Improving Quality, Efficiency, and Cost-Effectiveness of Data Validation through Increased Automation

Cindy Lee Westergard
Senior Scientist
HSW Engineering, Inc.

Suzan Hughes
Partner
Environmental Synectics, Inc.

DoD EMDQ Workshop
April 2015
Portland, Oregon

Presentation Overview

- Data Quality Assessment and Validation
  - Brief background
  - How can we improve?
- Example: Data Validation Project with Metrics
- Recommendations and Effect on
  - Quality
  - Efficiency
  - Cost-Effectiveness

HSW Engineering, Inc.

synectics connecting people and data

Data Quality Assessment and Validation

- Progress in automated data review through mid1990s – many labs delivered more QC elements in 1998 than they do today
- What stopped the progress?

HSW Engineering, Inc.

synectics connecting people and data

Data Quality Assessment and Validation

- What stopped the progress?
  - Counter intuitively – focus on standardization
  - 'Wait and See' attitude from providers concerned about choosing the wrong path
  - Contracting mechanisms associated with little incentive for 'better, faster, cheaper'

Presently, a combination of automated and manual procedures is used in the validation industry.

Can we increase the automation?

Data Quality Element	Manual	Electronic
Chain-of-Custody	x	
Condition Upon Receipt	x	
Case Narrative	x	
Holding Times		x
Sample Results		x
Field Blanks		x
Equipment Blanks		x
Trip Blanks		x
Field Duplicates or Replicates		x
Dilution Factors		x
Detection Limits		x
Quantitation Limits		x
Method Blanks		x
Laboratory Control Samples		x
Matrix Spike / Matrix Spike Duplicates		x
Laboratory Duplicates		x
Surrogates		x
Tracers or Labeled Analogs	x	
Post-Digestion Spikes	x	
Serial Dilutions	x	
Preparation Logs	x	
Instrument Logs	x	
Internal Standards	x	
Instrument Tunes	x	
Calibration Blanks	x	
Initial Calibrations	x	
Initial Calibration Verifications	x	
Continuing Calibration Verifications	x	
Sample Chromatograms	x	
Mass Spectra	x	
Raw Data	x	

HSW Engineering, Inc.

synectics connecting people and data

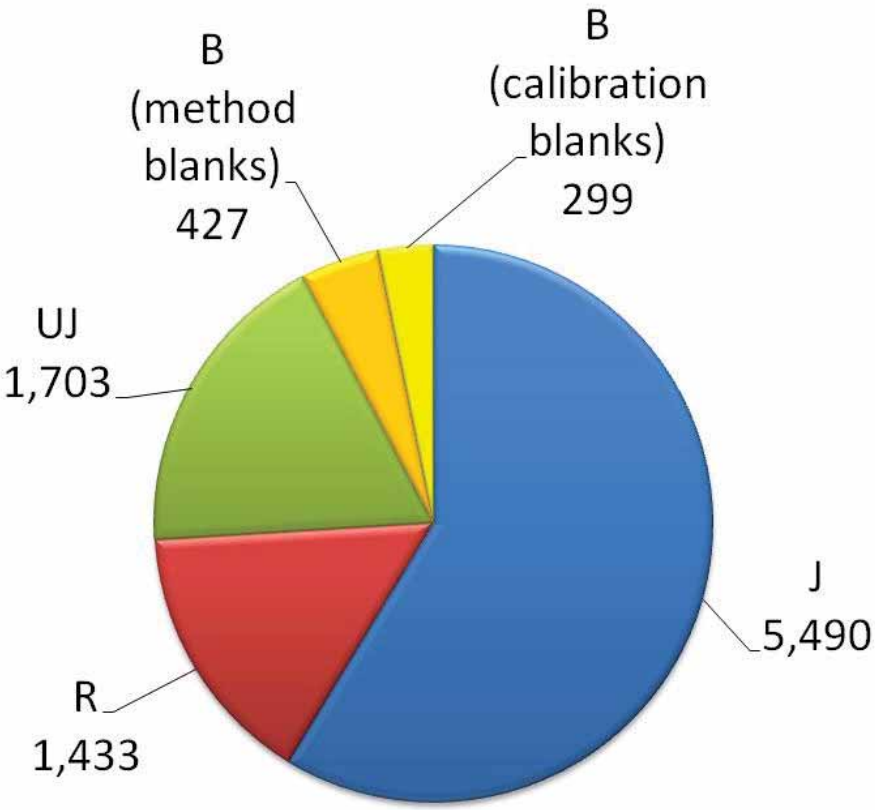
Example: Multi-Year DoD Remediation Project

