



Web Cures for the Environmental Data Blues

Web Cures for the *Blues*

- What are environmental data blues?
- How can web services help cure the blues?
- Who is working on the cure (Case Studies)?
- When will the patient be healed?

What Causes the Environmental Data Blues?

- **Dispersed Data** - Mission-critical environmental information dispersed across multiple silos of information and data stores.
- **Multiple Formats/Requirements** – Data stored in different programs, different valid values, different platforms, different owners
- **Multiple Ownership** – Multi-disciplinary project teams have to navigate the seas of data ownership, costing time and \$
- **Long Retention Times** – Regulations require long cleanups, monitoring, reporting, and retention
- **Multiple Sources** – Sources can range from laboratory data to automated real time monitoring networks

Why Locus knows the Blues?

- We manage environmental data for thousands of environmental sites via our online environmental database (EIM)
- We see the results of the blues every day
- We have to address the issues resulting from years of data collection from multiple sources on multiple platforms every day

Just how **b l u e** is it out there?

Number of EDD formats	>200 not including “hybrid” Excel formats
Number of databases or spreadsheet formats	Thousands
Number of different database designs	Gazillion
Number of people accessing and reporting with environmental data	>100,000
Number of years data must be managed	Low – 10-20 High - >100
Number of different database synonyms for TCE	65 (Source ChemFinder)

Who's got the Blues?

- **Government – (Data Owners and Users)** Federal, state, local, foreign and domestic. Many systems, many different formats and platforms
- **Business (Data Owners) –** Thousands of systems, years of data, multiple contractors, multiple sites
- **Consulting firms – (Data Users) –** Hundreds of clients, mixed internal systems, often little coordination or consistency between offices and clients

Example

- 3,000 sites
- 30-40 different database formats
- Dozens of consultants
- Hundreds of individuals
- Millions of records
- 1 client

Why Cure The Blues?

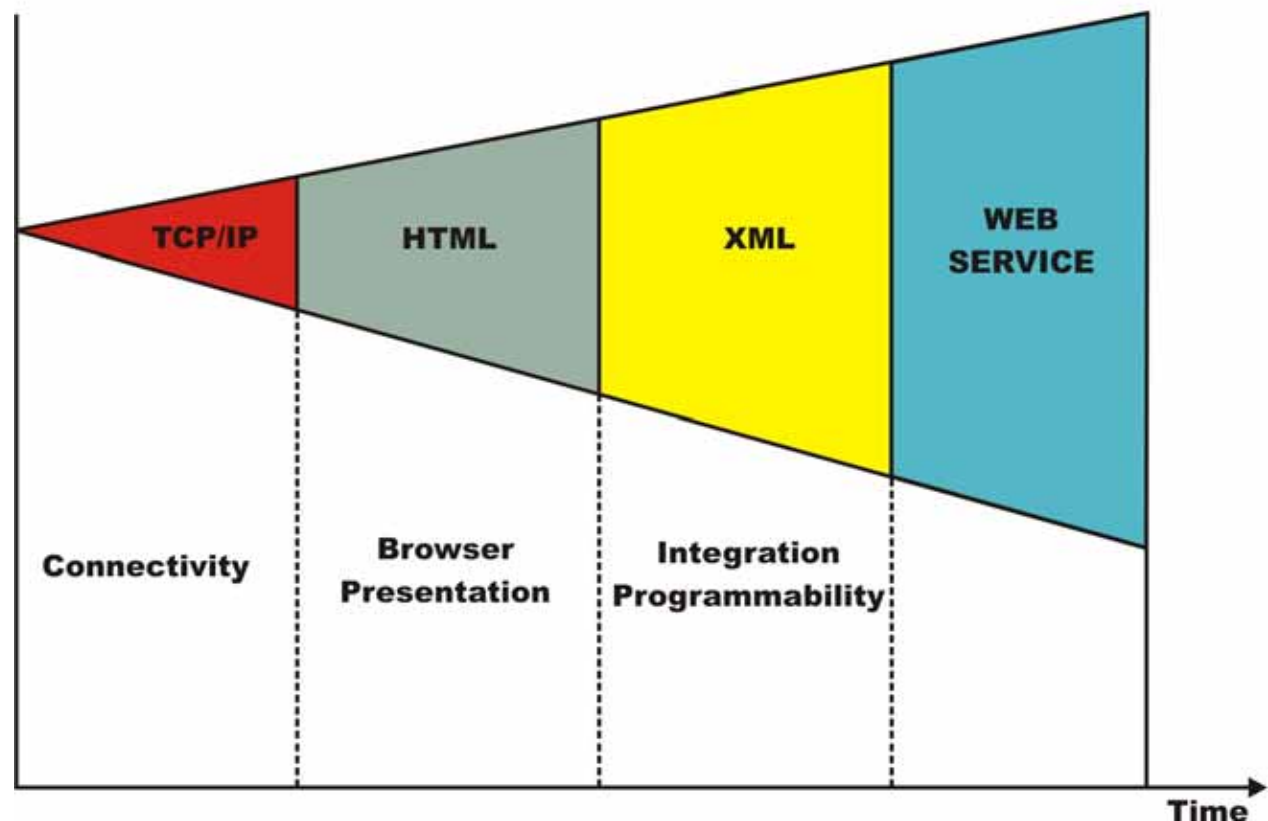
- **Save Money** - Data management is a long-term, ongoing cost - optimize to reduce costs
- **Manage Liability** – Better access and control of data allows better knowledge and management of liability
- **Improve Decision Making** – Access all your data more efficiently
- **Improve Governance** – Everyone is expected to know and understand their government and/or corporate responsibilities



How to Cure the Blues

Internet technology using **Web Services** and **XML** can bring disparate and dispersed data sources together for effortless tracking of and access to information.

