Invitational Meeting – Report on Project Status and Development of a New Roadmap

March 25-26, 2009
Orlando, Florida

DIGGS is being developed through the Transportation Pooled Fund Study (TPF 5(111)) coordinated by the Ohio Department of Transportation (ODOT). The focus of the TPF project is to compile the standards development work of the AGS, COSMOS, the University of Florida, and others to create a new international data exchange format. The project, "Development of Standards for Geotechnical Management Systems, Project TPF-5(111)," was approved and funded in the Summer of 2005 at a funding level of approximately $700k over three years to develop the first release of DIGGS.

The DIGGS development effort has been a challenging one over the past few years. There has been a great deal of progress and accomplishments to date, however, a substantial amount of work still remains. As we're approaching the end of the DIGGS v.1 release review period, it is apparent that the level of review and extent of participation during this phase has been limited and insufficient to assess the true readiness of the standard. Furthermore, many have raised concerns regarding the development of DIGGS, some related to technical data model aspects, while others related to organizational issues and how the work is being carried out. It will be important at this meeting come to some understanding on these key project-level issues:

- **Re-evaluate our goal for this project.** What constitutes success for us? When will we know we've achieved it? What do you expect to get in return from your investment of funds/resources/time in this project?
- **Establish a roadmap to achieve the goal.** Define the tasks, the milestones, their deliverables, and the costs. Identify how this will get done, and when.
- **Identify tools.** What types of tools do DIGGS stakeholders need to actively participate in the development and implementation process?
- **Reconsider the governance structure.** Are there other organizational approaches to enhance partnerships with Pooled Fund Team, AGS, COSMOS, and the commercial software partners to maximize leveraging opportunities and align common needs?
- **Reconsider the breadth of data types in DIGGS.** Are we trying to cover too many data types?
- **Consider enlisting a larger pool of technical experts.** Do we need more technical folks involved with the development of the data model and schema?

Invited Participants include members from the Geotechnical Management System (GMS), Geotechnical Data Coalition (GDC), Special Interest Group (SIG) Chairs, and selected industry partners.

Last Name	First Name	Organization	Phone	Email
Beach	**Kirk**	Ohio DOT	(614) 275-1342	Kirk.Beach@dot.state.oh.us
Benoit	**Jean**	UNH	(603) 862-1419	jean.benoit@unh.edu
Bobbitt	**John**	POSC	(713) 267-5174	john-bobbitt@sbcglobal.net
Bray	**Chris**	Keynetix		chris.bray@keynetix.com
Caronna	**Salvatore**	gINT	707-838-1271	scaronna@gintsoftware.com
Chandler	**Roger**	AGS	011-44-01-5276-8888	roger.chandler@key-systems.com
Dasenbrock	**Derrick**	MN		Derrick.Dasenbrock@dot.state.mn.us
Deaton	**Scott**	Dataforensics	(678) 406-0106	sdeaton@dataforensics.net
Fontaine	**Leo**	CT	860-594-3180	leo.fontaine@po.state.ct.us
Fritz	**Mike**	MO	573-526-4346	mike.fritz@modot.mo.gov
Gorman	**Laurel**	USACE	601-634-4484	Laurel.T.Gorman@erdc.usace.army.mil
Hoit	**Marc**	UF	(352) 392-1301	Marc_Hoit@ncsu.edu
Holmes	**Will**	KY	502-564-8900	will.holmes@ky.gov
Horhota	**David**	FL	352-955-2924	david.horhota@dot.state.fl.us
Jung	**Jay**	UF/BSI	352-392-9537	jchun@ce.ufl.edu
Lefchik	**Thomas**	FHWA	(614) 280-6845	thomas.lefchik@fhwa.dot.gov
McVay	**Mike**	UF	(352) 392-8697	mcm@ce.ufl.edu
Mohamed	**Khalid**	FHWA		Khalid.Mohamed@fhwa.dot.gov
Mulla	**Mohammed**	NC	919-250-4088	mmulla@dot.state.nc.us
Oliver	**Len**	TN	615-350-4130	Len.Oliver@state.tn.us
Patterson	**David**	UKHA	011-44-117-372-8399	david.patterson@highways.gsi.gov.uk
Ponti	**Daniel**	USGS	650-329-5679	dponti@usgs.gov
Power	**Chris**	Mott-MacDonald		Christopher.Power@mottmac.com
Roblee	**Cliff**	Caltrans	916-227-7183	cliff.roblee@dot.ca.gov
Spink	**Tim**	CIRIA	011-44-20-8774-2953	tim.spink@mottmac.com
Stagg	**Kim**	Delta	(914) 303-4611	kstagg@deltaenv.com
Turner	**Loren**	Caltrans	(916) 227-7174	loren.turner@dot.ca.gov
Walthall	**Steve**	Bechtel/AGS		sxwaltha@bechtel.com
Weaver	**Scot**	EarthSoft	(435) 554-3099	sweaver@earthsoft.com

Meeting Agenda (Final 3/23/09)

Day 1 – March 25, 2009

Topic	Speaker/ Facilitator	Time
Welcome – Meeting agenda, goals, and recent changes within the project	L. Turner	8:30 am – 8:50 am
Project accomplishments	T. Lefchik	8:50 am – 9:10 am
Technical overview of the latest release of DIGGS	C. Power	9:10 am – 9:30 am
US & UK DIGGS implementation efforts • AGS implementation efforts (20 min) • COSMOS implementation efforts (20 min)	 R. Chandler D. Ponti	9:30 am – 10:10 am
Break		**10:10 am – 10:20 am**
US & UK DIGGS implementation efforts (continued) • UK Transportation Agency Perspective (20 min) • US State DOT implementation efforts (20 min)	 D. Patterson D. Dasenbrock	10:20 am – 11:00 am
Geo-Software community perspective – is DIGGS ready? • Keynetix (15 min) • Earthsoft (15 min) • Dataforensics (15 min) • gINT (15 min)	 R. Chandler S. Weaver S. Deaton S. Caronna	11:00 am – 12:00 pm
Working Lunch (provided in the meeting room)		**12:00 pm – 12:30 pm**
Summary of technical issues	M. Hoit	12:30 pm – 12:50 pm
AGS experience in managing/maintaining a transfer standard	S. Walthall	12:50 pm – 1:10 pm
A strawman strategic plan for a successful DIGGS v.1 release	L. Turner	1:10 pm – 1:30 pm
Facilitated discussion on strategic planning and roadmap development	W. Holmes	1:30 pm – 5:00 pm (w/break)
GMS Meeting (private)		5:00 pm – 6:00 pm
Dinner (details to be announced)		**6:30 pm**

Day 2 – March 26, 2009

Topic	Speaker/ Facilitator	Time
Facilitated discussion on DIGGS governance and project management	M. Hoit	8:30 am – 11:30 am (w/break)
Working Lunch (provided in the meeting room)		**11:30 am – 12:00 pm**
Meeting summary and action items	C. Roblee/ J. Bobbit	12:00 pm – 12:15 pm
GMS Meeting (private) • Review of project expenditures • Status of TPF commitments and obligations • Resolutions and decisions	K. Beach	12:15 pm – 1:00 pm

DIGGS Successes
or
"You've come a long way baby!"

Thomas Lefchik, P.E.
Federal Highway Administration

Why we are here.

DIGGS
DATA INTERCHANGE FOR GEOTECHNICAL AND GEOENVIRONMENTAL SPECIALISTS

GMS Workshop June 2004

DIGGS
DATA INTERCHANGE FOR GEOTECHNICAL AND GEOENVIRONMENTAL SPECIALISTS

TPF-5(111)

- through the lack of a standard data definition for Geotechnical data, there exists significant difficulty in archiving, reusing and sharing data
- The establishment of standards for the development of geotechnical management systems will provide the means for state DOTs to efficiently capture, store, retrieve, and share geotechnical data and information internally as well as with external agencies and user groups.

TPF-5(111)

- the development of an open and flexible XML (GML compliant) based data structure and data dictionary geotechnical management systems.
- The data structure will define the form and content (alpha or numeric) of the data, the precision, the units, the field size, the type of data acquired, other data attributes, and the relationships between the attributes.

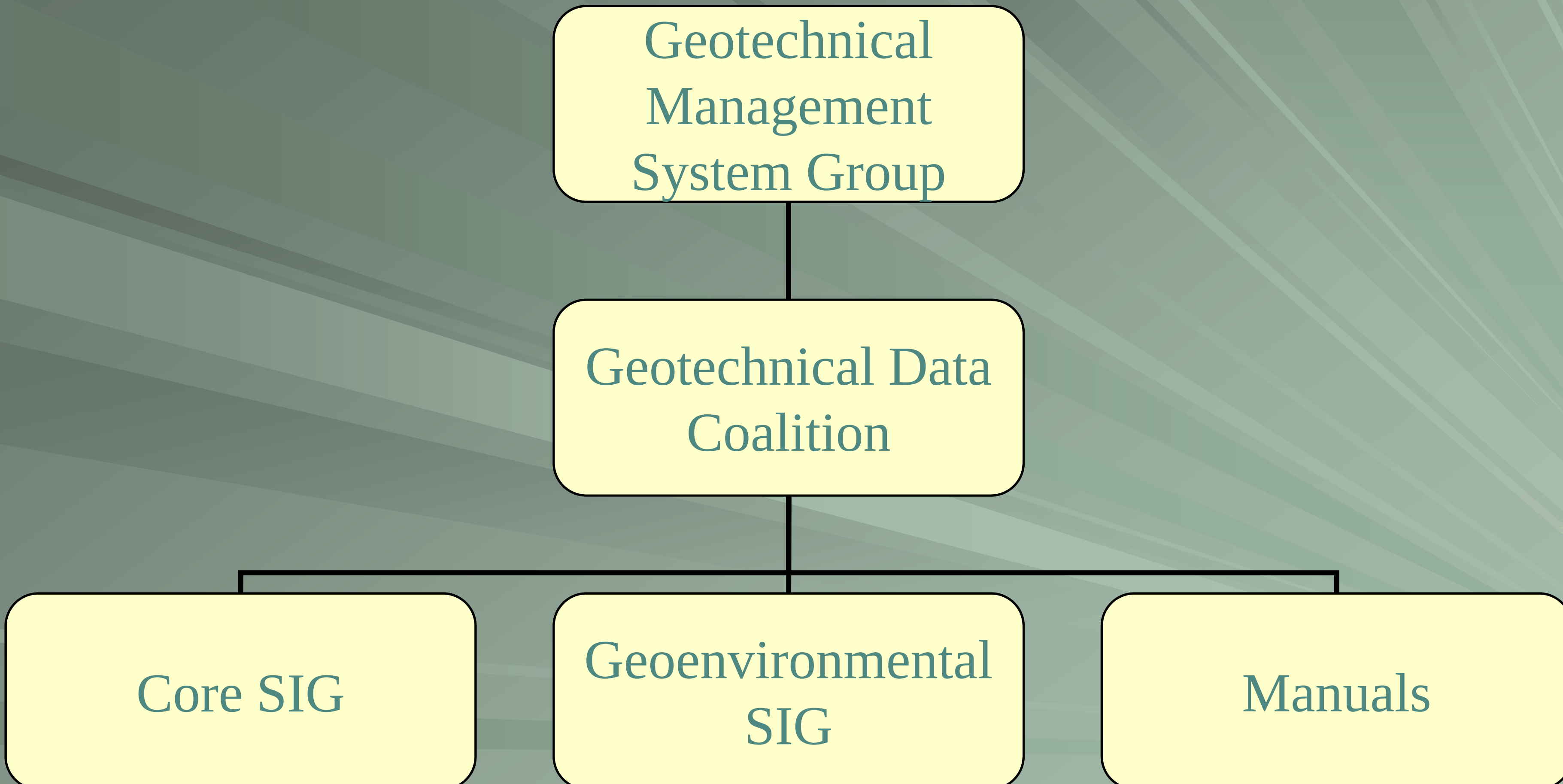
Pooled Fund Project TPF-5(111)

- Combine existing geotechnical data interchange standards
- Expand to include other data (i.e. geohazards, geotechnical assets)
- Survey state DOTs and others
- Finalize standards

Project Deliverables

- Data Dictionary
- Electronic data structure for data
- Electronic data structure for the metadata
- Allow local extensions and customizations

Organization



Collaboration Meeting

May 2005

- Association of Geotechnical and Geoenvironmental Specialists (AGS)
- Bridge Software Institute at the University of Florida
- California Department of Transportation
- Connecticut Department of Transportation
- Consortium of Organizations for Strong-Motion Observation Systems (COSMOS)
- Construction Industry Research and Information Association (CIRIA)
- Delta Environmental Consultants, Inc.
- Earthsoft
- Federal Highway Administration (FHWA) – Office of Federal Lands Highway
- Federal Highway Administration (FHWA) - Ohio Division Office
- Florida Department of Transportation
- Georgia Department of Transportation
- gINT Software Inc.
- Indiana Department of Transportation
- Kentucky Department of Transportation
- Keynetix Ltd
- Minnesota Department of Transportation
- Missouri Department of Transportation
- Mott MacDonald
- North Carolina Department of Transportation
- Ohio Department of Transportation
- Petrochemical Open Standards Consortium
- Tennessee Department of Transportation
- United Kingdom Highways Agency (UKHA)
- United States Army Corps of Engineers (USACE)
- United States Environmental Protection Agency (U.S. EPA)
- United States Geological Survey (USGS)
- United States Navy
- University of New Hampshire

Coordination & Review

- TransXML
- ASFE (formerly Associated Soil and Foundation Engineers) members are professional firms that provide "earth engineering" and related applied science services
- ASCE GeoInstitute
  – US Country Group

International Standard

- International cooperation in development
  - US federal agencies
  - AGS
  - CIRIA
  - COSMOS
- Joint Technical Committee 2
  - International Society of Rock Mechanics
  - International Society for Soil Mechanics and Geotechnical Engineering
  - International Association for Engineering Geology and the Environment

Geoenvironmental SIG

- SEDD – USEPA & USACE
  - Now a requirement for superfund sites
- AGS-E

DIGGS beta Version 1.0

- July 18, 2008
- DIGGS Version 1.0 includes:
  - Geotechnical ground investigation data
  - Geoenvironmental data
  - Deep foundations data
  - Borehole geophysical investigation data

Documents Drafted

- Introduction to DIGGS
- Data Dictionary
- Brochure

DIGGS
DATA INTERCHANGE FOR GEOTECHNICAL AND GEOENVIRONMENTAL SPECIALISTS

An introduction to the DIGGS electronic data transfer format

www.diggsml.org

What is DIGGS?

DIGGS (Data Interchange for Geotechnical and Geoenvironmental Specialists) is a standard international format for the electronic transfer of geotechnical and geoenvironmental data. DIGGS is software neutral and non-commercial. DIGGS can be used for transfer of all geotechnical and geoenvironmental data throughout all project stages offering enormous advantages in terms of workflow efficiency, data accuracy and validity, records retention and management, and consequently cost savings.

Who was involved in DIGGS development?

DIGGS was developed with the participation of:

- United States Federal Highway Administration (FHWA)
- United Kingdom Highways Agency (UKHA)
- Eleven United States Departments of Transportation
- United States Geological Survey (USGS)
- United States Army Corps of Engineers (USACE)
- United States Environmental Protection Agency (US EPA)
- United States Navy (USN)
- Construction Industry Research and Information Association (CIRIA)
- United Kingdom Association of Geotechnical and Geoenvironmental Specialists (AGS)
- Consortium of Organizations for Strong-Motion Observation Systems (COSMOS)
- The University of Florida
- The University of New Hampshire
- Petrochemical Open Standards Consortium (POSC)
- Major software vendors including Keynetix, gINT, and EarthSoft

Who is DIGGS designed for?

DIGGS is designed to assist anyone who wants to send or receive geotechnical or geoenvironmental information such as owner agencies, companies associated with software and databases, academic institutions, industry organizations, research organizations, etc.

DIGGS Version 1.0 includes:

- Geotechnical ground investigation data
- Geoenvironmental data
- Deep foundations data
- Borehole geophysical investigation data

Future developments will include:

- Geotechnical asset management
- Surface geophysical investigation data
- Electronically stored geotechnical documents
- Geotechnical instrumentation
- Geological, geotechnical and geoenvironmental hazard assessment and risk management

Why should I use DIGGS?

Currently, data can be transferred between parties to a project in many different formats, from paper reports to electronic documents, spreadsheets, etc.. DIGGS has much to offer that will improve the data and information workflow for everyone:

- Electronic data is more efficient than paper-based reports.
- Data transfer is faster and more efficient, and can be undertaken in close to real-time.
- Data validation is carried out using a set of rules that is the same for all parties.
- Data does not need to be re-entered at each project step, reducing time and mistakes.
- DIGGS data, which is geographically located, can be used in CAD and GIS packages

With DIGGS, data entered in the field can be sent, used with various software programs, stored, and reused.

Recent Accomplishments

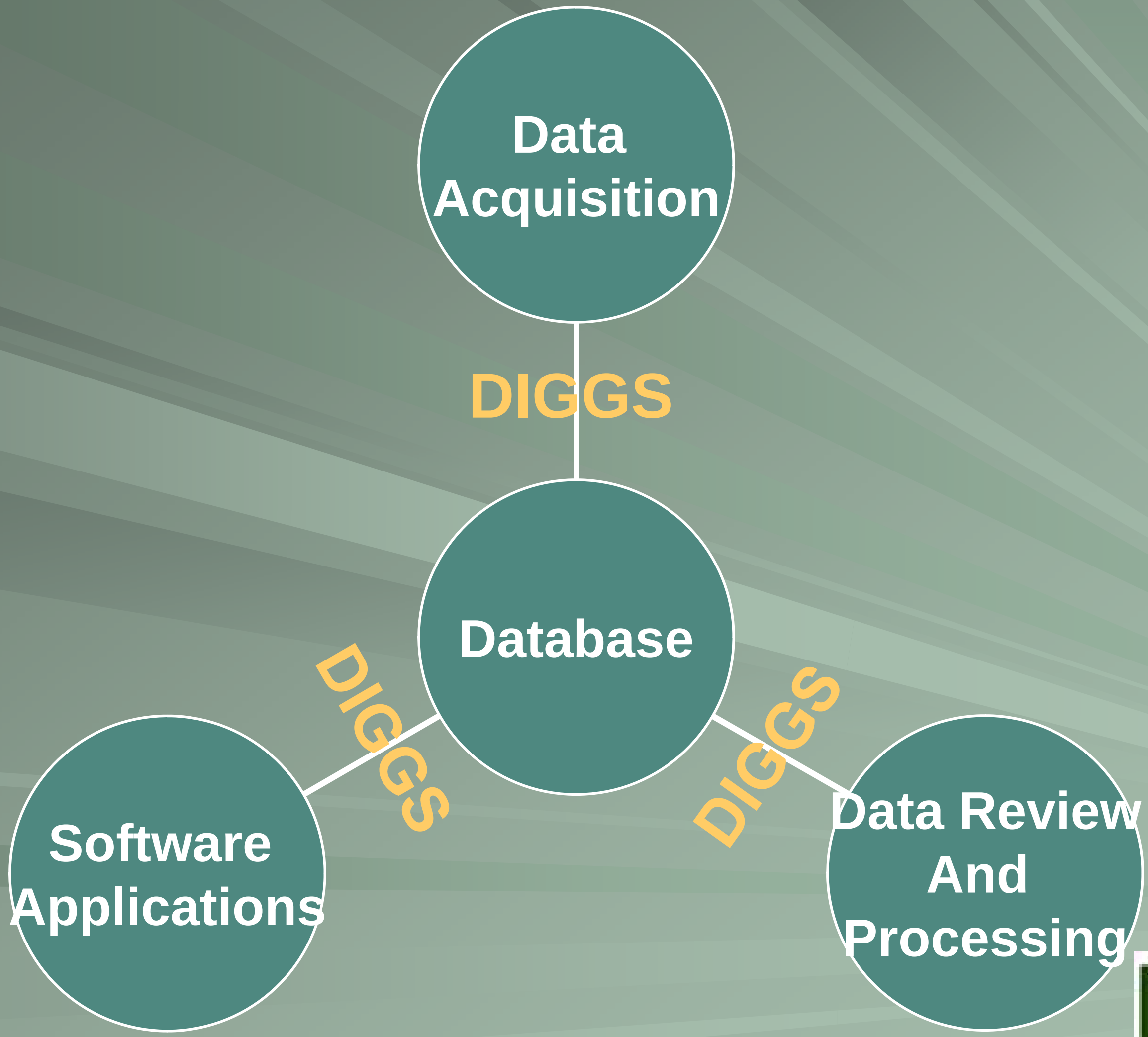
- Spring 2006 work began on AGS-DIGGS converter and examples
- July 2006 CIRIA issued report of practice.
- August 2007 Version 0.9
- August 2007 Logo

Presentations and articles

- Highway Geology Symposium October 2007
- CTIP Newsletter January 2006
- GeoCongress March 2008
- GeoCongress February 2006
- OTEC October 2008
- Midwest Geotechnical Conference September 2007
- AGS meeting June 2008
- Appalachian Coalition August 2006 & 2005

DIGGS DATA INTERCHANGE FOR GEOTECHNICAL AND GEOENVIRONMENTAL SPECIALISTS

The DIGGS Advantage



DIGGS
DATA INTERCHANGE FOR GEOTECHNICAL AND GEOENVIRONMENTAL SPECIALISTS


DIGGS
DATA INTERCHANGE FOR GEOTECHNICAL AND GEOENVIRONMENTAL SPECIALISTS

DIGGS
Moving us from the past to the future

- Past:
  - Paper management of data - fragmented, time consuming and expensive
  - Manual information manipulation and analysis
- Future:
  - Seamless electronic data transfer and management system - efficient, fast and economical
  - Unlimited electronic data manipulation and analysis


DIGGS
USDA
DIGGS
DATA INTERCHANGE FOR GEOTECHNICAL
AND GEOENVIRONMENTAL SPECIALISTS

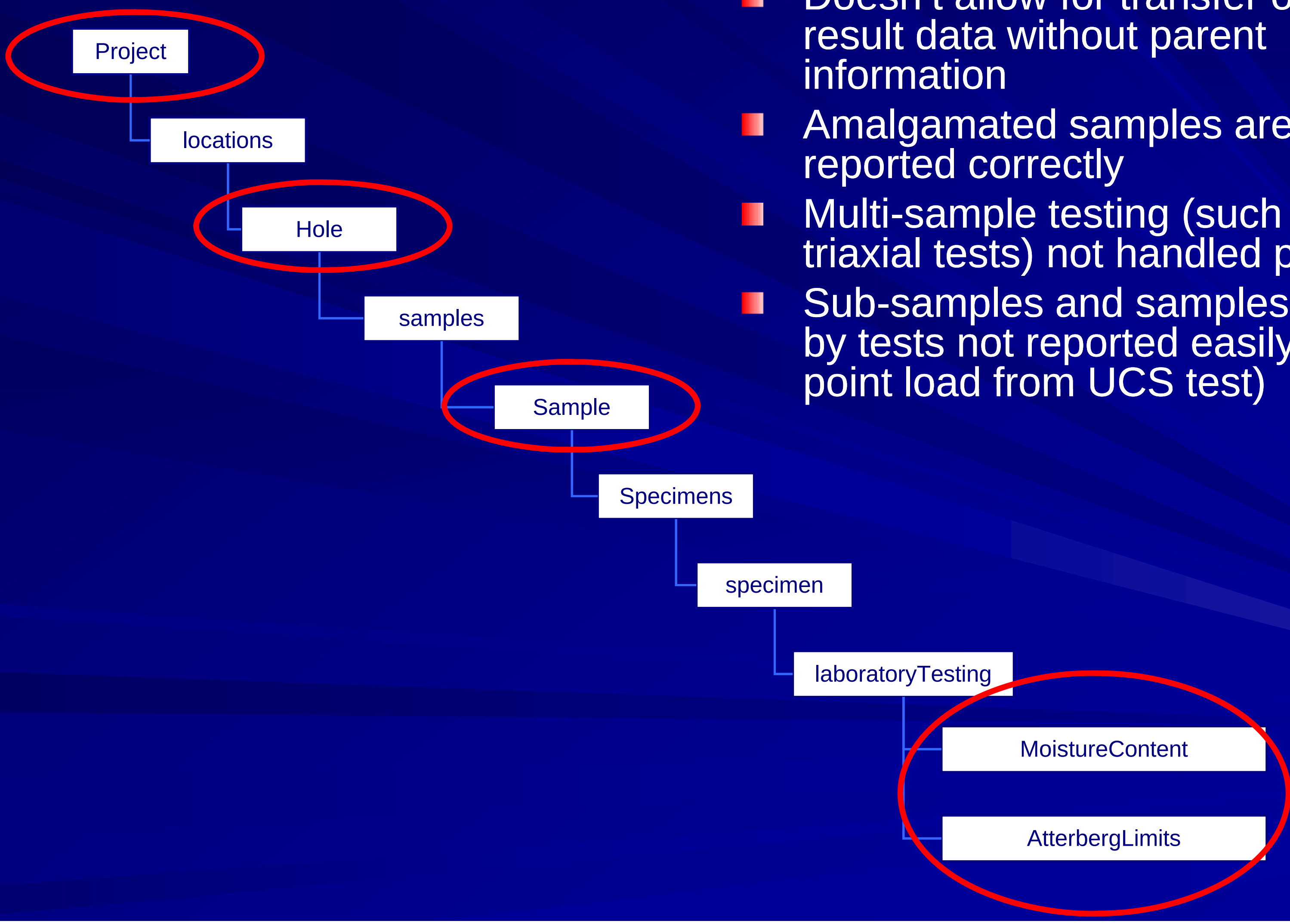
Technical overview of the latest release of DIGGS

Dr Chris Power
Mott MacDonald, UK
AGS Data Management Committee

Aims

- To outline the key concepts of the DIGGS data model
- To relate DIGGS to the ‘real world’
- To not mention XML
- To not lose the audience!

Early DIGGS – Based on UK AGS model



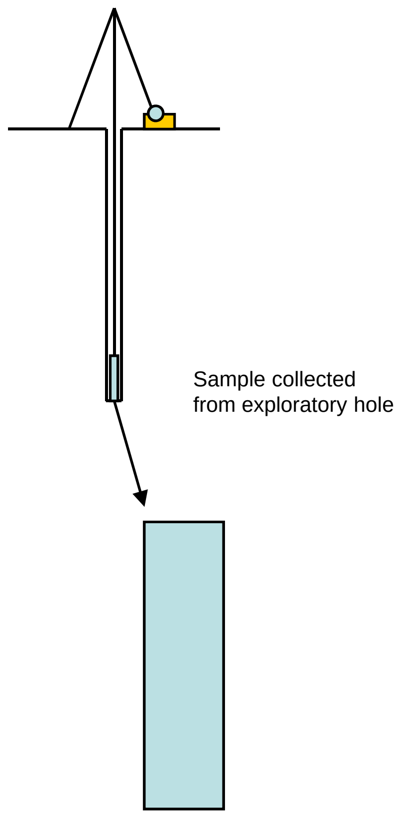
- Doesn't allow for transfer of test result data without parent information
- Amalgamated samples are not reported correctly
- Multi-sample testing (such as triaxial tests) not handled properly
- Sub-samples and samples created by tests not reported easily (eg point load from UCS test)

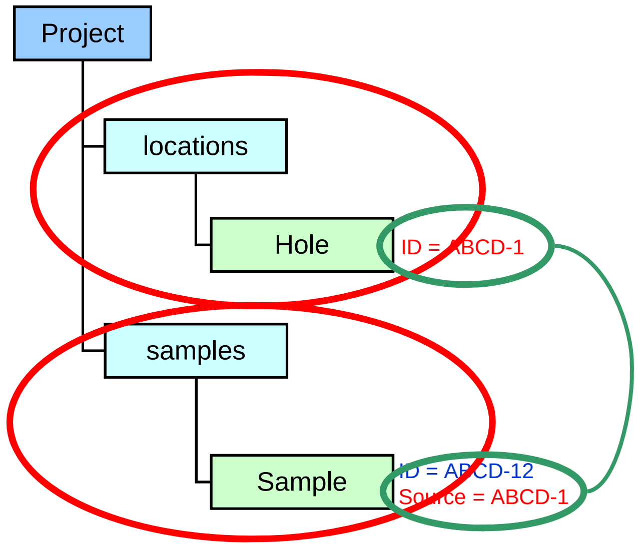
DIGGS development

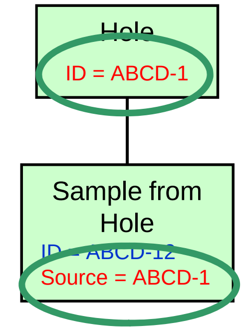
- Twelve versions of DIGGS have been created to date
- Latest release is v1.0a
- v1.0a was draft release for public consultation
- Following examples are based on v1.0a release

Example 1 – Sample Taken from an Exploratory Hole

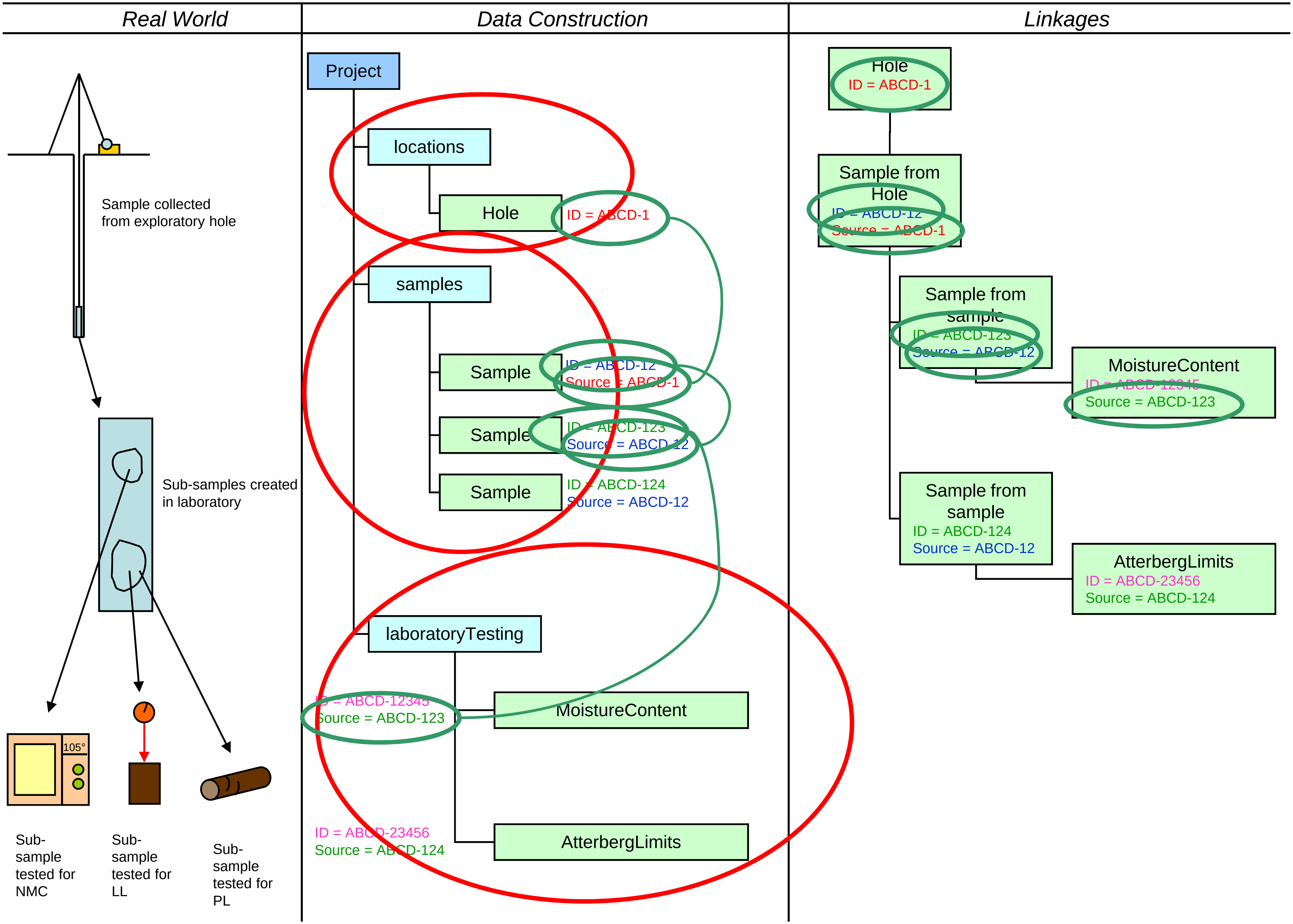
Real World	*Data Construction*	*Linkages*



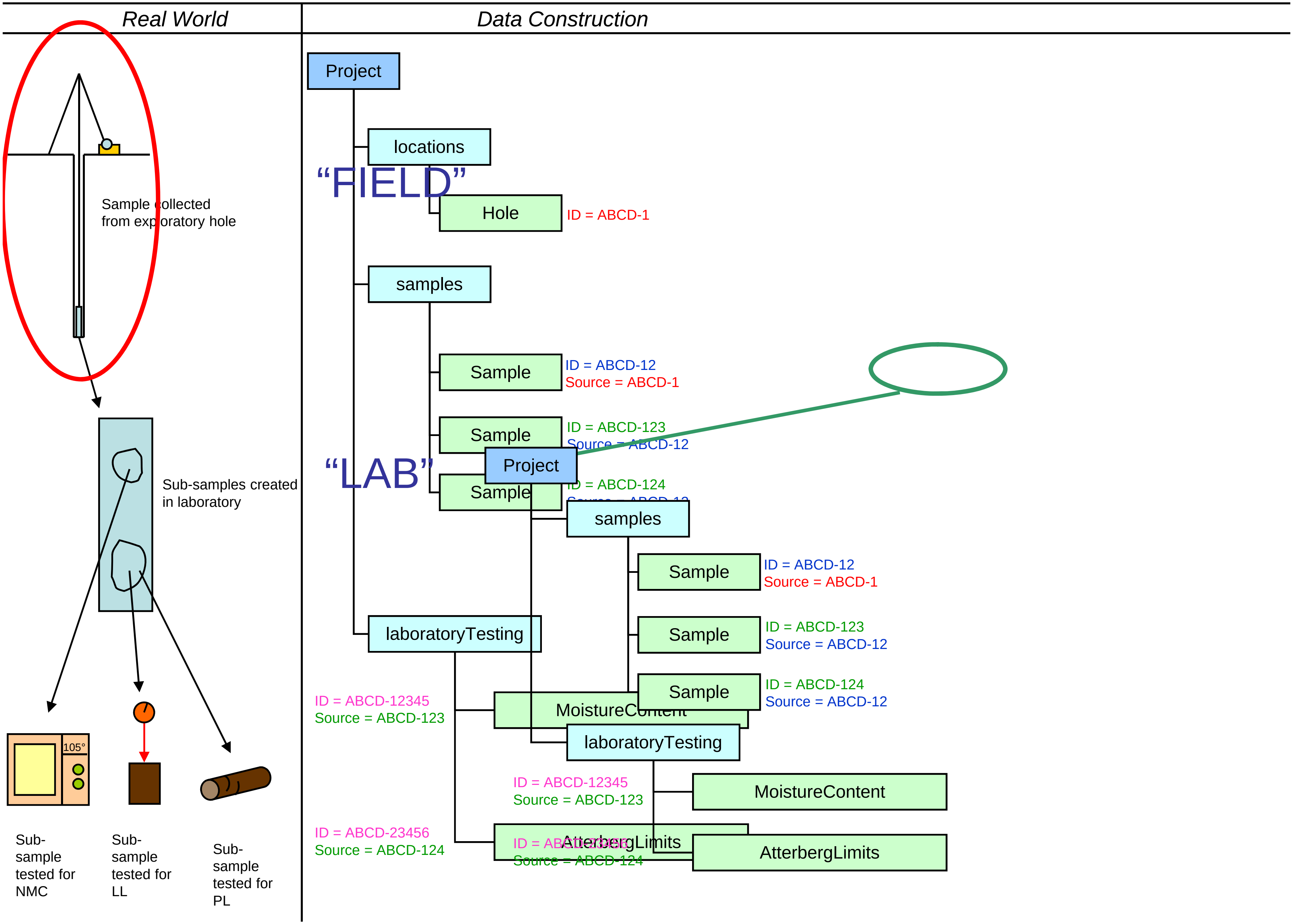




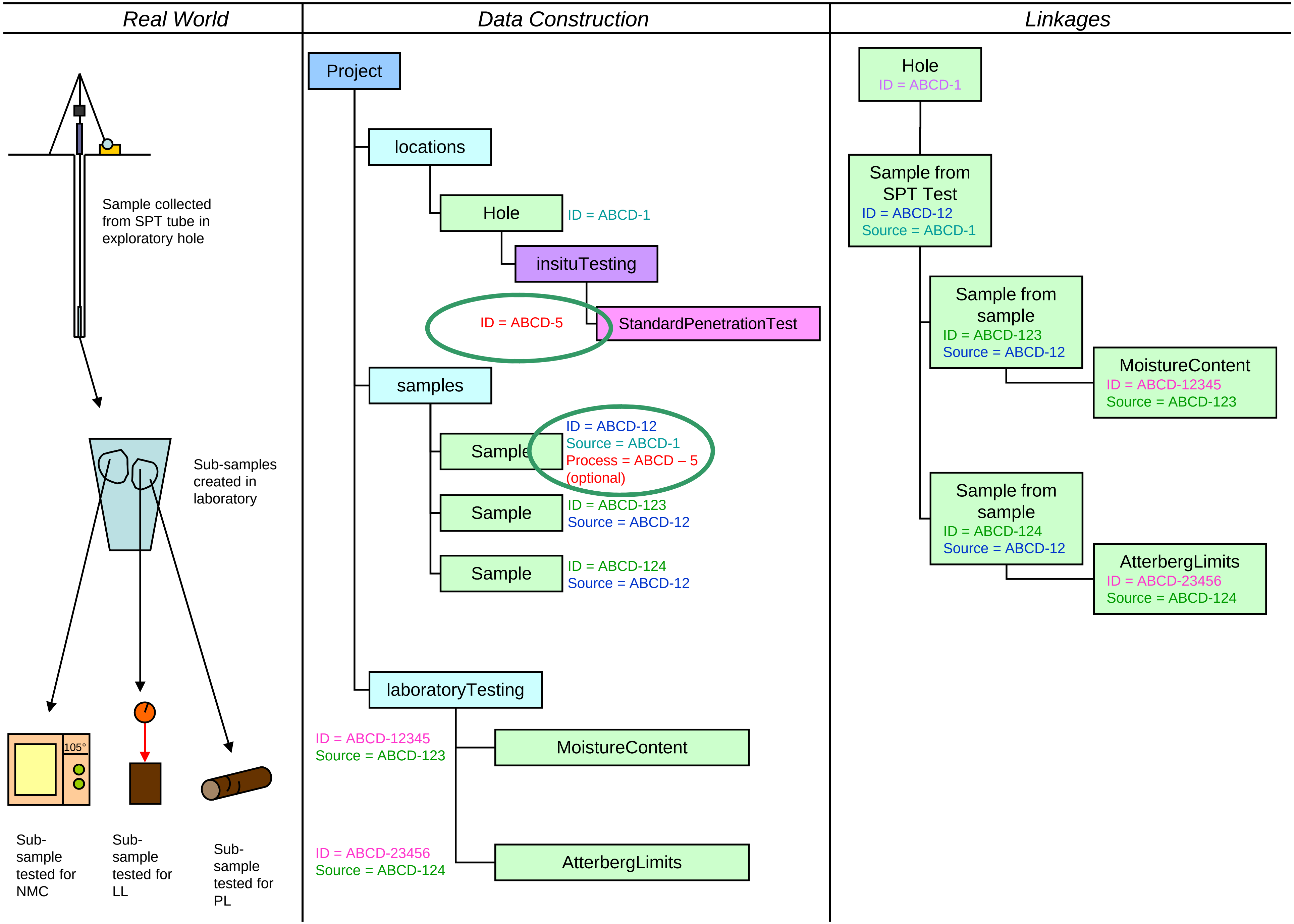
Example 2 – Sample Taken from an Exploratory Hole, tested for NMC, LL and PL