- 450 soil samples during first phase
- 70,250 analytical results
- 100% review of data deliverables
  - Manual review of general information
  - Automated review of available electronic data
  - Manual review of
    - instrument calibration data
    - additional method QC data for metals and organics

HSW Engineering, Inc.

synectics connecting people and data

Of 70,250 results, 9352 results (13%) were qualified



Calibration Blanks – Metals

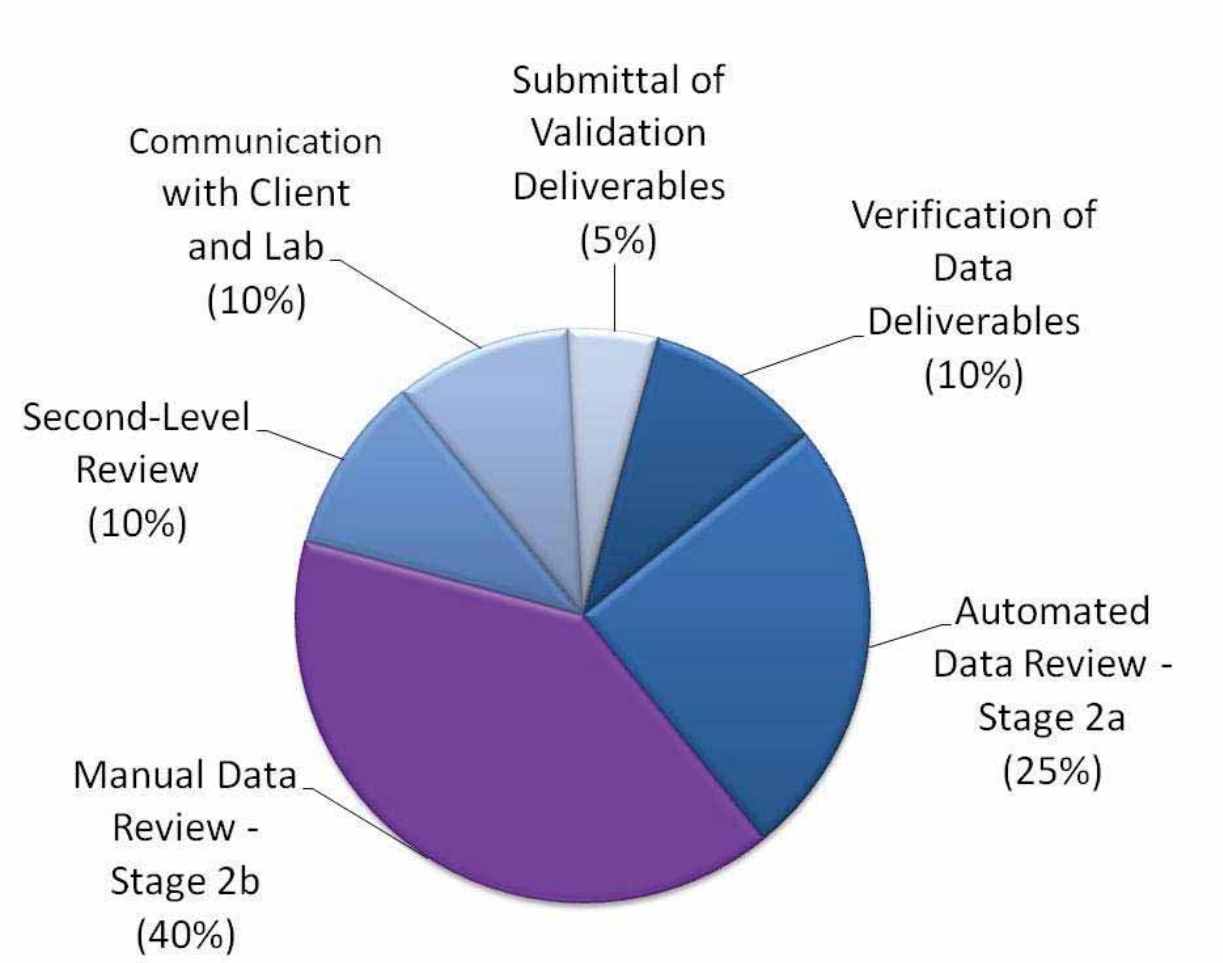
Analyte	Initial Calib. Blank (ug/L)	C	Continuing Calibration Blank (ug/L) 10	C	11	C	12	C	Prepa-ration Blank	C	M
Aluminum	10.000	U	10.000	U	41.100		7.600	J	0.541	J	P
Antimony	5.000	U	5.000	U	5.000	U	5.000	U	0.500	U	P
Arsenic	5.000	U	5.000	U	5.000	U	5.000	U	0.500	U	P
Barium	5.000	U	5.000	U	5.000	U	5.000	U	0.500	U	P
Beryllium	2.000	U	2.000	U	2.000	U	2.000	U	0.500	U	P
Cadmium	2.000	U	2.000	U	2.000	U	2.000	U	0.500	U	P
Chromium	5.000	U	5.000	U	5.000	U	5.000	U	0.500	U	P
Cobalt	5.000	U	5.000	U	5.000	U	5.000	U	0.500	U	P
Copper	5.000	U	5.000	U	5.000	U	5.000	U	0.500	U	P
Iron	200.000	U	200.000	U	71.100	J	200.000	U	50.000	U	P
Lead	5.000	U	5.000	U	0.780	J	5.000	U	0.054	J	P
Manganese	5.000	U	5.000	U	5.000	U	5.000	U	0.500	U	P
Molybdenum	5.000	U	5.000	U	5.000	U	5.000	U	0.500	U	P
Nickel	5.000	U	5.000	U	5.000	U	5.000	U	0.500	U	P
Selenium	-1.107		-0.840		-1.000		-1.300		0.500	U	P
Silver	5.000	U	5.000	U	5.000	U	5.000	U	0.500	U	P
Thallium	2.000	U	2.000	U	2.000	U	2.000	U	0.500	U	P
Vanadium	-2.543		[illegible]		[illegible]		[illegible]		[illegible]		
Zinc	5.000										