The Internet Evolution



What Are Web Services?

One of the primary advantages of the XML Web services architecture is that it allows programs written in different languages on different platforms to communicate with each other in a standards-based way (Microsoft, 2001)

- Small task-oriented applications that are accessible through the Internet
- Act as the bridge between different applications, computers, intranets, and database systems
- Make heavy use of XML (eXtensible Markup Language) for transferring information
- Let applications share data, and—more powerfully—invoke capabilities from other applications regardless of platform

A web service is a site intended for use by computer programs instead of by human beings (Microsoft, 2002)

Why Use Web Services?

- Simply put, Web Services is the most cost-effective way to bridge the various platforms, systems, users, owners, and formats seen at most environmental sites
- Web Services can address situations where data are:
 - Located on different PCs or networks
 - Run on different operating systems/programs
 - Managed by different consultants
 - May be in paper form, databases, e-mails, or other media



Some Examples of Web Services

CNN/SI Scoreboard

CNN SI Scoreboard
Jul 09
Major League Baseball

7th

Atlanta Braves 5
New York Mets 3

[Preview](#) [Details](#) [Box](#) [Recap](#)

Klesko hits 3 run homer to left

Money [refresh](#)

Dow	10,242.82 ▲ 46.91	Get Quote:
NASDAQ	1,978.62 ▲ 18.36	<input type="text"/> Go
S&P	1,122.51 ▲ 5.87	

Quotes delayed at least 20 minutes

Find a Broker: Ameritrade, ShareBuilder, Trade Now, TD Waterhouse

Find a Mortgage: LendingTree

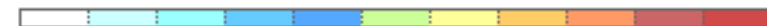
[Insurance](#) • [Loans](#) • [News](#) • [Markets](#) • [Credit Report](#)

Auckland Today

at: 7:58 am NZST

F° | [C°](#)

Currently: 54°	 Mostly Cloudy	Hi: 58
		Lo: 47



BBC as rain continues • Watchdog criticises [\[icon\]](#) [\[icon\]](#) [\[icon\]](#)

Web Services – Some Definitions

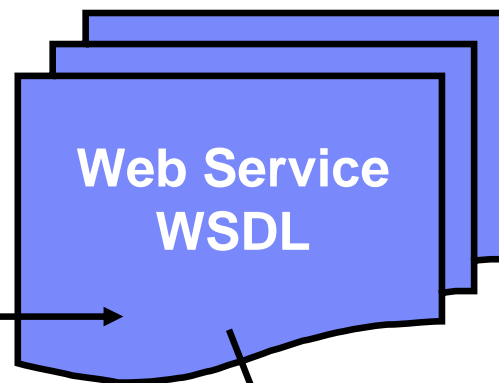
- **SOAP** (Simple Object Access Protocol) is the XML-based set of rules for the call-and-response communication between Web Service-enabled applications. SOAP is the glue that holds Web Services together by ensuring reliable delivery of messages.
- **WSDL** (Web Services Description Language) describes the design of a Web Service so a client can discover how to invoke and use it.
- **UDDI** (Universal Description, Discovery, and Integration) is the standard for registering all available Web Services in use. UDDI is like a phone book for locating a particular Web Service.
- **XML** (eXtensible Markup Language) is rapidly becoming the de facto standard for transferring data between databases and applications on the World Wide Web.

How a Sample Web Service Works:

TerraServer-USA .NET



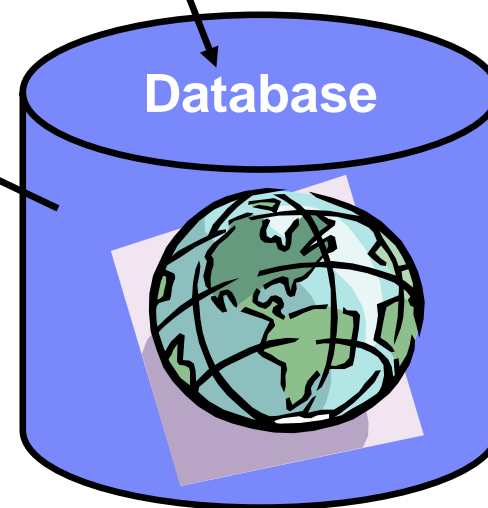
SOAP Request:
I Want a Photo of
Biltmore Estate NC



