Project, has its contact person with their contact information; this is a unifying concept. Thus, Category 1 is where the placeholders for all these bits of information (Organization ID, type, and function; contact person information; etc,) is contained. It also contains placeholders for linking entities, such as Organization ID, that also appear in other Categories in the list.

3.2 The Result

Figure 1 shows the organization of the data elements associated with the Results; it is a four-level hierarchical structure that spreads from the Category, through the Group and the Subject, into the individual data element. The structure shown in Figure 1 is common to all seven categories in the integrated list.

All monitoring activities are done to obtain Monitoring Results; this is the primary unifying concept, and it is always associated with the “What” (the characteristic). Most results also require a data element for “units” and entities that link them to the measurement system’s batch they were derived from. Placeholders for the accuracy and precision of each Result are a bonus. Result descriptors also include data elements that categorize each Result (i.e., Result type and –for a subset of these - Endpoint type); being able to categorize a result allows for including results of different types in any data management system. **Table 2** shows examples of Result types and endpoint types.

3.3 The Reason for Monitoring: Descriptors of the Study-Dataset

A Study-Dataset is a batch of Results that were collected in order to answer a specific study question and therefore share the same intent and design attributes. Selection of a sampling design principle (e.g., directed/ targeted, systematic, or random) is a part of any planning process; this is a unifying concept and a bit of information that is easy to share. The Study Intent data fields share the reason why monitoring was done at a given location and at a given time. Combined with information about Station type and sampling conditions, intent and design information provide “handles” for retrieving, sorting, pooling, and filtering data from a central database. **Table 3** shows examples of Pick Lists for spatial intent (why there) and temporal intent (why then), as well as the three fundamental sampling design principles plus other, non-deliberate sampling design options.

3.4 The Site Visit: Date and Time

A Station Visit represents a unique combination of time and space, and can generate numerous results for a variety of characteristics that are ‘packaged’ together and support each other. Thus, it is a natural unit of representativeness, and it becomes the natural unit of counting when considering the power of a dataset.

3.5 The Location

As the name of this category indicates, this group of data elements describes the “where”, i.e., the location of the monitoring Station.

3.6 The Field Activity

A Field Activity, a.k.a., Sample/Measurement/Observation, can be an action or an object. **Table 4** presents a list of Field Activities organized in categories and types. Whether the Result is a field measurement of dissolved oxygen or of a substrate particle size, the Data Elements describing the Activity describe the spatial sampling frame and communicate even the spatial component where the activity was done (for example, substrate particle size was measured at the third transect point of Transect F). Another important group of Activity descriptors is dedicated to the Sample, i.e., the chunk of medium that has been transferred into a container for analysis elsewhere.

3.7 The Measurement System

Measurement systems are devices and/or procedures used for quantitation or evaluation of environmental characteristics, including instruments used for field measurements, sampling & analysis processes, physical habitat assessments, and biological assessments. Many descriptors of the measurement system are common to all areas of inquiry, and constitute the bulk of Category 7. Data quality is an integral part of the measurement system, as are all the actions to affect and to check the quality of the data. **Table 5** shows bits of information that are relevant to calibrations and accuracy checks of water quality sensors. Table 5 also provides a good example of re-organization of information, from use of individual characteristics and repetitive features as data elements to a structure in which where a concept that unifies these characteristics becomes the data element (i.e., the data field) while the characteristics, which now are cell contents, become choices in a Pick-List.

Quality checks come in a variety of types even within one area of inquiry. They can be easily sorted out by category and type, as shown in **Table 6**. This organization allows for inclusion of quality checks from different areas of inquiry in the same category or type of check. Combined with a set of unifying data elements such as “value of the first observation”, quality checks details and outcomes can be summarized on the same page **(Table 7**).

4. DISCUSSION

4.1 Data Elements and their Pick-Lists

The new list differs from the Methods Board’s original WQDE list in two important structural aspects: (a) the data elements in the new list are, essentially, the column headers of a database table; in other words they are equivalent to Data Fields; (b) the possible contents of the database table cells under many of these Data Field are provided separately, as Pick Lists. Every data element and pick-list item are “database ready” because their status (a data field or cell content) is clearly defined. However, they are totally independent of any database structure, and could serve as the building blocks of any data models.

Pick lists are important for data standardization and, in many cases, are essential for data comparability. The Methods Board's original WQDE (NWQMC 2006) had offered an excellent start, and the new lists described in this paper expand it.