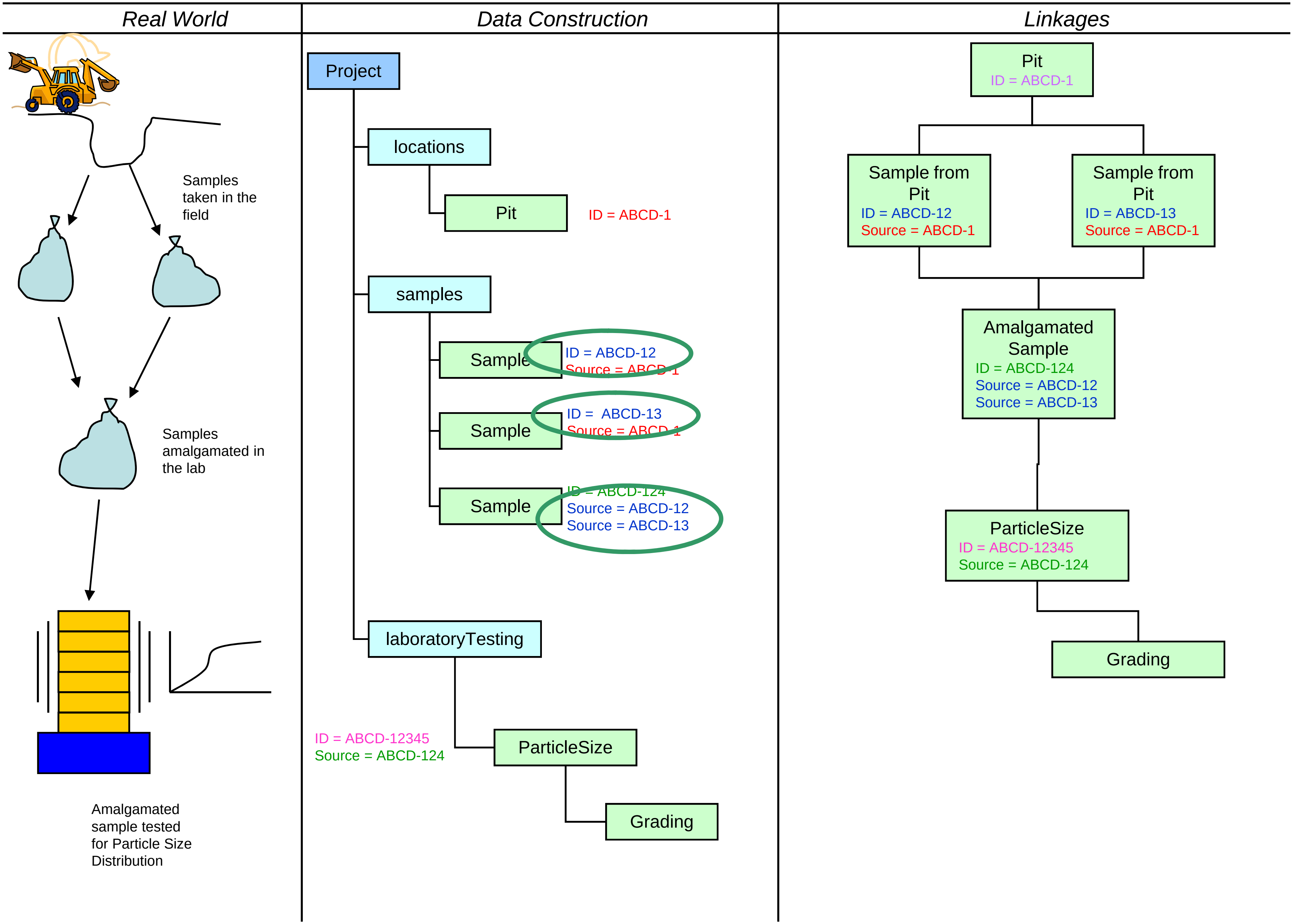
Example 2 – Sample Taken from an Exploratory Hole, tested for NMC, LL and PL



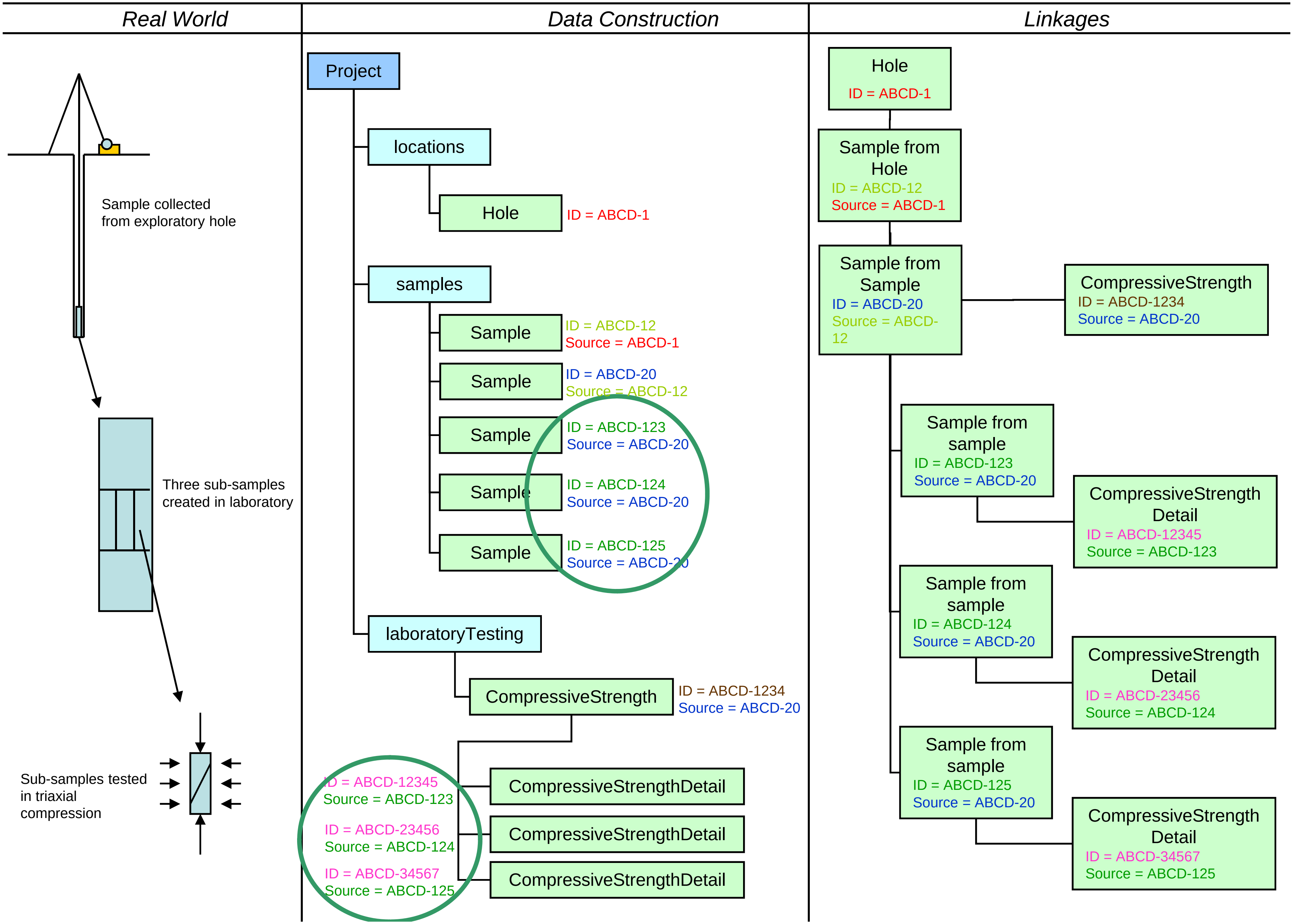
Example 3 – Sample Taken from an SPT in an Exploratory Hole, tested for NMC, LL and PL



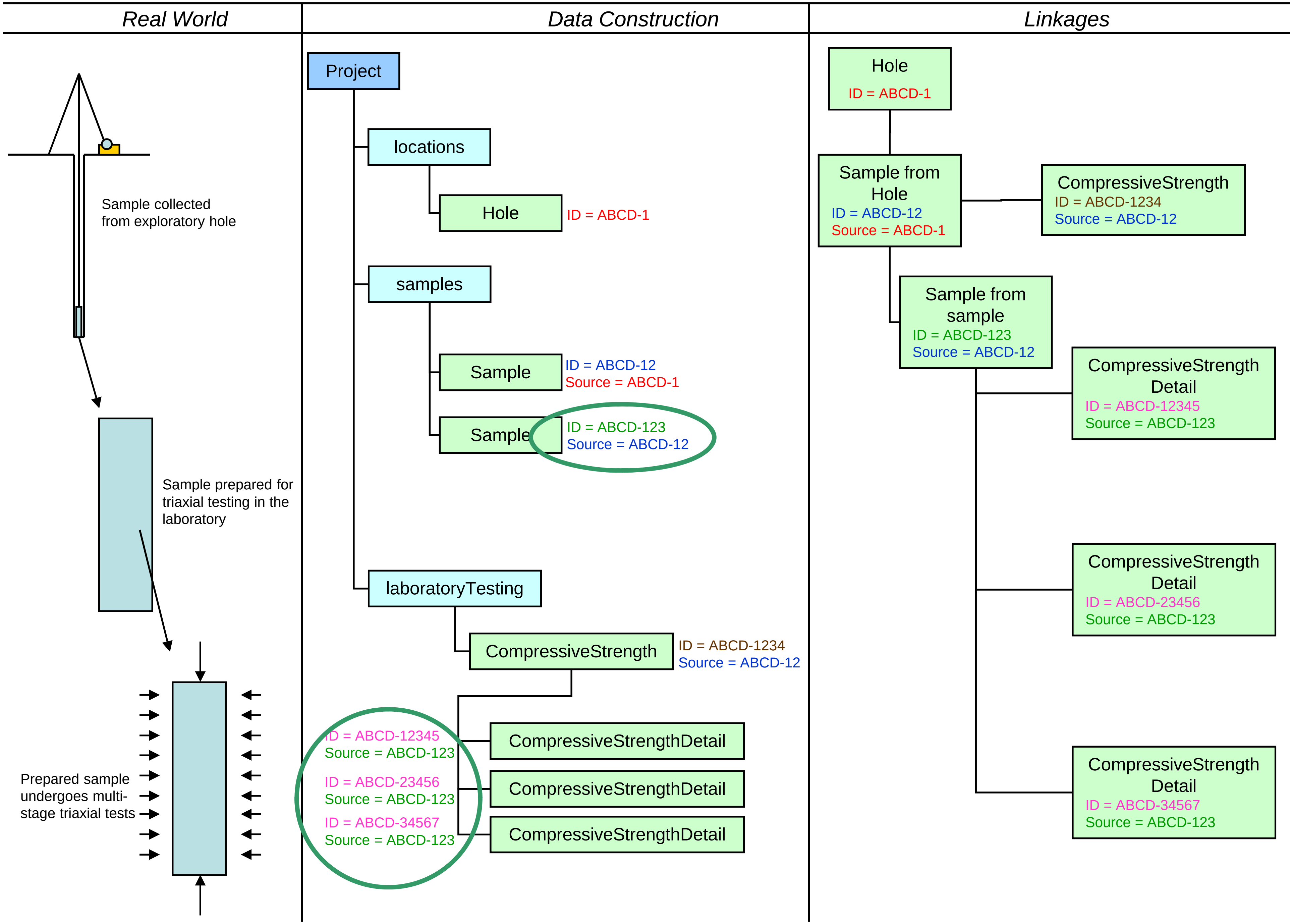
Example 4 – Two samples taken from trial pit, amalgamated and tested for PSD



Example 5 – Sample Taken from an Exploratory Hole, tested by 3 stage, 3 sample triaxial test

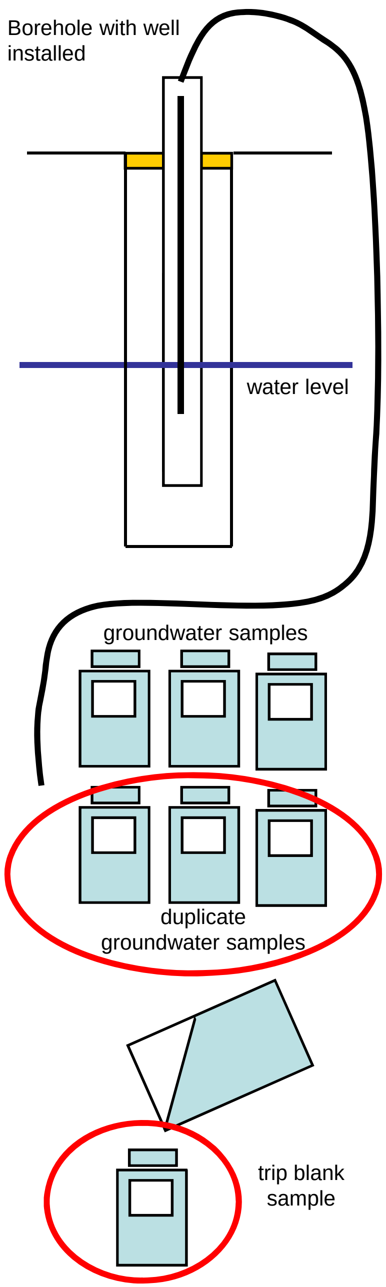


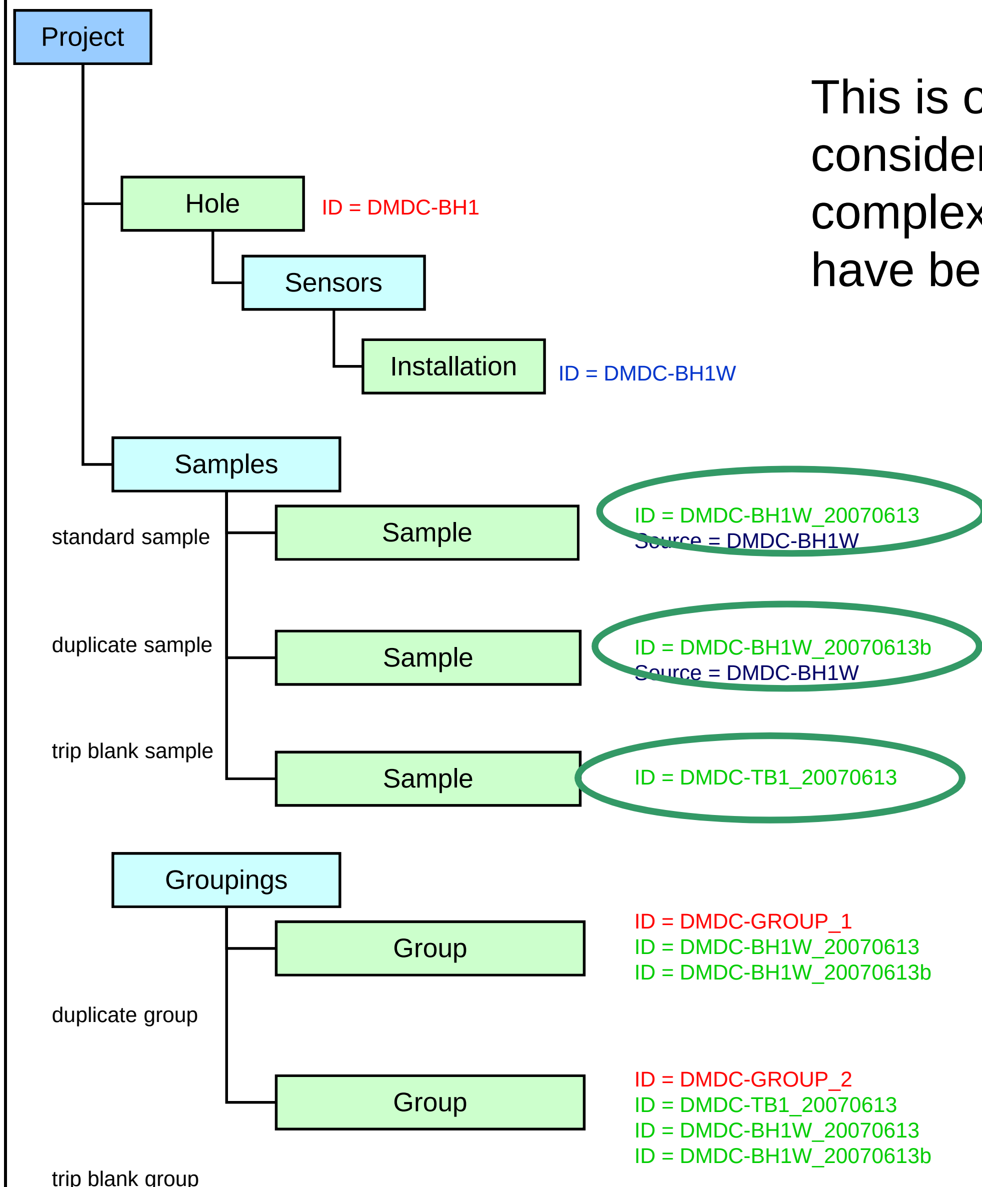
Example 6 – Sample Taken from an Exploratory Hole, tested by 3 stage, 1 sample triaxial test



Example 7 – Geoenvironmental: Field Quality Control Samples

Real World	*Data Construction*





This is one example of a considerable number of complex examples that have been considered

ASSOCIATION OF GEOTECHNICAL &
GEOENVIRONMENTAL SPECIALISTS

Feasibility Study into Implementing DIGGS in the UK

Roger Chandler

AGS Representative on DIGGS GDC

March 2009 - Orlando

Today's presentation

- Summarise DIGGS review work carried out by AGS committee
- Present the AGS's current position in relation to DIGGS and Eurocodes
- Highlight lessons learnt and areas of concern

Review 1

- Complete committee review

  2007 – 2008 (started before public release)
  Whole Committee
  (Consultants , Site Investigation Contractors,
  1 commercial + 2 in house development teams)

- Goal

  To understand DIGGS
  To ensure that it was AGS 3.1 compatible

Review 1 – Where did we start?

- How to get started

  No user documentation
  Few real example files
  Little or no knowledge of XML/GML and UML
  No tools to use, create, convert or display data
  Out of Excel comfort zone.

- Result = A lot of wasted time – here's how we wasted it …

Review 1 – Step 1

- Viewed the XML examples in a commercial XML tool (Oxygen)
  - Lots of questions about the schema by looking at the error messages generated
  - Quickly got very technical
  - Concern raised on level of complexity
- XML training day for committee members
  - Basics of XML and Schema definition
  - Compounded frustration and fear of the level of XML knowledge required.

AGS

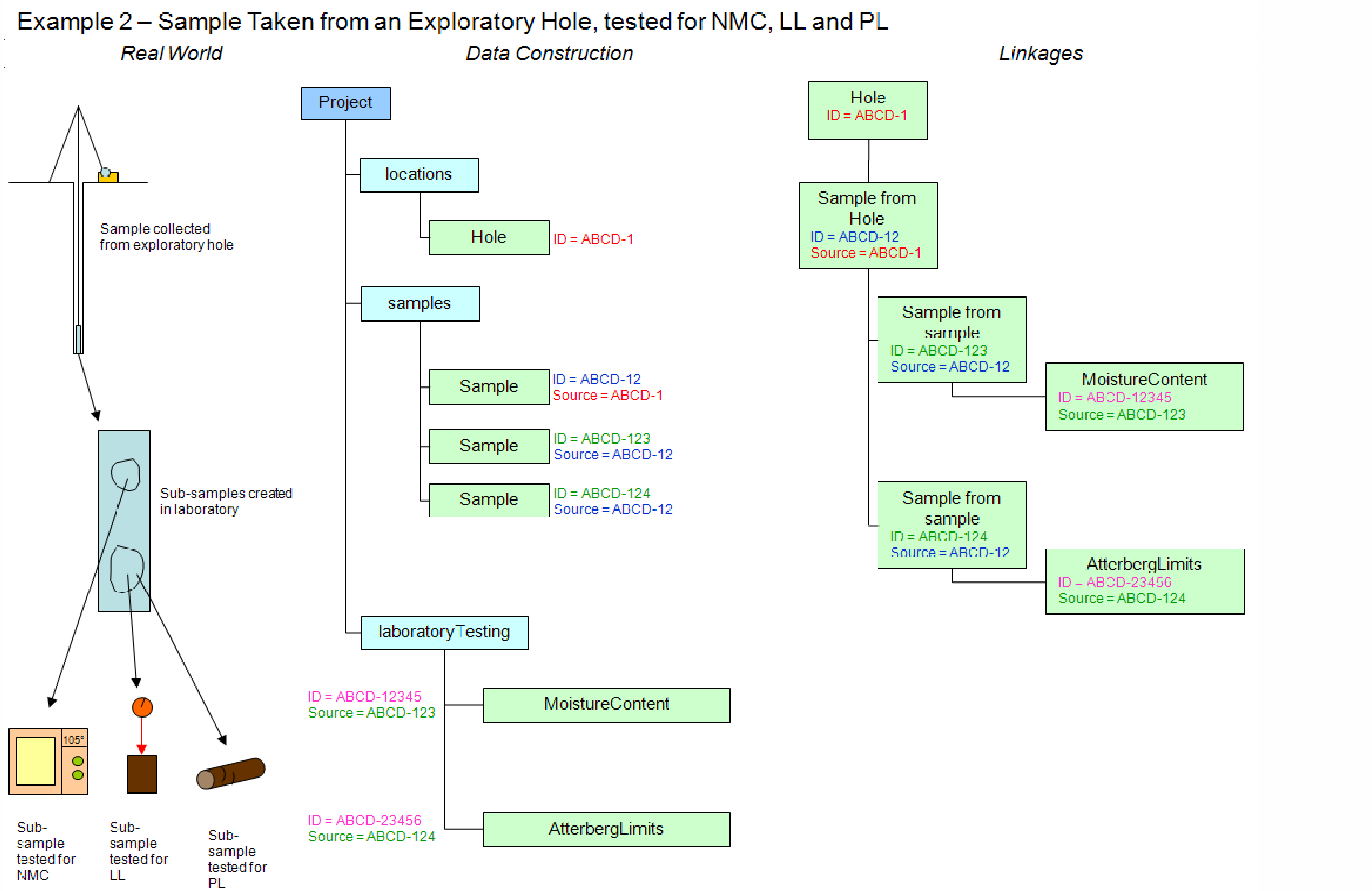
Review 1 – Step 3

- Formed small group

  Roger Chandler, Chris Bray - Keynetix
  Tim Spink , Chris Power – Mott MacDonald

- Chris Power learnt how to create example files using Oxygen – Tim stayed non technical
- Chris up to speed quickly and his understanding grew enough to write sections of the DIGGS user guide
- Banned the use of "XML" and "pointy brackets" in DIGGS presentations

Also produced from this group …



Slide format

- Very effective at communicating DIGGS structure
- Used at AGS User Group meeting June 2008
- Used for US and UK environmental SIGs
- Used in the User Guide extensively

- Good overview but not detailed enough to help the main committee review AGS 3.1 compatibility.

Lessons learnt in this review

- 1. The complex nature of the nested reference objects makes it difficult to present the schema or data in style that can be reviewed easily. This makes it difficult to fully understand what can and should be stored for each object without a custom application.
- 2. The complete UML diagram is overly complex and means little to most of the people we have shown it to. It would be much more useful if it was broken down to segments such as Geology, Wells, Lab testing etc.

Interoperability with Other Datasets

- 3. The “pointy brackets” view of XML examples should only be used as a last resort when trying to explain the format to engineers.
- 4. It is important to differentiate between the domain experts’ job and the Schema experts’ job to ensure that the members of each group understand their roles and that they do not need to fully understand the theory of the other group. However, some basic understanding of each other's roles is essential.

Interoperability with Other Datasets

- 5. The use of Hierarchy diagrams against real life examples is very useful and if used from the outset of assessing DIGGS could potentially save a lot of time.
- 6. Guided creation of data files assists significantly in the understanding of the format.

AGS and Eurocodes

- AGS committee has set itself the goal of revising the AGS format before the end of 2009 to accommodate additions in data transfer mandated by Eurocodes.
- The committee is currently reviewing what the additional requirements are.
- Should the release at the end of 2009 be CSV or XML?
- Sub group formed to answer this question.

Review 2 – Step 1

Romain Arnould – Fugro
Jackie Bland - Geotechnics/Fugro
Chris Bray – Keynetix
Salvatore Caronna – gINT
Roger Chandler – Keynetix
Chris Power – Mott MacDonald
Peter Whittlestone – Arup
Mark Bevan - Structural Soils

Review 2 - Goals

- Goals

1. Be able to open the schema in free software, not just proprietary xml viewers.
2. To test whether DIGGS can work with a local implementation.
3. To be able to understand the schema and explain it to others.
4. To justify complexity in the schema where we can't remove it from our implementation.
5. Ensure the schema and an AGS extension meets all AGS 4.0 requirements and is DIGGS compliant.
6. We don't/won't do politics.

AGS

Review 2 - Examples

Chunk No	Area of Interest	Equivalent AGS Table
1	Project Data	PROJ
2	Hole Data	HOLE
3	Sample Data	SAMP
4	SPT Data	ISPT
5	DP Data	DPRG/DPRB
6	Lab Testing Data	CLSS/TRIG/TRIX
7	Wells and Installations	MONP/MONR
8	Geology	GEOL
9	CPT Data	STCN

Review 2 - Software

- Lots tried – only a few could be used

1. XML Spear editor (http://www.donkeydevelopment.com) can be used to create and validate DIGGS files, this is free for personal use.
2. XML Copy Editor (http://xml-copy-editor.sourceforge.net) is also a simple-to-use validating XML editor that is free to use.
3. XML Marker (http://symbolclick.com) is a useful tree-view and text XML editor. It provides tools to browse the data hierarchy and is free to use.
4. Oxygen (http://oxygenxml.com) is the only commercial application that we used – AGS granted academic licences

Review 2 – Simple Validation

- Validation takes too long using online schemas for DIGGS
- Catalog files not supported by the majority of free XML software tools
- Local referencing of schemas enabled DIGGS to work with software not able to use Catalogs

Review 2 – Complex Validation

- Areas where we feel XML schema based validation is lacking

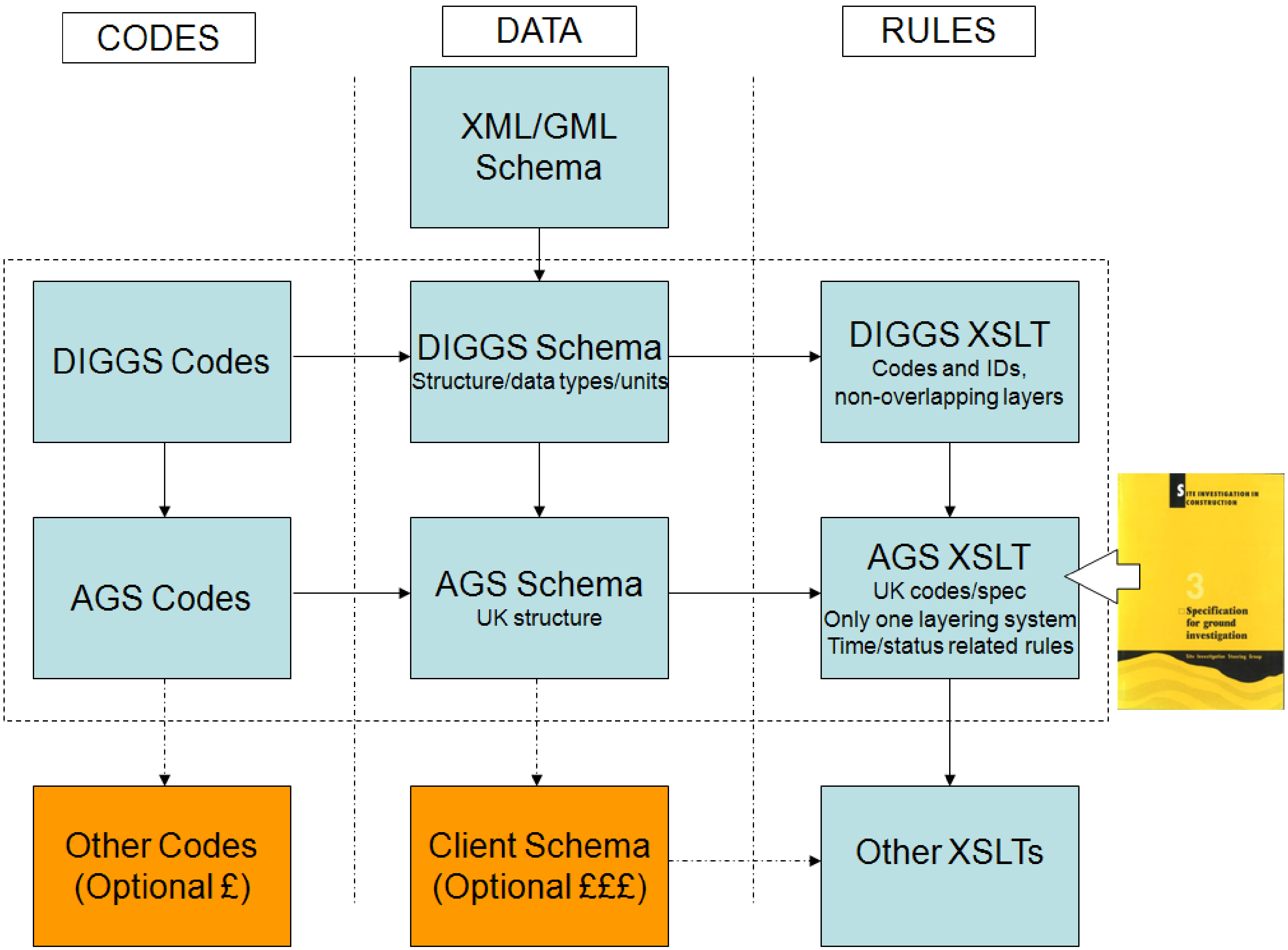
  IDs
  Non overlapping Layers
  codeSpace values

- Areas where the AGS would like to validate more in local implementation

  Only 1 layer system
  2 seating blows + 4 main blows for SPT
  Decimal Places and Significant Figures

AGS

DIGGS – How validation could work



DIGGS data checker

- DIGGS should commission an independent checker to check DIGGS data using Schematron rule definitions.

The rules files could be downloadable from the DIGGS, AGS or client's websites depending on whether they were international, local or client based rule definitions.

GML

- Appears to make schema far more complicated
- Couldn't get DIGGS files to work in 3 big GIS systems (AutoCAD Map, ArcMap, MapInfo) or free viewers such as Tatuk GIS viewer
- Did get the GML viewer from Snowflake Software to view the data but not attribute it
- Is the inclusion of GML sufficiently benefitial?

Schema Definition Problems

- Many found (sample below)

  Mandatory elements not tagged as such
  SPT and Dynamic Probes need complete review
  Lab testing allows for tests to be carried out on more than one specimen.
  CPT structure could be used for other data types such as pressuremeter. dialotometer, dynamic probes. Technical questions raised over current table structure
  Table by table check still required – many questions remain unanswered on the forum.

DIGGS Complexity

- Some areas have been identified that may bring added complexity with little benefit

  (Roles, Business Associates, Specification and Equipment, Remarks)

- Not easy to remove objects and properties from local implementation.

  Makes a local implementation more complex

- Too object orientated

  May create significant problems for some implementations

Lessons learnt

- Example creation has been extremely important on the education road to DIGGS.
- It would have been useful to have either an AGS to DIGGS converter or export routines from commercial software applications such as gINT and HoleBASE.
- Communication between LIG and DIGGS main committee should happen at an early stage and questions need to be answered quickly and independantly

Will AGS 4 be CSV or XML?

- The review committee will await the outcome of the DIGGS meeting on the 25th March in Florida before it makes any recommendations to the AGS data management committee.
- An unofficial vote at the last meeting indicated that in its current format DIGGS would not be used for AGS 4
- We believe there is still many months of technical and governance work to be completed by the DIGGS organisation.

AGS

DIGGS and the COSMOS/PEER-LL GEOTECHNICAL VIRTUAL DATA CENTER

DIGGS Invitational Meeting, March 25 and 26, 2009

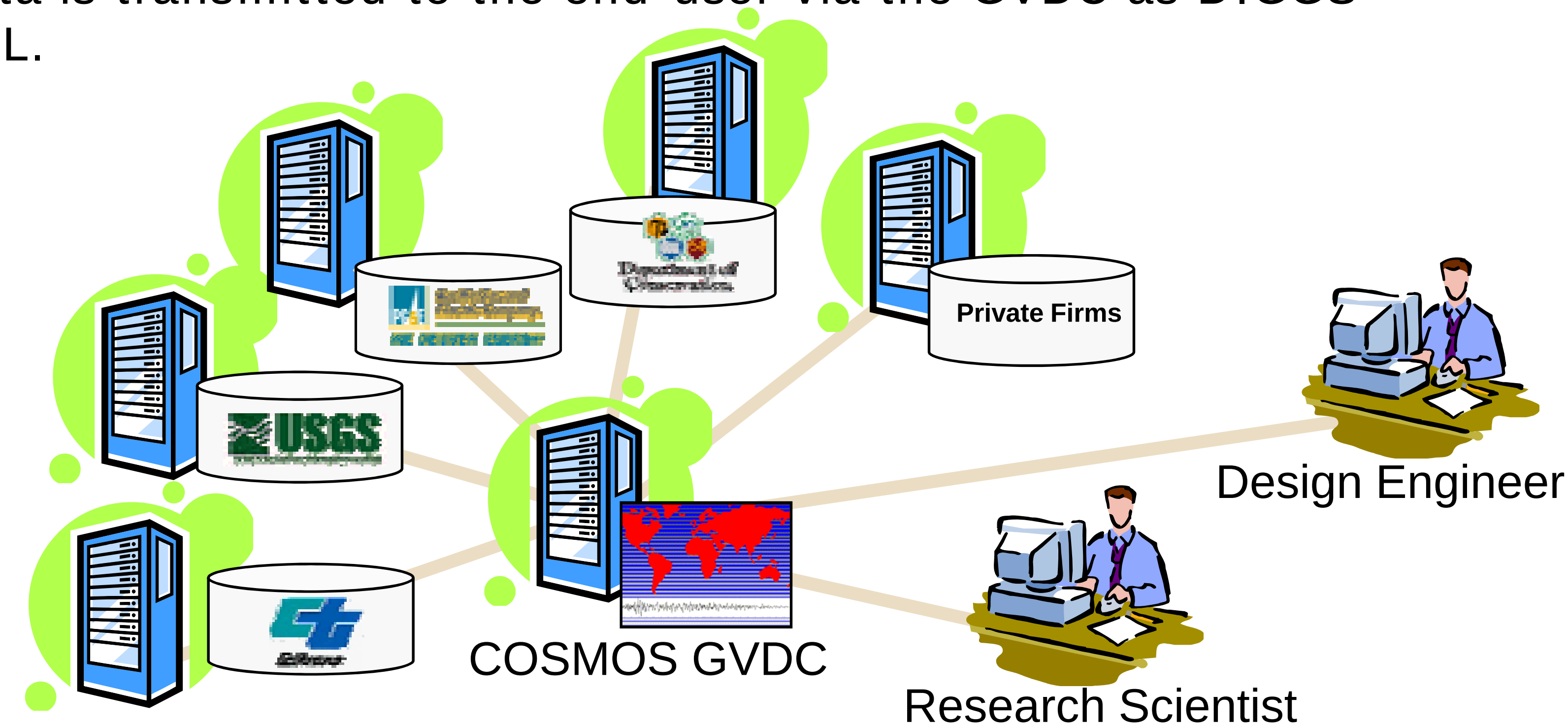
Daniel Ponti, USGS
Loren Turner, CALTRANS

Overview

- The COSMOS/PEER-LL Geotechnical Virtual Data Center (GVDC)
- User Experience
- Behind the Scenes
- GVDC Timeline

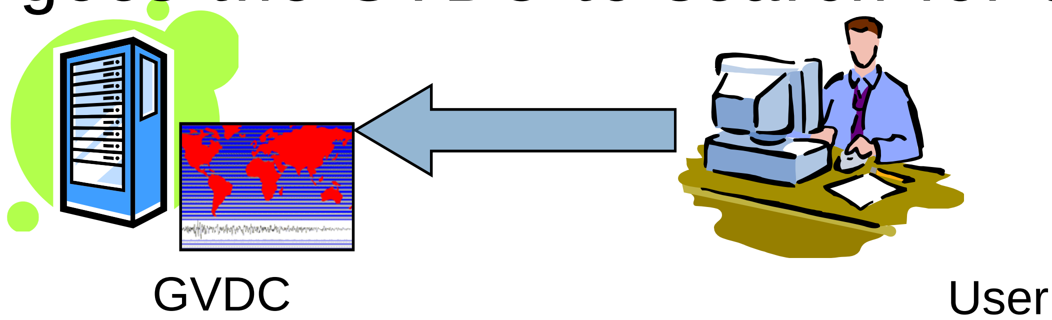
Geotechnical Virtual Data Center

- The GVDC is a web application that acts as a “broker” for geotechnical data. It is not a data repository.
- Data is held by registered data providers who maintain their data in their own proprietary systems, and make available to the GVDC only the data they choose.
- Data is transmitted to the end-user via the GVDC as DIGGS XML.