Web Service
WSDL

Data Request:
Get Photo of
Biltmore Estate NC

XML Response:
Send Photo of
Biltmore Estate NC



Database

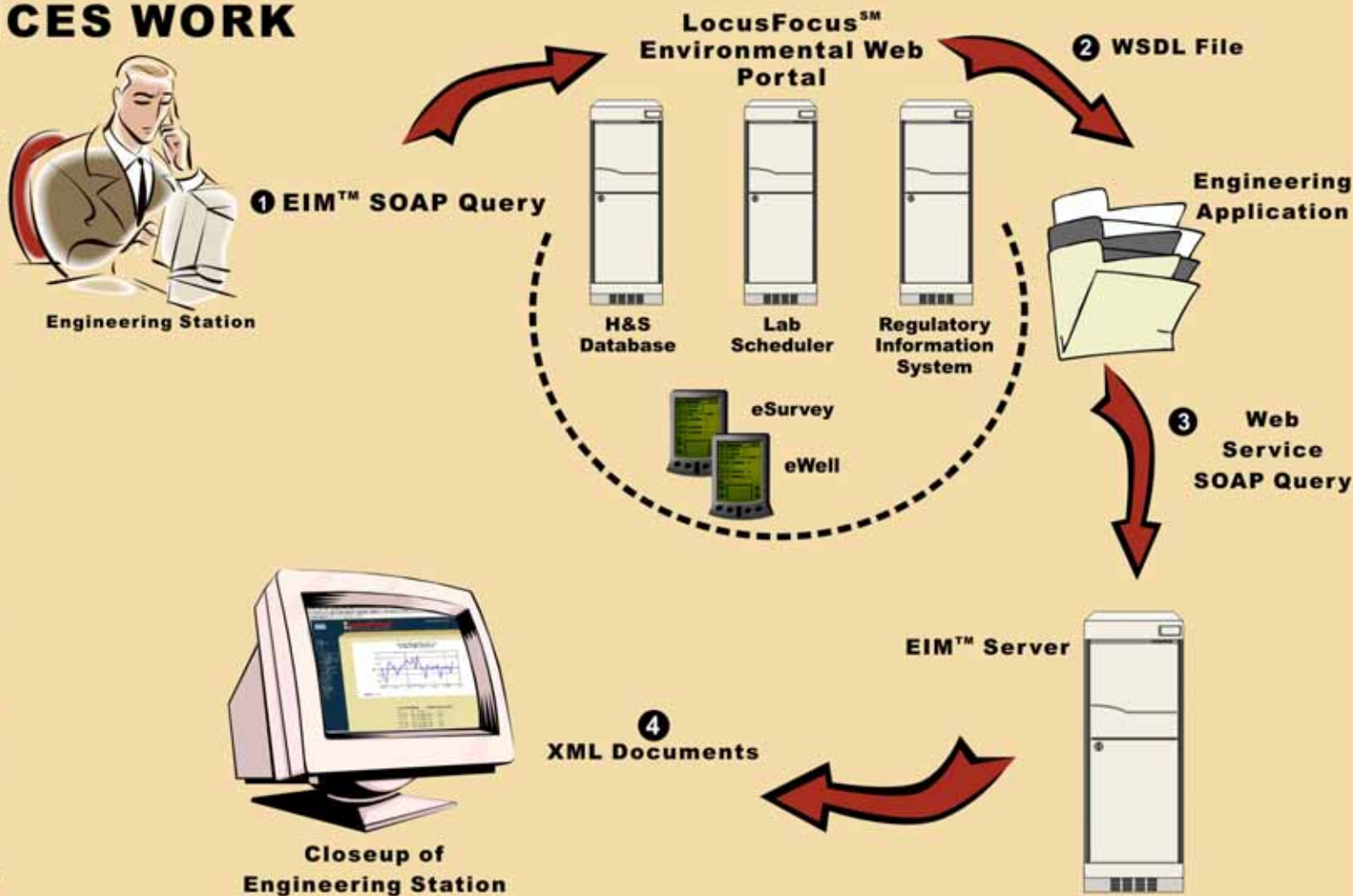
HOW WEB SERVICES WORK

1 An engineer enters a site name with proper User ID and password into the Web Services application, which sends a simple object access protocol (SOAP) query to the LocusFocusSM or other site server database, requesting its web services description language (WSDL) file.

2 The server sends the WSDL file, which describes all the available web services to the engineering application.

3 The application parses the WSDL file, selects a web service to provide site data, and sends a SOAP query (with the Site ID number) to another server, which verifies the security of the engineering application and compiles the requested data.

4 The server sends the information in XML format to the engineering application, which assembles and displays all the data, like sites in SVG format, analytical data, photos and documents available, and contract information for all site participants from various consultants to regulatory agencies' personnel.



Advantages of XML

- **XML is for structuring data** - it is not a programming language.
- **XML looks a bit like HTML**, but uses tags to describe data, rather than format text; the XML file contains data nodes with attributes and elements that define the data.
- **XML is self descriptive** – an XML file contains “metadata” that describes the data in the file; years after creating the file, a user can look at it and understand the file contents.
- **XML is revolutionizing how applications talk to other applications** – or more broadly, how computers talk to other computers – by providing a universal data format that lets data be easily adapted or transformed.
- **XML can use DTD files** – Document Type Definitions – to impose a data format and allow for data checking and verification.