4.2 The "Long List" versus the Core Data Elements

Another fundamental difference between the new integrated list and the Methods Board's original WQDE list lies in their target audiences: the new list is created for the data generators at the Project level, while the Methods Board's WQDE list has been 'customized' for the data user who mines the data from a central database. Thus, folks at the Project level use a "long list" with all the necessary detail for local processing and management of their data, while the data users the core data elements of the Methods Board's WQDE lists (i.e., only the bits of information that they want to have in order to use the data.

4.3 Key Entities Described by the Data Elements

The entire Data Elements List, with its 300-plus data elements, is designed to describe one Result. Indeed, many of the data elements call for contents that are specific to an individual Result, e.g., the detection limit of DDT, part of the organochlorine analytical suite, in a specific sample. However, many descriptors are shared by a large number of Results because, in order to describe one Result, a person has to describe the laboratory batch it came out of, or even the Study Dataset it belongs to. Each Result belongs to a given Sample and a given Dataset, and each Sample belongs to a given Dataset. A Station can serve different Study Datasets, and so can a Sample and even a Result. There may be several Study Datasets, or lines of inquiry, in one monitoring Project. In other words, one surface temperature Result, e.g., in the center of a lake, can be included in three Study Datasets: a vertical profile ("is the lake stratified?"), a spatial survey ("where is the hottest spot in the lake at 2 PM?"), and an impact assessment study ("is there temperature gradient moving away from the hydroelectric plant outfall?").

4.4 The Power of Categorization and Typing

Organizations, Results, Activities, Quality Checks, and many other entities that are common to all manner of environmental monitoring are unifying concepts. The efforts of categorizing and typing the variants contained in each of these entities has yielded matrices that clearly show how variants from different areas of inquiries belong together and can be placed on the same page in terms of information contents. These matrices can also provide the Pick-Lists for descriptors of these important entities.

4.5. Next Steps

The integrated data elements list is a living document, and it is meant to be updated and honed as new data comparability issues are encountered, as expected when we merge data from different areas of inquiry. The hierarchical structure of the list can easily be expanded to include new elements where they belong, without a disruption to its 'skeleton'. It is a modest start at this point, but it is hoped that the fundamental structure of the list will remain functional as the list grows over time.

Acknowledgements

This paper could not happen without the input and feedback provided by all the friends and colleagues who reviewed the list through the years; thank you all. Special thanks are due to LeAnne Astin, Dan Sullivan, Faith Fitzpatrick, and the other colleagues from the Methods Board's WQDE workgroup. Thanks go to Chuck Spooner of EPA's Office of Water for keeping me connected to the Data Standards efforts across the Nation in all their incarnations through the years. I also wish to thank Gayle Rominger of YSI for supporting the Sensors QA Initiative work that gave me the incentive to (finally) finalize and publish this work.

Disclaimer

The content of this paper does not necessarily reflect the views and policies of the Methods Board or the National Water Quality Monitoring Council, nor does mention of commercial products constitute endorsement or recommendation for use.

References

Environmental Data Standards Council (EDSC) 2006. A collection of Data Standards (including ESAR and others). www.envdatastandards.net

Katznelson, R. (2002). Letting Monitoring Data Speak for Themselves. in: Proceedings of the third National Monitoring Conference of the National Water Quality Monitoring Council, "Building a Framework for the Future", May 19-23, 2002, Madison, WI.

Katznelson, R. (2008). Unifying concepts in environmental monitoring (Abstract). in: Proceedings of the sixth National Monitoring Conference of the National Water Quality Monitoring Council, "Monitoring: Key to Understanding our Waters", May 18-22, 2008, Atlantic City, NJ.

National Water Quality Monitoring Council (NWQMC) 2006. Water Quality Data Elements: a User Guide. Tech. Report no. 3.
http://acwi.gov/methods/pubs/wdqe_pubs/wqde_trno3.pdf