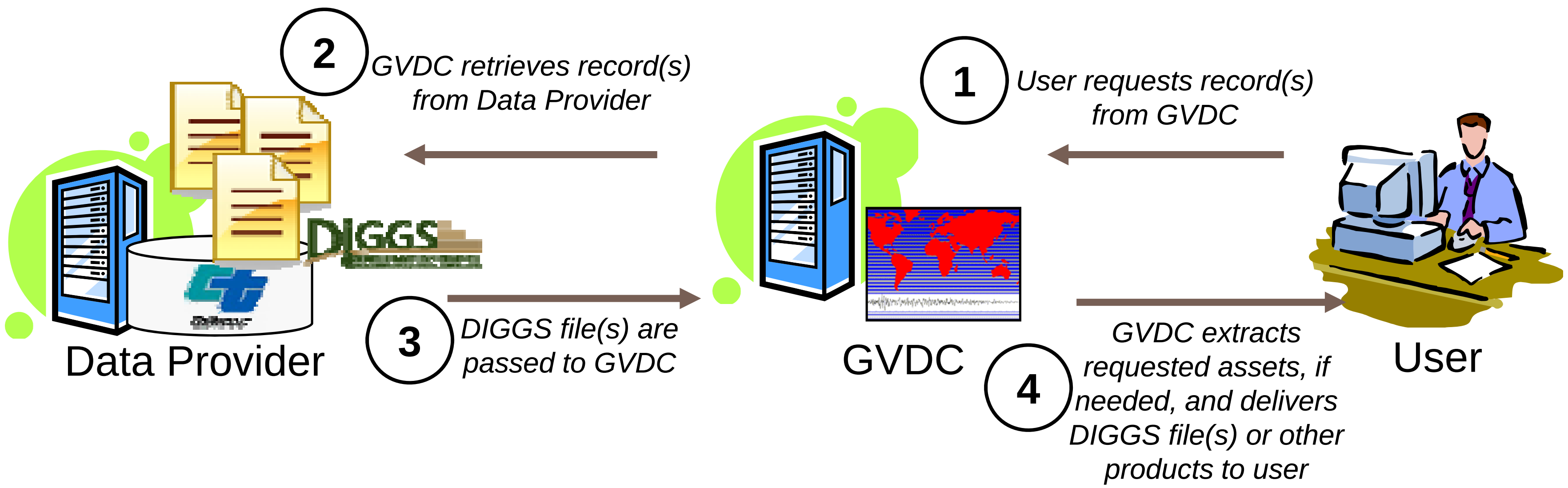


User Experience

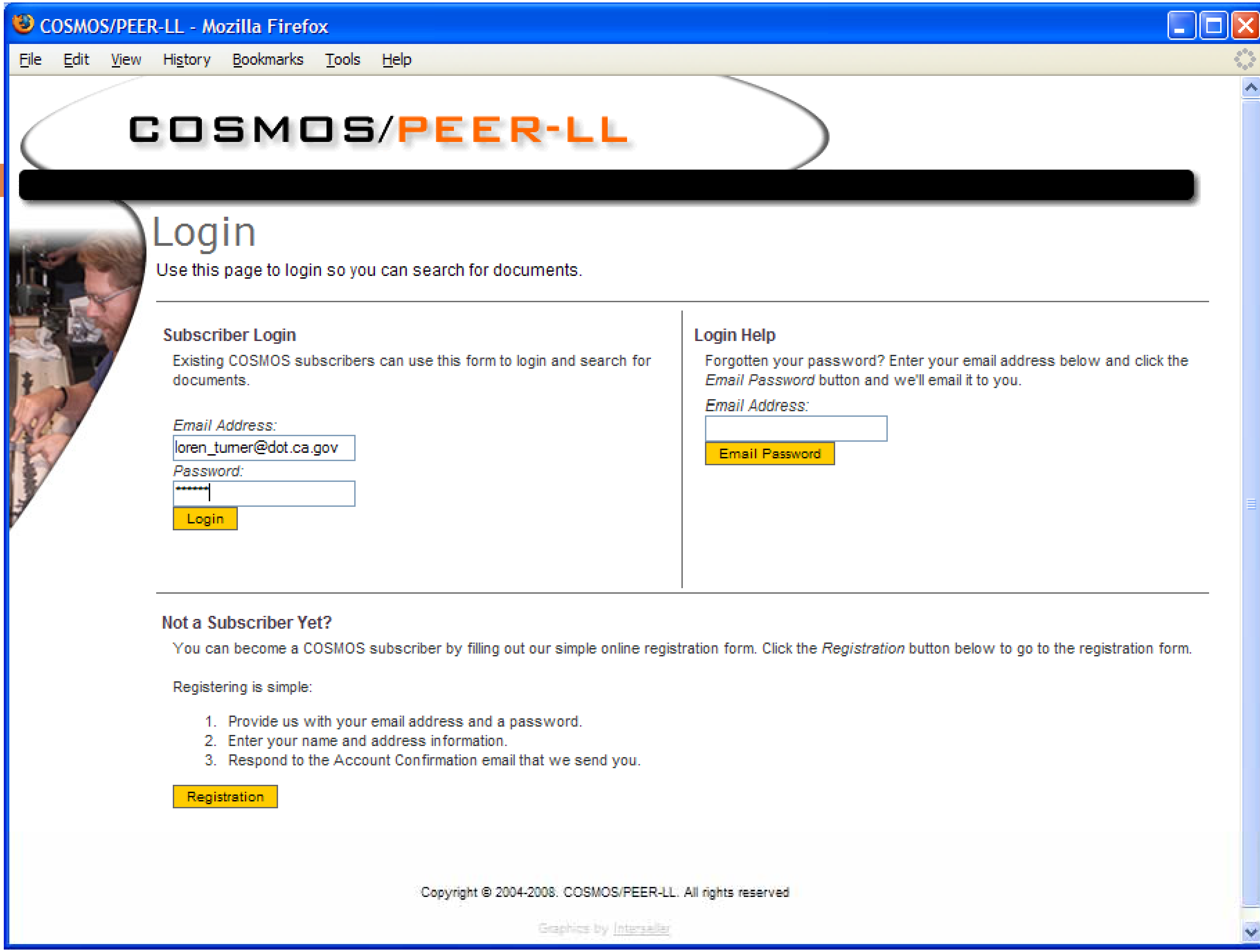
- A user goes the GVDC to search for data



- The user requests to download and/or preview the record(s) returned by the search process.




My Local PHP Index - Mozilla Firefox
File Edit View History Bookmarks Tools Help
COSMOS/PEER-LL
Geotechnical Virtual Data Center
PROJECT 2L03
HOME
PROJECT INFO
ABOUT
CALENDAR
USER SURVEY
FORUM
NEWS AND EVENTS
Geotechnical Virtual Data Center (GVDC)
The overall project is divided into a short-term and a long-term objective. The project we are now undertaking encompasses the short-term objective only, to develop a pilot web-based system linking the PG&E, Caltrans, CGS and USGS example geotechnical data sets. The long-term objective (a future project not yet funded) is to extend the pilot system and develop a web-based system linking multiple data sets... read more >>
Browse through documentation on the GVDC
Read More
Find geotechnical data from our map-based interface.
User Login
Please see the Project Workshop agenda , June 21-23 '04 in Newport Beach, CA. The results of the user scenario survey more >>
The objective is to develop consensus recommendations for classifying, archiving, and web dissemination of geotechnical data... more >>
COSMOS and the PEER Lifelines Program are coordinating additional workshops and establishing a pilot project leading to... more >>
Sponsored by
CalTrans
California Energy Commission
Pacific Gas & Electric
PEER-Lifelines Program
In Partnership with
Pacific Earthquake Engineering Research Center
United States Geological Survey
California Geological Survey
Implemented by
University of Southern California
Consortium of Organizations for Strong-Motion Observations Systems
Caltrans
PG&E
USGS
science for a changing world


COSMOS/PEER-LL - Mozilla Firefox

File Edit View History Bookmarks Tools Help

COSMOS/PEER-LL

Login

Use this page to login so you can search for documents.

Subscriber Login

Existing COSMOS subscribers can use this form to login and search for documents.

Email Address:

loren_turner@dot.ca.gov

Password:

Login

Login Help

Forgotten your password? Enter your email address below and click the *Email Password* button and we'll email it to you.

Email Address:

Email Password

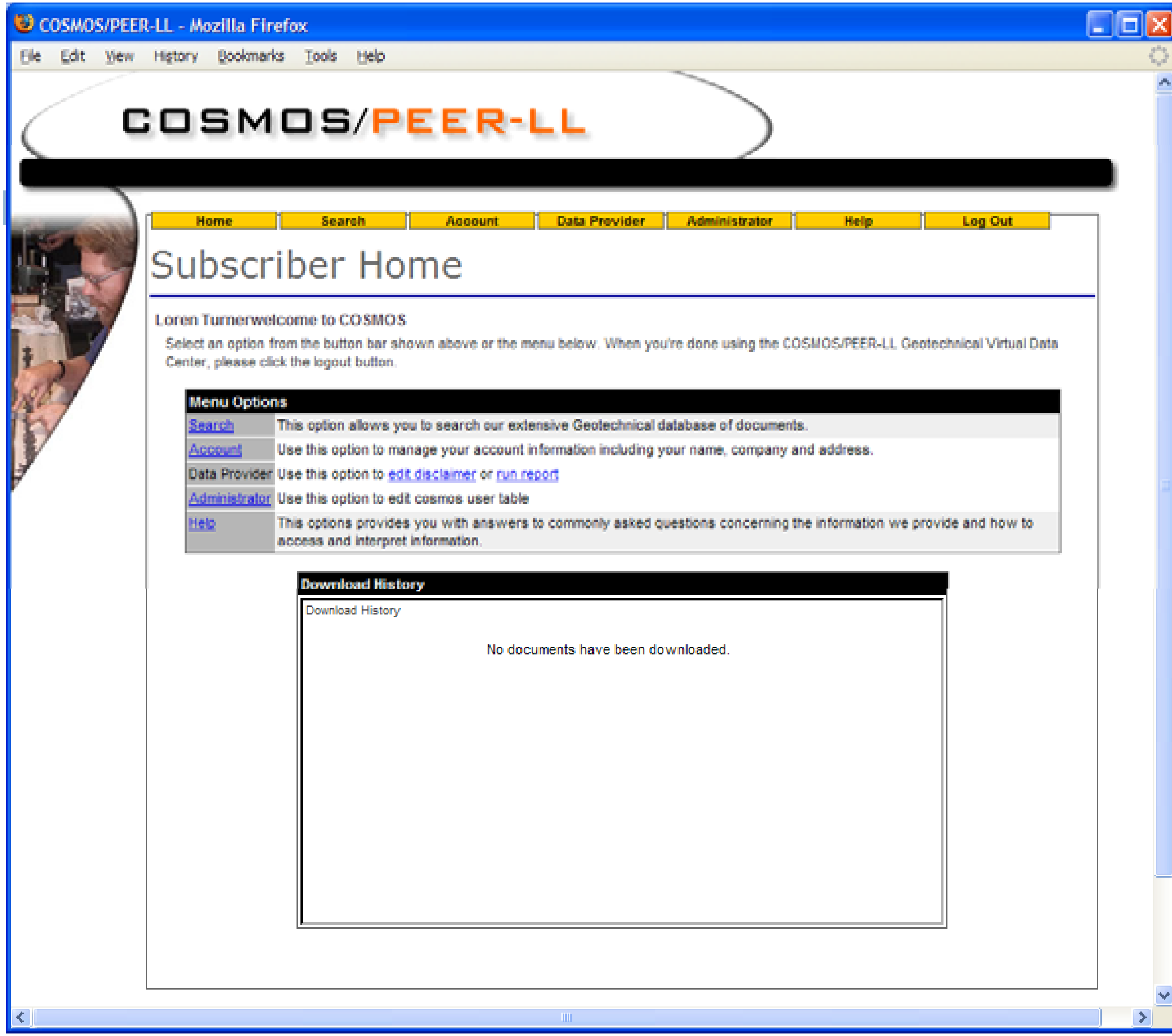
Not a Subscriber Yet?

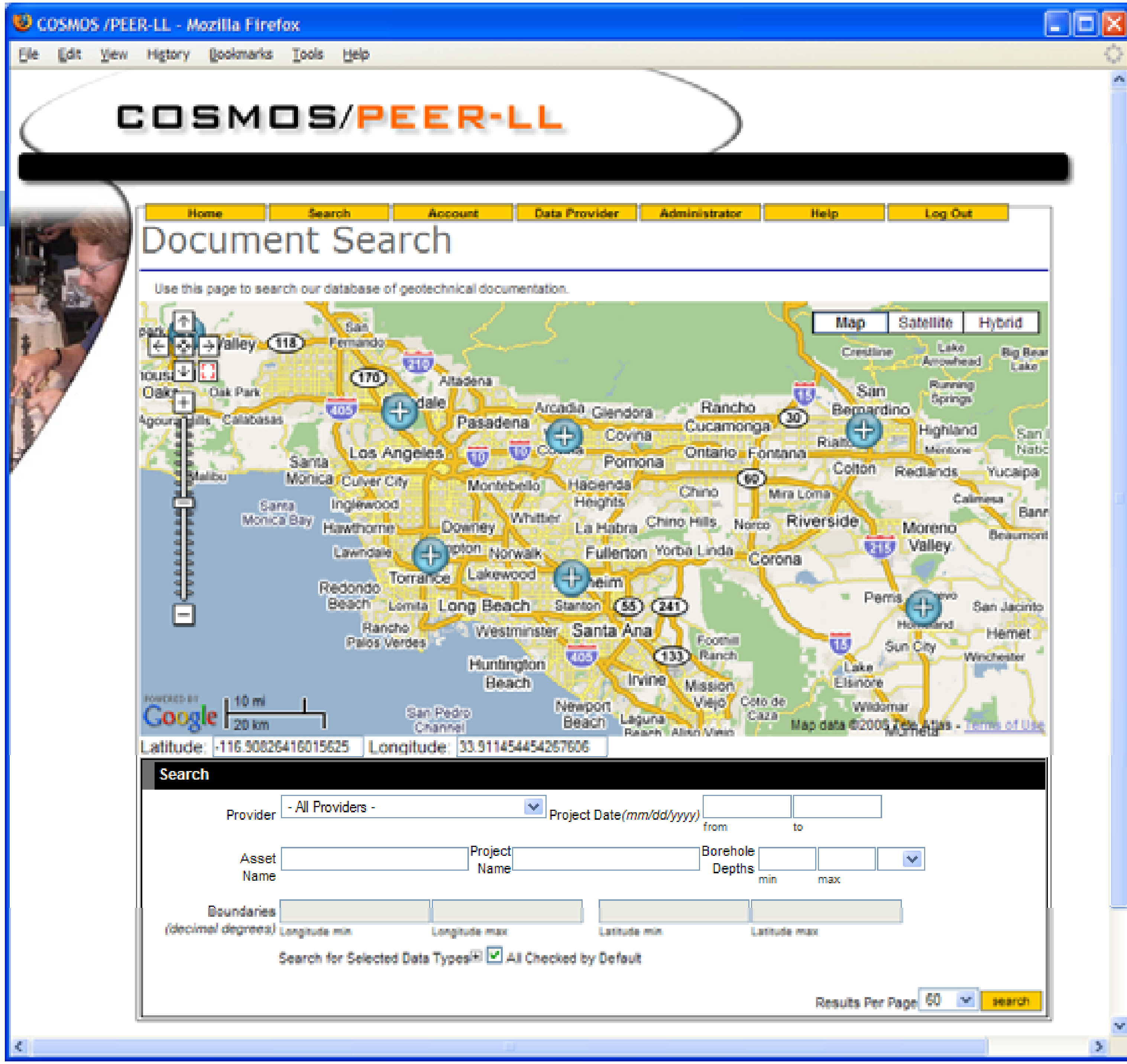
You can become a COSMOS subscriber by filling out our simple online registration form. Click the *Registration* button below to go to the registration form.

Registering is simple:

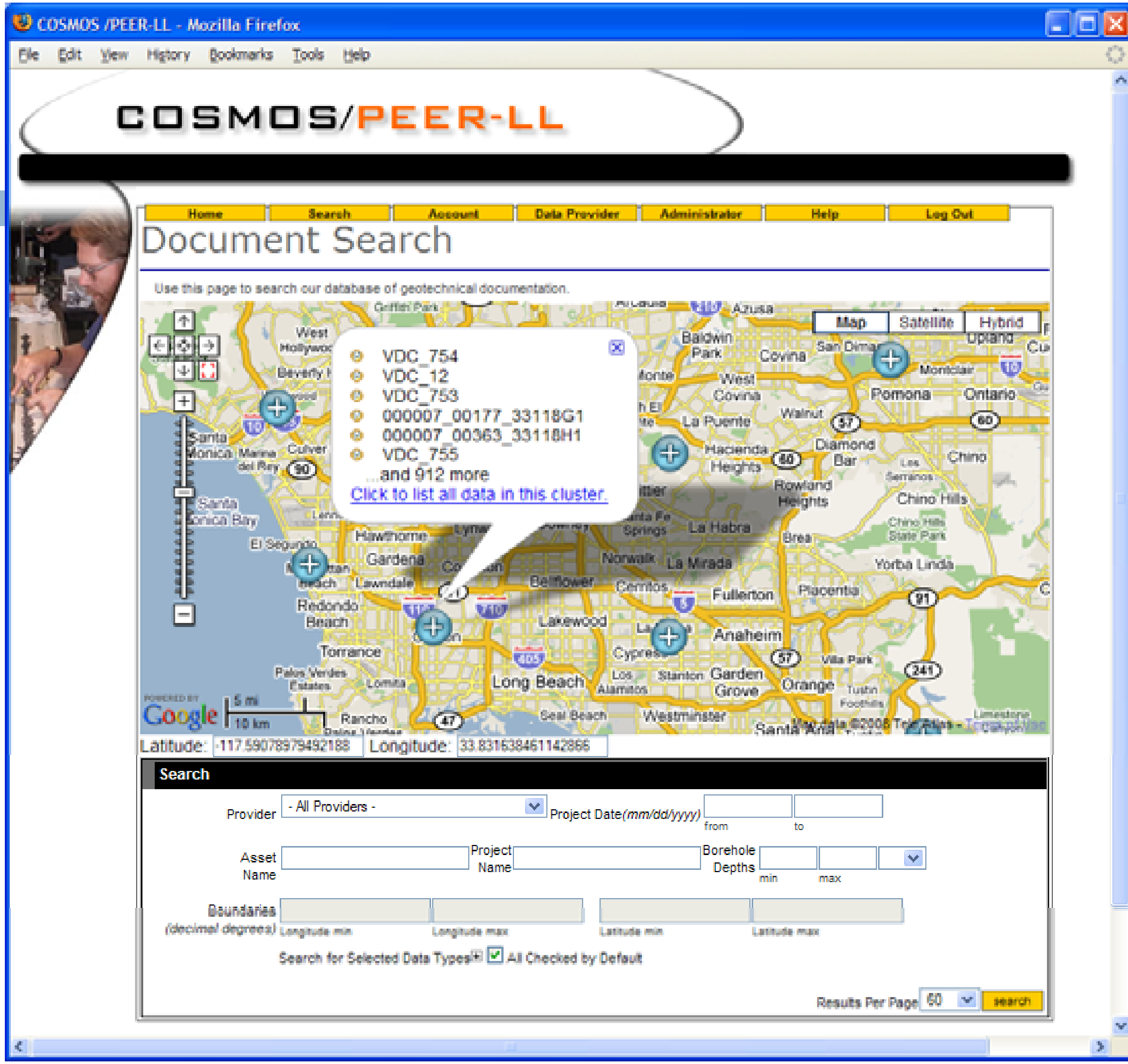
1. Provide us with your email address and a password.
2. Enter your name and address information.
3. Respond to the Account Confirmation email that we send you.

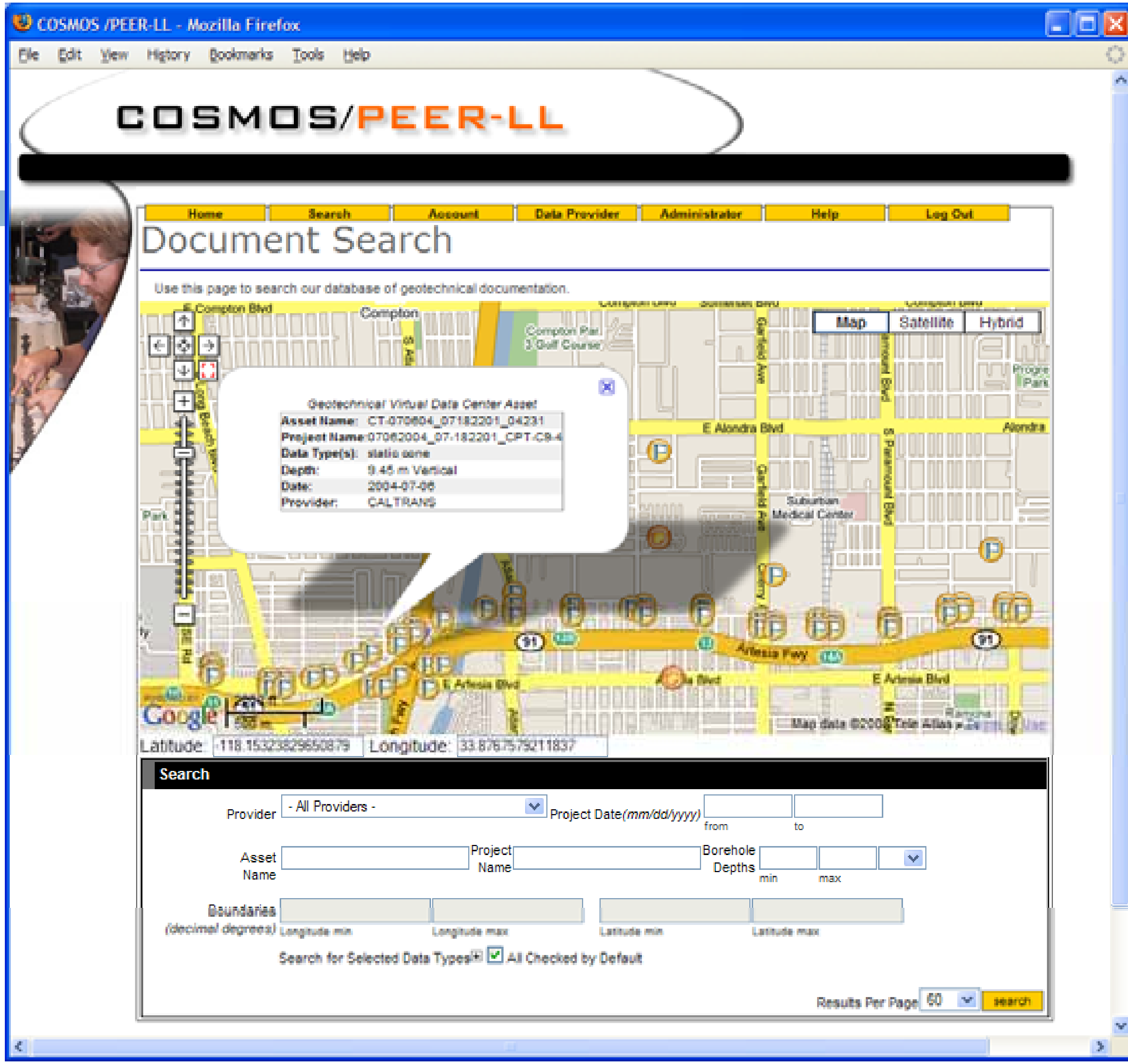
Registration

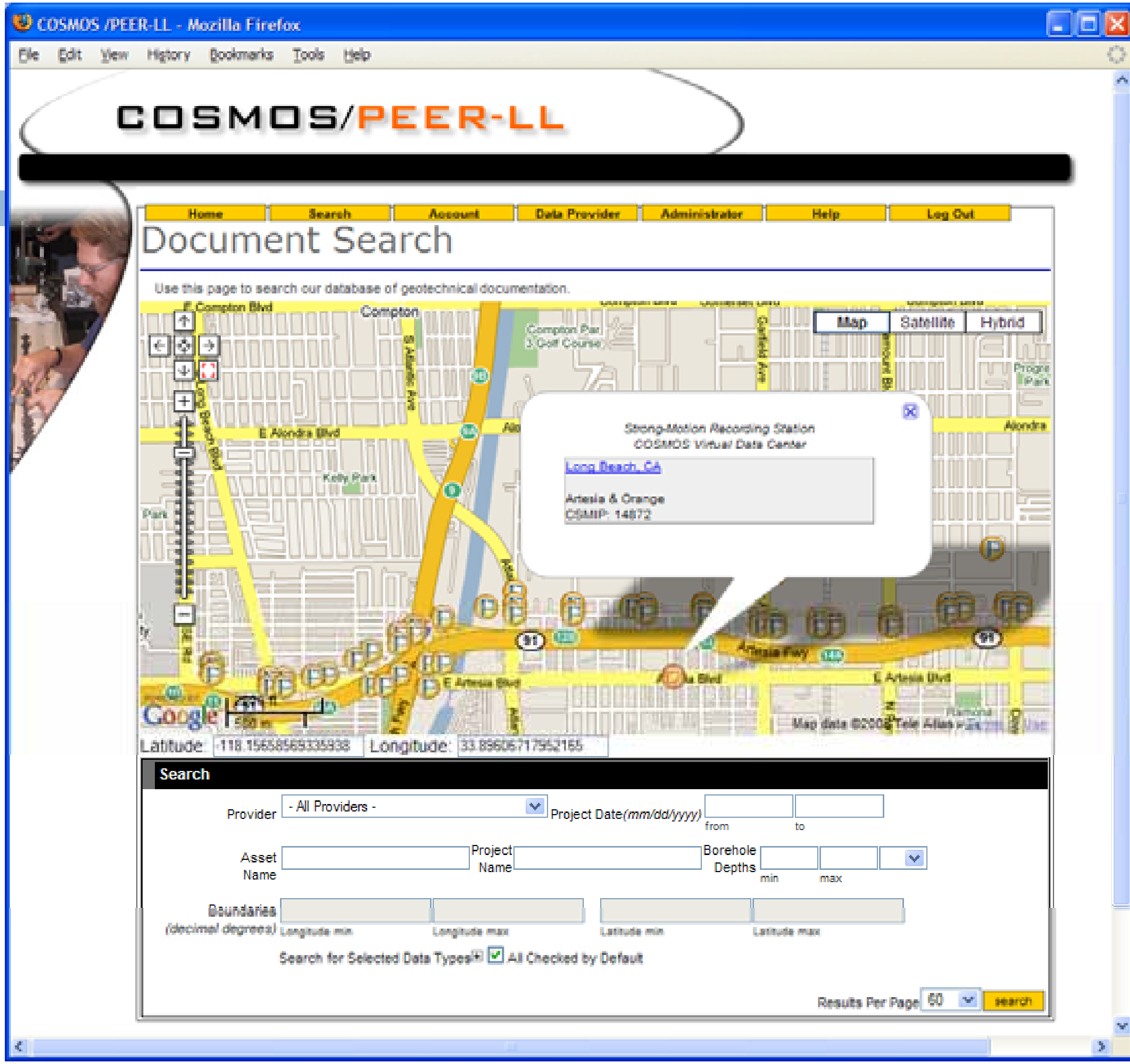

COSMOS/PEER-LL - Mozilla Firefox
File Edit View History Bookmarks Tools Help
COSMOS/PEER-LL
Home
Search
Account
Data Provider
Administrator
Help
Log Out
Subscriber Home
Loren Turnerwelcome to COSMOS
Select an option from the button bar shown above or the menu below. When you're done using the COSMOS/PEER-LL Geotechnical Virtual Data Center, please click the logout button.
Menu Options
Search This option allows you to search our extensive Geotechnical database of documents.
Account Use this option to manage your account information including your name, company and address.
Data Provider Use this option to edit disclaimer or run report
Administrator Use this option to edit cosmos user table
Help This options provides you with answers to commonly asked questions concerning the information we provide and how to access and interpret information.
Download History
Download History
No documents have been downloaded.

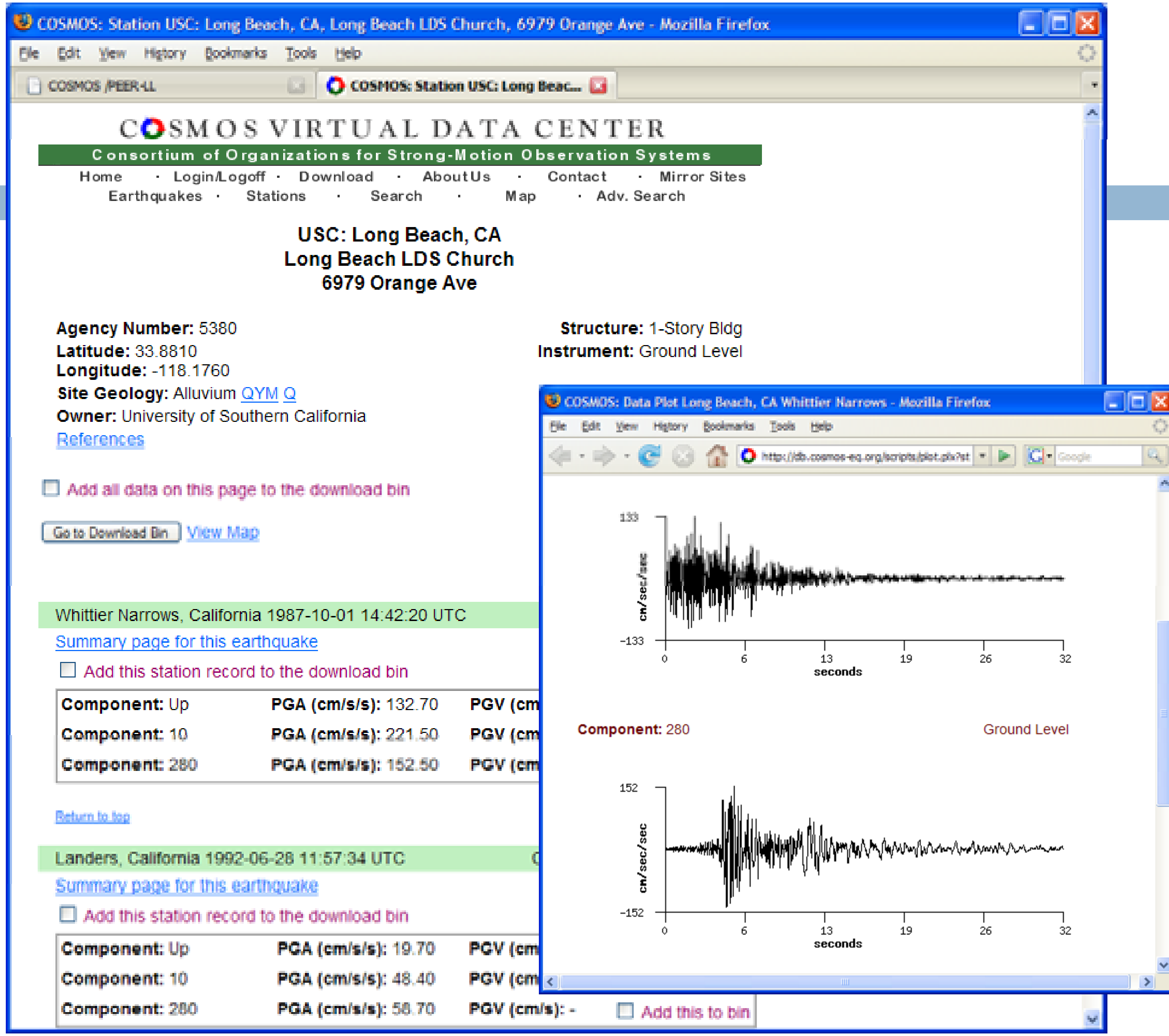


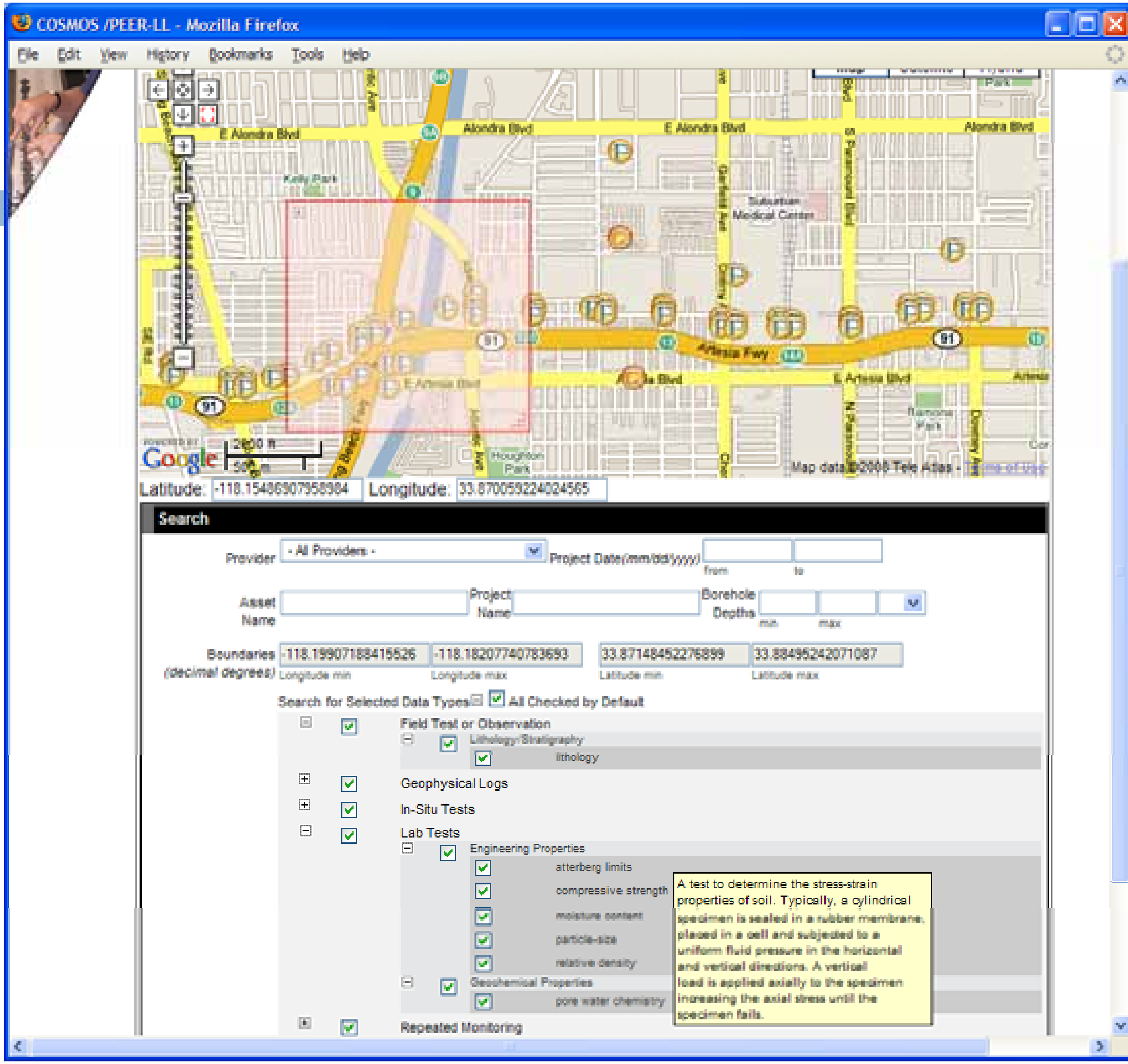
COSMOS /PEER-LL - Mozilla Firefox
File Edit View History Bookmarks Tools Help
COSMOS/PEER-LL
Home Search Account Data Provider Administrator Help Log Out
Document Search
Use this page to search our database of geotechnical documentation.
Map Satellite Hybrid
Latitude: -116.908264160156625 Longitude: 33.9114544542676O6
Search
Provider - All Providers -
Project Date(mm/dd/yyyy) from to
Asset Name
Project Name
Borehole Depths min max
Boundaries (decimal degrees) Longitude min Longitude max Latitude min Latitude max
Search for Selected Data Types All Checked by Default
Results Per Page 50 search

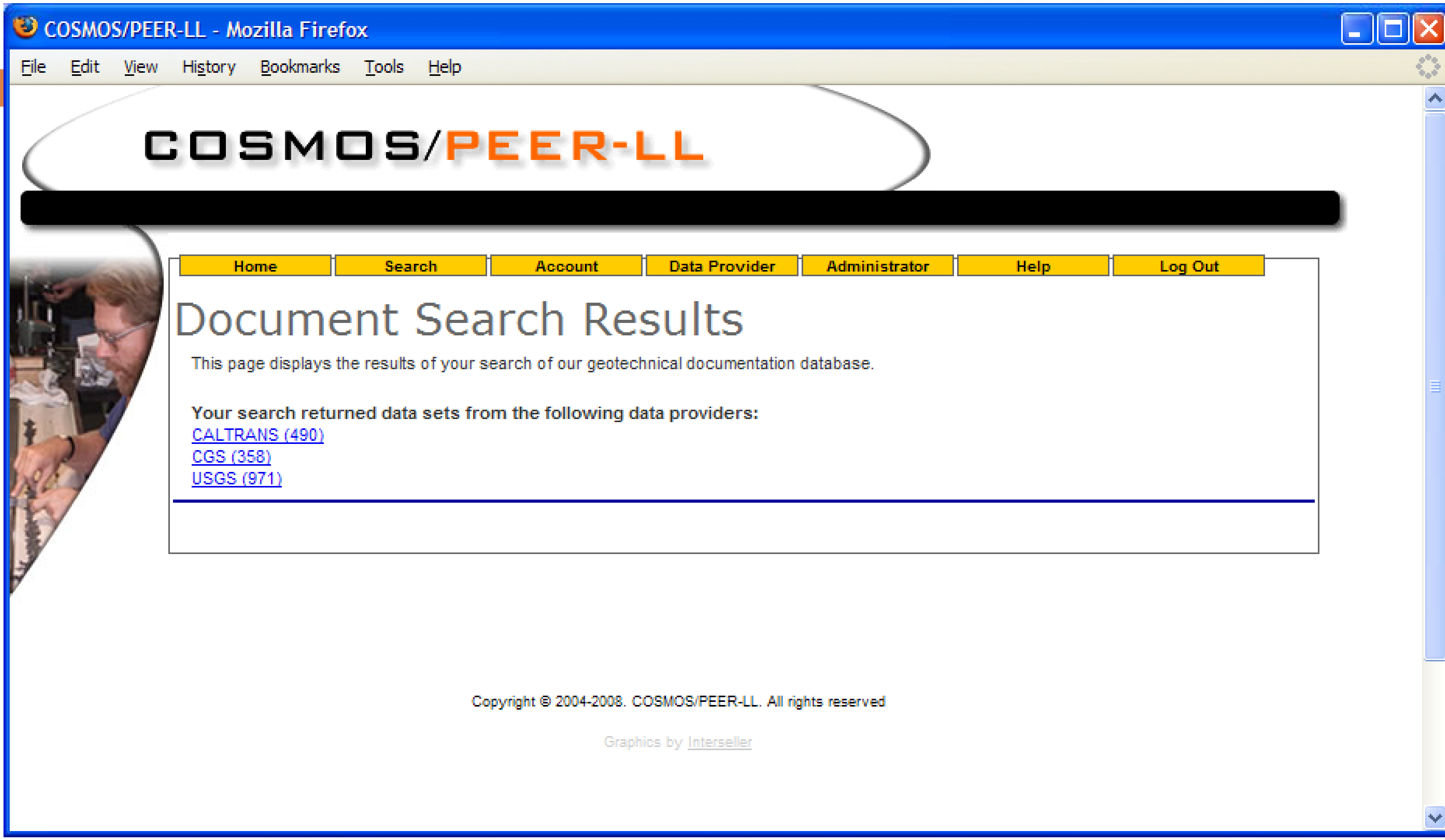

COSMOS /PEER-LL - Mozilla Firefox
File Edit View History Bookmarks Tools Help
COSMOS/PEER-LL
Home Search Account Data Provider Administrator Help Log Out
Document Search
Use this page to search our database of geotechnical documentation.
Map Satellite Hybrid
VDC_754
VDC_12
VDC_753
000007_00177_33118G1
000007_00363_33118H1
VDC_755
...and 912 more
Click to list all data in this cluster.
Latitude: -117.590789794921 88 Longitude: 33.831638461142866
Search
Provider - All Providers -
Project Date(mm/dd/yyyy) from to
Asset Name
Project Name
Borehole Depths min max
Boundaries (decimal degrees) Longitude min Longitude max Latitude min Latitude max
Search for Selected Data Types All Checked by Default
Results Per Page 50 search