EPA Sample No.	D/F	Time	% R	Analytes AL	SB	AS	BA	BE	CD	CA	CR	CO	CU	FE	PB	MG	MN	HG	MO	NI	K	SE	AG	NA	SR	TL	SN	TI	V	ZN
ICB-11	1	1708	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
ICV-11	1	1711	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
[redacted] 1	1	1720	0		X	X		X	X		X	X	X		X				X	X		X	X			X			X	X
[redacted] 2	1	1725	0		X	X		X	X		X	X	X						X	X		X	X			X			X	X
[redacted] 3	1	1727	0		X	X		X	X		X	X	X		X				X	X		X	X			X			X	
[redacted] 4	1	1729	0		X	X		X	X		X	X	X		X				X	X		X	X			X			X	X
[redacted] 5	1	1732	0		X	X		X	X		X	X	X		X				X	X		X	X			X			X	X
ICB-12	1	1737	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
ICV-12	1	1739	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
ICB-13	1	1741	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
ICV-13	1	1744	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
ZZZZZZ	1	1746	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
ZZZZZZ	1	1748	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
CSA	1	1751	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
CSAB	1	1753	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X

HSW Engineering, Inc.

synectics
connecting people and data

Percentage of Time Spent on Data Review and Validation Tasks



HSW Engineering, Inc.