XML – Sample File

```
<breakfast_menu>
  <food>
    <name>Belgian Waffles</name>
    <price>$5.95</price>
    <description>two of our famous Belgian Waffles with plenty of real maple syrup</description>
    <calories>650</calories>
  </food>
  <food>
    <name>Strawberry Belgian Waffles</name>
    <price>$7.95</price>
    <description>Belgian waffles covered with strawberries and whipped cream</description>
    <calories>900</calories>
  </food>
  <food>
    <name>Berry-Berry Belgian Waffles</name>
    <price>$8.95</price>
    <description>Belgian waffles covered with various fresh berries and whipped cream</description>
    <calories>900</calories>
  </food>
  <food>
    <name>French Toast</name>
    <price>$4.50</price>
    <description>thick slices made from our homemade sourdough bread</description>
    <calories>600</calories>
  </food>
</breakfast_menu>
```


How Do You Set Up a Web Service?

1. Create your Web Service application (using Visual Basic, Java, Cold Fusion, or some other coding language)
2. Create the WSDL (Web Services Description Language) file
3. Post the Web Service and the WSDL on your server in a virtual directory
 - *Register the Web Service with UDDI if you want others to find it*
 - *Let users develop applications that use the Web Service*

O
p
t
i
o
n
a
l

Two Case Studies – Using Web Services to solve environmental data *Blues*

- **Case Study 1 - Developing a Web Service to import EPA's new XML SEDD** – by developing a single web service “clearinghouse”
- **Case Study 2 - Developing a Web Service to save customer money** – by developing a single web service to reduce reporting costs



Case Study 1: Using Locus EIM and XML With EPA/USACEs XML SEDD



US Army Corps
of Engineers ®



Problem:

- Hundreds of EDD formats costing the industry time and money to perform a simple operation – load environmental data to databases and share data

Solution:

- Standard XML EDD (SEDD) to create a single universal format that can be imported to any database using Web Services

What is SEDD?

(Staged Electronic Data Deliverable)

- An inter-agency effort spearheaded by the U.S. Environmental Protection Agency (EPA) and the U.S. Army Corps of Engineers (USACE)
- A file format based on Extensible Markup Language (XML)
- A method to ensure a lab's data can be used by **any** Web Service or XML-based web application

What Is EIM?



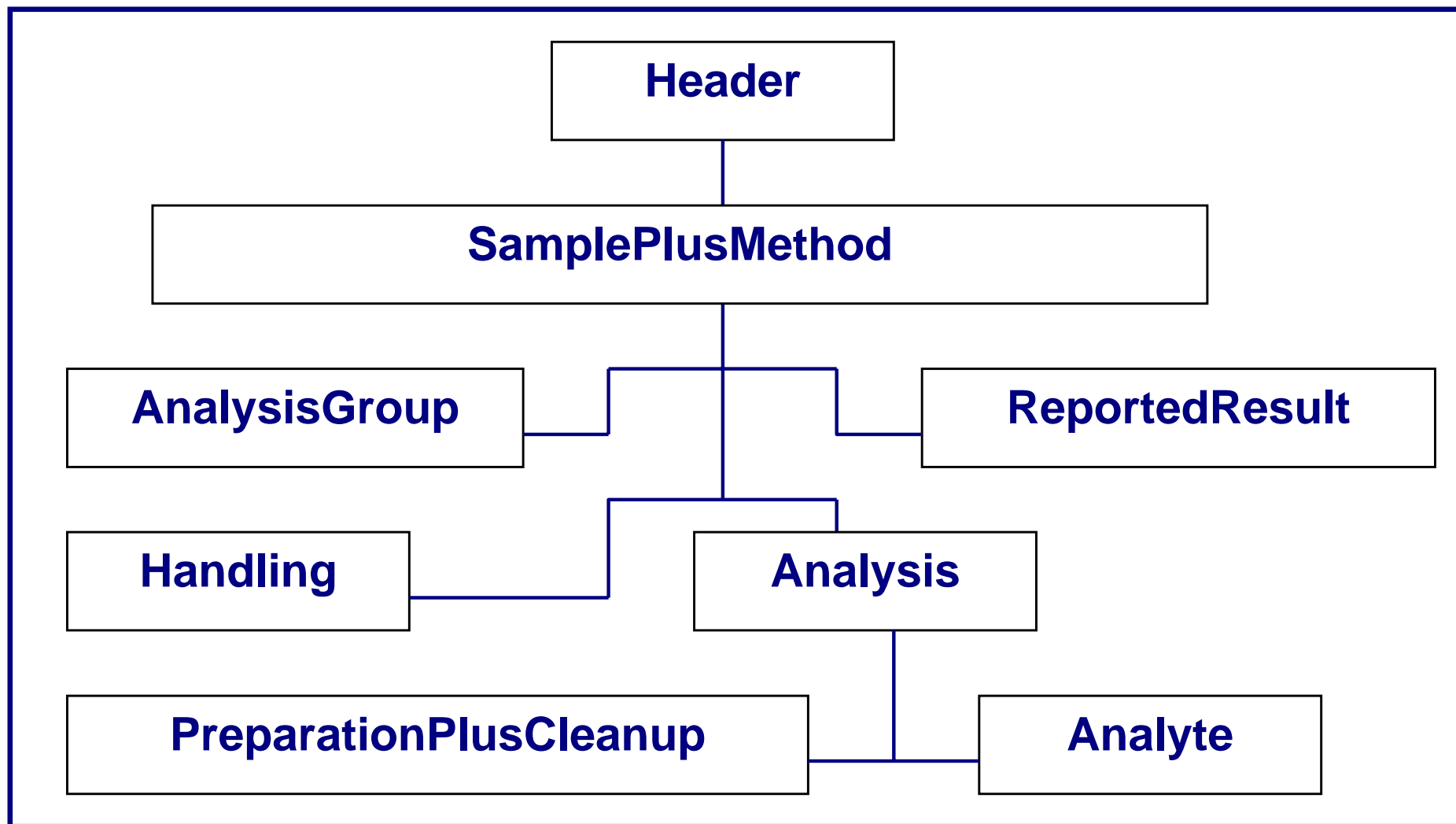
- Commercial web-based environmental database management system developed by Locus
- Used by more than 3,000 sites to manage analytical, field, survey, geologic, and other environmental data
- Built from the ground up to take advantage of web technologies to streamline data management