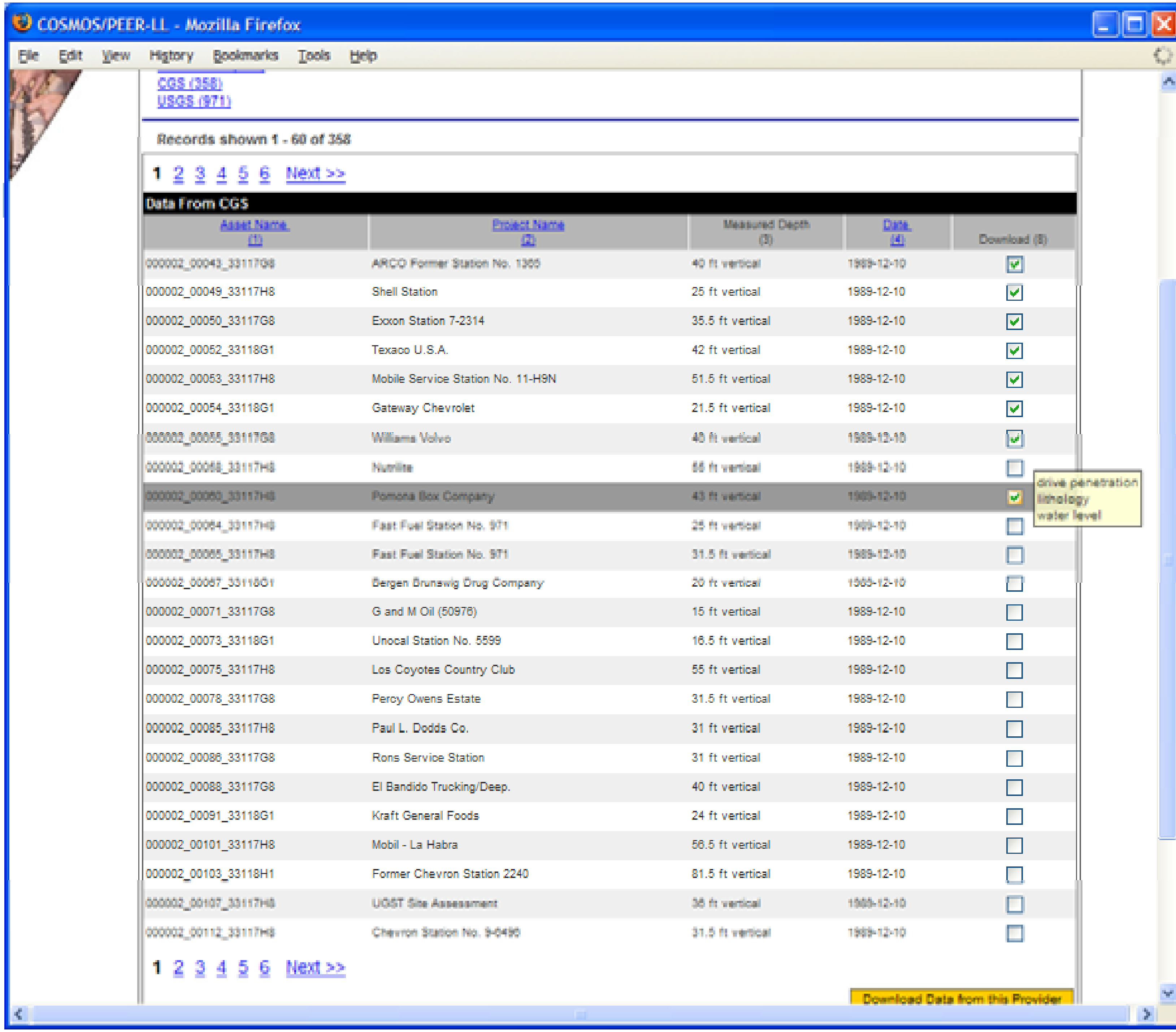

COSMOS /PEER-LL - Mozilla Firefox
File Edit View History Bookmarks Tools Help
COSMOS/PEER-LL
Home
Search
Account
Data Provider
Administrator
Help
Log Out
Document Search
Use this page to search our database of geotechnical documentation.
Map
Satellite
Hybrid
Geotechnical Virtual Data Center Asset
Asset Name: CT-070604_07182201_04231
Project Name: 07062004_07-182201_CPT-C8-4
Data Type(s): static cone
Depth: 9.45 m Vertical
Date: 2004-07-06
Provider: CALTRANS
E Alondra Blvd
E Artesia Blvd
Artesia Fwy
Suburban Medical Center
Latitude: -118.15323829650879
Longitude: 33.876757921837
Search
Provider
- All Providers -
Project Date(mm/dd/yyyy)
from
to
Asset Name
Project Name
Borehole Depths
min
max
Boundaries
(decimal degrees)
Longitude min
Longitude max
Latitude min
Latitude max
Search for Selected Data Types
All Checked by Default
Results Per Page
50
search


COSMOS /PEER-LL - Mozilla Firefox
File Edit View History Bookmarks Tools Help
COSMOS/PEER-LL
Home
Search
Account
Data Provider
Administrator
Help
Log Out
Document Search
Use this page to search our database of geotechnical documentation.
Map
Satellite
Hybrid
Strong-Motion Recording Station
COSMOS Virtual Data Center
Long Beach, CA
Artesia & Orange
CSMIP: 14872
Latitude: -118.15658569335938
Longitude: 33.89606717952165
Search
Provider
- All Providers -
Project Date(mm/dd/yyyy)
from
to
Asset Name
Project Name
Borehole Depths
min
max
Boundaries
(decimal degrees)
Longitude min
Longitude max
Latitude min
Latitude max
Search for Selected Data Types
All Checked by Default
Results Per Page
60
search


COSMOS: Station USC: Long Beach, CA, Long Beach LDS Church, 6979 Orange Ave - Mozilla Firefox
File Edit View History Bookmarks Tools Help
COSMOS /PEER-LL
COSMOS: Station USC: Long Beac...
COSMOS VIRTUAL DATA CENTER
Consortium of Organizations for Strong-Motion Observation Systems
Home · Login/Logoff · Download · About Us · Contact · Mirror Sites
Earthquakes · Stations · Search · Map · Adv. Search
USC: Long Beach, CA
Long Beach LDS Church
6979 Orange Ave
Agency Number: 5380
Latitude: 33.8810
Longitude: -118.1760
Site Geology: Alluvium QYM Q
Owner: University of Southern California
References
Structure: 1-Story Bldg
Instrument: Ground Level
Add all data on this page to the download bin
Go to Download Bin
View Map
Whittier Narrows, California 1987-10-01 14:42:20 UTC
Summary page for this earthquake
Add this station record to the download bin
Component: Up PGA (cm/s/s): 132.70 PGV (cm
Component: 10 PGA (cm/s/s): 221.50 PGV (cm
Component: 280 PGA (cm/s/s): 152.50 PGV (cm
Return to top
Landers, California 1992-06-28 11:57:34 UTC
Summary page for this earthquake
Add this station record to the download bin
Component: Up PGA (cm/s/s): 19.70 PGV (cm
Component: 10 PGA (cm/s/s): 48.40 PGV (cm
Component: 280 PGA (cm/s/s): 58.70 PGV (cm/s): -
Add this to bin
COSMOS: Data Plot Long Beach, CA Whittier Narrows - Mozilla Firefox
File Edit View History Bookmarks Tools Help
133
-133
cm/sec/sec
0 6 13 19 26 32
seconds
Component: 280
Ground Level
152
-152
cm/sec/sec
0 6 13 19 26 32
seconds


COSMOS /PEER-LL - Mozilla Firefox
File Edit View History Bookmarks Tools Help
E Alondra Blvd
Alondra Blvd
Kelly Park
Suburban Medical Center
Artesia Fwy
E Artesia Blvd
Houghton Park
Map data ©2008 Tele Atlas
Latitude: -118.15406907958984
Longitude: 33.870059224024565
Search
Provider
- All Providers -
Project Date(mm/dd/yyyy)
from
to
Asset Name
Project Name
Borehole Depths
min
max
Boundaries (decimal degrees)
-118.19907188415526
-118.18207740783693
33.87148452276899
33.88495242071087
Longitude min
Longitude max
Latitude min
Latitude max
Search for Selected Data Types
All Checked by Default
Field Test or Observation
Lithology/Stratigraphy
lithology
Geophysical Logs
In-Situ Tests
Lab Tests
Engineering Properties
atterberg limits
compressive strength
moisture content
particle-size
relative density
Geochemical Properties
pore water chemistry
Repeated Monitoring
A test to determine the stress-strain properties of soil. Typically, a cylindrical specimen is sealed in a rubber membrane, placed in a cell and subjected to a uniform fluid pressure in the horizontal and vertical directions. A vertical load is applied axially to the specimen increasing the axial stress until the specimen fails.


COSMOS/PEER-LL - Mozilla Firefox
File Edit View History Bookmarks Tools Help
COSMOS/PEER-LL
Home
Search
Account
Data Provider
Administrator
Help
Log Out
Document Search Results
This page displays the results of your search of our geotechnical documentation database.
Your search returned data sets from the following data providers:
CALTRANS (490)
CGS (358)
USGS (971)
Copyright © 2004-2008. COSMOS/PEER-LL. All rights reserved
Graphics by Interseller


COSMOS/PEER-LL - Mozilla Firefox

File Edit View History Bookmarks Tools Help

CGS (358)
USGS (971)

Records shown 1 - 60 of 358

1 2 3 4 5 6 Next >>

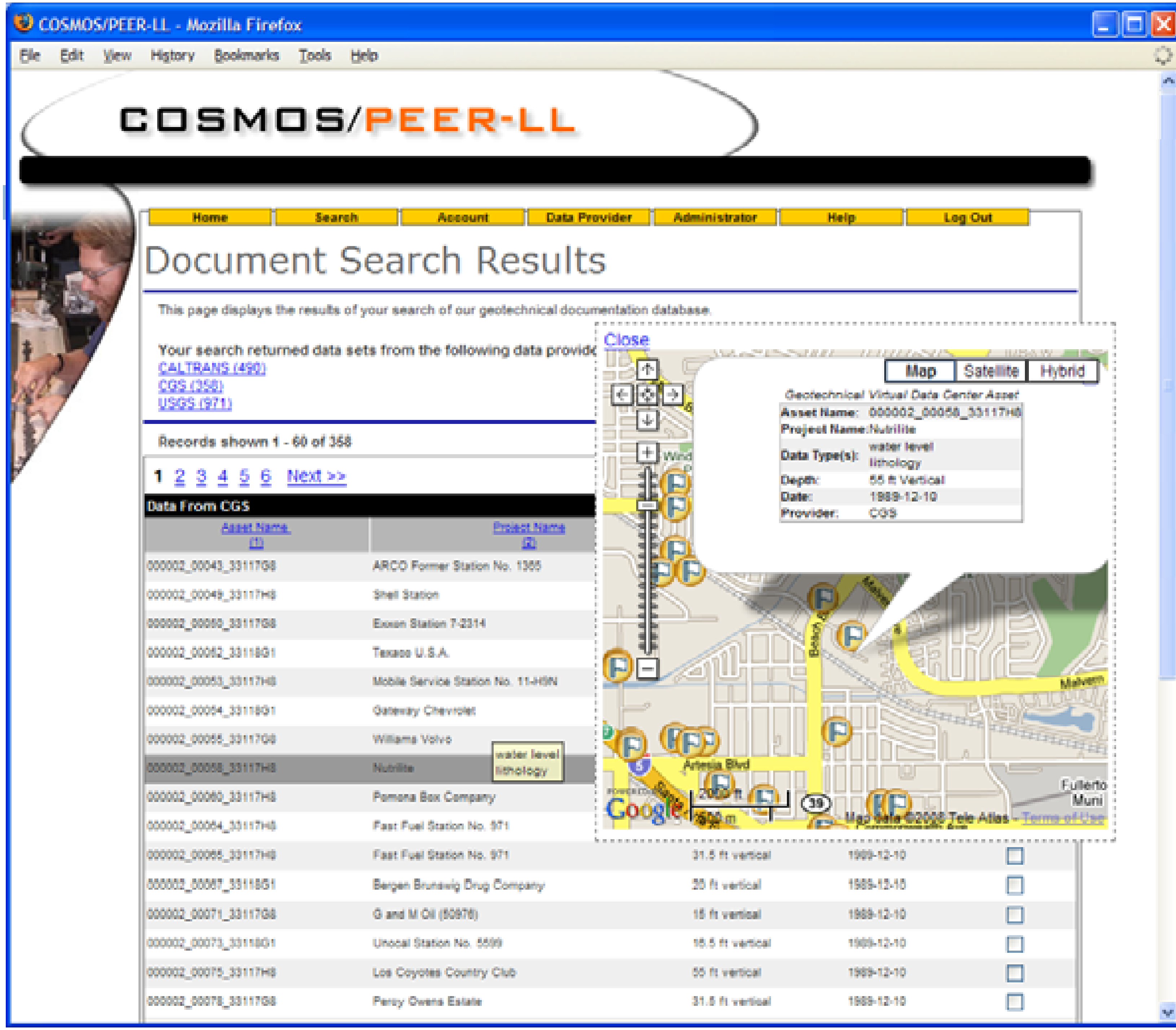
Data From CGS

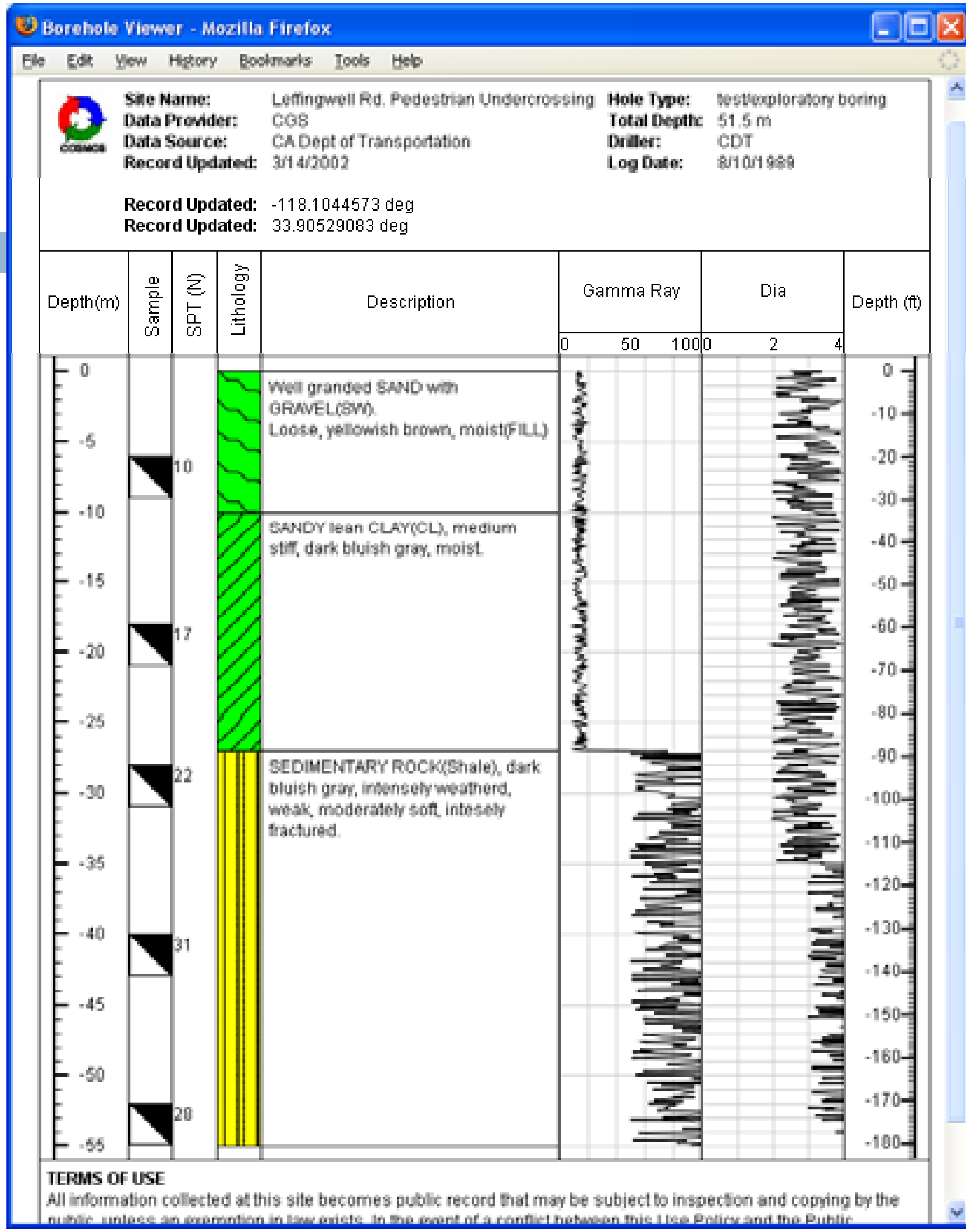
Asset Name (1)	Project Name (2)	Measured Depth (3)	Date (4)	Download (8)
000002_00043_33117G8	ARCO Former Station No. 1365	40 ft vertical	1989-12-10	☑
000002_00049_33117H8	Shell Station	25 ft vertical	1989-12-10	☑
000002_00050_33117G8	Exxon Station 7-2314	35.5 ft vertical	1989-12-10	☑
000002_00052_33118G1	Texaco U.S.A.	42 ft vertical	1989-12-10	☑
000002_00053_33117H8	Mobile Service Station No. 11-H9N	51.5 ft vertical	1989-12-10	☑
000002_00054_33118G1	Gateway Chevrolet	21.5 ft vertical	1989-12-10	☑
000002_00055_33117G8	Williams Volvo	40 ft vertical	1989-12-10	☑
000002_00058_33117H8	Nutrilite	55 ft vertical	1989-12-10	☐
000002_00060_33117H8	Pomona Box Company	43 ft vertical	1989-12-10	☑
000002_00064_33117H8	Fast Fuel Station No. 971	25 ft vertical	1989-12-10	☐
000002_00065_33117H8	Fast Fuel Station No. 971	31.5 ft vertical	1989-12-10	☐
000002_00067_33118G1	Bergen Brunswig Drug Company	20 ft vertical	1989-12-10	☐
000002_00071_33117G8	G and M Oil (50976)	15 ft vertical	1989-12-10	☐
000002_00073_33118G1	Unocal Station No. 5599	16.5 ft vertical	1989-12-10	☐
000002_00075_33117H8	Los Coyotes Country Club	55 ft vertical	1989-12-10	☐
000002_00078_33117G8	Percy Owens Estate	31.5 ft vertical	1989-12-10	☐
000002_00085_33117H8	Paul L. Dodds Co.	31 ft vertical	1989-12-10	☐
000002_00086_33117G8	Rons Service Station	31 ft vertical	1989-12-10	☐
000002_00088_33117G8	El Bandido Trucking/Deep.	40 ft vertical	1989-12-10	☐
000002_00091_33118G1	Kraft General Foods	24 ft vertical	1989-12-10	☐
000002_00101_33117H8	Mobil - La Habra	56.5 ft vertical	1989-12-10	☐
000002_00103_33118H1	Former Chevron Station 2240	81.5 ft vertical	1989-12-10	☐
000002_00107_33117H8	UGST Site Assessment	36 ft vertical	1989-12-10	☐
000002_00112_33117H8	Chevron Station No. 9-0490	31.5 ft vertical	1989-12-10	☐

drive penetration
lithology
water level

1 2 3 4 5 6 Next >>

Download Data from this Provider


COSMOS/PEER-LL - Mozilla Firefox
File Edit View History Bookmarks Tools Help
COSMOS/PEER-LL
Home Search Account Data Provider Administrator Help Log Out
Document Search Results
This page displays the results of your search of our geotechnical documentation database.
Your search returned data sets from the following data provide
CALTRANS (490)
CGS (358)
USGS (971)
Records shown 1 - 60 of 358
1 2 3 4 5 6 Next >>
Data From CGS
Asset Name (1)
Project Name (2)
Close
Map Satellite Hybrid
Geotechnical Virtual Data Center Asset
Asset Name: 000002_00058_33117H8
Project Name: Nutrilite
Data Type(s): water level lithology
Depth: 55 ft Vertical
Date: 1989-12-10
Provider: CGS
Artesia Blvd
Malvern
Fullerton Muni
Map data ©2008 Tele Atlas - Terms of Use


Borehole Viewer - Mozilla Firefox
File Edit View History Bookmarks Tools Help
Site Name: Leffingwell Rd. Pedestrian Undercrossing
Data Provider: CGS
Data Source: CA Dept of Transportation
Record Updated: 3/14/2002
Hole Type: test/exploratory boring
Total Depth: 51.5 m
Driller: CDT
Log Date: 8/10/1989
Record Updated: -118.1044573 deg
Record Updated: 33.90529083 deg
Depth(m)
Sample
SPT (N)
Lithology
Description
Gamma Ray
0 50 100
Dia
0 2 4
Depth (ft)
Well granded SAND with GRAVEL(SW). Loose, yellowish brown, moist(FILL)
SANDY lean CLAY(CL), medium stiff, dark bluish gray, moist.
SEDIMENTARY ROCK(Shale), dark bluish gray, intensely weatherd, weak, moderately soft, intesely fractured.
10
17
22
31
20
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
-55
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
-110
-120
-130
-140
-150
-160
-170
-180
TERMS OF USE
All information collected at this site becomes public record that may be subject to inspection and copying by the

COSMOS/PEER-LL - Mozilla Firefox

File Edit View History Bookmarks Tools Help

CGS (358)
USGS (971)

Records shown 1 - 60 of 358

1 2 3 4 5 6 Next >>

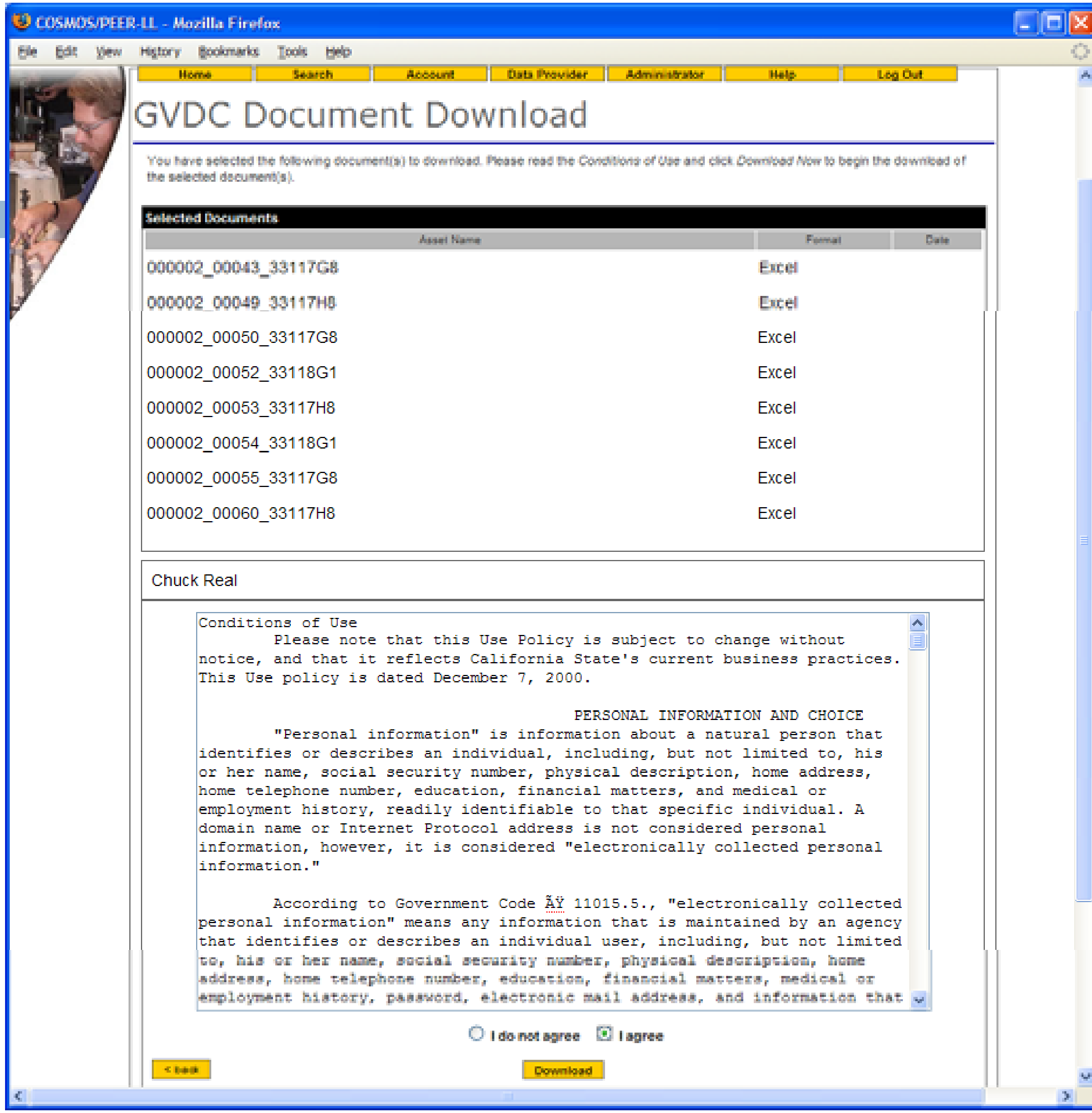
Data From CGS

Asset Name (1)	Project Name (2)	Measured Depth (3)	Date (4)	Download (8)
000002_00043_33117G8	ARCO Former Station No. 1365	40 ft vertical	1989-12-10	☑
000002_00049_33117H8	Shell Station	25 ft vertical	1989-12-10	☑
000002_00050_33117G8	Exxon Station 7-2314	35.5 ft vertical	1989-12-10	☑
000002_00052_33118G1	Texaco U.S.A.	42 ft vertical	1989-12-10	☑
000002_00053_33117H8	Mobile Service Station No. 11-H9N	51.5 ft vertical	1989-12-10	☑
000002_00054_33118G1	Gateway Chevrolet	21.5 ft vertical	1989-12-10	☑
000002_00055_33117G8	Williams Volvo	40 ft vertical	1989-12-10	☑
000002_00058_33117H8	Nutrilite	55 ft vertical	1989-12-10	☐
000002_00060_33117H8	Pomona Box Company	43 ft vertical	1989-12-10	☑
000002_00064_33117H8	Fast Fuel Station No. 971	25 ft vertical	1989-12-10	☐
000002_00065_33117H8	Fast Fuel Station No. 971	31.5 ft vertical	1989-12-10	☐
000002_00067_33118G1	Bergen Brunswig Drug Company	20 ft vertical	1989-12-10	☐
000002_00071_33117G8	G and M Oil (50976)	15 ft vertical	1989-12-10	☐
000002_00073_33118G1	Unocal Station No. 5599	16.5 ft vertical	1989-12-10	☐
000002_00075_33117H8	Los Coyotes Country Club	55 ft vertical	1989-12-10	☐
000002_00078_33117G8	Percy Owens Estate	31.5 ft vertical	1989-12-10	☐
000002_00085_33117H8	Paul L. Dodds Co.	31 ft vertical	1989-12-10	☐
000002_00086_33117G8	Rons Service Station	31 ft vertical	1989-12-10	☐
000002_00088_33117G8	El Bandido Trucking/Deep.	40 ft vertical	1989-12-10	☐
000002_00091_33118G1	Kraft General Foods	24 ft vertical	1989-12-10	☐
000002_00101_33117H8	Mobil - La Habra	56.5 ft vertical	1989-12-10	☐
000002_00103_33118H1	Former Chevron Station 2240	81.5 ft vertical	1989-12-10	☐
000002_00107_33117H8	UGST Site Assessment	38 ft vertical	1989-12-10	☐
000002_00112_33117H8	Chevron Station No. 9-0490	31.5 ft vertical	1989-12-10	☐

drive penetration
lithology
water level

1 2 3 4 5 6 Next >>

Download Data from this Provider

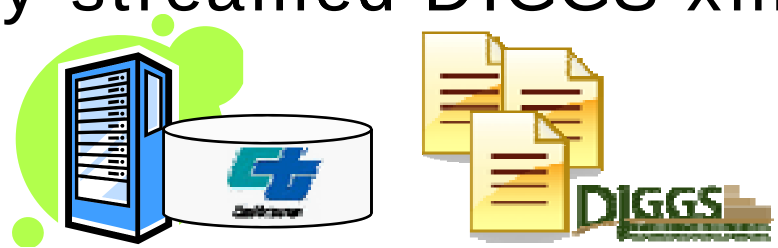

COSMOS/PEER-LL - Mozilla Firefox
File Edit View History Bookmarks Tools Help
Home Search Account Data Provider Administrator Help Log Out
GVDC Document Download
You have selected the following document(s) to download. Please read the Conditions of Use and click Download Now to begin the download of the selected document(s).
Selected Documents
Asset Name Format Date
000002_00043_33117G8 Excel
000002_00049_33117H8 Excel
000002_00050_33117G8 Excel
000002_00052_33118G1 Excel
000002_00053_33117H8 Excel
000002_00054_33118G1 Excel
000002_00055_33117G8 Excel
000002_00060_33117H8 Excel
Chuck Real
Conditions of Use
Please note that this Use Policy is subject to change without notice, and that it reflects California State's current business practices. This Use policy is dated December 7, 2000.
PERSONAL INFORMATION AND CHOICE
"Personal information" is information about a natural person that identifies or describes an individual, including, but not limited to, his or her name, social security number, physical description, home address, home telephone number, education, financial matters, and medical or employment history, readily identifiable to that specific individual. A domain name or Internet Protocol address is not considered personal information, however, it is considered "electronically collected personal information."
According to Government Code ÂŸ 11015.5., "electronically collected personal information" means any information that is maintained by an agency that identifies or describes an individual user, including, but not limited to, his or her name, social security number, physical description, home address, home telephone number, education, financial matters, medical or employment history, password, electronic mail address, and information that
I do not agree
I agree
< back
Download

Behind the Scenes

- Technology Framework on GVDC
  - Windows Server 2003
  - Apache
  - PostgreSQL/PostGIS
  - PHP
  - Java Servlets
  - Javascript & AJAX
  - Canvas
  - XSLT transforms
- GoogleMaps
  - Clustering
  - Selection box tool
  - Map overlays

Behind the Scenes

- Requirements of a GVDC data provider:
  - Digital repository of their geotechnical data
  - Web or ftp server that is Internet accessible to GVDC server
  - Register as a data provider with the GVDC
- The data provider must supply DIGGS files to the GVDC. This requires a mapping application that produces either static files or dynamically-streamed DIGGS xml from database



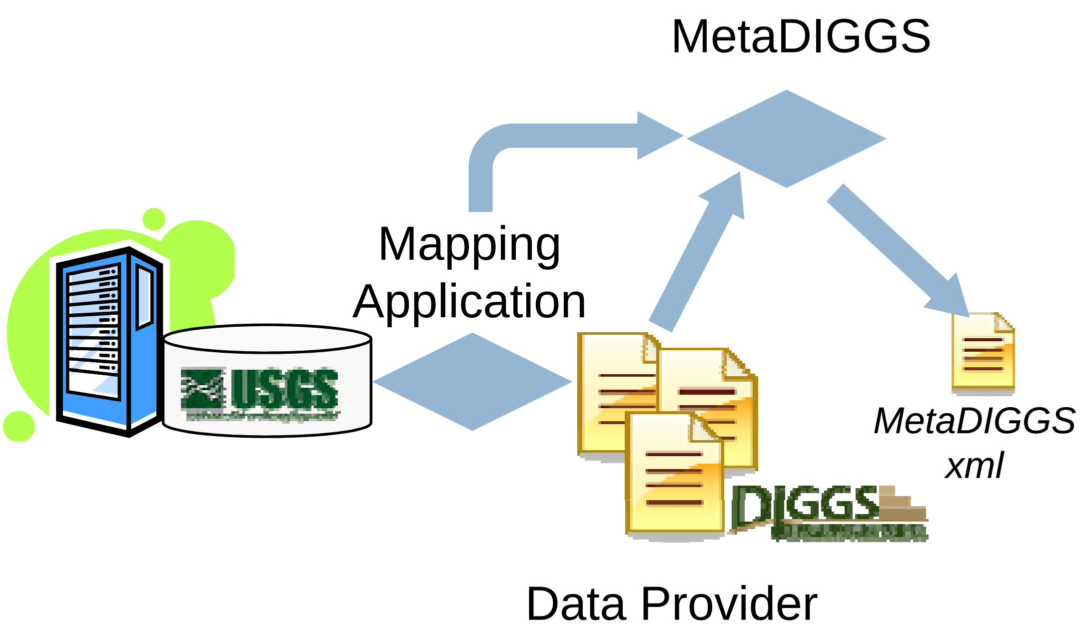
Data Provider

Behind the Scenes

- The GVDC provides a Java application (metaDIGGS) that extracts metadata from the DIGGS files or database for harvesting by the GVDC
- The GVDC then “harvests” the MetaDIGGS file and stores this information in its database

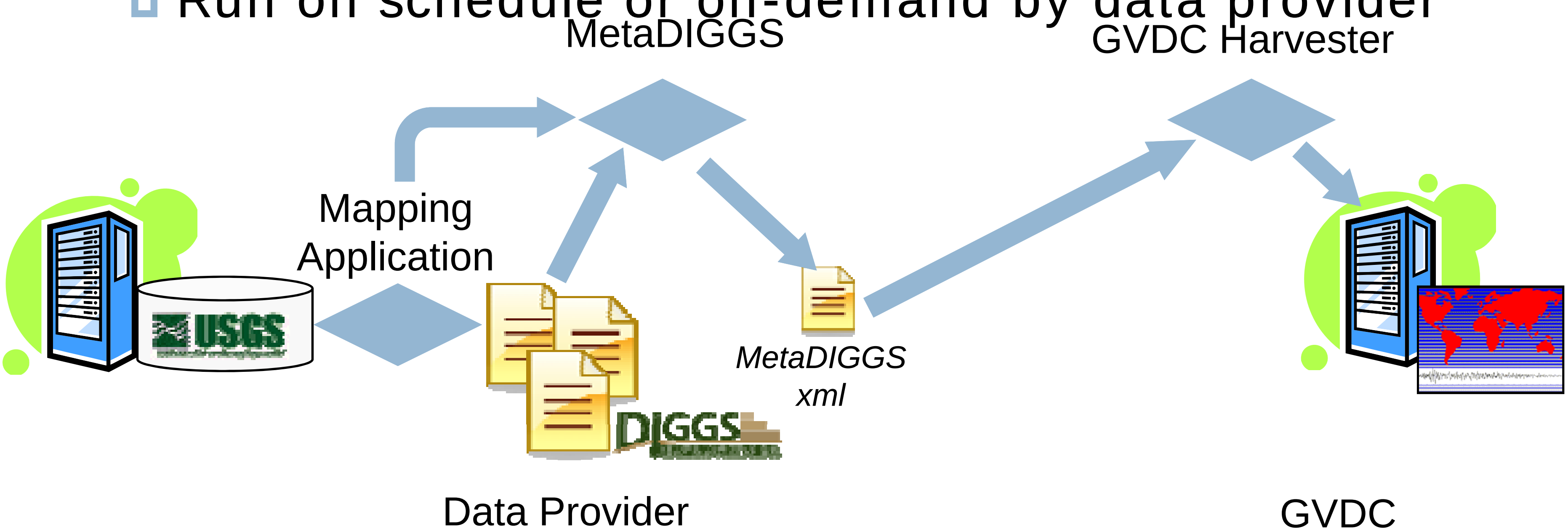
Behind the Scenes

- MetaDIGGS application
  - XSLT file (DIGGS –> MetaDIGGS)
  - Java wrapper (GUI and scheduler)
  - Configuration file



Behind the Scenes

- GVDC Harvester Application
  - Reads MetaDIGGS xml and populates GVDC database
  - Simple dump/replace
  - Run on schedule or on-demand by data provider



MetaDIGGS Schema

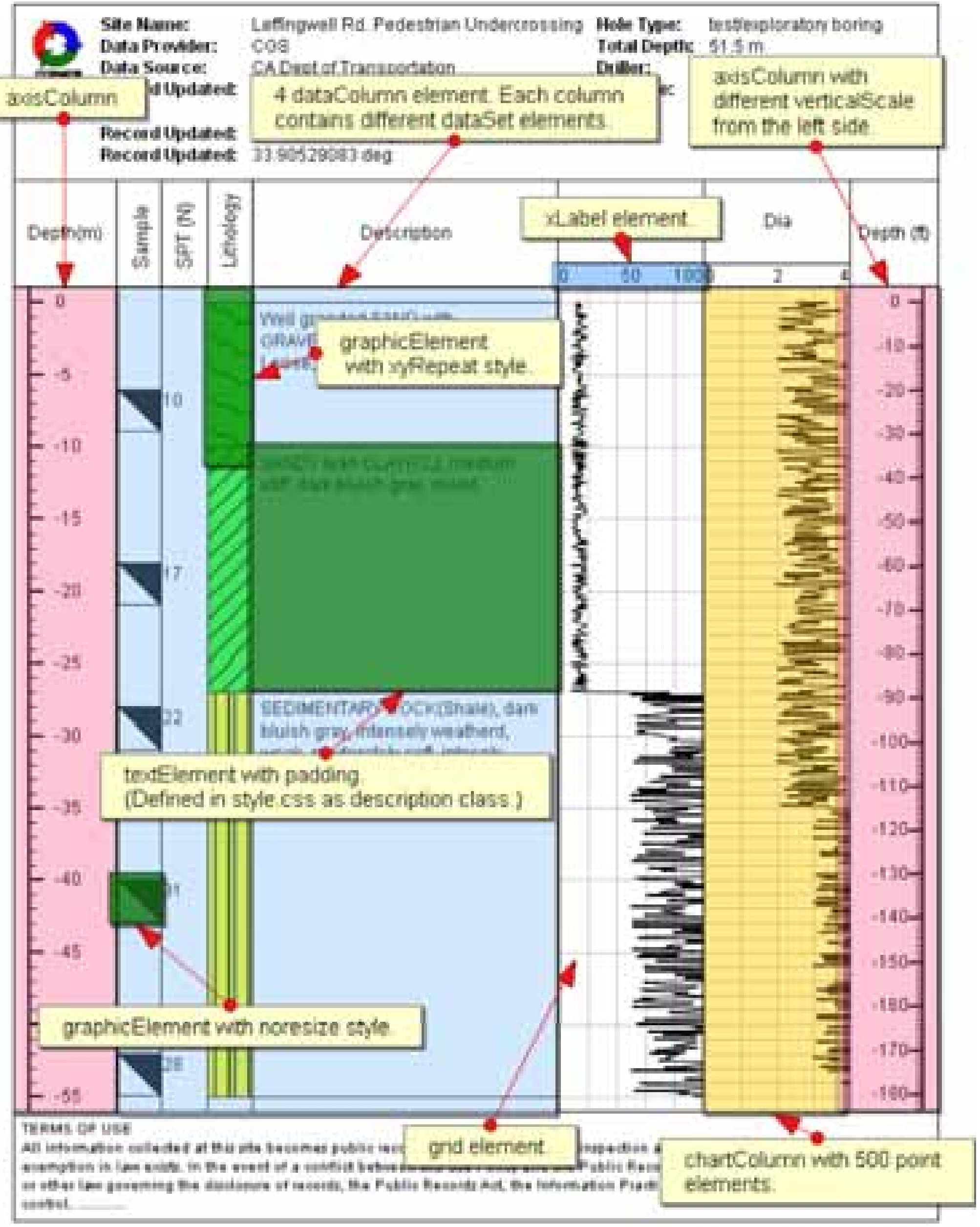
- asset
  - id, name, project_name, project_id
  - primary_asset_type (location, hole, etc.)
  - start_date, end_date
  - lat_northing, lon_easting, srid
  - depth, depth_uom, deviated
  - project_purpose, data_source, last_modified
  - xml_url
- asset_test
  - test_id
  - data_availability, file_reference
  - depth_top, depth_bottom, depth_uom

Borehole Previewer

- Borehole log drawn dynamically on browser:
  - CosmosLog XML file
    - Contains data to be plotted (derived from DIGGS xml), organized in specific element types and order of log columns
  - Dictionary XML file
    - Contains paths to graphic symbols used on log
    - COSMOS standard, but each dataprovider could provide their own dictionary and symbols
  - Viewer XSL and CSS to convert CosmosLog XML to html on browser.
    - One XSL and CSS used for COSMOS previewer
    - Uses JavaScript to draw data columns within <canvas> elements (HTML 5)

Borehole Previewer

- Viewer xsl, css, and javascript completed
- Viewer Application (in development)
  - DIGGS xml ->XSLT -> CosmosLog xml
  - Possible user-customized vertical scale
  - Other user-customizations possible but not planned



GVDC Timeline

- Web site search/retrieve functionality completed
- MetaDIGGS and CosmosLog schemas completed
- MetaDIGGS application engine mostly complete and in testing
- April: Harvester Application Development
- April-May: DIGGS->CosmosLog transform
- May: Download servlet (multi-hole ->single hole extraction and zipping)
- May-June: DIGGS->Excel transform
- Summer: Site roll-out with limited datasets

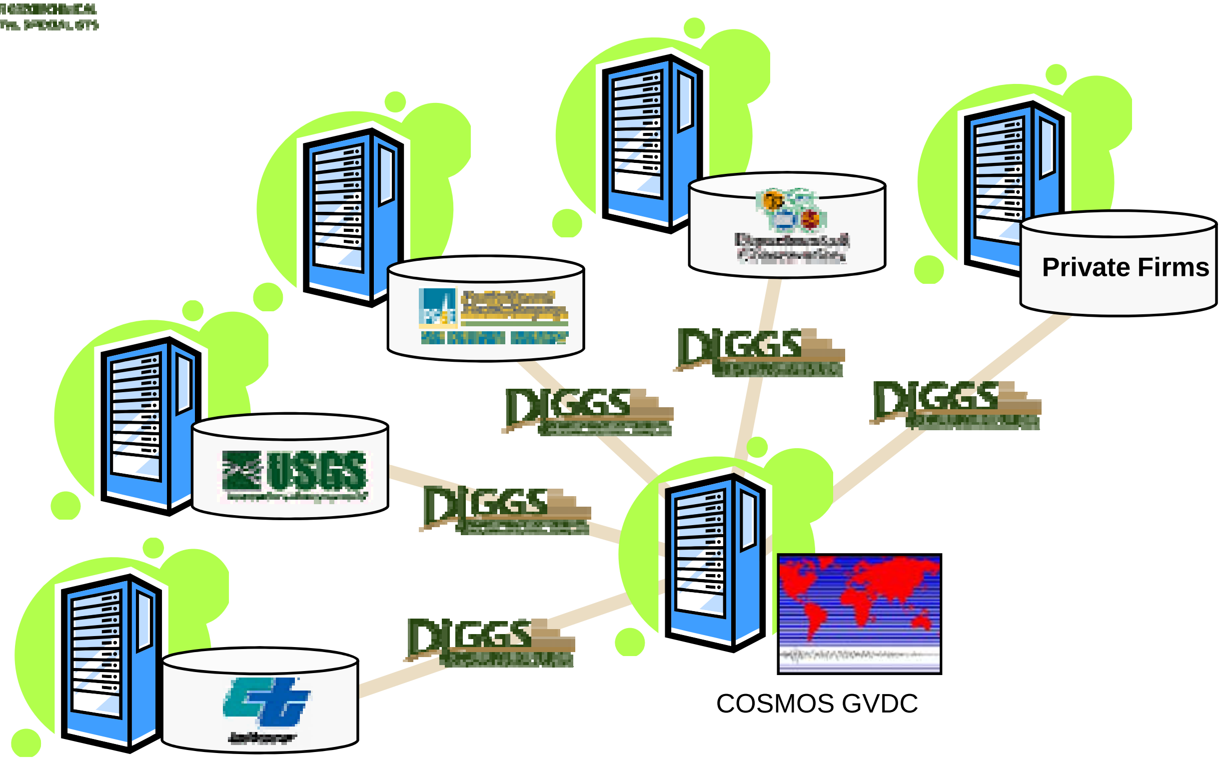
DIGGS and the COSMOS/PEER-LL GEOTECHNICAL VIRTUAL DATA CENTER

DIGGS Invitational Meeting, March 25 and 26, 2009

Daniel Ponti, USGS
Loren Turner, CALTRANS

Geotechnical Virtual Data Center

- Virtual gateway to data repositories from multiple agencies.
- Uses DIGGS for standardized data exchange



Project History

1992 NSF/FHWA sponsors the *National Geotechnical Experiment Sites* (NGES) that publishes geotech research data.

1996 The *Resolution of Site Response in the Northridge Earthquake* (ROSRINE) project pioneers web dissemination of geotechnical data.