synectics
connecting people and data

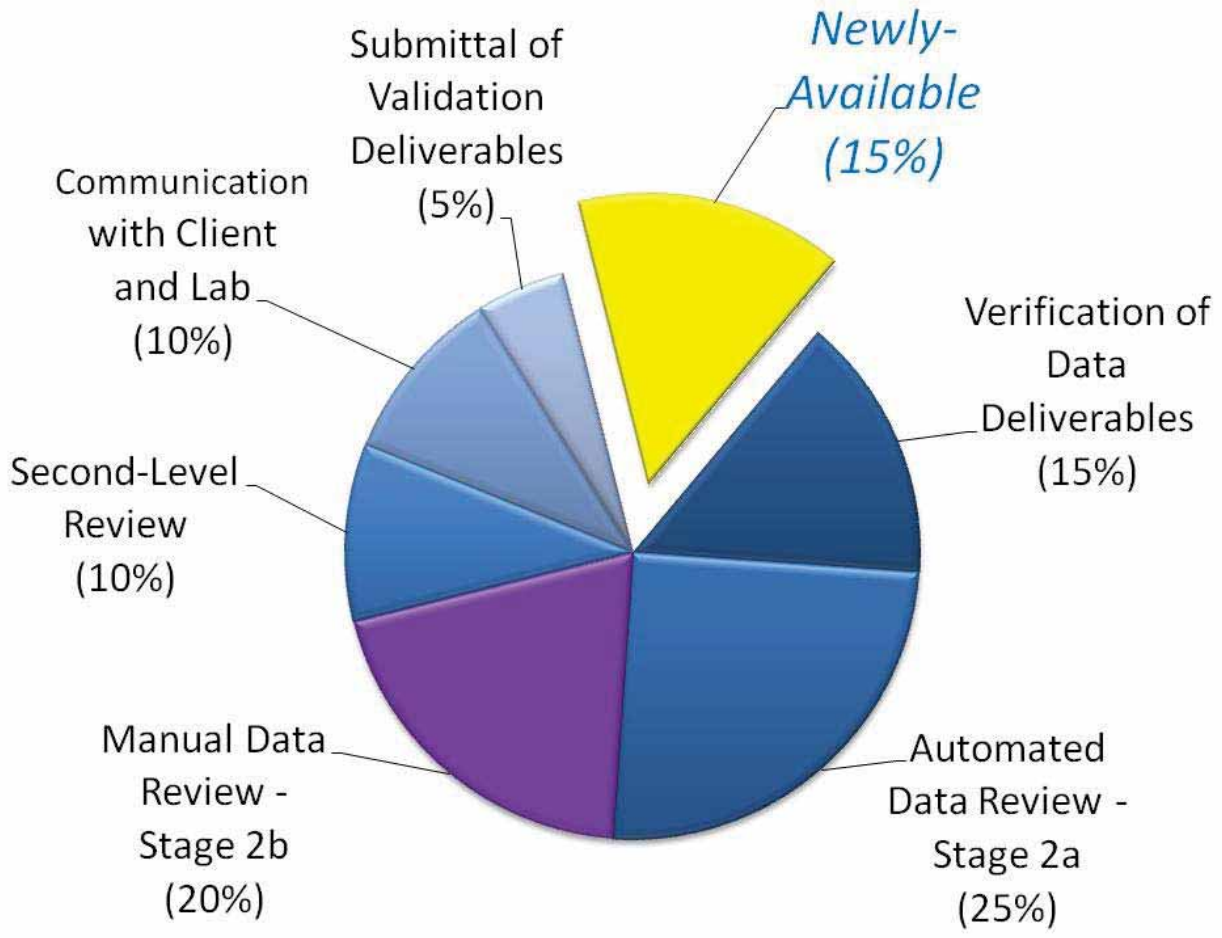
What if even a subset of these numeric instrument QC elements could be assessed electronically?

Data Quality Element	Manual	Electronic
Chain-of-Custody	x	
Condition Upon Receipt	x	
Case Narrative	x	
Holding Times		x
Sample Results		x
Field Blanks		x
Equipment Blanks		x
Trip Blanks		x
Field Duplicates or Replicates		x
Dilution Factors		x
Detection Limits		x
Quantitation Limits		x
Method Blanks		x
Laboratory Control Samples		x
Matrix Spike / Matrix Spike Duplicates		x
Laboratory Duplicates		x
Surrogates		x
Tracers or Labeled Analogs	x	
Post-Digestion Spikes		x
Serial Dilutions		x
Preparation Logs	x	
Instrument Logs	x	
Internal Standards		x
Instrument Tunes	x	
Calibration Blanks		x
Initial Calibrations		x
Initial Calibration Verifications		x
Continuing Calibration Verifications		x
Sample Chromatograms	x	
Mass Spectra	x	
Raw Data	x	

HSW Engineering, Inc.