Table 1: Outline of the new Integrated Data Elements List

Category number	Category name (& WQDE module)	Group number	Group name	Information contents
1	Monitoring Project ("Contact")	1	Project Identifiers	
		2	Organization identifiers	ID, type, function
		3	Project Contacts	
2	Result ("Result")	1	Result (how much?)	
		2	Characteristic (of what?)	Analyte name, CAS#, taxon, life-stage
		3	Result descriptors	Type, qualifiers, MDL, accuracy, precision
3	Study-Dataset ("Reason")	1	Identifier	ID, scenario
		2	Intent (Why?)	why here, why now
		3	Design	Sampling design principle
		4	Power (to see change)	# of Station-Visits, frequency, intervals
		5	Representativeness	Compositing principle
4	Site Visit ("Date/time")	1	Trip	ID, crew, dates
		2	Station-Visit	ID, date, time, conditions
5	Location ("Location")	1	Site identity	IDs, Gauge #, monument IDs
		2	Site descriptors	Station type, drainage
		3	Watershed/aquifer	names
		4	Site Location	county, city, directions
		5	Geospatial information	lat/long, elevation, benchmark
6	Activity-Field ("sample collection")	1	Activity descriptors	ID, type, operator, SOP
		2	Activity spatial component Location	Spatial sampling frame, components in frame
		3	Activity - how performed?	fixtures, access, devices
		4	Activity Records	scribe, logger files, COCs
		5	Sample *(sensu strictu)* Description	medium, assemblage, preservation, shipping
7	Measurement System ("Sample Analysis")	1	Methodology	Name, function, citation
		2	Evaluative Methods	categorization protocols
		3	Instruments (Field)	type, description
		4	Field 'Tailgate' analyses	reagent kits, ELISA
		5	Laboratory	ID, type, certifications
		6	Analytical chemistry	preparation, batch, date
		7	Biota Counts, Measurements, and Tests	taxonomy, biota & microbial counts, toxicity,
		8	Physical assessments	habitat, geomorphology
		9	Actions to Affect Data Quality	training, calibration, standards,
		10	Actions to Check Quality	quality check type, batch, expected/observed values, standards & voucher collections, QC outcomes

Table 2: Monitoring Result Types and Endpoint Types

Result Type	Endpoint Type	Examples
(txt) **Verbal category**	Not applicable	water murkiness=murky; flow conditions=isolated pools
(txt) **Numeric range category**	Not applicable	upper canopy cover=10-40%
(num) **Individual value**	Not applicable	Water Temperature=18 C
Estimated number or numeric range	Not applicable	Flow discharge is about 3 cfs
Count	Not applicable	
Score (individual characteristic)	Not applicable	
Derived Endpoint (for **1 data point**, or sample, or test)	Simple Arithmetic Endpoint	survival (in a toxicity test)=90%; fecal coliforms = 325 cfu/100 mL; % EPT=31; Q=5 cubic ft/sec;
	Qualtile of Concentration or size	Cdub48hLC50=0.4 ug/L Diazinon; Ppro7dEC15=33% sample; Pebbles d50=46 mm
	Effect category	NOEC: Endpoint of a toxicity test: No Observed Effect Concentration
	Most Probable Number (MPN)	fecal coliforms = 234 MPN/100 mL
	Complex Endpoint	TUa=3.2 ; Tolerance-related BMI Metrics;
Aggregated Endpoint (for **1 data point**, or sample, or spatial entity)	Simple aggregate	percent gravel in Reach = 32%
	Compound Endpoint	mean PEC toxicity quotient; Index of Biological Integrity (IBI)
	Aggregated score	habitat value=18 (of 25)
	Rating	
	Grade	
	(more)	
Descriptive statistic (for **many data points**)	arithmetic mean (avg)	mean global temperature for June 1st 2009=12.4 C
	median	median copper conc.=13 ug/L
	minimum	
	maximum	
	geometric mean of five consecutive weeks	
	moving weekly average	
	(more)	
Rank (for many data points)		third in the Nation!