1998 USC Workshop highlights growing need for geotechnical data management and exchange.

2001 Phase 1 – PEER-LL sponsors a workshop to assess user needs and build consensus to develop a Geotechnical Virtual Data Center (GVDC).

2004 Phase 2 – PEER-LL tasks COSMOS to develop a pilot GVDC that demonstrates the feasibility of the technology.

2005 Phase 3 – PEER-LL tasks COSMOS to update GVDC to integrate DIGGS, and revise/simplify system architecture.

Sponsors and Partners

Sponsored by:

- CalTrans
- California Energy Commission
- Pacific Gas & Electric
- PEER-Lifelines Program

In Partnership with:

- Pacific Earthquake Engineering Research Center
- United States Geological Survey
- California Geological Survey

Implemented by:

- University of Southern California
- Consortium of Organizations for Strong-Motion Observations Systems

Project Team

- **Carl Stepp (PI),** Consortium of Organizations for Strong-Motion Observation Systems (COSMOS)
- **Jean Benoit**, University of New Hampshire
- **John Bobbit**, Petrotechnical Open Standards Consortium (POSC)
- **Sean Devlin**, U.S. Geological Survey
- **Dan Ponti**, U.S. Geological Survey
- **Charles Real**, California Geological Survey
- **Toru Saito**, Saito Statistics
- **Jennifer Swift**, University of Southern California
- **Loren Turner**, Caltrans
- **Yang Zhu**, Caltrans

http://geodata2.usc.edu:8084/diggs_files/2003 CPT Data/05-407801/CT-071503_05407801_03334.xml - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://geodata2.usc.edu:8084/diggs_files/2003%20CPT%20Data/05-407801/CT-071503_05407801_03334.xml


```xml
<?xml version="1.0" encoding="UTF-8" ?>
- <Diggs xmlns="http://www.diggsml.org/0.10" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns:diggs_env="http://www.diggsml.org/0.10/environmental" xmlns:diggs_geo="http://www.diggsml.org/0.10/geotechnical"
    xmlns:diggs_mon="http://www.diggsml.org/0.10/monitoring" xmlns:diggs_pil="http://www.diggsml.org/0.10/piling"
    xmlns:gml="http://www.opengis.net/gml" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:smil20="http://www.w3.org/2001/SMIL20/"
    xmlns:smil20lang="http://www.w3.org/2001/SMIL20/Language" xmlns:witsml="http://www.witsml.org/schemas/131"
    xsi:schemaLocation="http://www.diggsml.org/0.10 C:/Schemas/diggs/complete.xsd">
  - <businessAssociates>
    - <BusinessAssociate>
        <gml:name>Gem-Yeu Ma</gml:name>
        <id>CT-PM</id>
      - <address>
          <city>Sacramento</city>
          <state>California</state>
          <country>USA</country>
          <postalCode>95819</postalCode>
        </address>
        <phoneNumber>(916)</phoneNumber>
        <emailAddress>gem-yeu_ma@dot.ca.gov</emailAddress>
      </BusinessAssociate>
    - <BusinessAssociate>
        <gml:name>TOTH/MACIAS</gml:name>
        <id>CT-OP</id>
      </BusinessAssociate>
    </businessAssociates>
  - <equipments>
    - <Equipment>
        <id>CT-CPT</id>
        <serialNumber>618TC</serialNumber>
        <class />
      </Equipment>
    </equipments>
  - <projects>
    - <Project>
        <gml:description>County-Route-PM: SLO -101-none 45.5/none -999, Bridge Number: , Bridge Name:</gml:description>
        <gml:name>07152003_05-407801_CPT-4-03</gml:name>
      - <roles>
        - <Role>
            <rolePerformed>CalTrans Project Manager</rolePerformed>
          - <associates>
              <Ref path="//*[./id='CT-PM']" />
            </associates>
          </Role>
        </roles>
        <id>CT-071503_05407801</id>
```


Done Internet

http://geodata.usc.edu:1100/gvdc_demo/scripts/GenerateXLS.asp?URL=http://geodata.usc.edu:8080/u - Microsoft Internet Explorer

File Edit View Insert Format Tools Data Go To Favorites Help

A1 ID

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	ID	CODESPA	HOLE_ID	CORE_ID	SOURCE	TOP	TOP_UOM	TOP_QUA	BASE	BASE_UO	BASE_QU	CLASS_SYSTE	PRIMARY	SECONDA	SECONDA
2	38495	USGS_FOQUS_ID				0	ft		128	ft		USGS FOQUS	unknown/not observed		
3	38497	USGS_FOQUS_ID				128	ft		155	ft		USGS FOQUS	clay		
4	38498	USGS_FOQUS_ID				155	ft		175	ft		USGS FOQUS	sand		
5	38499	USGS_FOQUS_ID				175	ft		225	ft		USGS FOQUS	clay		
6	38500	USGS_FOQUS_ID				225	ft		249	ft		USGS FOQUS	sand		
7	38501	USGS_FOQUS_ID				249	ft		1008	ft		USGS FOQUS	shale		
8	38502	USGS_FOQUS_ID				1008	ft		1032	ft		USGS FOQUS	sandstone		
9	38503	USGS_FOQUS_ID				1032	ft		1170	ft		USGS FOQUS	shale		
10	38504	USGS_FOQUS_ID				1170	ft		1180	ft		USGS FOQUS	sandstone		
11	38505	USGS_FOQUS_ID				1180	ft		1636	ft		USGS FOQUS	shale		
12	38506	USGS_FOQUS_ID				1636	ft		1641	ft		USGS FOQUS	sandstone		
13	38507	USGS_FOQUS_ID				1641	ft		1648	ft		USGS FOQUS	shale		
14	38508	USGS_FOQUS_ID				1648	ft		1671	ft		USGS FOQUS	sandstone		
15	38509	USGS_FOQUS_ID				1671	ft		1680	ft		USGS FOQUS	shale		
16	38510	USGS_FOQUS_ID				1680	ft		1684	ft		USGS FOQUS	sandstone		
17	38511	USGS_FOQUS_ID				1684	ft		1718	ft		USGS FOQUS	shale		
18	38512	USGS_FOQUS_ID				1718	ft		1770	ft		USGS FOQUS	sandstone		
19	38513	USGS_FOQUS_ID				1770	ft		1777	ft		USGS FOQUS	shale		
20	38514	USGS_FOQUS_ID				1777	ft		1805	ft		USGS FOQUS	sandstone		
21	38515	USGS_FOQUS_ID				1805	ft		1860	ft		USGS FOQUS	shale		
22	38516	USGS_FOQUS_ID				1860	ft		1895	ft		USGS FOQUS	sandstone	shale	
23	38517	USGS_FOQUS_ID				1895	ft		1912	ft		USGS FOQUS	shale		
24	38518	USGS_FOQUS_ID				1912	ft		1932	ft		USGS FOQUS	sandstone		
25	38519	USGS_FOQUS_ID				1932	ft		1950	ft		USGS FOQUS	shale		

Site / Hole / Layer / Component

Safe roads, Reliable journeys, Informed travellers

HIGHWAYS AGENCY

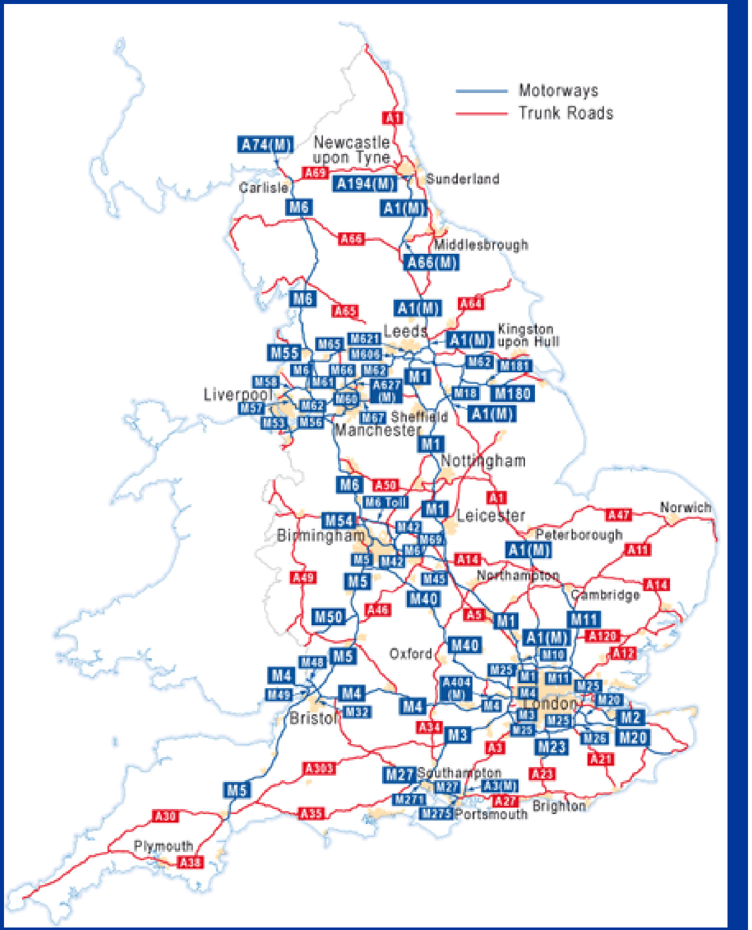
DIGGS Implementation efforts – UK Transportation Agency Perspective

David Patterson, Highways Agency, UK

DIGGS Invitational Meeting – Orlando, Florida, March 25-26, 2009

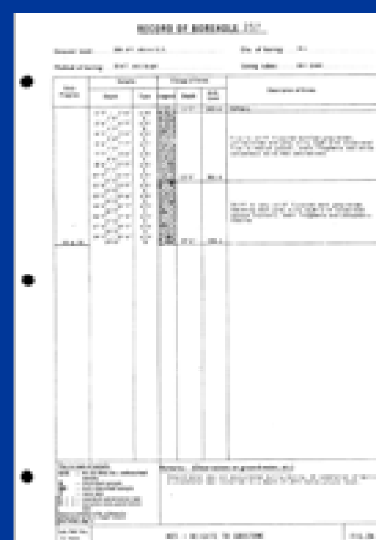
The UK Highways Agency

- The HA is an Executive Agency of the Department of Transport
- Managed by 14 Managing Agents
- 4500 mile network
- Approx 2500 miles of cuttings, embankments and bunds (berms)
- Performance is strongly related to:
  - age
  - geological conditions
  - drainage

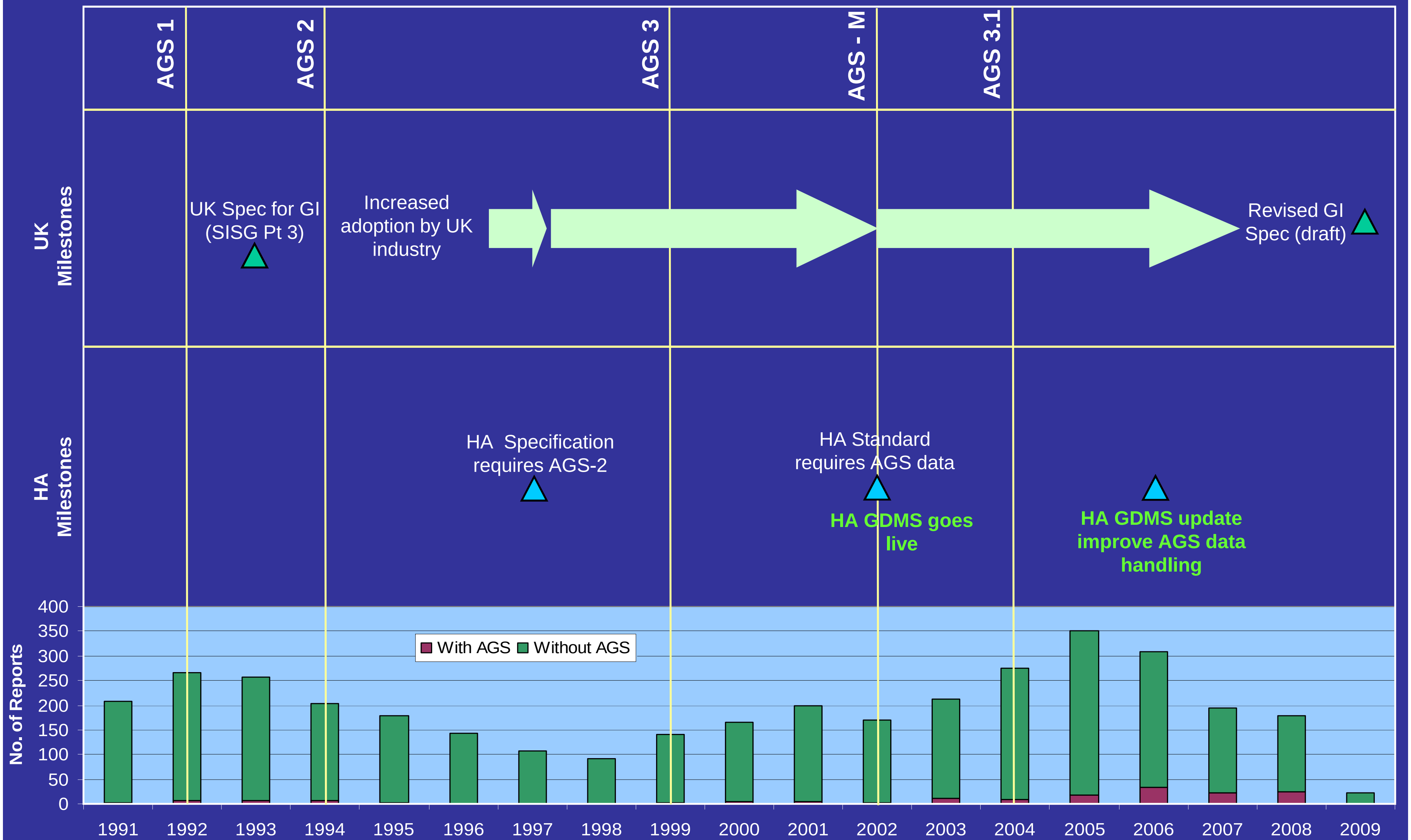


The HA Geotechnical Data Management System (HA GDMS)

- Internet-based GIS
- Stores data on:
  - spatial context (mapping and aerial photos)
  - assets
  - reports
  - boreholes
- Supports UK AGS data transfer format
  - data storage/retrieval
  - summary logs
  - summary test sheets



AGS Implementation Timeline



Executive Summary of
HA assessment of DIGGS implementation

March 2009

- Governance
- National implementation
- Data quality and validation
- Software
- Extensibility
- Future vision
- Key requirements for implementation

Governance

- Strong, enduring governance structure required
- HA will have strong reliance of UK Local Implementation Group (AGS)
- Management and maintenance of the standard must be robust
- Lessons can be learnt from the experience of the AGS
- Governance of included third-party elements also required (eg GML, POSC etc)

National implementation

- AGS data transfer still not fully embraced in the HA supply chain
- Will DIGGS reset the clock to zero? HA cannot afford for this to happen
- Local implementation of DIGGS very important – by AGS?
- HA to coordinate with other asset groups in the UK, and the AGS to work on single implementation?

Data quality and validation

- Pros
  - fixed schema potentially removes validation errors
  - Non-ambiguous checking facility (unlike AGS)
  - More consistent data held in HAGDMS
- Cons
  - schema alone does not fully validate data
  - orphans can exist in a valid DIGGS file

Software

- Pros
  - Removing *ad hoc* spreadsheets improves data quality
  - Requirement for databases potentially removes orphans
  - Use of XML promotes increase in available software
- Cons
  - Total reliance on software suppliers (including QA)
  - AGS is simple, DIGGS is complex
  - Impact on the supply chain
    - new software
    - training
    - increased cost
  - Will available software increase?
    - might monopolies develop?
  - Changes required to HAGDMS

Extensibility

- Pros
  - Potential for extensions into highway specific areas
  - Potential for an HA specific schema
  - Controlled extension better than current AGS methodology
- Cons
  - Client specific schemas place burden on supply chain
  - Ease of extensibility promotes development of ‘non-standard standards’
  - DIGGS allows the same thing to be done in several ways
    - How does HA specify what to do?
    - How can this be validated in the data?

Future vision

- Increased topics
  - environmental
  - asset management
  - piling/foundations
  - geophysics
  - 2-D and 3-D geometry
- Increased usage across HA supply chain
  - AGS currently restricted to investigation/design
  - DIGGS potential to increase to:
    - feasibility
    - tendering
    - construction
    - operation/maintenance
    - decommissioning
- Encourage a more uniform approach across client bodies

Key requirements for implementation

- Multiple DIGGS compliant tools available
- Robust governance
- Acceptable level of impact on supply chain
- Schema and adoption rules accepted by UK industry
- Assured maintenance and longevity
- Stronger data validation (to remove potential for orphan records)
- Removal of inherent DIGGS flexibility to reflect UK application
- Demonstrable cost benefits to HA business

DIGGS 2009

State DOT Geo-Data Systems Implementation Efforts

Summarizing Responses From:
California
Connecticut
Florida
Minnesota
Ohio
Tennessee

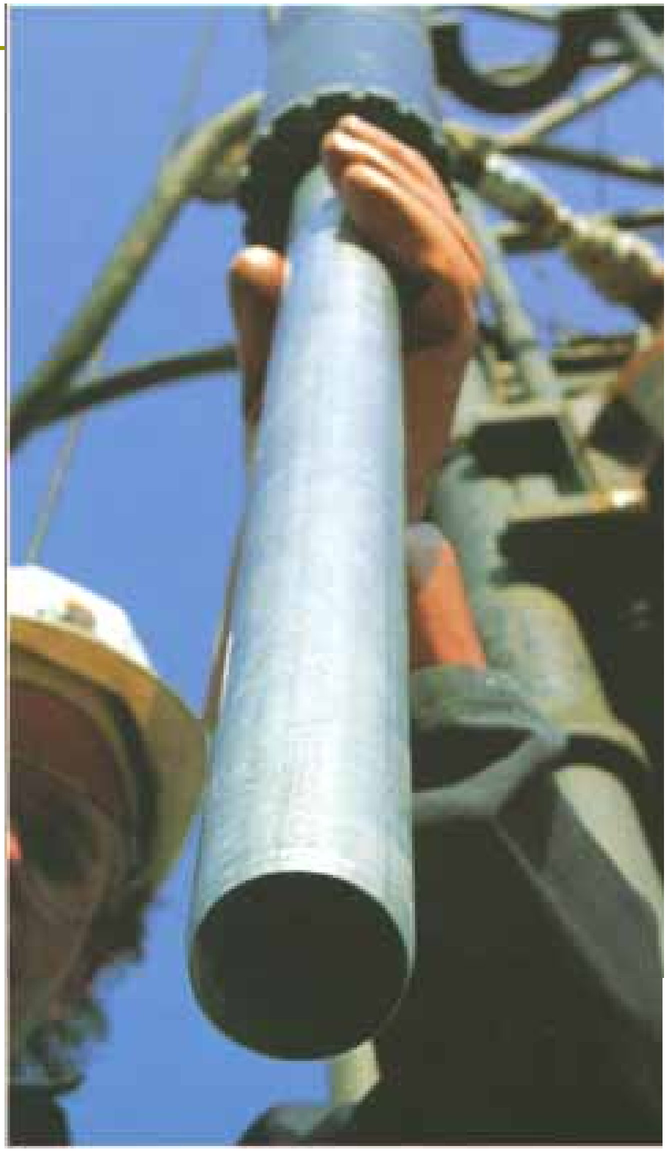
Derrick Dasenbrock P.E., Mn/DOT (for TPF member states), 25 March 2009

Key Activities

- Borehole Data
  - Point Location
  - Drilling Operations

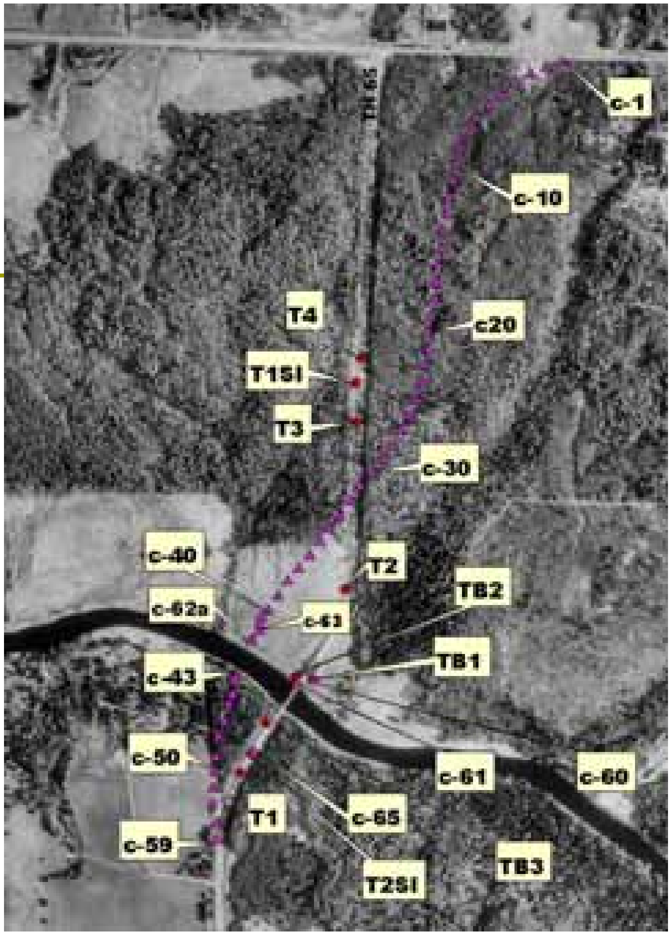
Key Activities

- Borehole Data
  - SAMPLES!



Data TRANSFER

- Site Information
- Depth Information
  - Field
  - Lab Testing
  - Soil and Rock

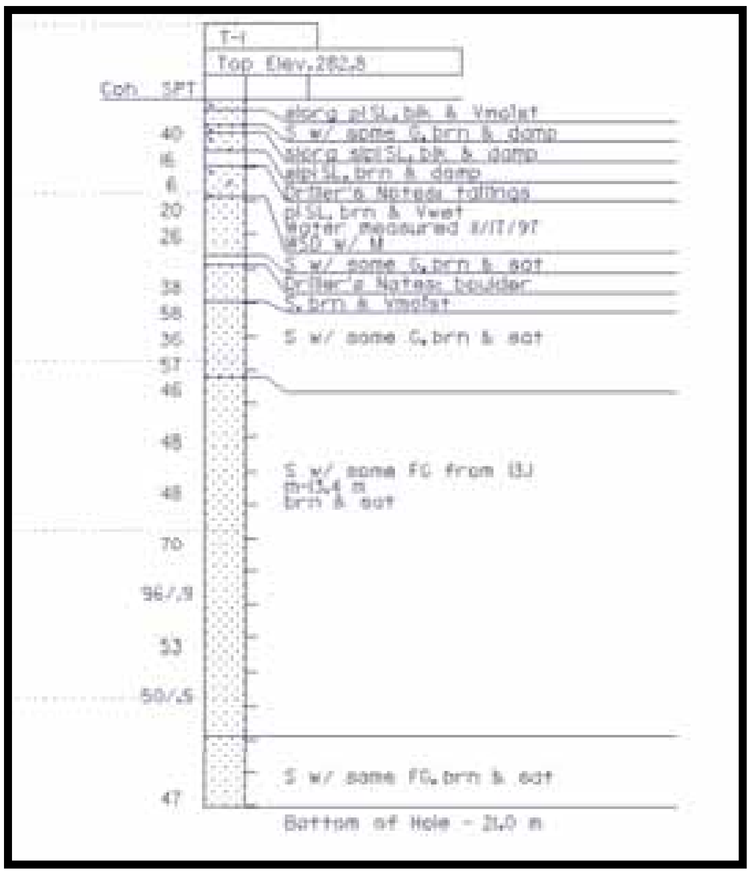


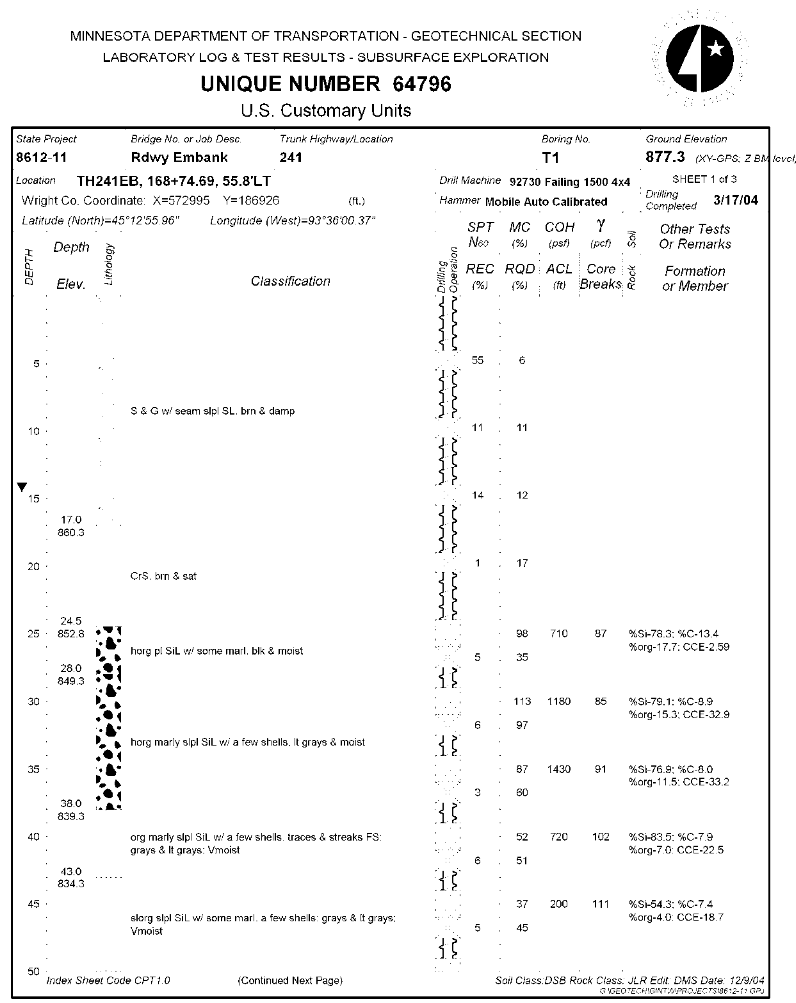
Data TRANSFER

- Lab Data (results and/or test data)

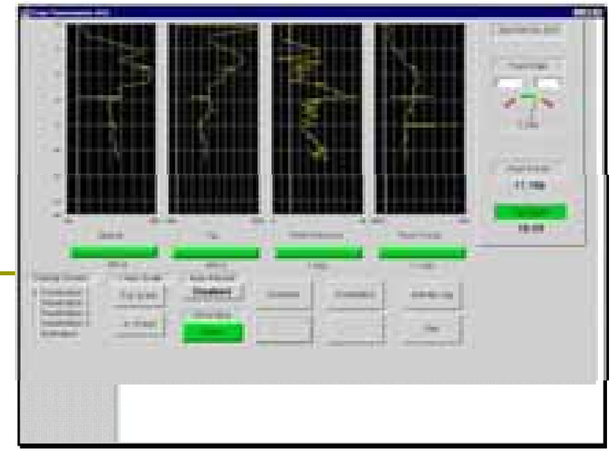
Logs and Log Data

- Electronic / Paper





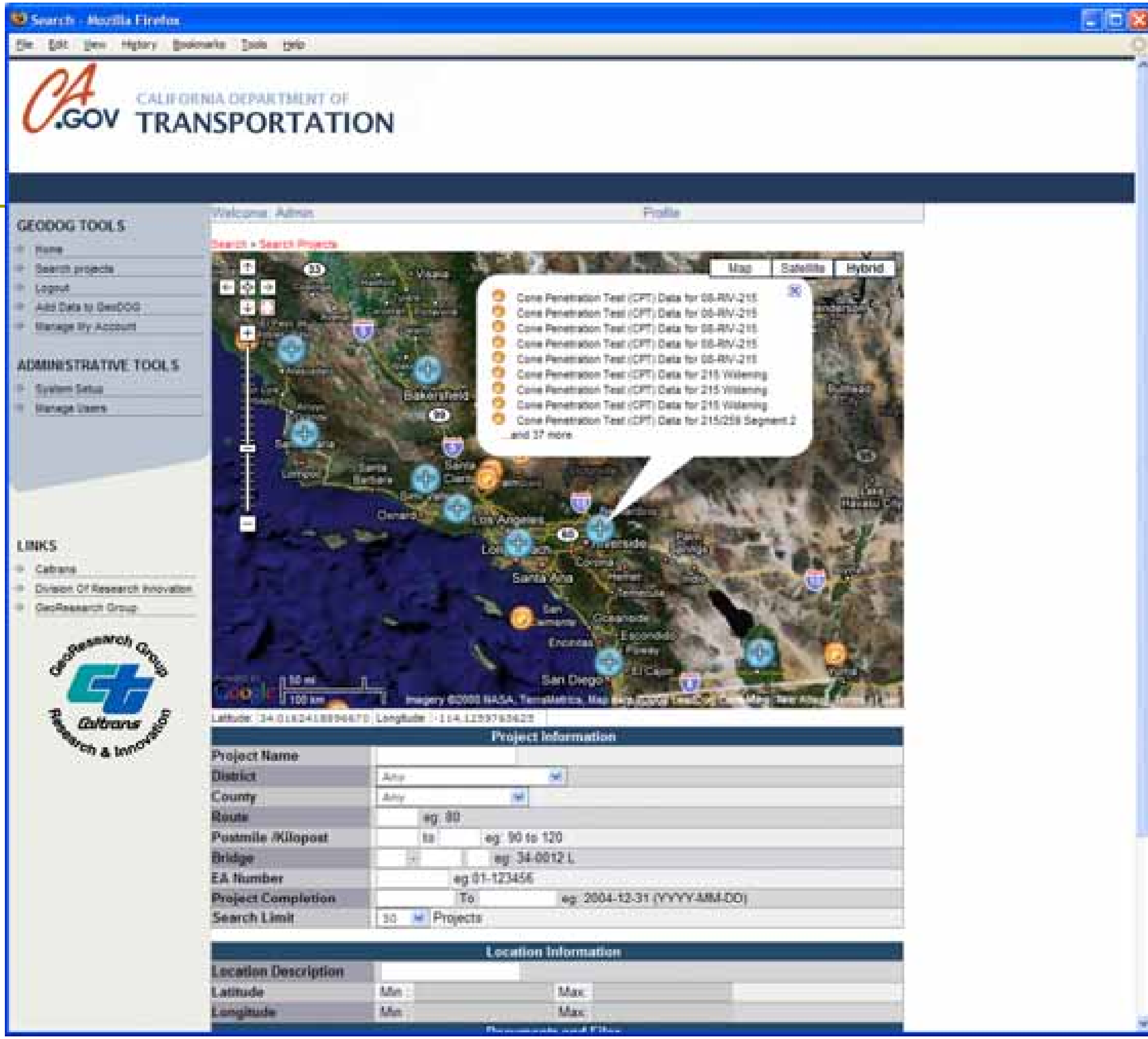
Data TRANSFER



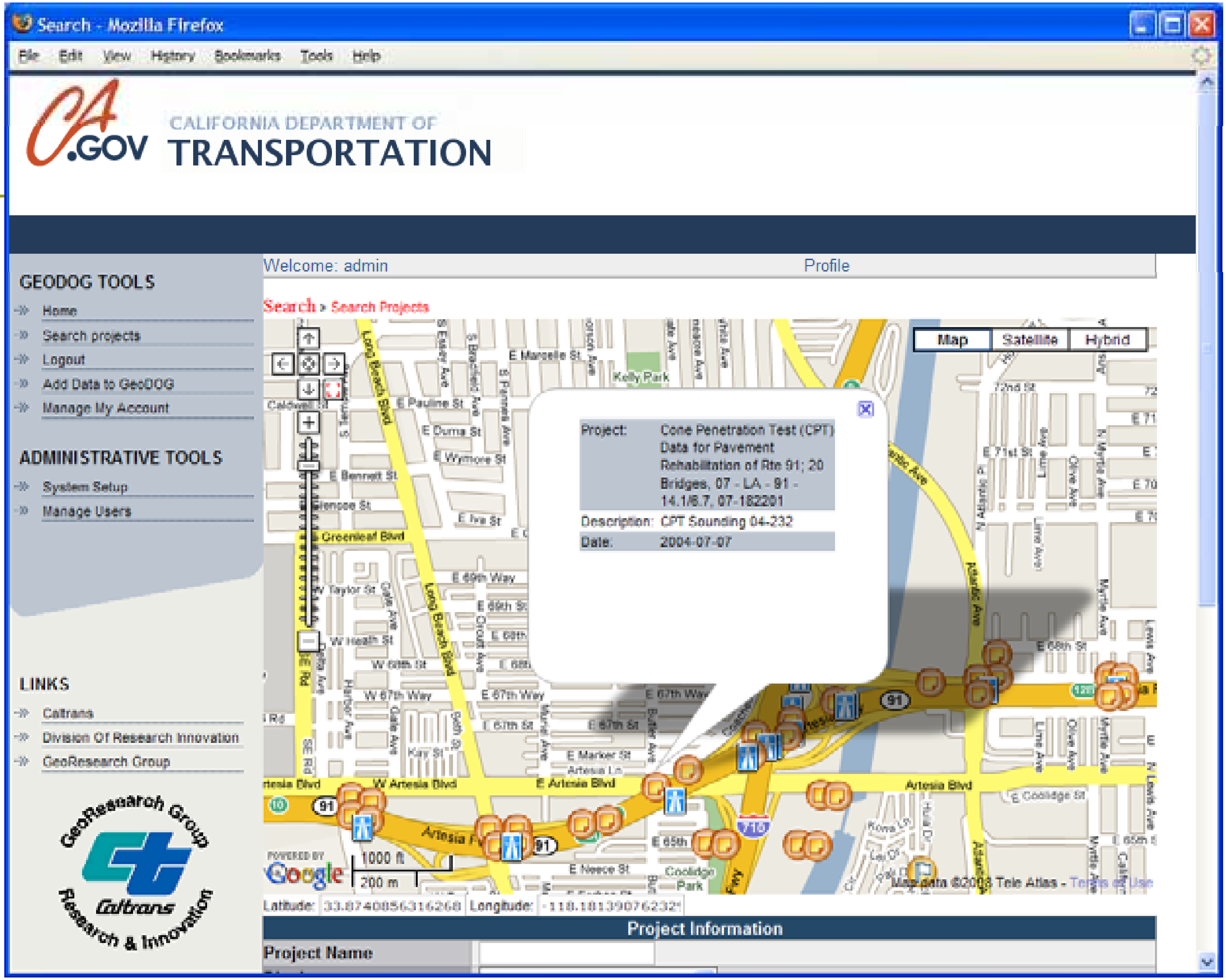
- Borehole data
  - From field to office
  - Intraoffice (among software)
  - Interoffice (among staff)
  - From office to External

California

- GeoDOG
  - **D**igital Repository **O**f **G**eotechnical Services
  - Central data repository on the Caltrans web for Department’s geotechnical documents and data.
  - Uses web-based map interface (GoogleMaps API) for browsing/searching
  - Facilitates data exchange between soils lab, engineer, and drafting services
  - Supports upload/download of DIGGS files
- [COSMOS Virtual Data Center]