synectics
connecting people and data

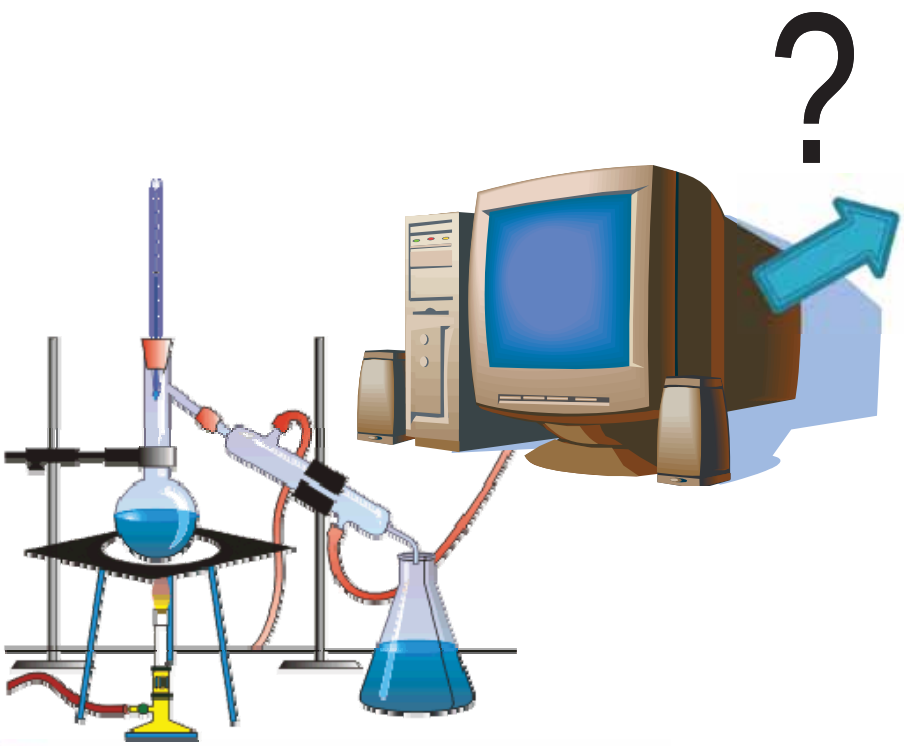
Estimated Percentage of Time with Enhanced Electronic Deliverables



HSW Engineering, Inc.

synectics connecting people and data

So if an increased level of automated review is a good thing, how do we get there?