Table 3: Examples of Pick-Lists for Intent and Design of the Study-Dataset

Subject	Field Name	Pick-List Item Name	Pick-List Item Definition
3.2.1 Spatial intent			
	3.2.1.1 Station Selection Intent		
	3.2.1.1.1	fixed Station for long term monitoring	Monitoring at the same spot each time to create a long-term record of conditions
	3.2.1.1.2	permit compliance monitoring	Monitoring for the purpose of comparison with water quality benchmark specified in a discharge permit to check if that discharge is in compliance
	3.2.1.1.3	source Identification	Identifying the source of a given constituent within a river network or land use activities
	3.2.1.1.4	characterization of refuge areas	Identifying and characterizing habitat areas having the best-case-scenario in term of extreme conditions; i.e., the least impacted habitats in a reach
	3.2.1.1.5	impact assessment	Monitoring to determine whether an impact to a given ecosystem has occurred; often involves selection of stations upstream and downstream of the disturbance
		[[more]]	
3.2.2 Temporal Intent			
	3.2.2.1 Sample Timing Intent		
	3.2.2.1.1	routine monitoring	Repeated monitoring at fixed time intervals to provide long-term data
	3.2.2.1.2	snapshot	One-time monitoring of multiple Stations
	3.2.2.1.3	dry weather discharge	Monitoring during dry weather to characterize non-storm flow
	3.2.2.1.4	storm runoff monitoring	Monitoring storm runoff events at different water levels and phases during the event
	3.2.2.1.5	worst case scenario	Monitoring during the times anticipated to represent the most critical or the most extreme conditions within the natural fluctuations.
	3.2.2.1.6	identification of specific conditions	Monitoring during the times anticipated to represent certain conditions in a waterbody, such as summer stratification.
		[[more]]	
3.3.1 Spatial Design			
	3.3.1.1 Reach Selection Design		
	3.3.1.1.1	systematic	Deterministic approach, points selected deliberately at fixed-intervals of area, length, or time
	3.3.1.1.2	directed (to environment)	Deterministic approach, points selected deliberately based on knowledge of their attributes of interest as related to the environment monitored; also known as "targeted", "judgmental", "authoritative", "knowledge-based" etc.
	3.3.1.1.3	stratified random	Probabilistic approach, deliberate, points selected at random from a population stratified by specific attributes
	3.3.1.1.4	(deliberate to operations)	Deterministic approach, points selected deliberately based on operational requirements or logistical constraints
	3.3.1.1.5	non-deliberate (anecdotal)	Non-of-the-above, non-deliberate; points selected causally or whenever/wherever, or by given constraints, or opportunistically

Table 4: Categories and Types of Field Activities

Activity Category (see Notes below)	Activity Type	Activity Type Definition or Description
Evaluative	**Categorical Observation**	the operator chooses 'the most appropriate' option from a list of several; this includes choosing a numeric range category, Typing (e.g., Rosgen channel type) and Scoring (e.g., level of impairment)
	Numeric (Range) Estimate	the operator reports his/her impression as an estimated number, or range
	Taxonomic Identification	the operator identifies the taxonomic affiliation of biota (e.g., ID of individual electroshocked fish, or of BMI in net during streamside assessments)
Measured	**Discrete Field WQ Measurement**	operator puts instrument in water and records reading
	Continuous/Time series Field WQ Measurement	operator deploys and retrieves sensors with data loggers; this yields a set of sequential results
	Morphology Survey	operators use tape, compass, stadia rod, and level to measure slope, bearing, and elevations along thalweg profile or cross section, etc. for creation of a 3D picture
	Discharge (Flow)	operator measured stream dimensions and current velocity to calculate total flow discharge
	Biota dimensions/weight	operators catch biota and measure their dimensions or weight, then release or process them further
	Count	operator counts the number of items in relevant groups (abiotic or biotic)
Collected	**Sample (abiotic media)**	operator collects an environmental medium (water, sediment, soil) into a container for analysis elsewhere
	Biota sample	operator captures living organisms and places them in a container, with or without preservatives, for analysis elsewhere

Notes
"Activity" is something you do in the field to generate data or start a data generation process. It can be an action, or an object
Evaluative: Generating Result using Eyes, Brain and Experience
Measured: Generating Results in situ, using an Instrument (probe, kit, rod, tape, etc.),
Collected: Transferring medium of interest into containers for processing/analysis elsewhere

Table 5: Conversion of Individual Characteristics and Repetitive Features into a Unified Data Element

A. Individual characteristics & Features

Field name	Value	Pick-List examples
Instrument ID Characteristic		pH, DO, Temp
Temperature (C) of Standard Material	x	
Barometric Pressure at Calibration (mmHg)	x	
Salinity at Calibration (ppt)	x	

pH electrode voltage	x	
pH electrode voltage acceptable range		
pH electrode voltage within range?		yes, no
DO electrode voltage	x	
DO electrode voltage acceptable range		
DO electrode voltage within range?		yes, no
Conductivity cell constant	x	
(more)		

B. Unified Concepts w Pick-lists

Field name	Value	Pick-List examples
Instrument ID Characteristic		pH, DO, Temp
Calibration environment property		**Temperature. Barometric Pressure, Salinity**
Calibration environment property Unit		C, mmHg, ppt
Calibration environment property Value	x	

Instrument Performance Diagnostic Attribute		**electrode voltage, conductivity cell constant**
Instrument Performance Diagnostic Attribute Value	x	
Instrument Performance Diagnostic Attribute acceptable range		
Performance within range?		yes, no