The Four SEDD Stages

SEDD is being implemented in stages for maximum flexibility.

- **Stage 1** includes **basic analytical data** elements (such as the sample ID, analyte, result, and qualifier) to convey results to the end user.
- **Stage 2a** adds method quality control data to Stage 1.
- **Stage 2b** adds instrument quality control data to Stage 2a.
- **Stage 3** adds additional measurement data to Stage 2b to allow for independent recalculation of reported results.
- **Stage 4** adds the raw instrument data files to Stage 3.

SEDD Stage 2a File Structure



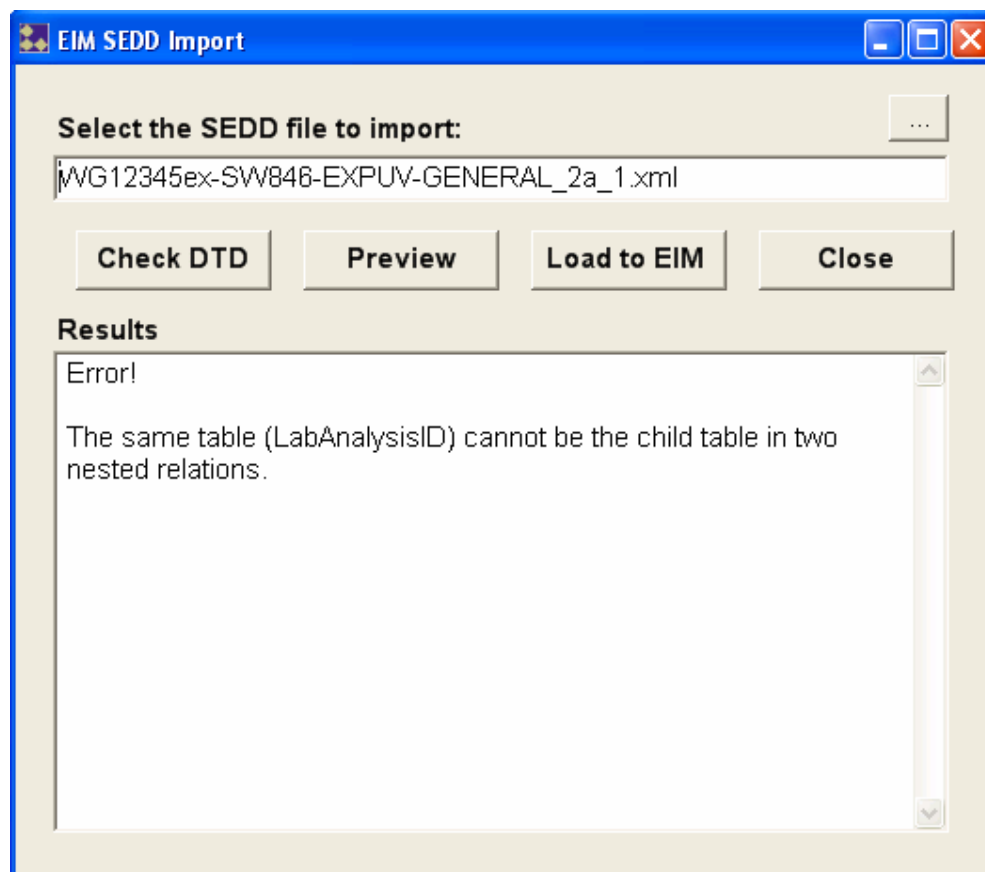
SEDD – Sample File (abridged)

```

<Header>
<EDDID>SEDD</EDDID>
<EDDVersion>Draft 5.0</EDDVersion>
<EDDIImplementationID>GENERAL_2a</EDDIImplementationID>
<LabDataPackageName>SW846 S2a VOA Pilot</LabDataPackageName>
<LabDataPackageVersion>1</LabDataPackageVersion>
<LabReportedDate>10/20/2003</LabReportedDate>
<SamplePlusMethod>
  <ClientMethodID>SW846_8260B</ClientMethodID>
  <ClientMethodType>GCMS_Internal_Standard</ClientMethodType>
  <ClientSampleID>Sample-01</ClientSampleID>
  <LabSampleID>C3J030277001</LabSampleID>
  <MatrixID>Water</MatrixID>
  <CollectedDate>10/02/2003 12:09</CollectedDate>
  <LabID>TLab</LabID>
  <LabName>Test Lab</LabName>
  <LabReceiptDate>10/03/2003 10:15</LabReceiptDate>
  <Analysis>
    <ClientMethodID>SW846_8260B</ClientMethodID>
    <LabAnalysisID>C3J030277001</LabAnalysisID>
    <LabFileID>C3J030277001</LabFileID>
    <AnalyzedDate>10/07/2003 23:07</AnalyzedDate>
    <Analyte>
      <CASRegistryNumber>17060-07-0</CASRegistryNumber>
      <AnalyteName>1,2-Dichloroethane-d4</AnalyteName>
      <AnalyteType>Surrogate</AnalyteType>
      <LabQualifiers />
      <Result>11</Result>
      <ResultUnits>ug/L</ResultUnits>
    </Analyte>
  </Analysis>
</SamplePlusMethod>