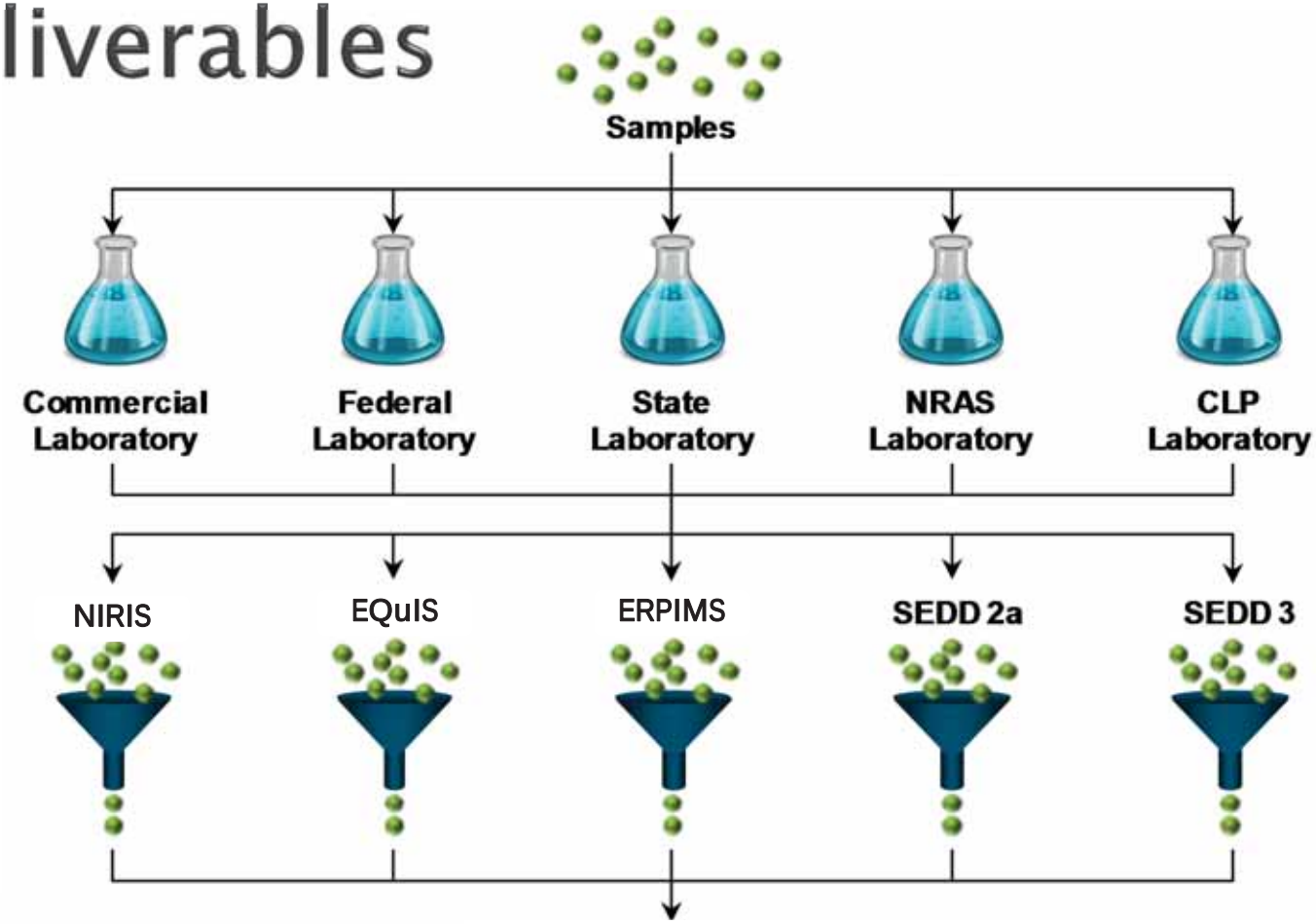
Calibration | Soil | Water

Analyte Label	Analyte Name	ICAL RSD	ICAL Correlation Coefficient	ICV Lower Estimation	ICV Upper Estimation	CCV Lower Estimation	CCV Upper Estimation
630-20-6	1,1,1,2-TETRACHLOROETHANE	30.00	0.99	-20.00	20.00	-20.00	20.00
71-55-6	1,1,1-TRICHLOROETHANE	30.00	0.99	-20.00	20.00	-20.00	20.00
79-34-5	1,1,2,2-TETRACHLOROETHANE	30.00	0.99	-20.00	20.00	-20.00	20.00
79-00-5	1,1,2-TRICHLOROETHANE	30.00	0.99	-20.00	20.00	-20.00	20.00
76-13-1	1,1,2-Trifluoro-1,2,2-trichloroethane [Fr						
75-34-3	1,1-DICHLOROETHANE	30.00	0.99	-20.00	20.00	-20.00	20.00
75-35-4	1,1-DICHLOROETHENE	30.00	0.99	-20.00	20.00	-20.00	20.00
563-58-6	1,1-DICHLOROPROPENE	30.00	0.99	-20.00	20.00	-20.00	20.00
87-61-6	1,2,3-TRICHLOROBENZENE	30.00	0.99	-20.00	20.00	-20.00	20.00
96-18-4	1,2,3-TRICHLOROPROPANE	30.00	0.99	-20.00	20.00	-20.00	20.00
526-73-8	1,2,3-TRIMETHYLBENZENE						
120-82-1	1,2,4-TRICHLOROBENZENE	30.00	0.99	-20.00	20.00	-20.00	20.00
95-63-6	1,2,4-TRIMETHYLBENZENE	30.00	0.99	-20.00	20.00	-20.00	20.00
96-12-8	1,2-DIBROMO-3-CHLOROPROPANE	30.00	0.99	-20.00	20.00	-20.00	20.00
106-93-4	1,2-DIBROMOETHANE	30.00	0.99	-20.00	20.00	-20.00	20.00
95-50-1	1,2-DICHLOROBENZENE	30.00	0.99	-20.00	20.00	-20.00	20.00
107-06-2	1,2-DICHLOROETHANE	30.00	0.99	-20.00	20.00	-20.00	20.00
540-59-0	1,2-Dichloroethene						
78-87-5	1,2-DICHLOROPROPANE	30.00	0.99	-20.00	20.00	-20.00	20.00
354-23-4	1,2-Dichlorotrifluoroethane [Freon 123a						

HSW Engineering, Inc.

synectics connecting people and data

Currently, labs produce over 100 different types of electronic deliverables



Slide courtesy of
Sara Goehl, U.S. EPA

Isn't One Format the Answer?