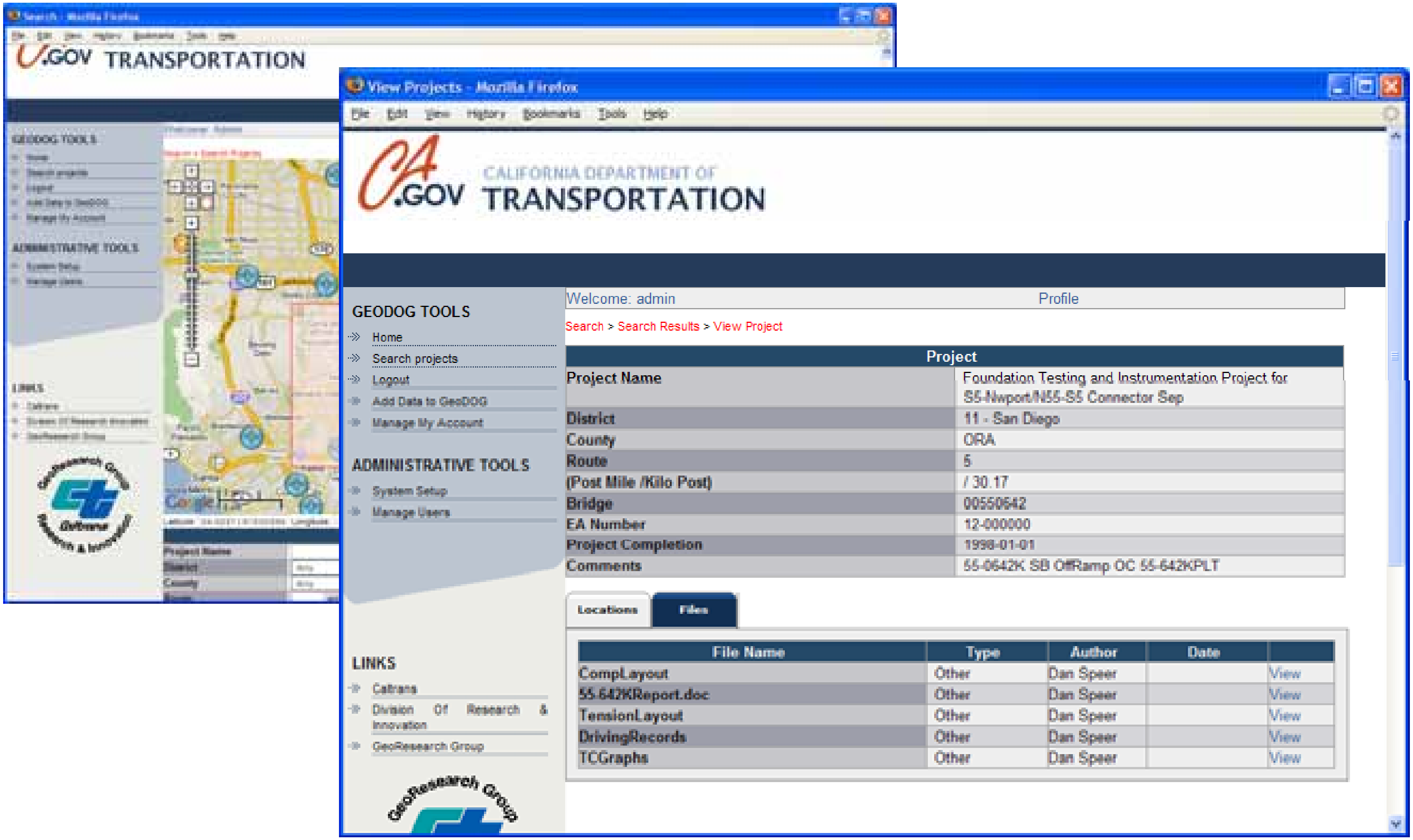


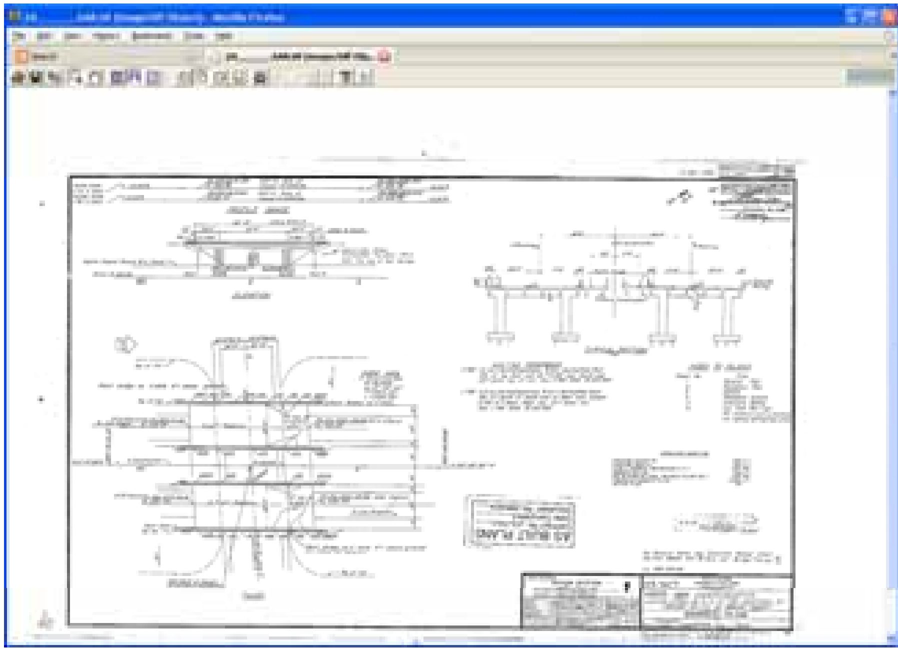
Search - Mozilla Firefox
File Edit View History Bookmarks Tools Help
CA.GOV
CALIFORNIA DEPARTMENT OF
TRANSPORTATION
GEODOG TOOLS
Home
Search projects
Logout
Add Data to GeoDOG
Manage My Account
ADMINISTRATIVE TOOLS
System Setup
Manage Users
LINKS
Caltrans
Division Of Research Innovation
GeoResearch Group
GeoResearch Group
Caltrans
Research & Innovation
Welcome Admin
Profile
Search » Search Projects
Map
Satellite
Hybrid
Cone Penetration Test (CPT) Data for 08-RIV-215
Cone Penetration Test (CPT) Data for 215 Widening
Cone Penetration Test (CPT) Data for 215/259 Segment 2
...and 37 more
Bakersfield
Santa Barbara
Oxnard
Los Angeles
Long Beach
Santa Ana
Riverside
San Diego
Project Information
Project Name
District
County
Route
eg: 80
Postmile /Kilopost
to
eg: 90 to 120
Bridge
eg: 34-0012 L
EA Number
eg 01-123456
Project Completion
To
eg: 2004-12-31 (YYYY-MM-DD)
Search Limit
50
Projects
Location Information
Location Description
Latitude
Min
Max
Longitude
Min
Max

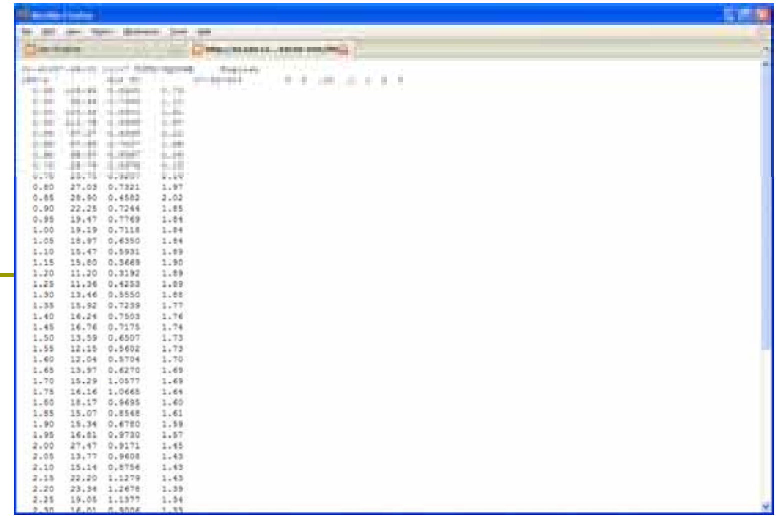


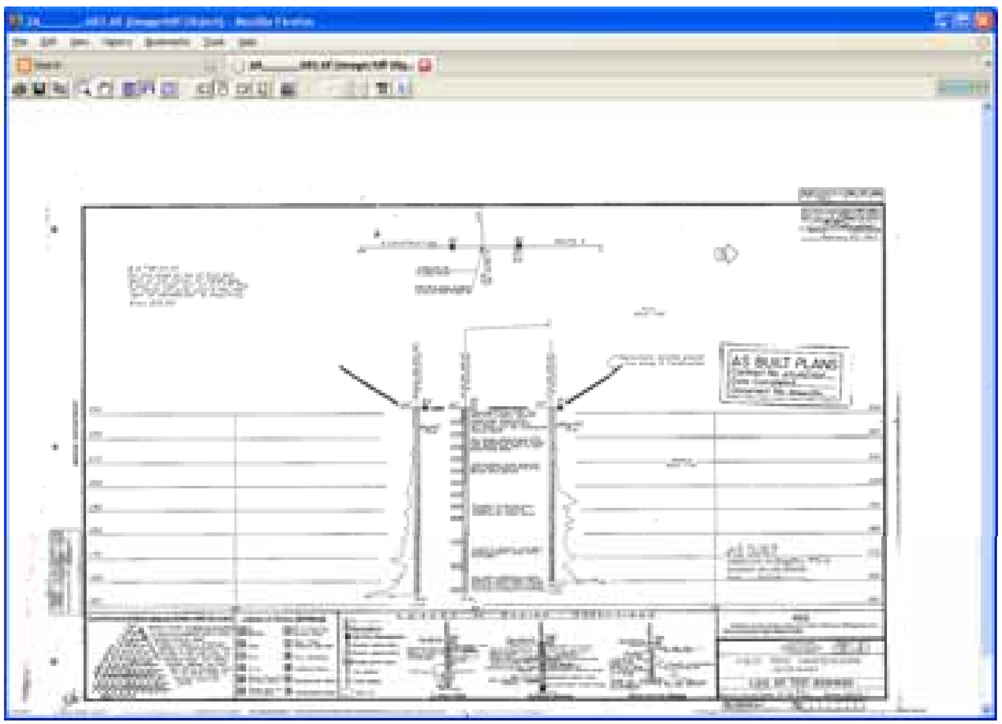
Search - Mozilla Firefox
File Edit View History Bookmarks Tools Help
CA.GOV
CALIFORNIA DEPARTMENT OF
TRANSPORTATION
Welcome: admin
Profile
GEODOG TOOLS
Home
Search projects
Logout
Add Data to GeoDOG
Manage My Account
ADMINISTRATIVE TOOLS
System Setup
Manage Users
LINKS
Caltrans
Division Of Research Innovation
GeoResearch Group
GeoResearch Group
Caltrans
Research & Innovation
Search » Search Projects
Map Satellite Hybrid
Project: Cone Penetration Test (CPT) Data for Pavement Rehabilitation of Rte 91; 20 Bridges, 07 - LA - 91 - 14.1/6.7, 07-182201
Description: CPT Sounding 04-232
Date: 2004-07-07
Latitude: 33.87408856316268 Longitude: -118.18139076232
Project Information
Project Name

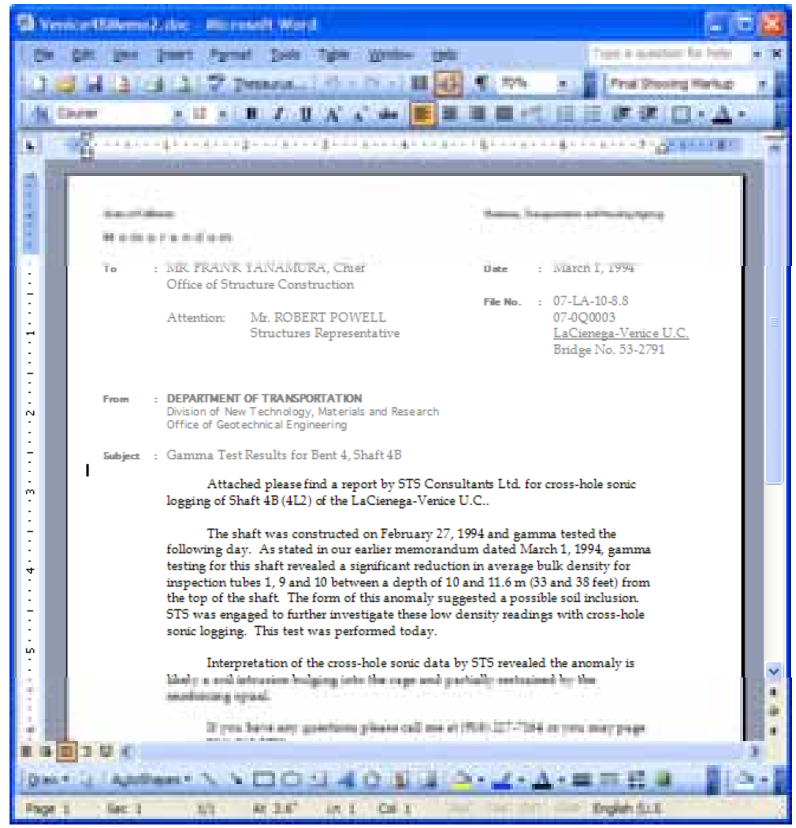
Caltrans GeoDOG







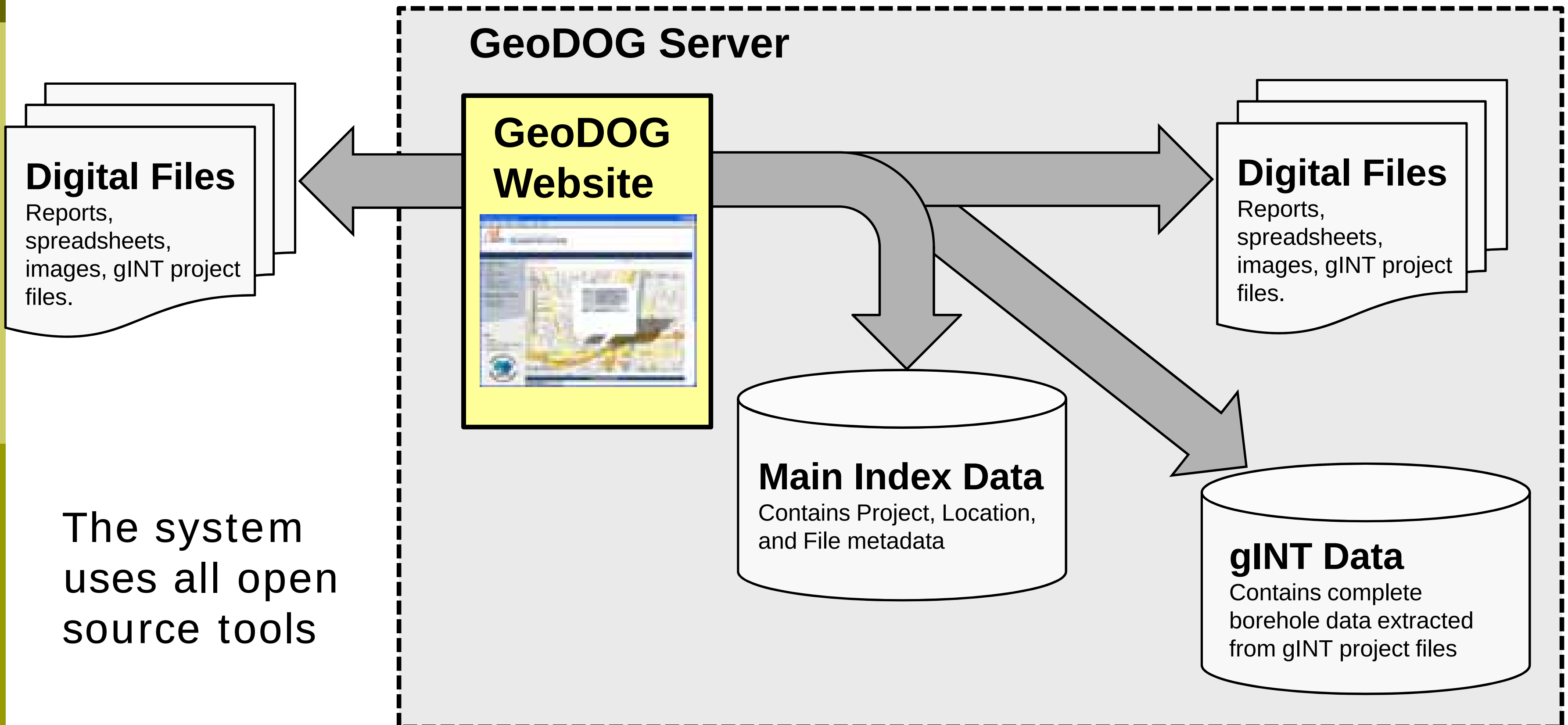




To : MR. FRANK YANAMURA, Chief
Office of Structure Construction
Attention: Mr. ROBERT POWELL
Structures Representative
Date : March 1, 1994
File No. : 07-LA-10-8.8
07-0Q0003
LaCienega-Venice U.C.
Bridge No. 53-2791
From : DEPARTMENT OF TRANSPORTATION
Division of New Technology, Materials and Research
Office of Geotechnical Engineering
Subject : Gamma Test Results for Bent 4, Shaft 4B
Attached please find a report by STS Consultants Ltd. for cross-hole sonic logging of Shaft 4B (4L2) of the LaCienega-Venice U.C..
The shaft was constructed on February 27, 1994 and gamma tested the following day. As stated in our earlier memorandum dated March 1, 1994, gamma testing for this shaft revealed a significant reduction in average bulk density for inspection tubes 1, 9 and 10 between a depth of 10 and 11.6 m (33 and 38 feet) from the top of the shaft. The form of this anomaly suggested a possible soil inclusion. STS was engaged to further investigate these low density readings with cross-hole sonic logging. This test was performed today.
Interpretation of the cross-hole sonic data by STS revealed the anomaly is

Caltrans GeoDOG

- The server hosts the GeoDOG web interface, maintains an index of available data, stores all files, and parses/stores gINT data.



The system uses all open source tools

Connecticut

- Working on developing a program to import/export DIGGS data to/from our enterprise database
  - Not yet completed
  - Effort is limited to borehole data (location, drilling details, sample info, etc.)
  - Will expand to other data types in time
  - Will start requiring all our outside consultants to provide a DIGGS file w/project geotechnical data; currently receive MS Access files

Connecticut

- "...no plans to go independent of DIGGS provided the software vendors **develop the necessary tools**. If the tools aren't developed, that would change things."

Florida

- FDOT Geotechnical Database
- Bridge Software Institute (BSI) has developed three unique pieces of software that can access the database
  - FB-Deep
  - Pile Technician
  - Database Spreadsheets

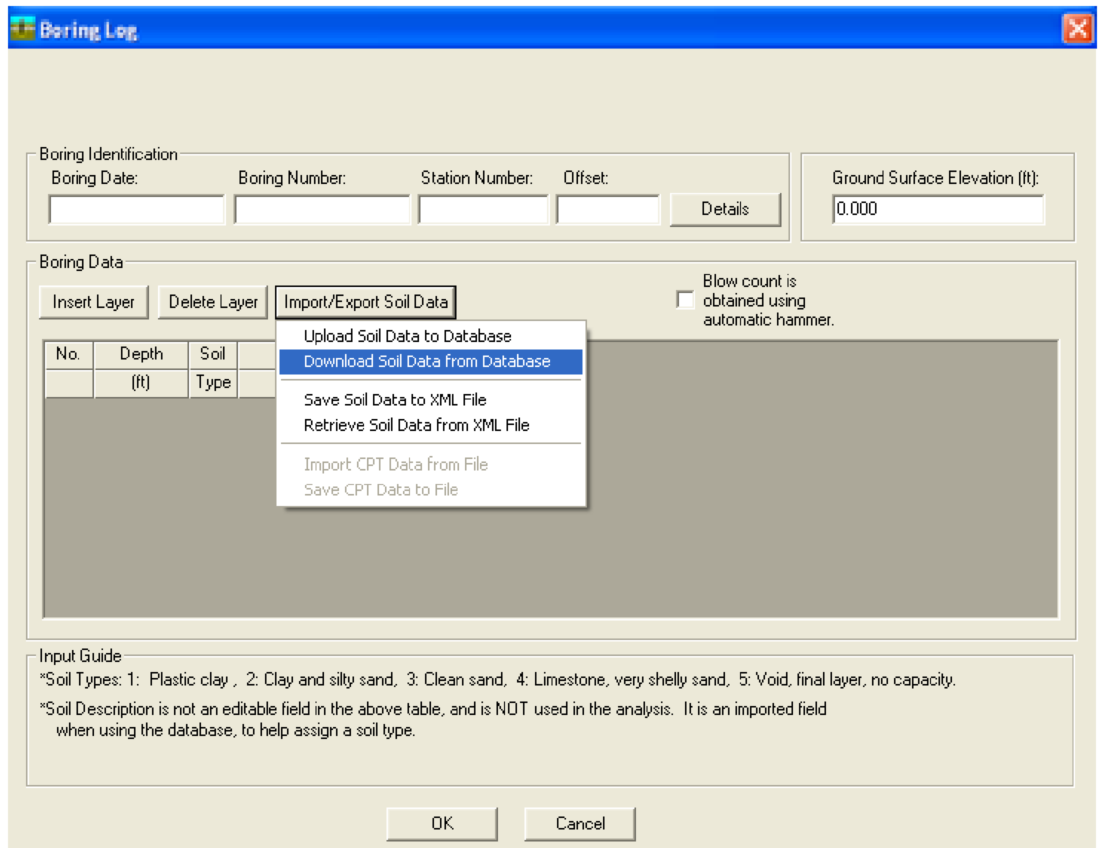
FB-Deep

- A computer program that computes
  - Static Axial Capacity of driven piles and drilled shafts,
  - Using SPT analysis for drilled shafts and SPT and CPT analysis for driven piles

- From the FB-Deep interface, users can:
  - a) upload soil data to the database
  - b) download soil data from the database

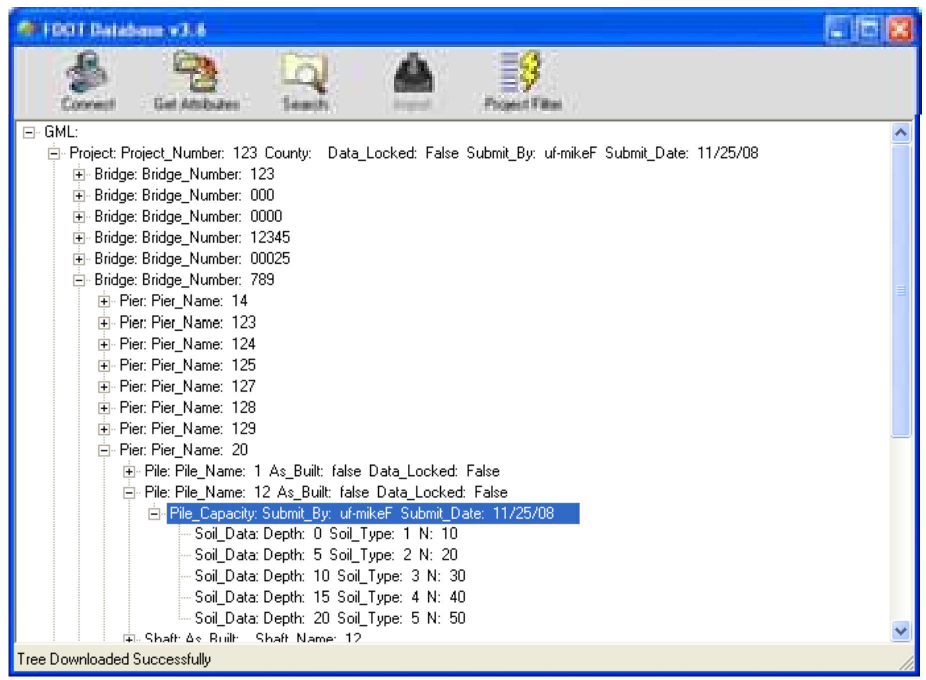
FB-Deep Download Example

On the Boring Log, select the "Download Soil Data from Database" menu option. This menu is launched by clicking the "Import/Export Soil Data" button.



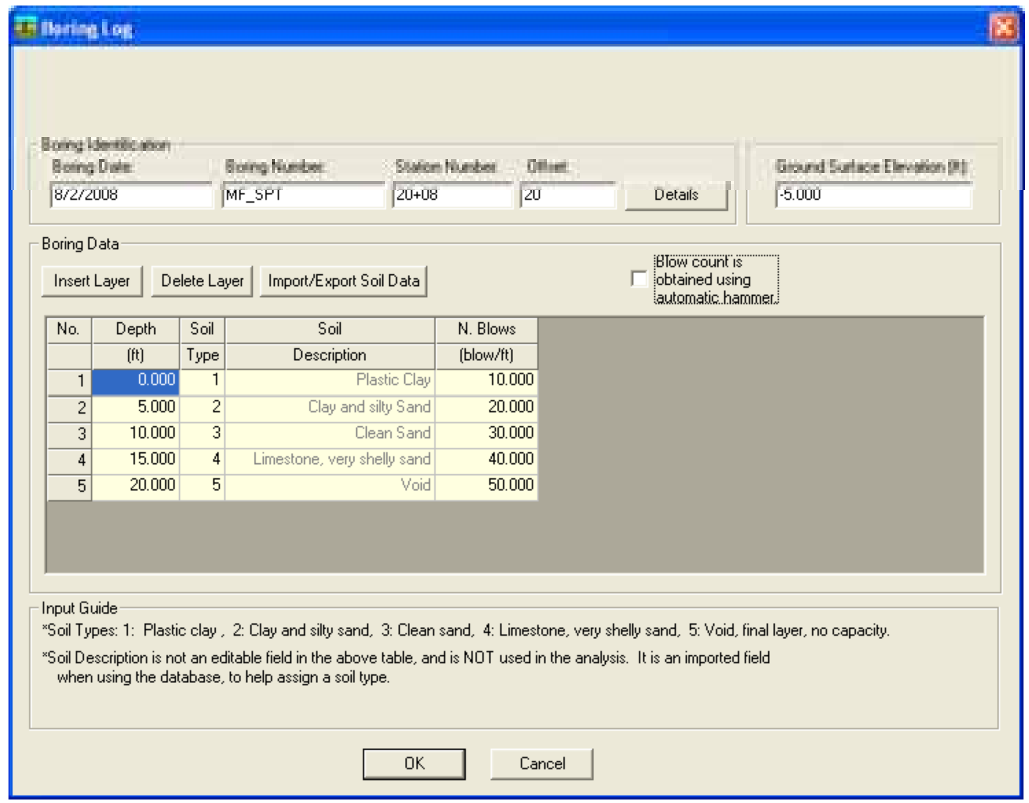
Database Tree

- Search the projects in the database tree for the desired soil data
  - To import the data into FB-Deep, highlight the desired section and click the “Import” button on the toolbar.



Download Complete

- The downloaded soil data appears in the FB-Deep interface.
  - It can be used in a pile or shaft analysis, or saved to a file for future use.

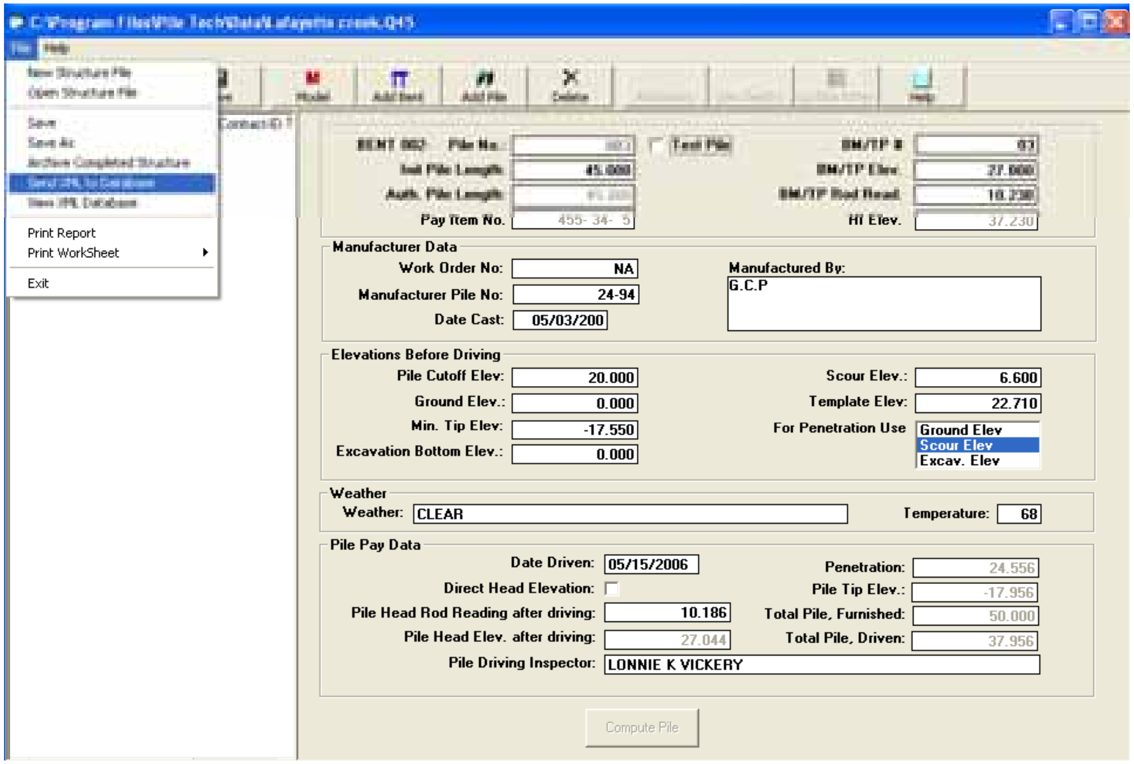


Pile Technician

- Was developed for the FDOT to provide a fast and efficient manner of entering Pile data to calculate payment for work preformed by the contractor.
- Pile driving history and the necessary documentation can be uploaded to the Database.

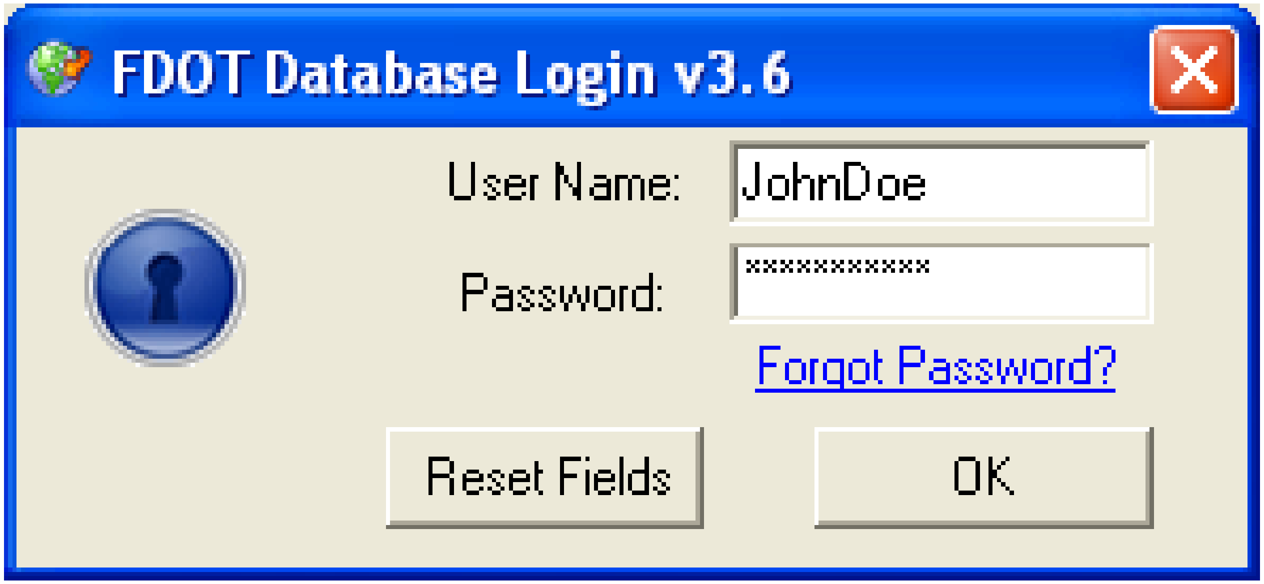
Pile Technician Upload Example

- After data input is complete, select “Send XML to Database” from the file menu.



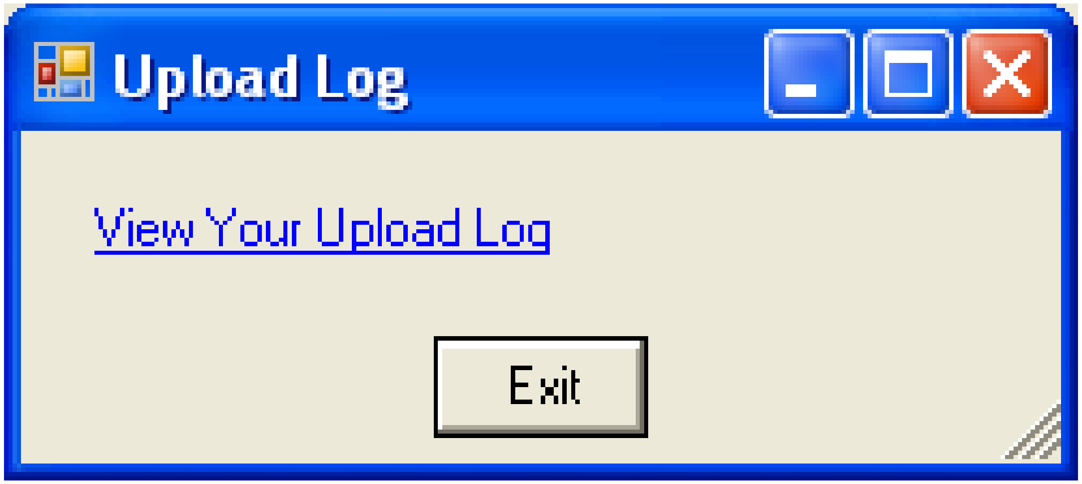
Log-In Screen

- Enter a User Name and Password
  - Only authorized users will have established permissions to view and utilize the database



View the Upload Log

- After the upload is completed, the Upload Log screen appears.
- To view the results of the uploading process, click the “View Your Upload Log” link.



Upload Log

- The Upload Log shows what data has been updated in the database.

```
Your DLL Version: 3.6
UploadFileName_withPath: C:\DOCUME~1\cmi\LOCALS~1\Temp\print.tif
UploadFileCode:BridgeReport
UploadFileGUIDName:10_4f14c031_print.tif
UploadFileLocation:BridgeReport\10_4f14c031_print.tif


Projects- Project_Number='STEST'
    Not Updated
Bridges- Bridge_Number='600172'
    Not Updated
Piers- Pier_Name='BENT 002-'
    Not Updated
Piles- Pile_Name='003'***AsBuilt=1
    Not Updated
Pile_Driving_Info - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
Pile_Driving_Log - Updated
```

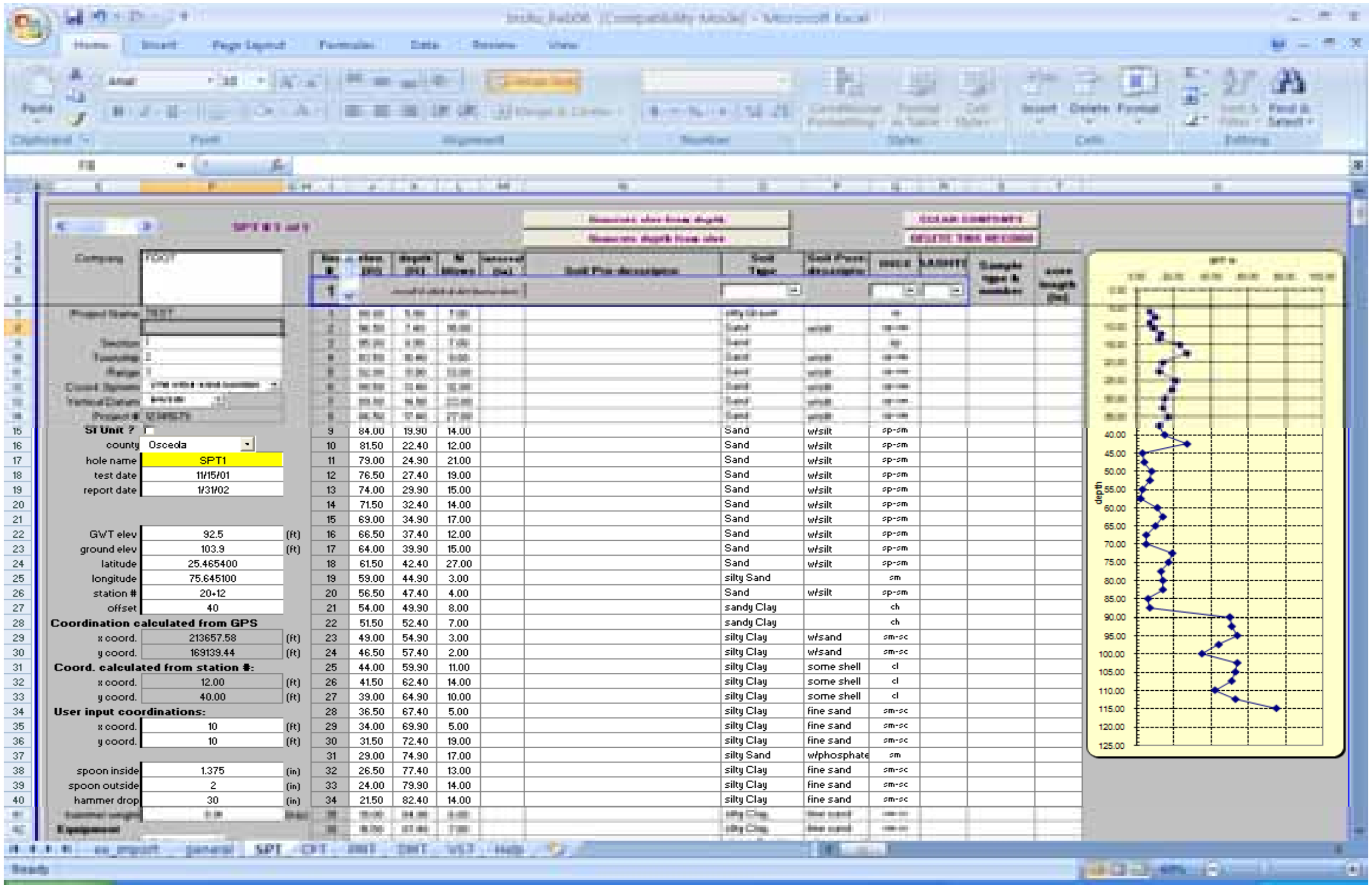
Database Spreadsheets

- Database spreadsheets are excel files to assist in the transferring of data to the FDOT Database.
- The spreadsheets also offer graphing features to help distinguish data.