```


How EIM Uses Web Services to Import SEDD



- The EIM SEDD Import Web Service allows users to connect to EIM from their local PC and import SEDD files.
- After login, the user selects a SEDD file and uses the Web Service to check if the file matches the DTD file.
- If the file fails, the user is given diagnostic information.

How EIM Uses Web Services to Import SEDD (con't)

If the file passes the DTD check, the user can preview the file and load it into EIM for data checks.

EIM SEDD Import

Select the SEDD file to import:

WG12345ex-SW846-VOAMS-GENERAL_2a_1.xml

Check DTD

Preview

Load to EIM

Results

Document is valid!

SEDD File Preview

Header:		EDDID:	SEDD	EDDVersion:	Draft 5.0	EDDImplementationID:	GE		
SamplePlusMethod:		ClientMethodID:	SW846_8260B	ClientMethodType:	GCMS_Internal_Standard	ClientSampleID:	Sa		
CAS	AnalyteName	Analyt	LabAnalysisID	Result	ResultType	ResultUnits	ExpectedR	ExpectedR	PercentRec
75-71	Dichlorodifluorome	Spike	WG12345ex/0801	59.5	=	ug/L	50.0	ug/L	119
74-87	Chloromethane	Spike	WG12345ex/0801	49.9	=	ug/L	50.0	ug/L	99.9
75-01	Vinyl Chloride	Spike	WG12345ex/0801	54.9	=	ug/L	50.0	ug/L	110
74-83	Bromomethane	Spike	WG12345ex/0801	51.8	=	ug/L	50.0	ug/L	104
75-00	Chloroethane	Spike	WG12345ex/0801	54.4	=	ug/L	50.0	ug/L	109
75-69	Trichlorofluoromet	Spike	WG12345ex/0801	55.2	=	ug/L	50.0	ug/L	110
67-64	Acetone	Spike	WG12345ex/0801	54.0	=	ug/L	50.0	ug/L	108
75-35	1,1-Dichloroethene	Spike	WG12345ex/0801	50.6	=	ug/L	50.0	ug/L	101
75-09	Methylene Chlorid	Spike	WG12345ex/0801	54.0	=	ug/L	50.0	ug/L	108
75-15	Carbon Disulfide	Spike	WG12345ex/0801	51.8	=	ug/L	50.0	ug/L	104
156-6	trans-1,2-Dichloro	Spike	WG12345ex/0801	48.9	=	ug/L	50.0	ug/L	97.8
75-34	1,1-Dichloroethane	Spike	WG12345ex/0801	49.7	=	ug/L	50.0	ug/L	99.3
78-93	2-Butanone	Spike	WG12345ex/0801	54.7	=	ug/L	50.0	ug/L	109
594-2	2,2-Dichloropropa	Spike	WG12345ex/0801	51.3	=	ug/L	50.0	ug/L	103
156-5	cis-1,2-Dichloroeth	Spike	WG12345ex/0801	49.2	=	ug/L	50.0	ug/L	98.4
67-66	Chloroform	Spike	WG12345ex/0801	51.1	=	ug/L	50.0	ug/L	102
74-97	Bromochlorometha	Spike	WG12345ex/0801	52.0	=	ug/L	50.0	ug/L	104
71-55	1,1,1-Trichloroetha	Spike	WG12345ex/0801	54.5	=	ug/L	50.0	ug/L	109
563-5	1,1-Dichloropropo	Spike	WG12345ex/0801	54.1	=	ug/L	50.0	ug/L	108
56-23	Carbon Tetrachlori	Spike	WG12345ex/0801	51.6	=	ug/L	50.0	ug/L	103
107-0	1,2-Dichloroethane	Spike	WG12345ex/0801	50.5	=	ug/L	50.0	ug/L	101
71-43	Benzene	Spike	WG12345ex/0801	50.4	=	ug/L	50.0	ug/L	101

EIM Notifies User SEDD file is uploaded

The Web Service e-mails a site manager a notification of an uploaded SEDD file.