"Absolutely! As long as it's ours!"

- -Everyone

HSW Engineering, Inc.

synectics
connecting people and data

If Standardizing Isn't Happening…

- Laboratories now have IT tools in place to provide a wide range of EDDs
- EDD formats received from the lab are related to final deliverables only in terms of content
  - Work with your laboratories to obtain additional quality control elements in the format that works with your data management and Automated Data Review systems
  - Export validated data to your government clients in their required format

The Result…..?

HSW Engineering, Inc.

synectics connecting people and data

Improvements in the Inter-related Components of Quality, Efficiency, and Cost-Effectiveness

HSW Engineering, Inc.

synectics
connecting people and data

Improvements in *Quality*

- Reduces potential for data validator errors
  - Less chance of manually entering an incorrect qualifier
  - Less likely to overlook items that require data qualification
- Allows for more consistent evaluation
  - By single validator over extended project
  - By multiple validators involved in single project

HSW Engineering, Inc.

synectics connecting people and data

Gains in *Efficiency*

- Can quickly and easily evaluate entire data set rather than a subset
- Can involve lower-level staff to a greater extent

HSW Engineering, Inc.

synectics
connecting people and data

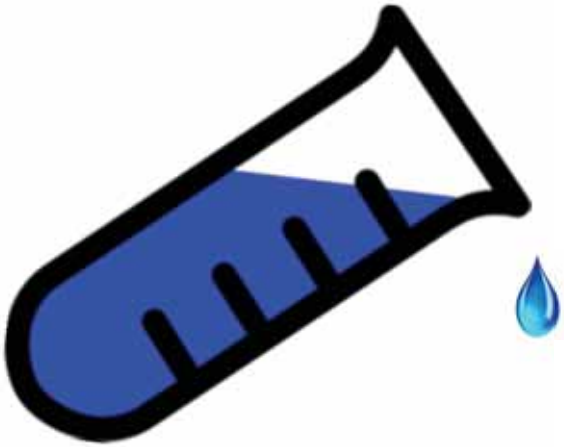
Greater *Cost-Effectiveness*

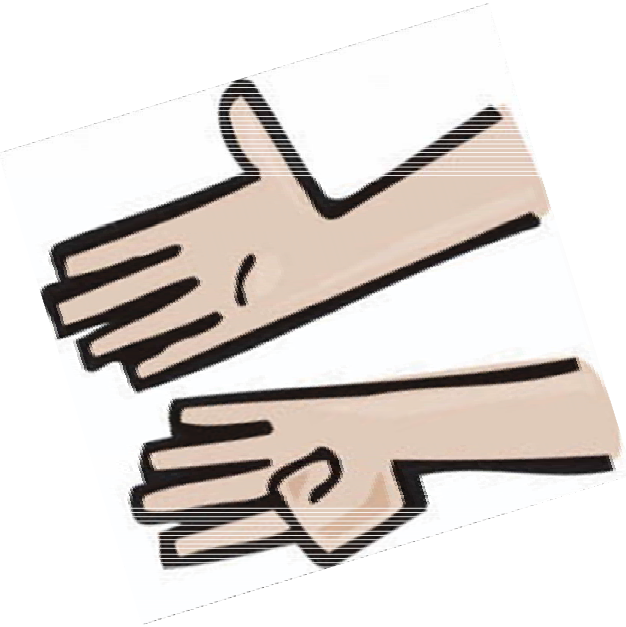
- Based on the example used in this presentation, validation costs could be reduced by an estimated 15% or more
- Reduced manual review time results in shortened turnaround times for validation deliverables and, thus, for project decision-making

HSW Engineering, Inc.

synectics connecting people and data

Ask the lab what they can do…and we'll do our part to make it work!





HSW Engineering, Inc.

synectics
connecting people and data

Acknowledgements

- Sara Goehl, EPA Office of Superfund Remediation and Technology Innovation Analytical Services Branch
- Steve Maxwell, TechKNOWLEDGEy Strategic Group
- DoD Contractors
- Commercial environmental laboratories

HSW Engineering, Inc.

synectics connecting people and data

Thank you! Questions?

Cindy Lee Westergard
HSW Engineering, Inc.
cwestergard@hsweng.com
(813) 968-7722 x227

Suzan Hughes
Environmental Synectics, Inc.
suzan.hughes@synectics.net
(916) 737-4013