Table 6: Categories and Types of Quality Checks

Quality Check Category	Quality Check Type
Comparison to a 'Standard'	Accuracy Check of measurement (a.k.a post calibration check) Taxonomic ID check
Survey Loop	Loop closure
Repeats	repeated field measurement repeated estimate (one number) repeated categorical observations field duplicates lab replicates split samples MS/MSD
Inspections/verifications	sample custody seal sample in cooler temperature arrival temperature storage temperature holding time
Blanks	bottle blank equipment rinsate field blank trip blank method blank filter blank reagent blank GFC filter weight loss blank
Spikes	LCS CRM Surrogate matrix spike Internal standard Field spike
Positive/negative controls	reference toxicant test bacterial culture

Table 7: Examples of Quality Check Summaries

Data Field	Note	Example 1	Example 2	Example 3	Example 4
Q check date		4/5/2006	3/4/1997	4/5/2006	9/1/1998
Q Check category		Spike	Survey loop	Repeat	Blank
Q Check type		CRM	loop closure	repeated field measurement	Field Blank
spiked test medium	1	lab water	n/ap	n/ap	n/ap
Batch Type	2	analytical lab batch	survey run	instrument	sample batch [one team one Trip]
Batch Entity Name	3	Lab Batch ID	survey run ID	Instrument ID	Trip ID
Batch Entity ID		TRI-NH343	SurvDE56	EC-SLC03	SW2-SM-4
Activity ID		n/ap	PR07-V1surv	WD-e34	SM-4-FB
Date /period		n/ap	3/4/1997	3/8/06 to 4/5/06	9/1/98 to 9/3/98
DQ aspect addressed		accuracy	accuracy	precision	sample integrity - lack of contamination
characteristic		ammonia as N	elevation	specific conductance	naphthalene
Result unit		mg/L	dec.ft	uS	ng/g
Expected or 1st Value Type	**4**	nominal conc.	survey loop origin	Primary measurement result	zero
Expected or 1st Result		0.56	532.32	560	<1
Observed or 2nd Value type		measured conc.	survey loop closure	Repeated measurement result	analytical result
Observed or 2nd Result		0.49	532.29	590	1.2
differential or drift		0.07	0.03	30	n/ap
QC outcome computation		percent recovered	subtract Observed from Expected	differential as percent of average	(narrative: dirty container!)
QC Outcome		87	0.03	5	(flags on all Trip's Results)
QC Outcome unit		%	dec.ft	% RPD	n/ap
Data Quality Indicator (EPA QAPP)		accuracy	n/ap	precision	n/ap

Notes:
1. Pick-List examples for spiked test medium: Env sample, lab water, clean sand
2. Batch Type examples: sample, analytical lab batch, tailgate kit, instrument, survey run, toxicity test
3. Batch Entity Name examples: Lab Batch ID, IDEXX run ID, ELISA run ID, Instrument ID
4. Examples for "Expected or 1st Value Type": internal plus nominal spike, NIST thermometer reading,

Category number	Category name	Group number	Group name	Subject number	Subject name	Field ID	Field name	Pick-List examples
2	Result	1	Result (how much?)	2.1.1	Value	2.1.1.1	Result Value	7.8 (pH), murky, 8 (liter/sec)
2		1		2.1.2	Unit	2.1.2.1	Result Unit	mg/L, uS,
2		1				2.1.2.2	Measurement Basis	wet weight; ash-free dry weight
2		2	Characteristic (of what?)	2.2.1	Name	2.2.1.1	Characteristic	Temp, DO, Sp.Cond., pH, Turb
2		2		2.2.2	Identifier - Chemical	2.2.2.1	Substance Unique ID	
2		2				2.2.2.2	Substance Registry System Name	CAS, WQX code,
2		2		2.2.3	Identifier - Micro/tox/ biological	2.2.3.1	Taxon Name	
2		2				2.2.3.2	Life Stage	Juvenile, 2nd instar,
2		2		2.2.4	Identifier - Physical	2.2.4.1	(placeholder)	
2		3	Result descriptors	2.3.1	Result Type	2.3.1.1	Result Type	
2		3				2.3.1.2	Endpoint Type	
2		3		2.3.2	Result Qualifiers	2.3.2.1	Result Qualifier Code	
2		3				2.3.2.2	Batch Qualifier Code	J, R
2		3		2.3.3	Result-Specific procedures /outcomes	2.3.3.1	MDL	
2		3				2.3.3.2	Reporting Limit or PQL	

Figure 1: Organization of the Results and its Descriptors in the Integrated Data Elements List