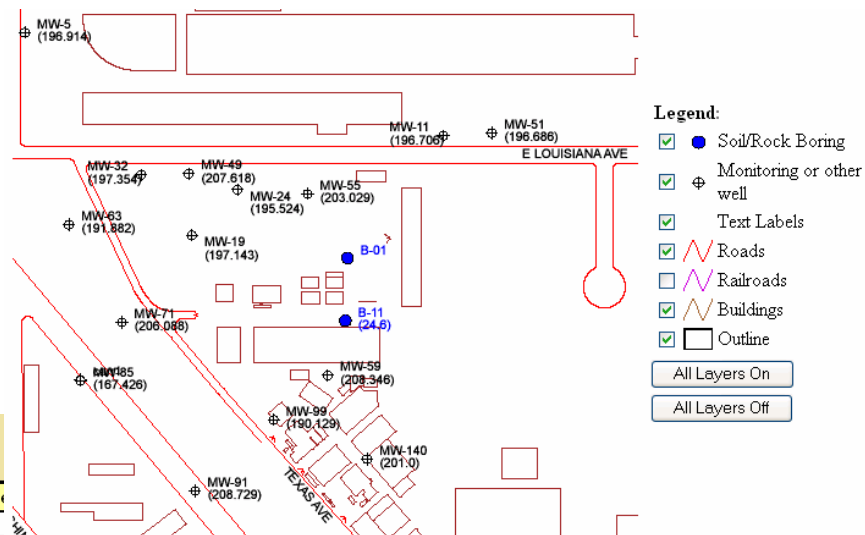
Send EDD Upload Notification

From*:	<input type="text" value="John Lab"/>
From Email Address:	<input type="text" value="john@lab.com"/>
To*:	<input type="text" value="Mike Manager"/>
To Email Address*:	<input type="text" value="mike@site.com"/>
Send CC copy to Email Address:	<input type="text"/>
Subject:	<input type="text" value="EDD 123.txt is now loaded into EIM site 456A"/>
Message*:	<div>The EDD 123.txt has been loaded to EIM site 456A as Dataset No. 13512. The EDD is ready for your review.</div>

Data is now available for analysis and reporting

Once the data is fixed and saved to EIM, it can be queried into tables, charts and maps.

Destination Table	No. Of Records Inserted
LAB_QC_PARTNER_DERIVED	3
LAB_SAMPLE	5
FIELD_SAMPLE_RESULT	48
LAB_SAMPLE_RESULT	41

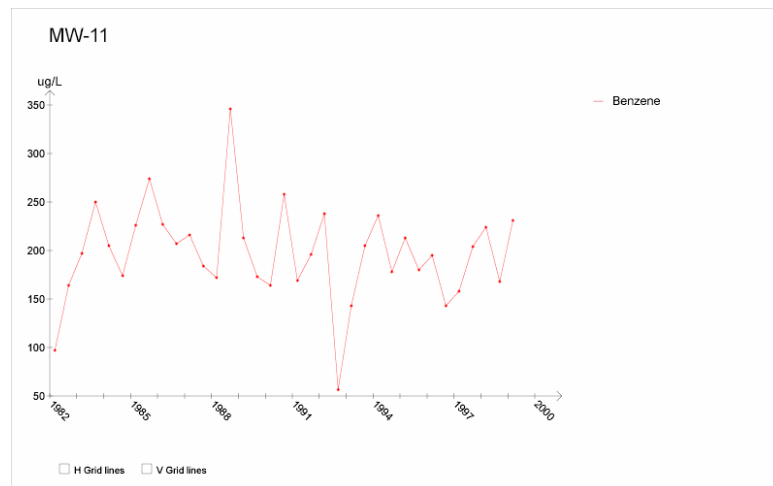


Data For Selected Locations:

Location ID	Parameter Name	Date Sampled	Report Result	Report Units	Lab Qualifier	Detected	Sample Purpose	Field Sample ID	Analytical Method
MW-11	1,2-Dichloroethane	05-12-2003	0.092	ug/l	U	N	REG	Sample-01	SW846_8260B
MW-11	Tetrachloroethene	05-12-2003	0.24	ug/l	U	N	REG	Sample-01	SW846_8260B
MW-11	cis-1,2-Dichloroethene	05-12-2003	4.9	ug/l		Y	REG	Sample-01	SW846_8260B
MW-11	trans-1,2-Dichloroethene	05-12-2003	0.65	ug/L	J	Y	REG	Sample-01	SW846_8260B
MW-11	1,1,1,2-Tetrachloroethane	05-12-2003	0.13	ug/l	U	N	REG	Sample-01	SW846_8260B
MW-11	Chloroform	05-12-2003	0.14	ug/l	U	N	REG	Sample-01	SW846_8260B
MW-11	1,1,1-Trichloroethane	05-12-2003	0.16	ug/L	U	N	REG	Sample-01	SW846_8260B
MW-11	Chloromethane	05-12-2003	0.075	ug/l	U	N	REG	Sample-01	SW846_8260B
MW-11	Chloroethane	05-12-2003	0.18	ug/l	U	N	REG	Sample-01	SW846_8260B
MW-11	Vinyl Chloride	05-12-2003	0.17	ug/l	U	N	REG	Sample-01	SW846_8260B
MW-11	Methylene Chloride	05-12-2003	0.31	ug/l	J	Y	REG	Sample-01	SW846_8260B
MW-11	1,1-Dichloroethane	05-12-2003	0.16	ug/l	U	N	REG	Sample-01	SW846_8260B
MW-11	1,1-Dichloroethene	05-12-2003	0.21	ug/l	U	N	REG	Sample-01	SW846_8260B
MW-11	1,1,2-Trichloroethane	05-12-2003	0.12	ug/l	U	N	REG	Sample-01	SW846_8260B
MW-11	Trichloroethene	05-12-2003	35	ug/l		Y	REG	Sample-01	SW846_8260B