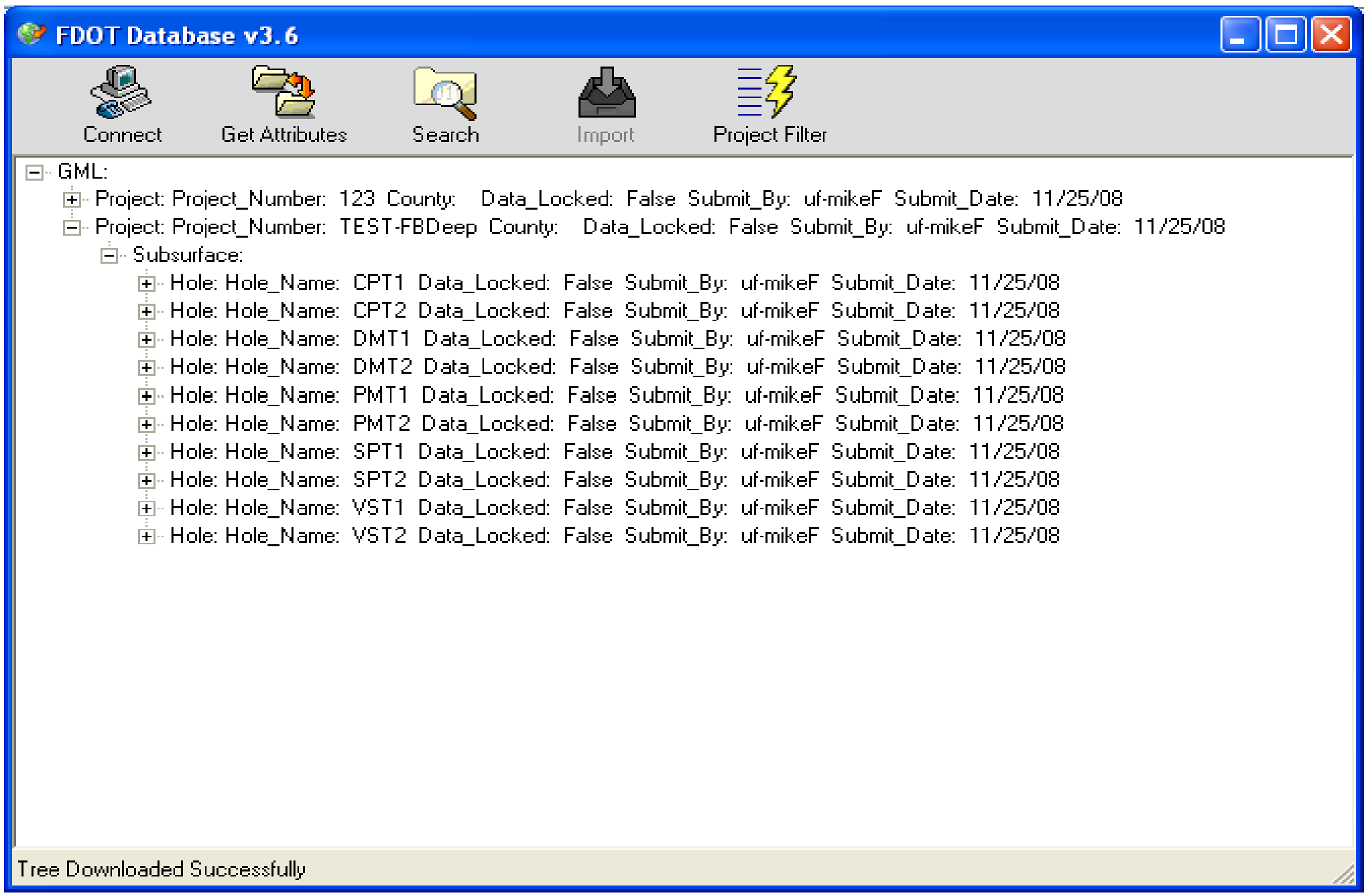
Data Types

- Inisitu -
  - **SPT, CPT, PMT, DMT, VST**
- Lab -
  - **Rock Strength, Triaxial, Sieve Analysis, Oedometer, Consolidation**
- Design -
  - **Driven Piles and Drilled Shafts**
- Load Test -
  - **Static, Osterberg, Statnamic**

Example of In-situ spreadsheet



Data is uploaded and can be found under its related project

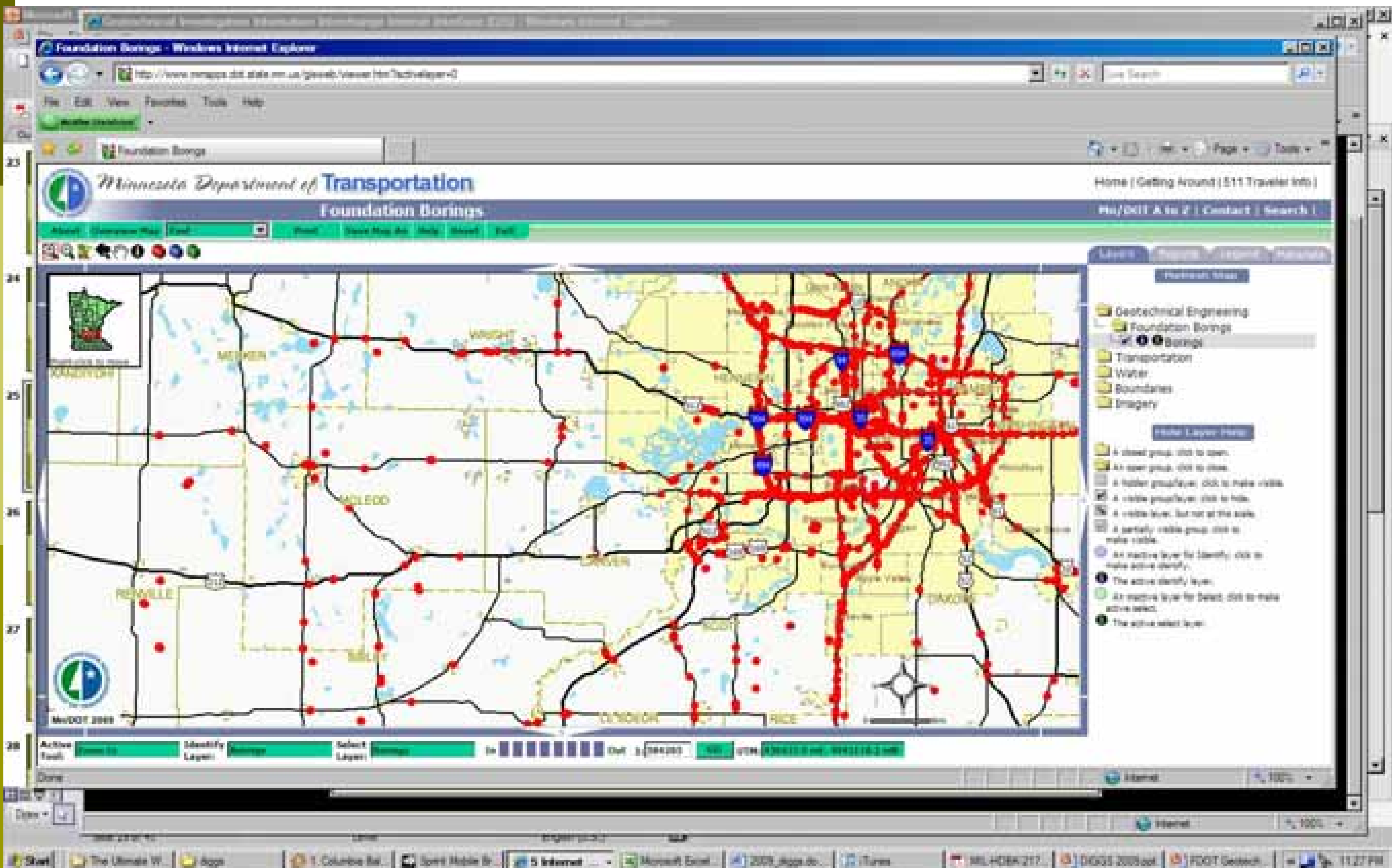


BSI
Bridge Software Institute

Minnesota

- Using gINT since 1993 for logs
- Point data for more than 30,000 records is available for searching
- Extensive use of automated importing/exporting
  - Trimble/Vertek/Microstation/gINT
- On-line ArcIMS database is operational
  - Boring location and summary data
  - Static PDF files of boring logs
  - http://www.mrr.dot.state.mn.us/geotechnical/foundations/Gis/gi5_splash.html

Mn/DOT "Gi5" Application



Minnesota (2004)

- **Monday - December 13, 2004 8:27 AM**

- Break the effort into smaller pieces with definable goals and outcomes. I would find it a whole lot easier to support a project with a lower price tag and a much shorter time frame.
- Pick the piece that would give the most chance for success (I would suggest boring information). I could support an effort that would produce a useable product in a 6 month time period, possibly a year.
- I have seen several efforts by Mn/DOT that set lofty goals and tried to satisfy every possible user flounder and fail because the technology changed before everything was developed. A couple of those are still floundering.

- Start with an already developed application. e.g. start with what the UK has developed (unless it is protected) and modify as needed to fit where we are trying to go.

Minnesota (2004)

- Develop the application so the next steps can be added
- States need to keep in mind that even if this pooled fund study proceeds, they will still need to spend a lot of time and or money getting their data into a useable form. The alternative is to forget the historical data and start from today.

- Although I believe this effort to be worthwhile, I am having a hard time justifying in my own mind that the proposal as it currently exists is worth supporting.
- "I need to see something with a much shorter time frame and lower price tag. **Start with one year and show me some results and we can go from there**."

Minnesota (2004)

- *Strongly* suggested is a prototype or 'pilot project' to be implemented first and that the whole project be constructed in phases, with the ability to change and expand the project built-in from the beginning.
- In this way, as new components of functionality are added, they can be built on the foundations *and* lessons learned from the previous project components.
- Perhaps a series of files for individual data components?

Minnesota (2004)

- Also suggested: We start with a set of data that is available and whose incorporation into the system has a high rate of return on the benefit side of the effort. In this case it would likely be soil borings, geotechnical asset infrastructure, or hazard/maintenance issues
- In this way, participants could be encouraged with short term, lower cost, investments and a tractable positive outcome. Both the time line and costs could be significantly condensed- with the additional benefit that during this time individual pooled fund contributors could identify their interests for the next steps and additions to the system.

Minnesota (2004)

- It was suggested by a database designer on review of the DIGGS plan- that the plan, as is [2004] , was **“ ‘typical of government,’ particularly defense, and although not ‘doomed to failure’, neither a blueprint for great success.”**

Minnesota (2004)

- Foresee first uses of DIGGS:
  - Exchanging data with Universities/out-of-state consultants
  - Uploading data from:
    - Inclinometers, Piezometers, and other monitoring equipment
    - Lab Testing equipment
    - Field in-situ tests (CPT, DMT, PMT, etc...)
      - Direct import into direct CPT/DMT methods?
    - Roadway borings
      - Direct import into Microstation/Geopak
- Exchanging Pile/Construction QC/QA PDA data for calculations
  - Direct import into WEAP/CAPWAP
- Leading to larger Int/Ext Data Warehousing?

Ohio

- Web-based Drilling Request System (DRS) will transfer drilling and sampling instructions to PC tablets
- Testing pLOG and IDEF for field data collection (drilling and sampling) with PC tablets; data transfer will be based on DIGGS

Ohio

- Contract with U. of Akron for detailed assessment of geotechnical laboratory; computer integrated work stations developed as pilot; detailed assessment defines schema and dictionary of development of a LIMs system; LIMs will be integrated with EQUIS using Oracle 10g as enterprise system; DIGGS exchange standard will be used for data transfer between LIMs/EQUIS

- Contract with EarthSoft for the development of a geotechnical/geoenvironmental data management system using EQUIS; data exchange using DIGGS standard

Ohio

- Established a document management system using Falcon; all scanned geotechnical documents have been indexed; requested modifications to initiate use of XML exchange format of associated data and information

Ohio

- Completed Scanning of boring logs; also scanned all other Geotechnical reports/other information
- GPS coordinates required for the past 2.5 yeas;
  - Prior to this, we are georeferencing plan sheets to aerial photograph of the same time period and digitizing boring locations to obtain coordinates.
- Using hand-held Windows-mobile data collectors since 2004 and entering information into databases or GIS systems
  - Started with GeoMedia and switched to ESRI; currently, migrating to use 3D Analyst and Spatial Analyst.

Ohio

- Development of the exchange standard will continue; anticipate annual maintenance agreement to keep up with changes in DIGGS standards and improvements in software.

Tennessee

- Tennessee DOT is not currently engaged in any direct effort related to DIGGS at this time.
- We have been using Word as our means of developing Boring Records and we have looked at programs such as gINT but we were waiting on the outcome of DIGGS before making any big investment in a geotechnical data program.

Tennessee

- We have the means to collect additional data such as GPS coordinates for our borings from survey data but we do not currently put that information on our Boring Records.

- We have several GIS initiatives related to Geohazards such as rockfalls, sinkholes, landslides, etc.
  - We have databases set up for these.

Tennessee

- In the process of looking into the Materials and Laboratory Management modules of the Site Manager Enterprise program and this may affect the final outcome of some of our data storage and exchange issues.

Tennessee

- My biggest concern with the DIGGS effort right now is that we increased the initiative beyond the original scope (this was/is a state pooled fund study) which was to develop a means for state DOT geotechnical units to easily exchange information.

Tennessee

- "If it takes a lot more money to complete the project then I am afraid **the additional funding will not be available from the states** and no other entity will be willing to fund the effort and DIGGS will "die on the vine". I hope this is not the case and maybe we can use the upcoming meeting to make sure this does not happen."

And The Rest…

- Georgia
  - Embroidering data on cotton textiles
- Indiana
  - 4th grade school students memorizing logs
- Kentucky
  - Keypunching onto Hollerith cards
- Missouri
  - Monks creating “illustrated manuscripts”
- North Carolina
  - Stored by surplus Enigma cipher machines

After Borehole Data?

- Construction Verification
- In-situ sensor data
  - Time Domain Data
- 2D fences, geophysics
- 3D
- Geophysical Data
- Geoenvironmental Data
- Asset Quality/Management
  - Time Domain Data

QC/QA Testing

- PDA/CAPWAP
- O-Cell/SLT

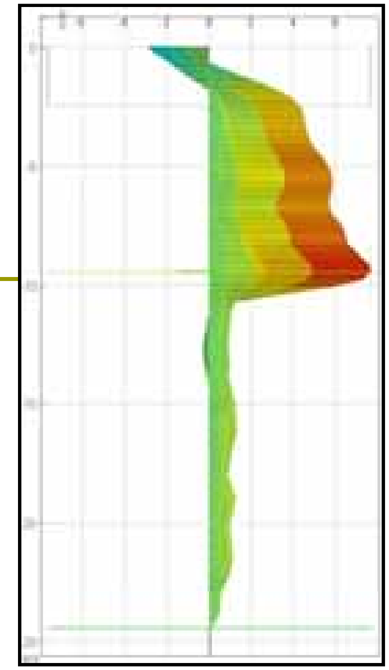
QC/QA Construction Monitoring

- Contractor Rig
  - Vertical Drains
  - DMM
  - Auger Cast Piles



Inclinometers

- Traditional
- In-place

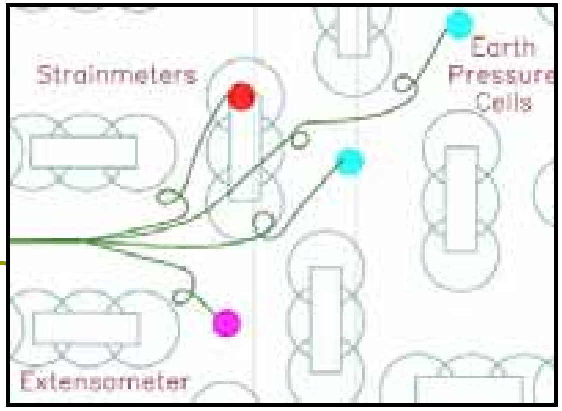


Piezometers

- Field Data Transfer
  - From logger to Databases

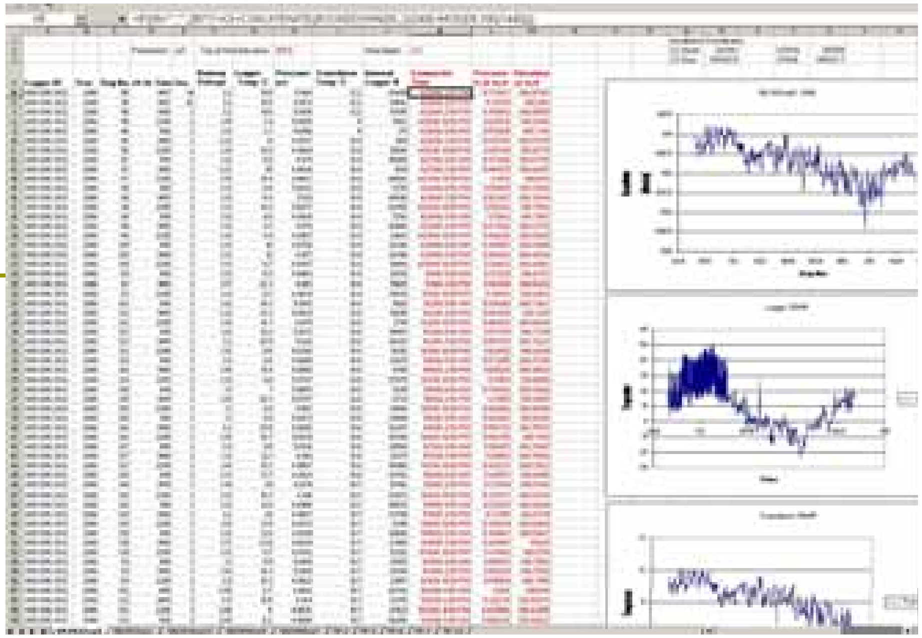
Sensors

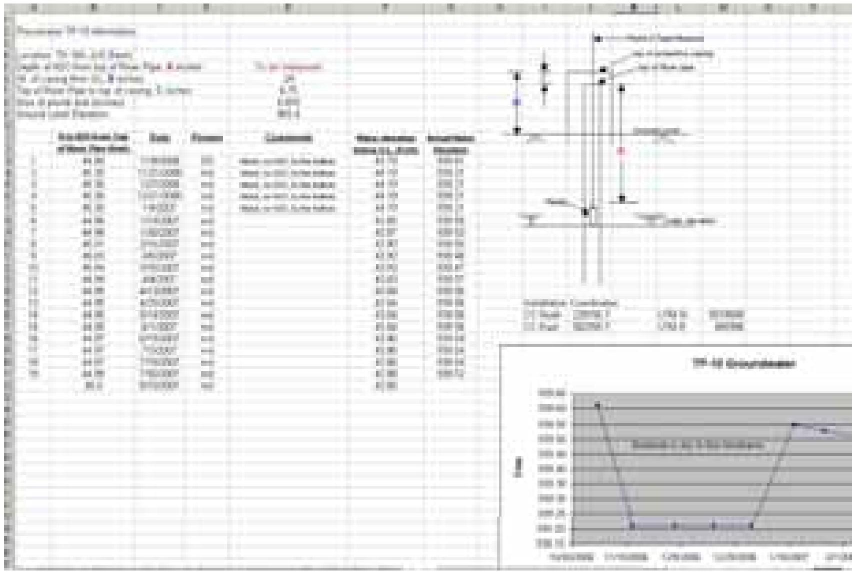
- Earth Pressure Cells
- Settlement Cells
- Strain Gages
- Load Cells

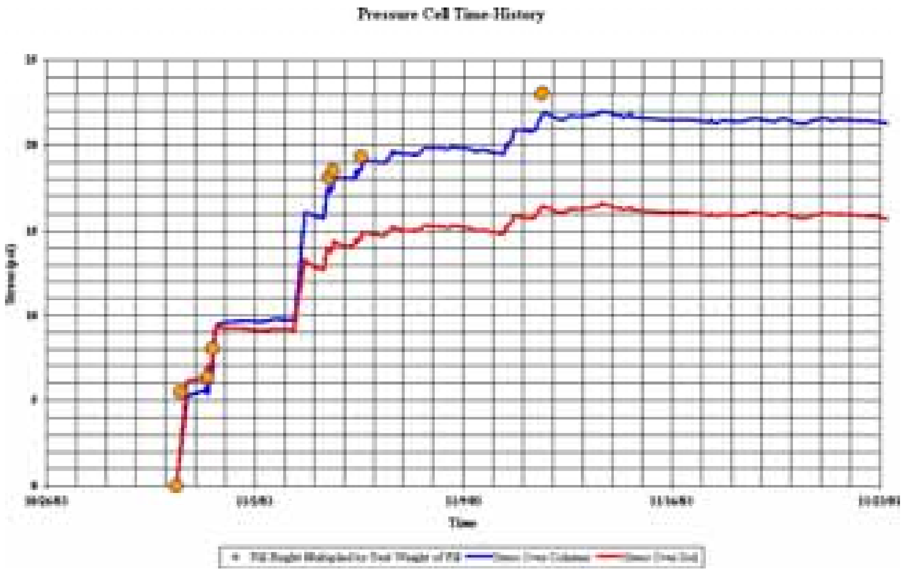


Sensor Data

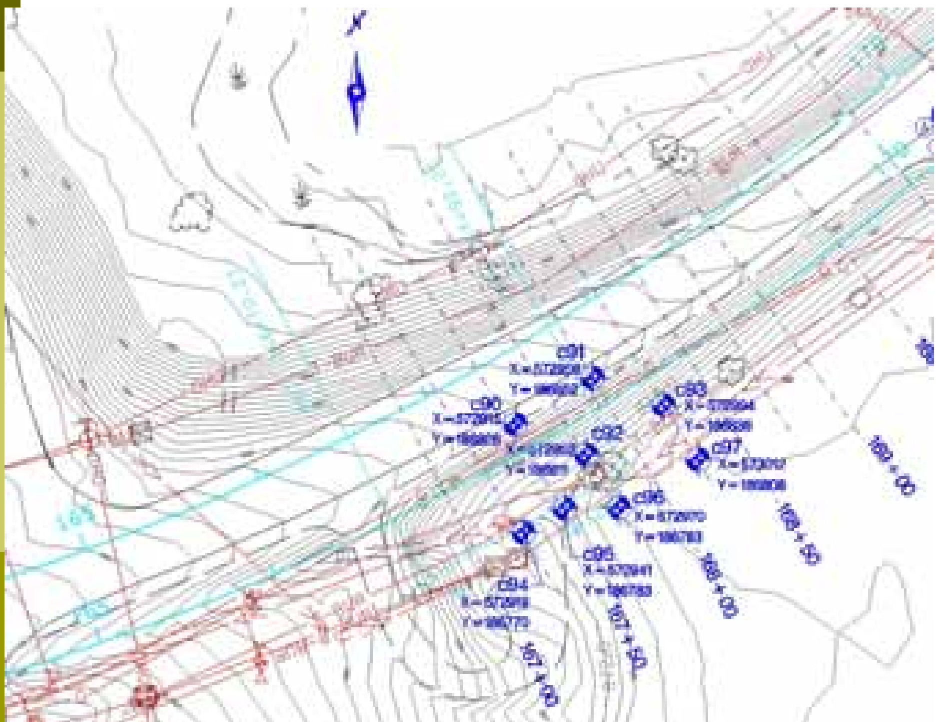
- Manual
- Automated







DIGGS



Commentary

- Needs to Work
  - A functional product **with applications** is critical
- Needs to be Simple (Adoptable)
  - Instrumentation vendors and DOT's
  - Correspondence Files or similar 'one-to-one' relations
- Needs to be in a small package
  - Small file sizes
- Needs to be supported by:
  - Software providers/Hardware MFG

Commentary

- Needs to meet user needs/fit user expectations
  - Exchange data among ‘.mdb’, ‘.xls’, ‘.csv’
- Does not need to be ‘all-inclusive’ to start
  - Extensible- perhaps develop “plug-ins” for new parts?
    - Start with lab ‘results’ and test type
  - Multiple formats (e.g. mp1, mp2, mp3, mp4)
  - Perhaps akin to many available data/photo formats?
- Needs to be **useful** to those paying for it- in addition to those benefiting from it

DIGGS

- Current databases and electronic (MS Excel or similar) files **contain** most geotechnical information we have been discussing
- It seems reasonable that it should be possible to move it from one place to another accurately

Geotechnics in the "real world"

There are challenges, and few things are as simple as they would seem to be at the start

Introduction to Keynetix

- UK's leading Supplier of Geotechnical and Geoevironmental Data management software
- Products ranging from handheld data collection (PocketSI) and Lab management (KeyLAB) borehole logging (HoleBASE) and AutoCAD presentation.
- Developers of the UK Highways Agency's Geotechnical Data Management System (www.HAGDMS.com), in partnership with Mott MacDonald

What We've Done with DIGGS data

- DIGGS exporter for HoleBASE 3.1
  - – Any project
  - – Any number of holes
  - – Any DIGGS objects
- Been extremely useful for internal example creation – now available to any HoleBASE software user

What We’ve Done with DIGGS data

- DIGGS Importer for Excel
  – Select DIGGS file
  – Structure flattened into excel spreadsheets – one sheet per object type
  – Displays links between samples and parents
- Usefulness increasing – available on request at the moment as it is an alpha release
- Took Keynetix senior developer 3 days to write

What We've done with DIGGS

- AGS – DIGGS converter
  - Shown at AGS meeting in June 2008
  - Only converts part of the AGS structure at present
  - Available on request as it is currently an alpha release

What We Think

- DIGGS gives us greater power to produce software that can move our industry forward
- We can work with DIGGS in its current format.
- However we feel that there are areas that could be simplified but it is equally important not to over simplify them
- It’s not changing the format that is important, as maybe you can live with things you don’t like, its making it a standard that everyone uses and supports that should be the goal

A Live Demonstration of Geotechnical Data Transfer Using DIGGS

Scot D. Weaver, M.S.E.
EarthSoft, Inc.

What is EQuIS?

Environmental Quality Information System

- The most widely used system in the world for managing technical sample data:
  - Groundwater
  - Surface Water (Stream or Lake/Reservoir)
  - Geology / Geotechnical
  - Meteorological
  - Air
- Data Quality first, then Data Usability
- Open System

Who is EarthSoft?

EarthSoft, Inc.

- EarthSoft founded (1994) as a Software Company
- ~ 40 Software and Environmental Professionals
- Award-winning 10-person Help Desk
- Same Management Team for 10 Years
- Revenue from licenses nearly tripled in 2008
- Latest Version is EQuIS 5

Who is EarthSoft?

EQuIS is used by…

- Industrials
  - Four of Top 10 BusinessWeek Global 1000
- Consultants and Labs
  - Nine of ENR Top 10, 70% of ENR Top 100
  - Over 400 labs globally
- Government
  - Seven of 10 US EPA Regions; almost 20 states
  - Cities, counties, port authorities, water/waste management districts, ...

Who is EarthSoft?

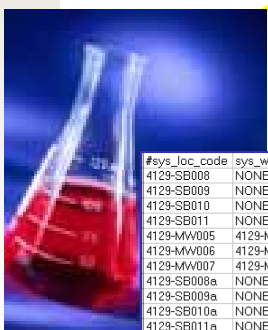
Why does this matter?

- Volume means...
  - Extensive reinvestment in development
  - Safety in numbers
- Longevity with experience and stability
- Not just a software vendor, but expertise, “Knowledgebase” resource

Field Data Collection

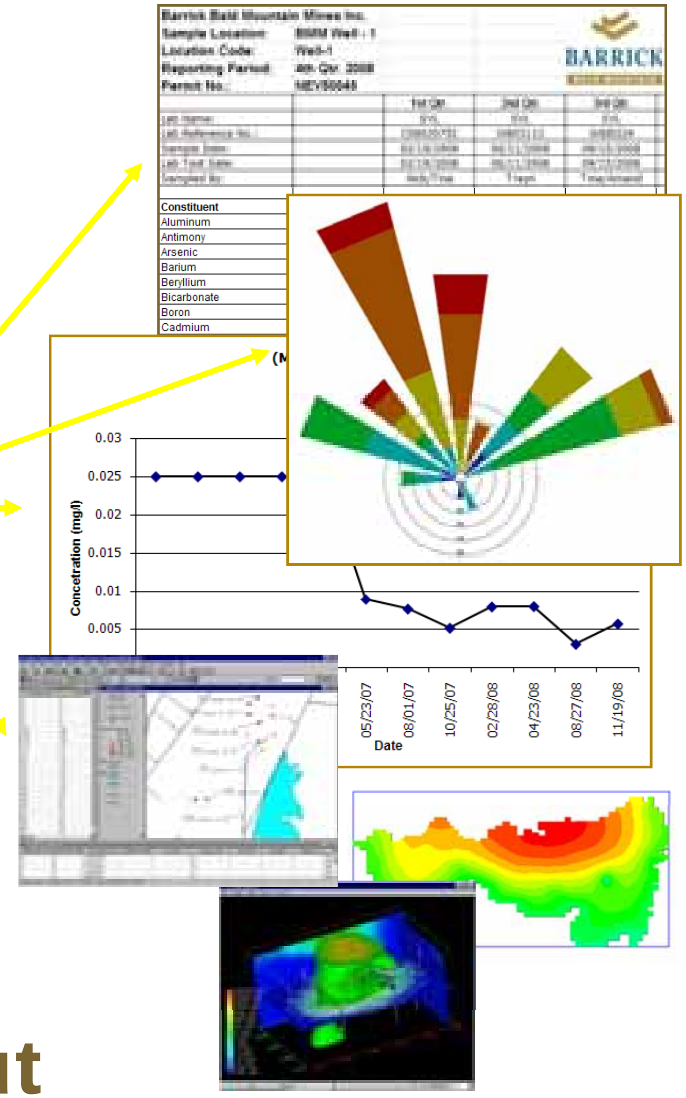
Monitoring / Instrumentation

EQuIS 5



#sys_loc_code	sys_well_code	X_coord	Y_coord
4129-SB008	NONE	620071.4941	3339792.757
4129-SB009	NONE	620085.495	3339809.48
4129-SB010	NONE	619666.5208	3339949.574
4129-SB011	NONE	619505.707	3339908.768
4129-MW005	4129-MW005	619443.7651	3339953.166
4129-MW006	4129-MW006	620133.2989	3340195.292
4129-MW007	4129-MW007	620162.5517	3340127.655
4129-SB008a	NONE	620071.4941	3339792.757
4129-SB009a	NONE	620085.495	3339809.48
4129-SB010a	NONE	619666.5208	3339949.574
4129-SB011a	NONE	619505.707	3339908.768
4129-MW005a	4129-MW005	619443.7651	3339953.166
4129-MW006a	4129-MW006	620133.2989	3340195.292
4129-MW007a	4129-MW007	620162.5517	3340127.655

Laboratory EDDs



Data In, Information Out

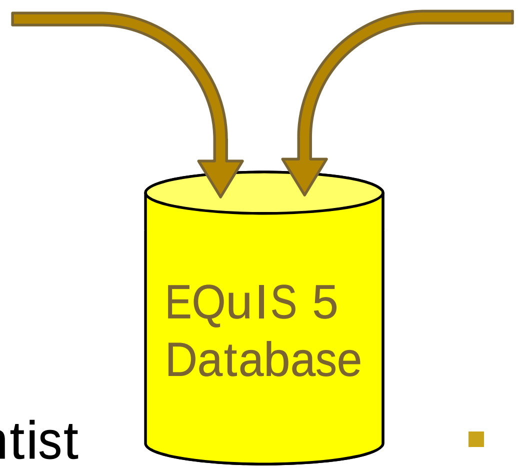
EarthSoft

EQuIS 5 Architecture

- EQuIS 5 Database (SQL Server or Oracle)
- EQuIS 5 Professional
  - Requires software installation
  - Editing data, reports; advanced analysis, etc.
- EQuIS 5 Enterprise
  - Access anywhere via the Web
  - Automated processing (input and output)
  - Simple, quick access to data

EQuIS 5 Architecture

EQuIS 5 PROFESSIONAL



EQuIS 5 ENTERPRISE

- Data manager, Scientist
- Windows app (installation)
- Data importing, editing
- Advanced analysis, modeling*
- Ultimate flexibility

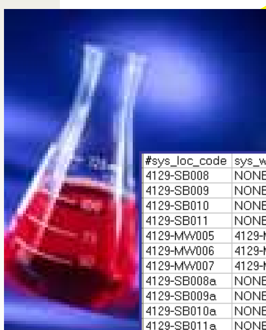
- Manager, Auditor, Executive
- Web browser (no install req.)
- Same database, same data
- Simple, quick, easy to use
- Automation

* May require additional third-party software

EarthSoft

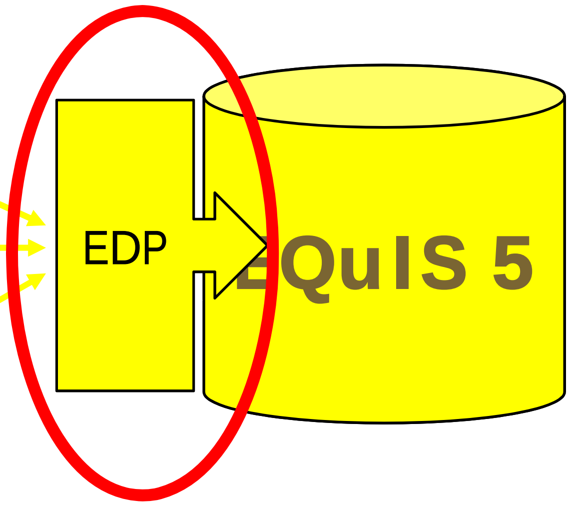
Field Data Collection

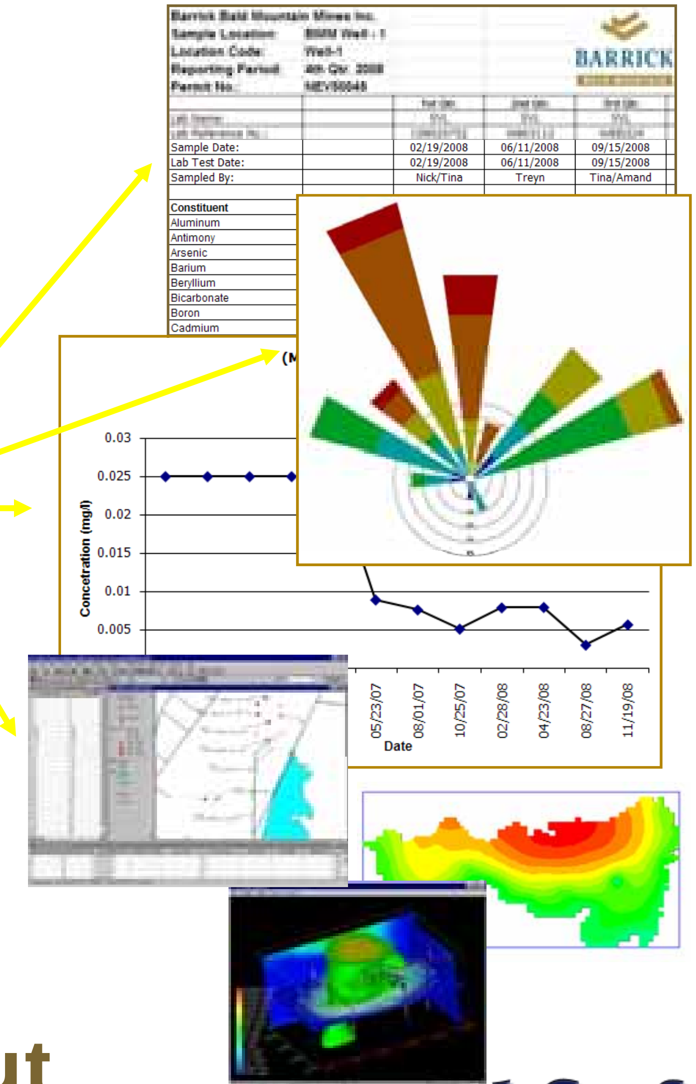
Monitoring / Instrumentation



#sys_loc_code	sys_well_code	X_coord	Y_coord
4129-SB008	NONE	620071.4941	3339792.757
4129-SB009	NONE	620085.495	3339809.48
4129-SB010	NONE	619666.5208	3339949.574
4129-SB011	NONE	619505.707	3339908.768
4129-MW005	4129-MW005	619443.7651	3339953.166
4129-MW006	4129-MW006	620133.2989	3340195.292
4129-MW007	4129-MW007	620162.5517	3340127.655
4129-SB008a	NONE	620071.4941	3339792.757
4129-SB009a	NONE	620085.495	3339809.48
4129-SB010a	NONE	619666.5208	3339949.574
4129-SB011a	NONE	619505.707	3339908.768
4129-MW005a	4129-MW005	619443.7651	3339953.166
4129-MW006a	4129-MW006	620133.2989	3340195.292
4129-MW007a	4129-MW007	620162.5517	3340127.655

Laboratory EDDs





Data In, Information Out

EarthSoft

What is DIGGS today?

Keys to success

- Tools that generate data in DIGGS format
- Tools that consume data in DIGGS format
- Tools that check for DIGGS correctness/ completeness
- Enterprise-level implementation (Automated Workflow)
- Added value: application of validation rules
- Added value: facilitation/simplification of data transfer

Dataforensics DIGGS Review

Scott L. Deaton, Ph.D.

President & Chief Software Architect

Dataforensics

Key Questions

- What has Dataforensics done with DIGGS?
- Is DIGGS ready to implement?
- What needs to be accomplished before it can be implemented?

What Dataforensics has done with DIGGS

- Dataforensics began reviewing DIGGS upon public release (August 2008)
- Initial reaction – it's very complex
- Review initially focused on CPT and DMT
  - Focus enabled us to begin to understand the fundamental concepts underlying DIGGS
- Review expanded to a cursory review of the geotechnical portion of schema
- Posted questions on DIGGS forum

Minor Problems Encountered

- Opening a DIGGS file in various XML editors took ~3 minutes
  - Non-DIGGS XML files do not cause same behavior
  - Need to remap schema files to local computer
  - Simplify namespaces (>200 files referenced)
- Schema diagrams too large to print
  - When printed the text is too small to read
  - Diagrams should be broken into smaller subsets (refer to ESRI object models)

Significant Problems Encountered

- Documentation inconsistent with examples
- Inconsistent terminology within schema
- Structural inconsistencies within schema
- Structural problems within schema
  - Validation problems
  - Self referencing tabular data
  - Inheritance overused
  - Recursion

Documentation inconsistencies

- “A Table containing delimited data. A default table is delimited by "." for decimals, By "," for columns and by " " for rows. Hence it's data block would look like this

  "12.345 12.345 12 12.345,67.890 67.890 67 67.890”

```
0.010,0.1300,0.40,0.0000,0.0013;
0.020,0.2400,0.40,0.1.0a,0.0078;
0.030,0.5500,0.40,0.0040,0.0126;
0.040,0.6800,0.40,0.0070,-0.0017;
0.050,0.7800,0.30,0.0120,-0.0121;
0.060,0.9000,0.30,0.0150,-0.0161;
0.070,0.9600,0.40,0.0200,0.0191;
0.080,1.0400,0.40,0.0240,-0.0120;
0.090,1.0700,0.30,0.0270,-0.0129;
1.0a0,1.1000,0.30,0.1.0a,-0.0123;
1.0a0,1.1300,0.40,0.0350,-0.0176;
```

- “Data validation is carried out using a set of rules that is the same for all parties in the data exchange. There can be no interpretation of the rules; therefore mistakes are much reduced“

KWA3

Table Object Problems

- Tables used to reduce file size
- Tables allow using a comma as the decimal separator (Europe) instead of decimal point
- Cannot validate tables
- Proposed alternative

  <Depth 0.01 0.02 0.03 0.04…</Depth>

  <Tip 0.13 0.24 0.55 0.68…</Tip>

Slide 7

KWA3 Questions seem aggressive/blaming. If you ask, they may have an answer. Or else they just didn't realize it and do not have an answer. Maybe not. Perhaps put out a statement demonstrating the inconsistencies.

Katie Aguilar, 3/3/2009

KWA1

Terminology Inconsistencies

- A device is named CPTCone whereas the test data is named StaticConeTest
  - Refer to Dilatometer and Dilatometer Detail for consistent naming convention
- “Index” is used to indicate a Depth value in the Table object for a Static Cone Test

Slide 8

KWA1 Questions seem aggressive/blaming. If you ask, they may have an answer. Or else they just didn't realize it and do not have an answer. Maybe not. Perhaps put out a statement demonstrating the inconsistencies.
Katie Aguilar, 3/3/2009

Structural Inconsistencies

- DMT Detail vs Static Cone Test
  - DMT doesn't use table object, Static Cone does
  - This inconsistency applies throughout schema with respect to lab testing (i.e. hydrometer, proctor, sieve, etc.)
- Detectors are defined for DMT but not for CPT. (pore pressure, conductivity, resistivity, shear wave velocity, HFFD, LFFD, etc).
- Offset distances, areas of sensor, etc are defined in the Static Cone Test object yet the Detectors are not definitive in the associated table object.
  - If these values are important for one sensor why are they not important for all the sensors?

KWA2

Structural Problems

- There are two basic concepts necessary for an interchange standard to be usable:
  - A data dictionary that we can agree upon so we know how to exchange data
  - Rules for how to parse the file that is being exchanged
- Users should not be responsible for naming the fields, measurement types, etc.
  - Each time a user receives a DIGGS file from a different user may require mapping it to their structure
  - The flexibility of the language definition may require multiple mappings for the same file for different languages/codelists

Slide 10

KWA2 Start with the point - "user's shouldn't be.." then show the proof/why. Builds a stronger case.
Katie Aguilar, 3/3/2009

Structural Problems

- It doesn’t appear to be possible to determine measurement types included in a CPT/Static Cone Test
  - Each language/person can name the detectors (tip, sleeve, pore pressure…)

```
<equipments>
	<diggs_geo:CPTCone xmlns="http://schemas.diggsml.com/1.0a/geotechnical">
		<diggs:id>DIGGSINC-CPT-CONE-1</diggs:id>
		<detectors>
			<diggs_mon:Detector>
				<diggs:id>DIGGSINC-CPT-CONE-1-RES</diggs:id>
				<diggs_mon:measurand codeSpace="agsCodeList_V1.xml">ConeResistance</diggs_mon:measurand>
			</diggs_mon:Detector>
			<diggs_mon:Detector>
				<diggs:id>DIGGSINC-CPT-CONE-1-COND</diggs:id>
				<diggs_mon:measurand codeSpace="agsCodeList_V1.xml">Conductivity</diggs_mon:measurand>
			</diggs_mon:Detector>
			<diggs_mon:Detector>
				<diggs:id>DIGGSINC-CPT-CONE-1-LSFR</diggs:id>
				<diggs_mon:measurand codeSpace="agsCodeList_V1.xml">LocalSideFrictionResistance</diggs_mon:meas
			</diggs_mon:Detector>
			<diggs_mon:Detector>
				<diggs:id>DIGGSINC-CPT-CONE-1-PWP</diggs:id>
				<diggs_mon:measurand codeSpace="agsCodeList_V1.xml">PorewaterPressure</diggs_mon:measurand>
			</diggs_mon:Detector>
		</detectors>
	</diggs_geo:CPTCone>
```

CPT Table Definition

```
<StaticConeTest>
        <diggs:id>DIGGSINC-SCT23423</diggs:id>
        <!-- need to update this to use a table -->
        <diggs_geo:tabularData>
                <diggs:Table>
                        <diggs:columns xmlns="http://schemas.diggsml.com/1.0a">
                                <diggs:Column index="1">
                                        <dataType>xs:double</dataType>
                                        <meaning>Index</meaning>
                                        <uom>m</uom>
                                </diggs:Column>
                                <diggs:Column index="2">
                                        <dataType>xs:double</dataType>
                                        <meaning>Measure</meaning>
                                        <uom>MN/m2</uom>
                                        <source>
                                                <Ref xlink:href="#DIGGSINC-CPT-CONE-1-RES" />
                                        </source>
                                </diggs:Column>
                                <diggs:Column index="3">
                                        <dataType>xs:double</dataType>
                                        <meaning>Measure</meaning>
                                        <uom>uS/cm</uom>
                                        <source>
                                                <Ref xlink:href="#DIGGSINC-CPT-CONE-1-COND" />
                                        </source>
                                </diggs:Column>
                                <diggs:Column index="4">
                                        <dataType>xs:double</dataType>
                                        <meaning>Measure</meaning>
                                        <uom>kN/m2</uom>
                                        <source>
                                                <Ref xlink:href="#DIGGSINC-CPT-CONE-1-LSFR" />
                                        </source>
                                </diggs:Column>
                                <diggs:Column index="5">
                                        <dataType>xs:double</dataType>
                                        <meaning>Measure</meaning>
                                        <uom>kN/m2</uom>
                                        <source>
                                                <Ref xlink:href="#DIGGSINC-CPT-CONE-1-PWP" />
                                        </source>
                                </diggs:Column>
                        </diggs:columns>
                        <diggs:blockSeparator>;</diggs:blockSeparator>
                        <diggs:decimalSeparator>.</diggs:decimalSeparator>
                        <diggs:tokenSeparator>,</diggs:tokenSeparator>
                        <diggs:data>
```

Structural Inconsistencies

- “Importantly, it should be understood that DIGGS is a format for the transfer of results, it is not intended to facilitate the transfer of what could be termed as raw data”
- DIGGS transmits p0, p1, p2 for DMT – these are interpreted values. The analogy within the CPT would be transmitting q_t not q_c.
  - Transmit A, B and C readings for DMT
  - Transmit q_c, f_s, u_2 for CPT
  - This is not raw data. This is data which is fundamental to several correlations and evaluations – the raw data is voltage

Structural Problem?