Next 15 Records

Total No. of Records Retrieved=48



Case Study 2: Providing Locus EIM Web Service To Solve Customer Reporting Needs

PROBLEM:

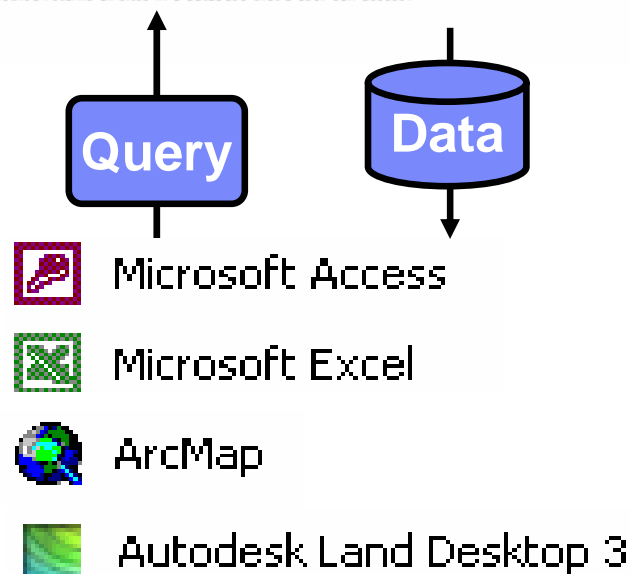
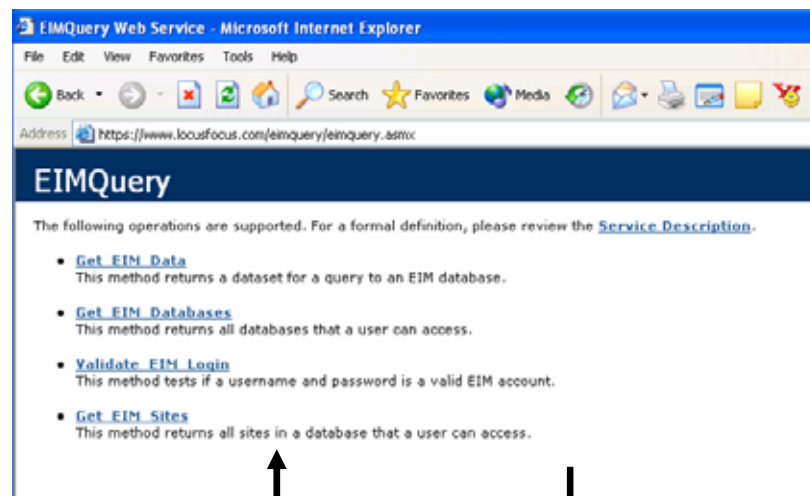
- Customer uses EIM for all data, but EIM did not have the “right” reports
- Customer want to use existing programs to generate reports but wants it dynamically tied to EIM data



Case Study 2: Providing Locus EIM Web Service To Solve Customer Needs (Con't)

Solution – Web Services

- **Create web service** - Locus creates web service to accept user login and SQL query; data returned as XML file to user
- **Access web service** - Client modifies application to query web service and receive data
- **Business as usual!**
 - Client keeps reports and process
 - Locus maintains one generic web service instead of multiple client-specific reports
 - Other clients develop applications using same web service
- **Everyone's happy!**



How Web Services may help solve some of the environmental data Blues

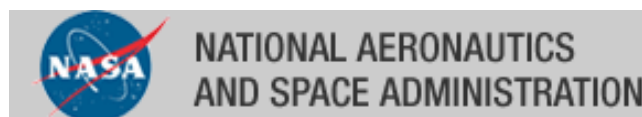
- **Link to other applications** – Locus links to ArcView, AutoCad, Access and Excel via web services to give customers more tools, preserve investments, and more choices.
- **Consolidate data** - Grab data from sensors, monitors, equipment, and wirelessly transmitting the information to a single location for automated reporting; grab data from various databases and consolidate into one.
- **Foster standardization across the industry** - XML SEDDs are the first step, other agencies are standardizing as well, using Web Services to make it easy, reasonable and affordable.



Who is Using Web Services for Environmental Data?



FRS



For More Information

ESRI: <http://www.esri.com/software/arcwebservices/index.html>

World Wide Web Consortium: <http://www.w3.org>

UDDI: <http://www.uddi.org>, <http://services.xmethods.net/>

XML: [http:// www.xml.org](http://www.xml.org)

EPA: <http://www.epa.gov/cdx>

SEDD: <http://www.epa.gov/superfund/programs/clp/sedd.htm>

Microsoft: <http://www.microsoft.com/net/basics/webservices.asp>

TerraServer-USA.NET: <http://terrasservice.net/default.aspx>

Locus and EIM: [http:// www.locustec.com](http://www.locustec.com)