- The insituTesting element can be placed:
  <diggs:Project>
  <diggs:locations>
  <diggs_geo:Hole>
  <diggs:insituTesting>
  and:
  <diggs:Project>
  <Location>
  <diggs:insituTesting>
- Symptom of larger issue – DIGGS allows files that are statically valid, yet functionally invalid and ambiguous

Is DIGGS ready to implement?

- No – fundamental technical problems are prevalent throughout DIGGS
- The items discussed herein are only the tip of the iceberg – additional discussion is included in the handout
- Too much flexibility
- Too much complexity
- Design for 80-90% of all scenarios and all users to expand schema

What needs to be accomplished in order to implement DIGGS?

- Structural issues need to be eliminated
- Terminology should be consistent
- Documentation needs to be correct
- Example files need to be more robust
- Once a version of DIGGS is released that mitigates these issues it can then be tested by software vendors as well as individual users

What needs to be accomplished in order to implement DIGGS?

- Management and development of DIGGS needs to be transparent
  - Not biased by software vendors and their specific desires
  - DIGGS should not be developed by the software vendors
- Forum needs to be an active place for users, software vendors, and DIGGS committee/developers to communicate

What needs to be accomplished in order to implement DIGGS?

- Get equipment manufacturers to support it – not just software vendors (CPT, DMT, instrumentation, environmental monitoring, lab testing, etc.)
- DIGGS must be able to be opened in a XML editor in a timely fashion
- Generate an empty XML file from schema
- Need to be able to generate database structure from schema

Conclusions

- DIGGS must make people's work easier – not more complicated
- If it adds complexity (and therefore cost) to their business process today it will not be adopted

Thank you.

Dataforensics Review of DIGGS Version 1.0

Scott L. Deaton, Ph.D.

It has been Dataforensics' experience that, for a user to adopt something, whether new software, a data interchange standard or even a new route to work, it must make the user's life simpler with very little inconvenience. The harder it is to learn or use the longer it will take for people to adopt. If there is not a viable economic benefit for the user, it will not be adopted. At this time

That being said, DIGGS, like many things, is a good idea. However for it to be widely adopted, it must not complicate the users work nor require they change their business process and it must be economically viable for the users, particularly in these economic conditions. In order for DIGGS to be ready to be released for general usage, the committee must make a concerted effort to strike a balance between flexibility and ease of use.

Dataforensics began reviewing DIGGS following the version 1 public release in August 2008. Initially the review focused solely on CPT and DMT and was expanded to include the geotechnical portion of the schema. During the review many questions and problems surfaced. Various posts were submitted to the DIGGS forum and to date nearly all of the questions remain unanswered. This document serves as a summary of many of the specific technical problems Dataforensics personnel encountered during the review of DIGGS. CPT and DMT are used to highlight the problems and inconsistencies that are prevalent throughout DIGGS.

Minor Problems

If you have ever opened a file that takes a long time to open, you know that most computer users have little patience, Dataforensics had a similar experience opening the DIGGS file "Example 02 – CPT Final.xml". When attempting to open the file it took more than three minutes to open. The file was opened with XML notepad, a freely available XML editor from Microsoft. This is attributed to the schema files being located on a remote web server. In order for DIGGS to be adopted, opening a DIGGS file cannot take such a long time. A simple method for downloading the schema files and re-mapping the files to a local folder structure should be implemented. Specifically, the main schema file should have an absolute path, and then all files that it references can use relative paths so remapping to local folder structure would be simple.

The last minor problem is not a functional problem; rather it is a documentation problem. A schema should be legible and well organized so that users can read and find information they need. The DIGGS schema diagrams are too large to print and when printed on a full size plotter the text is too small to read. These documents should be broken into smaller subsets such that the user can look at specific items such as In-situ testing, lab testing, boreholes, wells, etc. to allow for someone to look at a specific subset of interest at a reasonable scale.

Significant Problems

As Dataforensics personnel reviewed the documentation and example files, there were significant technical problems that were noted.

Each of the following items will be discussed herein within the framework of the CPT and DMT objects:

- The documentation is inconsistent with example files
- The schema uses inconsistent terminology
- The schema has structural inconsistencies

- The schema has structural problems
    - o User defined tabular data types
    - o Over usage of inheritance
    - o Recursion within the schema
- Validation problems

Documentation Inconsistencies

DIGGS states that “Data validation is carried out using a set of rules that is the same for all parties in the data exchange. There can be no interpretation of the rules; therefore mistakes are much reduced.” This statement seems to be violated by many of the example files provided.

During its review, Dataforensics saw several inconsistencies between documentation about DIGGS and the examples provided. Although no one expects a first release to be perfect, the example shown is important. It is recommended that documentation be reviewed by a third party prior to any release to minimize the potential for significant errors and confusion that can thwart adoption.

A table is defined in the DiggsLS.pdf file as “A Table containing delimited data. A default table is delimited by "." for decimals, by "," for columns and by " " for rows. Hence it's data block would look like this "12.345 12.345 12 12.345,67.890 67.890 67 67.890"

However, as shown in Figure 1, the table object of the CPT data from the “Example 02 - CPT Final.xml “ file, the table structure is inconsistent with the DiggsLS.pdf file. The table in Figure 1 shows that the column separator is a “,” and the row separator is a “;”.

```
0.010,0.1300,0.40,0.0000,0.0013;
0.020,0.2400,0.40,0.1.0a,0.0078;
0.030,0.5500,0.40,0.0040,0.0126;
0.040,0.6800,0.40,0.0070,-0.0017;
0.050,0.7800,0.30,0.0120,-0.0121;
0.060,0.9000,0.30,0.0150,-0.0161;
0.070,0.9600,0.40,0.0200,0.0191;
0.080,1.0400,0.40,0.0240,-0.0120;
0.090,1.0700,0.30,0.0270,-0.0129;
1.0a0,1.1000,0.30,0.1.0a,-0.0123;
1.0a0,1.1300,0.40,0.0350,-0.0176;
```

Figure 1 - Example Table Object for CPT Data (Example 02 - CPT Final.xml)

Another example of an inconsistency is also present in Figure 1. Note how the 10th and 11th rows begin with “1.0a0”. These values are in violation of the data type used in defining the table (shown in Figure 2 and 3). Each of the fields is supposed to be a double precision number.

This highlights how DIGGS is not taking advantage of the capabilities inherent with XML. It doesn’t appear to be possible to validate the data stored in the table object. It seems the complexity of the table object has been introduced for two reasons:

1. To reduce file sizes by eliminating the large number of data tags
2. To facilitate the usage of commas as the decimal separator

This complexity and flexibility has been introduced at the expense of validation within the schema and therefore requires using industry specific software to validate the file, not a standard XML editor. Conversely, if the following approach were utilized it would allow simple validation within the schema itself and users would not have to rely on the industry specific software.

For example, the same table structure shown above could be represented as follows where each data type would have a data tag and within the data tag and array of space delimited values.

<Depth> 0.01 0.02 0.03 0.04…</Depth>
<TipResistance> 0.13 0.24 0.55 0.68…</TipResistance>
<SleeveFriction> 0 0.001 0.004 0.007…</ SleeveFriction >

This can be easily validated by the schema using a simple regular expression. The space is used to parse the individual items and the schema can then check that each item within the tags is the particular data type. This approach has two drawbacks both of which are trivial.

1. This approach does not support text data (non-numeric) without using an escape character such as %20 commonly used within html. This is not a problem because these types of tabular data should not be text data types anyway.
2. This approach does not support having a decimal separator as something other than a decimal point (i.e. a comma).

Fundamentally, data should be stored in a common format (using a decimal point). UNIX systems as well as commonly used software within the geo-industry (Surfer, Rockware, ArcGIS, gINT, SQL Server, Access, etc) require data to be stored using a decimal point. Data should be stored in a common format and if users want it formatted differently, that is a reporting or presentation issue that can be left to the local settings of the user's computer.

Inconsistent Terminology

In order for an interchange standard to be intuitive, the naming convention should be self consistent, meaning each part should be logically consistent with the rest. This is something the AGS lacked as well. DIGGS should take the opportunity to learn from these mistakes and correct them. However, DIGGS is starting off with several inconsistencies that seem to be a result of its AGS lineage.

An interchange standard should be simple to understand. If an interchange standard uses naming conventions for data types that have no meaning, such simplicity is lost. This is one reason why the AGS is not intuitive – they use naming conventions that have no meaning to the average user. DIGGS is in danger of repeating this error and losing simplicity that can aid in its adoption. For example, in the table object of the Static Cone Test, the depth value is defined as "Index". Using the terminology, "index" could be an auto number, it could be text, it could be a date or time, it could be a GUID. Without searching and inferring the meaning, the user does not know what index is, and has to spend unnecessary time figuring it out. Conversely, naming it "Depth" provides users with instant knowledge of what the data type is. Another problem with "Index" is that, apparently, it cannot be validated as "1.0a0" was not flagged as a problem by the schema.

Additionally, having a field named Index used throughout the schema representing different data types (time, depth, etc) tends to make the schema less intuitive and more cumbersome to use. Having appropriately named data types is critical.

Consistent terminology is essential for ease of use by the user and to help with adaptation. For example, the device is named CPTCone whereas the test data is named StaticConeTest. Unless someone knows the AGS there is no intuitive relation. Dataforensics personnel had to search to find the table that held the CPT data. Having consistent naming convention is beneficial so when reviewing documentation and examples it is obvious that StaticCone and StaticConeTest or CPTCone and CPTConeTest are related. Additionally, consistency can save users time and effort, thus making adoption easier. As an example, the DIGGS creators were consistent with Dilatometer and Dilatometer Detail.

```
<StaticConeTest>
        <diggs:id>DIGGSINC-SCT23423</diggs:id>
        <!-- need to update this to use a table -->
        <diggs_geo:tabularData>
                <diggs:Table>
                        <diggs:columns xmlns="http://schemas.diggsml.com/1.0a">
                                <diggs:Column index="1">
                                        <dataType>xs:double</dataType>
                                        <meaning>Index</meaning>
                                        <uom>m</uom>
                                </diggs:Column>
                                <diggs:Column index="2">
                                        <dataType>xs:double</dataType>
                                        <meaning>Measure</meaning>
                                        <uom>MN/m2</uom>
                                        <source>
                                                <Ref xlink:href="#DIGGSINC-CPT-CONE-1-RES" />
                                        </source>
                                </diggs:Column>
                                <diggs:Column index="3">
                                        <dataType>xs:double</dataType>
                                        <meaning>Measure</meaning>
                                        <uom>uS/cm</uom>
                                        <source>
                                                <Ref xlink:href="#DIGGSINC-CPT-CONE-1-COND" />
                                        </source>
                                </diggs:Column>
                                <diggs:Column index="4">
                                        <dataType>xs:double</dataType>
                                        <meaning>Measure</meaning>
                                        <uom>kN/m2</uom>
                                        <source>
                                                <Ref xlink:href="#DIGGSINC-CPT-CONE-1-LSFR" />
                                        </source>
                                </diggs:Column>
                                <diggs:Column index="5">
                                        <dataType>xs:double</dataType>
                                        <meaning>Measure</meaning>
                                        <uom>kN/m2</uom>
                                        <source>
                                                <Ref xlink:href="#DIGGSINC-CPT-CONE-1-PWP" />
                                        </source>
                                </diggs:Column>
                        </diggs:columns>
                        <diggs:blockSeparator>;</diggs:blockSeparator>
                        <diggs:decimalSeparator>.</diggs:decimalSeparator>
                        <diggs:tokenSeparator>,</diggs:tokenSeparator>
                        <diggs:data>
```

Figure 2 - CPT Table Object Definition – (Example 02 - CPT Final.xml)

```
<equipments>
        <diggs_geo:CPTCone xmlns="http://schemas.diggsml.com/1.0a/geotechnical">
                <diggs:id>DIGGSINC-CPT-CONE-1</diggs:id>
                <detectors>
                        <diggs_mon:Detector>
                                <diggs:id>DIGGSINC-CPT-CONE-1-RES</diggs:id>
                                <diggs_mon:measurand codeSpace="agsCodeList_V1.xml">ConeResistance</diggs_mon:measurand>
                        </diggs_mon:Detector>
                        <diggs_mon:Detector>
                                <diggs:id>DIGGSINC-CPT-CONE-1-COND</diggs:id>
                                <diggs_mon:measurand codeSpace="agsCodeList_V1.xml">Conductivity</diggs_mon:measurand>
                        </diggs_mon:Detector>
                        <diggs_mon:Detector>
                                <diggs:id>DIGGSINC-CPT-CONE-1-LSFR</diggs:id>
                                <diggs_mon:measurand codeSpace="agsCodeList_V1.xml">LocalSideFrictionResistance</diggs_mon:meas
                        </diggs_mon:Detector>
                        <diggs_mon:Detector>
                                <diggs:id>DIGGSINC-CPT-CONE-1-PWP</diggs:id>
                                <diggs_mon:measurand codeSpace="agsCodeList_V1.xml">PorewaterPressure</diggs_mon:measurand>
                        </diggs_mon:Detector>
                </detectors>
        </diggs_geo:CPTCone>
```

Figure 3 - Static Cone Test – Detector Definition

Structural Inconsistencies

Table Objects

The usage of Table objects within DIGGS is the source of the single largest inconsistency. Presumably table objects are used in places where encoding data using tags for each item would result in enormous data files. The example provided for CPT data uses the table object. This is a sensible approach. However, there are several types of geotechnical data which also lend themselves to table objects to decrease the file size and make the interchange standard self consistent. These include "detail" data types such as sieve analysis, hydrometer, Proctor, DMT data, etc. Using similar structures for similar types of data makes sense to the user (chances are they are used to seeing the data in a table format), makes the interchange standard self consistent and would result in smaller files.

Moreover, detectors are defined explicitly for Dilatometer (DMT) but not for the Static Cone Test (CPT). Any auxiliary sensor that can be put on a CPT can be placed on a DMT (i.e. pore pressure, conductivity, resistivity, shear wave velocity, HFFD, LFFD, etc.). It is inconsistent that the structures for CPT and DMT are different with regards to detectors. They should be identical.

Another inconsistency noted is items in the Static Cone Test object. Offset distance and area of sensor, friction reducer, friction sleeve area, net area ratio correction, piezocone type, pore capacity, porous element type are defined in the Static Cone Test object. However, the sensors are not defined in the associated table object. This can lead to confusion by the user.

One last inconsistency is that the StaticConeTest and DMT Detail both have a top and bottom depth. Both CPT and DMT soundings may have a starting depth of the test that may differ from ground surface, but the ending depth is duplicate information that should be defined by the final reading within the tabular data portion of the sounding. Having both bottom depth and the tabular data defining the depth extent of the test can result in transmitted data that is inconsistent.

Object Inconsistencies

The Static Cone Test has the following elements: Distance tip to sleeve, friction reducer, friction sleeve area, net area ratio correction, piezocone type, pore capacity, porous element type, pushrodtype, saturation fluid, saturation method, sleeve capacity, surface capacity, tip apex angle, tip area, tip capacity. These items are hardcoded within the schema. However, the types of the measurands are not defined for the sensor itself. Either these items should not be defined in the Static Cone Test or the measurands should be defined in order to be self consistent.

DIGGS asserts that it is "a format for the transfer of results, it is not intended to facilitate the transfer of what could be termed as raw data". However, there are some object items which contradict this statement. From a CPT testing standpoint, it is typically unnecessary to transmit the distance tip to sleeve because this data is supposed to be processed while converting from voltage to tip/sleeve stress by the original processing software. This is true for all sensors. Therefore the various measurements within the tabular data are always at the same depth for a particular row of data in the table object. This metadata is typically only necessary within the original software that processes the raw voltage data and is unnecessary if you are only transmitting the results of the test.

Although it is unnecessary to transmit the distance from tip to sleeve, if it is decided that the distance from tip to sleeve is important then there should be a distance from the tip to every sensor on the cone: pore pressure, distance to conductivity sensor, distance to soil moisture sensor, as well as all of their associated attributes. As noted above, DIGGS has opted to include distances only from tip to sleeve.

Another inconsistency is with definitions of objects. For example, nowhere within the documentation is the “Cone Resistance” actually defined. Is the tip resistance q_c (uncorrected tip resistance) or q_t (corrected tip resistance)? Along the idea of DIGGS intent to transfer results and not raw data, it should be q_c (uncorrected tip resistance) because q_t is a calculated/derived parameter based on the net area ratio of the cone penetrometer. However, that is a dangerous assumption that all DIGGS users will know to enter q_c, and that no guidance is necessary.

A similar inconsistency with transferring results occurs within the DMT module. For the DMT, the A, B and C readings for the dilatometer should be transmitted because those are the test results, not the interpreted values of p0, p1, and p2. p0, p1 and p2 are calculated/corrected/interpreted values based on the A, B and C readings, and should be at the discretion of the end-user not the data producer.

Furthermore, some data items necessary for analysis for a DMT are missing: the Delta A and Delta B measurements before and after the sounding as well as the zero offset is necessary. Having this data allows the end user to determine the validity of the test. If there are large differences between Delta A or Delta B before and after the test, it allows the user to disregard the test results because of problems that occur during testing or make another correction. Without providing this type of information there is no way to determine validity of test.

Moreover, Dilatometer should not transmit a coefficient of consolidation value. This is not a result of the test, but a calculation/correlation from test results. This begs the question, at what point does DIGGS start or stop including correlations. There are 10-15 parameters that many people calculate from DMT data and more than 40 different CPT based correlations. Either all the correlations should be included or no correlations should be included. By the stated mission of DIGGS, it seems that since only results are supposed to be included, then any correlations would not be transmitted.

Structural Problems

User Defined Tabular Data

There should be two fundamental concepts present for an interchange standard to be usable:

1. A data dictionary that we can agree upon so we know how to exchange data
2. Rules for how do parse the file that is being exchanged.

These two fundamental concepts are not being utilized.

The flexibility being provided in DIGGS results in the ability of the user to define to names of many of the parameters included in the file. However, users should not be responsible for defining something as simple as a field name. It should be defined within the schema. If the users have the choice for defining the naming conventions then it is possible that each time a user receives a DIGGS file, they will have to create a mapping to import the data. Fundamentally, this is no better than the approach people are using today throughout the U.S. and much of the world as mappings are already necessary between databases. So, fundamentally all DIGGS has done today is make the process more complicated for users by complicating the mapping process compared with what is required today. Creating such mappings for each different recipient/source of DIGGS data is a large and unnecessary burden to a user.

Accordingly, field names and data types for commonly accepted data types should be in the schema instead of having to look these values up through the complex referencing system put in place within the file that also references Codelists which don't really tell us what the data is. The user should not have to map standard data items within an interchange standard. There should be "fields" or objects for all standard types of measurements that are commonly used. If it is something unusual, the user can extend the schema which then would require the user to specify their data during data exchange. The important point here is that everyone should call a particular measurement type the same thing. If common field names are not defined in the schema then the user must define a new import mapping for every DIGGS file. The user should be able to create an import mapping for the base schema once and only have to deal with custom mappings with rare, user-defined, schema extensions.

In taking from above where cone resistance was undefined, it would then be defined somewhere as "Uncorrected Tip Resistance" so that the field has a specific meaning within geotechnical engineering.

Within a CPT sounding, a large number of possible measurements are possible (tip, sleeve, pore pressure, conductivity, resistivity, HFFD, LFFD, inclination, etc.). Currently in DIGGS, anyone can name them whatever they like. Based on our knowledge of the CPT, we can visually rationalize what they are based on their detector names, names and their measurand. However, it is well known that a single test may have several different names. There must be a standard key accessible to all users.

Rules for how to parse the data are necessary to help the user and, again, prevent mappings for every data exchange. From looking at figures 2 and 3 above or even the schema document, it is unclear how a user or software would know what the data in the table object is without some kind of key (set of rules).

Inheritance

Inheritance is a useful concept, however it seems that inheritance has been used to an extreme within DIGGS. Nearly every object has the following elements – lang, equipment, roles, specification references, status. Having language defined at each object can, theoretically, result in multiple languages within a project. It seems that language should be defined in a high level entity that is not inherited everywhere. If the purpose of DIGGS is to allow people to transmit data in multiple languages within the same file, then this is accounting for no projects within the US and probably very few types of projects worldwide. It seems to be unnecessary flexibility/complexity.

Furthermore, as discussed previously, the user can name the detectors whatever they want. Assuming that they can be named differently in different languages, then a user can name detectors in their native tongue. Software handling the interchange cannot reconcile that "tip resistance" defined for one sounding is "Resistencia del Punto" in another sounding. The processing software needs some way of identifying this. This would require multiple data mappings within a single file in order to identify that both of these terms are actually uncorrected tip resistance. A simpler idea is to define tip resistance within the data dictionary.

DIGGS Validation Problems

DIGGS allows users to create files that are statically correct, but functionally invalid due to recursion or other underlying design flaws. For example, there doesn't seem to be any reference from CPT to a hole in the Schema document. This seems counterintuitive since it is fundamentally impossible to have a cone penetration test that is not associated with a hole. gINT Software identified and Dataforensics verified that it is possible to move the insitu testing object to various places within a DIGGS file and the files validate properly in XML editors. For example, the insituTesting element can be placed as a child of a hole or as a child or Location not associated with a hole as shown below or as a child of project not associated with a Location.

<diggs:Project>
<diggs:locations>
<diggs_geo:Hole>
<diggs:insituTesting>

and:
<diggs:Project>
<Location>
<diggs:insituTesting>

This identifies an underlying fundamental flaw that needs to be investigated further

Conclusions

Fundamental technical problems are prevalent throughout the DIGGS schema. Dataforensics used the CPT and DMT objects to illustrate terminology inconsistencies, structural inconsistencies and structural problems with the DIGGS schema. In order for DIGGS to be adopted, structural issues must be eliminated, terminology should be consistent, documentation should be correct, examples files need to be more robust and depict real world scenarios. Once these issues have been addressed then software vendors and individual users can test DIGGS to ensure the data interchange standard is flexible yet robust.

Furthermore, the management and development of DIGGS needs to be transparent. It should be guided by the software vendors and the technical committee however it should not be developed by the software vendors themselves. This would help eliminate bias imposed by individual vendors and would not require vendors to share trade secrets. The forum should be an active place for users, software vendors and the DIGGS committee/developers to communicate, whereas today it appears to be a mostly unmonitored archive of questions.

Lastly, DIGGS must be able to be opened by an XML editor in a timely fashion. Large files (10MB or more) should be tested to verify that they can be opened in various off the shelf editors. Finally, the off the shelf XML editors should be able to generate an empty XML file from the schema and generate a database structure from the schema. These capabilities are necessary to providing the user the ability to map to and from their database structures to and from DIGGS.

gINT Software and DIGGS

Salvatore Caronna

25 March 2009

DIGGS Meeting, Orlando

gINT software

Overview

- Progress to date
- Our approach
- Why our approach doesn't work
- Some other issues
- Reality check
- A way forward

Progress to date

- Autumn of 2006, we submitted an AGS to DIGGS conversion utility to the committee.
- In 2007 significant changes were made to the schema and we started work on the upgrade to the utility.

gINT® software

Progress to date

- Further schema changes were made and it appeared that significant changes were yet to be made and we halted work.
- In mid-2008 a release candidate schema was made available for review. At that time we began a concentrated effort to implement DIGGS support in gINT.

Progress to date

- Our main design specification was that we would create robust support for XML files, not for DIGGS, and that there was not to be knowledge of the DIGGS format in the code.
- After hundreds of hours we have succeeded in putting in place XML support but, to date, we have not been able to make DIGGS work.

gINT® software

Our approach

- “Flatten” the schema to appear to be a database, that is, a series of fields stored in tables.
- The correspondence file facility then allows the creation of mapping files.

Our approach

- XML Import
  - The source XML file is flattened with the data. This is written to a temporary database.
  - From this point, all the code is in place in the program to complete the import.

gINT® software

Our approach

- XML Export
  - An empty XML is created from the XSD.
  - The data in the gINT database is written to a temporary database based on the instructions in the mapping file which generates the appropriate table and field names.
  - The mapped data is inserted into the empty XML.
  - Structures with no data are removed.

Our approach

- The above approach only requires two capabilities: The ability to flatten the schema and the ability to generate an empty XML file from the schema.
- This process works with a number of XML formats, but not DIGGS.

gINT® software

Why our approach doesn't work

- The schema cannot be flattened.
- An empty XML file cannot be generated from the schema.
- Using Altova's XML Spy program:
  - Creating a database from the schema failed (equivalent to flattening).
  - Generating an empty XML file from the schema produced a corrupt file with many objects missing (including the Hole object).

Some other issues

- Unique Identifiers
  - Carrying unique record identifiers in source and target databases solves the problem of the problem of renaming of key fields during staged data transfers.
  - Adds to the complexity and requires that users of DIGGS change their database structures.

Some other issues

- Unique Identifiers
  - Unique identifiers in some objects are necessary to fulfill the requirement of identifying relationships.
  - These identifiers can be arbitrary and only need to be unique within each project.
  - The receiving software can set up their own structures and can discard the original identifiers.

Some other issues

- Key Fields
  - Key fields, except for unique identifiers, have been removed from DIGGS.
  - It is possible to have two holes with the same borehole name, samples of the same hole, depth, extent, type, and number, etc.
  - Many software publishers will need to drastically rework their programs.
  - Users work with key fields constantly. These are physical attributes that they understand, not unique identifiers.

Some other issues

- Object-Oriented Design
  - Good approach to programming.
  - It may be at the root of many of the problems in the schema.
  - At the very least it makes the schema complex and makes understanding the schema difficult.

gINT® software

Some other issues

- Object-Oriented Design
  - The current inheritance methodology is too coarse.
  - Too many objects are inheriting too many inappropriate items.
  - If the object-oriented structure remains, a finer grained approach is necessary.

gINT® software

Some other issues

- GML
  - Two advantages that are theoretically possible with the DIGGS schema:
    - Any GML compliant program could display DIGGS GML objects.
    - Web-based tools could convert coordinate systems.
  - Granted that these appear to be very nice side benefits but they add layers of complexity.

Some other issues

- GML
  - Is it a significant advantage to be able to see boreholes, samples, business associates (yes; they have GML tags as well), etc. in ArcGIS?
  - Perhaps, but is it worth the added complexity?
  - At the very least the scope of the GML tags need to be reduced dramatically.

Some other issues

- GML
  - Coordinate conversion seems like a good feature.
  - How often will this feature be necessary?
  - There are many general purpose programming tools that do the job today and they are making their way into database programs.
  - Is it worth the added complexity?

Some other issues

- GML
  - I don’t believe these facilities have been tested successfully.
  - If they remain, it has to proven that they do indeed work.

gINT® software

Some other issues

- Units
  - Another theoretical advantage is the ability for Web-based tools to convert unit systems.
  - This requires units be associated with each item of data, instead of with each field.

Some other issues

- Units
  - Except for contaminant data (which needs a separate units field), we have never seen the need for units of a field to vary with each data item.
  - Nor do I know of any program that supports this.

Some other issues

- Units
  - If this is to be left in the schema, a rule should be implemented that requires that units cannot change for a field in a project.
  - Alternatively, units can be associated with fields with a few exceptions (like contaminant data).

Some other issues

- Units
  - My understanding is then we lose the ability to use automated units conversion utilities.
  - How necessary is units conversion, does it actually work, and is it worth the added complexity?

Some other issues

- Invalid Validation
  - One of the main selling points to me for XML was the self-validating nature of the format.
  - Both the AGS4 subcommittee and we have found numerous instances of invalid files passing validation.

Some other issues

- Language
  - Language is an attribute of data records in many DIGGS objects.
  - This is just one case of more information than is called for throughout the schema.

Reality check

- Moving an industry in a new direction requires:
  - The new direction be significantly better than the current situation
  - Strong support by governing bodies
  - Strict specifications
  - The switch must be simple

Reality check

- An example in the US of a new direction that failed was the change to metric units in road projects in the ‘90s.
- Although our scope is much smaller, we are proposing something much more complex.

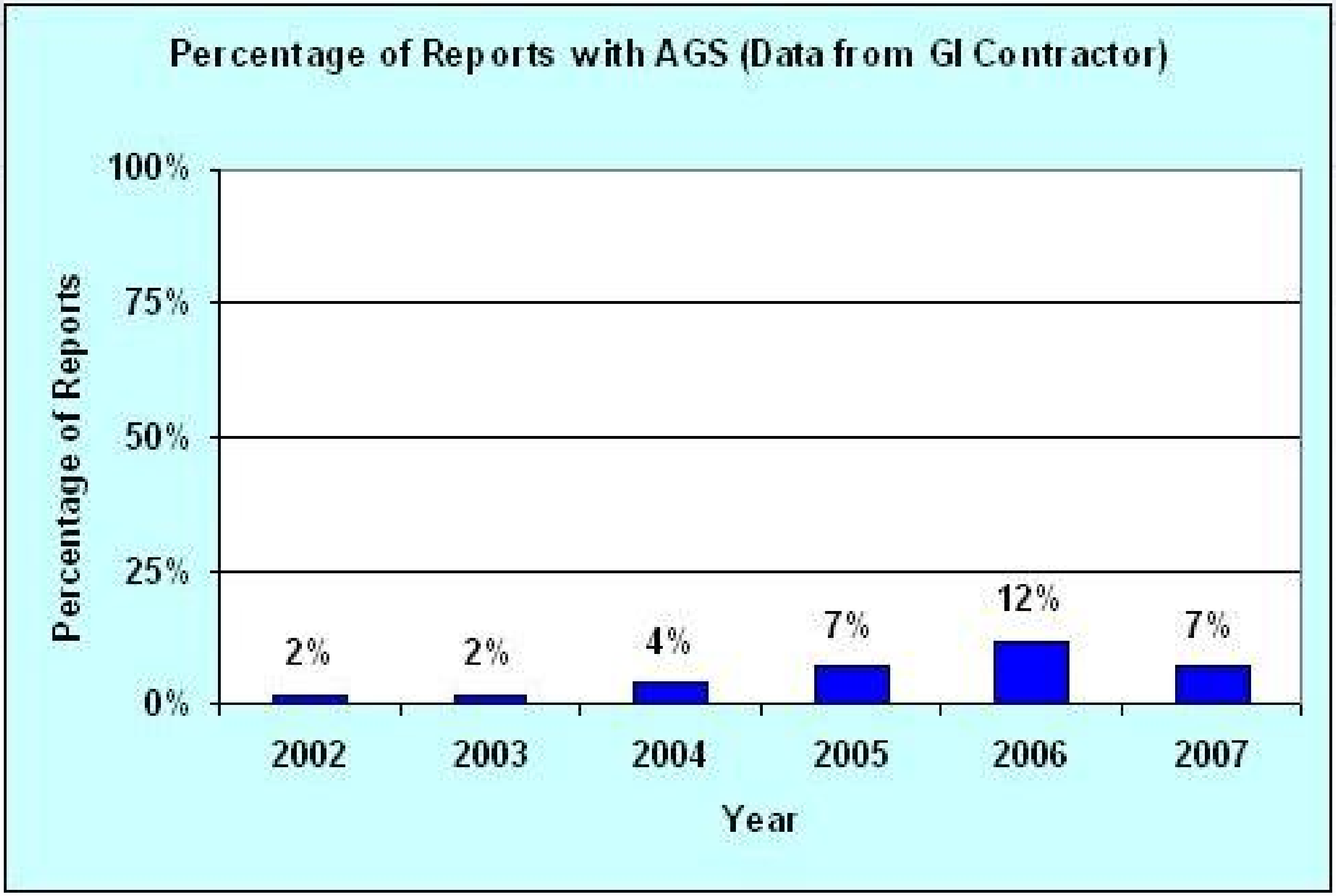
Reality check

- It required years before the AGS format gained traction in the UK.
- Making the jump from AGS 3.0 to AGS 3.1, which just added more groups and variables, required years as well.

Reality check

- 17 years after the initial introduction of the AGS format, a small percentage of geotechnical projects use the format in the UK.
- The following chart is from a medium size Ground Investigation contractor (from David Patterson, AGS Meeting presentation; June 2008).

Reality check



Reality check

- We are looking at a monumental change to jump to XML.
- For the US market, and other markets that are not used to working with interchange standards, this is a bigger change than in the UK.

Reality check

- I believe this is a necessary change but it must be as simple as possible.
- Let’s solve 80% of the issues, not 99%.

gINT® software

A way forward

With any project, the fundamental steps are:

1. Make it work
2. Make it right
3. Make it complete, efficient, and elegant

I believe our fundamental problem was that we tried jumping directly to step 3.

gINT® software

A way forward

- Data Dictionary
  - Ensure that the new AGS4 data dictionary can be mapped to DIGGS.
  - Put the DIGGS data dictionary in an Excel, CSV, and/or ACCESS format for review.
- Schema
  - Make decisions on the high level issues such as GML, unique identifiers, key fields, etc.
  - Rework the schema so that it can be flattened and a proper, empty XML file can be generated.

gINT® software

gINT Software and DIGGS - Salvatore Caronna

Introduction

gINT Software has been involved with the DIGGS standard since its inception. We have worked on the data dictionary side of the working committee and have gone through a number of iterations attempting to implement support for DIGGS. This paper presents our experiences and recommendations.

Progress to Date

In the autumn of 2006, we submitted an AGS to DIGGS conversion utility to the committee based on the format at that time. In 2007 significant changes were made to the schema and we started work on the upgrade to the utility. Further changes were made and it appeared that additional significant changes were to be made and we halted work.

In mid-2008 a release candidate schema was made available for review. At that time we began a concentrated effort to implement DIGGS support in gINT. Our main design specification was that we would create robust support for XML files, not for DIGGS, and that there was not to be knowledge of the DIGGS format in the code. After hundreds of hours we have succeeded in putting in place XML support but, to date, we have not been able to make DIGGS work.

Our Approach

gINT does not have a fixed database. With a few constraints, our clients can create whatever structure they require. This is typical of many programs in this niche (LogPlot, LogDraft, WinLog, etc.). Therefore, we need tools for our clients to easily map to and from DIGGS. These tools have been in operation in gINT since the early '90s and currently gINT supports about a dozen file formats, including some XML formats.

Our approach to XML support is the following:

- "Flatten" the schema to appear to be a database, that is, a series of fields stored in tables. The relationships between these objects are irrelevant at this stage.
- With a flattened structure, the gINT correspondence file facility will then allow the creation of mapping files, one for import and one for export (they can be very different).
- On import, the source XML file is flattened with the data. This is written to a temporary database. From this point, all the code is in place in the program. It has had support for importing from a database through a mapping file for years.
- On export, an empty XML is created from the XSD. The data in the gINT database is written to an intermediate, temporary database based on the instructions in the mapping file which generates the appropriate table and field names (this has again been in place for many years). The data in the temporary database is then inserted into the empty XML file with additional records being spawned as necessary. Finally, structures with no data are removed.

This process works with a number of XML formats, but not DIGGS.

Why Our Approach Doesn't Work

The above approach only requires two capabilities: The ability to flatten the schema and the ability to generate an empty XML file from the schema. These two basic characteristics of the schema are crucial for successful implementation. We could not perform these two tasks with DIGGS.

We work with the XML tools built into Visual Studio.NET. After weeks of unsuccessful coding we tried using the tools in Altova XML Spy to generate a database from the schema (equivalent to flattening) and generating an empty XML file from the schema. Altova also failed these tasks. Speaking with the USGS, CalTrans, and the University of Florida, they also gave up using general purpose XML tools and resorted to hard-coding. This is an unacceptable approach for a public interchange standard.

Just having the ability to flatten the schema would allow our clients to perform the mapping to and from their database structures which would generate a proper review of the schema.

With the ability to flatten and generate an empty XML file, an added advantage is that working with DIGGS files in Excel becomes much simpler since the flattened structure is effectively a grid configuration.

The case can be made that we can hard-code support for DIGGS and not try to treat it generically. There are four problems with this approach:

- Common, well known, and simple to use programming tools cannot be used.
- Developers would be faced with writing multiple import/export filters instead of one for XML.
- Any change in the schema would require significant recoding. This would slow down significantly rolling out any revisions to the schema.
- It has been tried by us, USGS, CalTrans, and the University of Florida. After many programmer months of effort none of this group has completed the implementation.

Some Other Issues with the Format

Following are a few high-level issues with the format. Others have identified more detailed issues.

Unique Identifiers

The requirement of carrying unique record identifiers in the generator's and consumer's databases solves the problem of renaming of key fields during staged data transfers but adds to the complexity and requires that users of DIGGS change their database structures.

With the non-hierarchal nature of some of the structures in DIGGS, unique identifiers in some objects are necessary. However, to fulfill the requirement of identifying relationships these identifiers can be arbitrary and only need to be unique within each project. The receiving software can then set up their own structures and can discard the original identifiers.

Key Fields

Key fields, except for unique identifiers, have been removed from DIGGS. It is therefore possible to have two holes with the same borehole name, samples of the same hole, depth, extent, type, and number, etc. Key fields must be reinstated with proper validation or many software publishers will need to drastically rework their programs. Further, end users work with key fields constantly. These are physical attributes that they understand, not unique identifiers.

Object-Oriented Design

The object-oriented approach to programming is a good one. However, it may be at the root of many of the problems in the schema. At the very least it makes the schema complex and makes understanding the schema difficult.

One problem with the current inheritance methodology is it is too coarse. Too many objects are inheriting too many inappropriate items. If the object-oriented structure remains, a finer grained approach is necessary whereby more base classes of varying structures are used appropriately.

GML

Two advantages that are theoretically possible with the DIGGS schema are:

- The ability for any GML compliant program to display DIGGS GML objects.
- The ability for Web-based tools to convert coordinate systems.

Granted that these appear to be very nice side benefits but they add layers of complexity.

DIGGS is an interchange format that will be generally used with dedicated systems that understand the data. Is it a significant advantage to be able to see boreholes, samples, business associates (yes; they have GML tags as well), etc. in ArcGIS? Perhaps but is it worth the added complexity? At the very least the scope of the GML tags need to be reduced dramatically.

The coordinate conversion seems like a good feature. However, how often will this feature be necessary? There are many general purpose programming tools that do the job today and they are making their way into database programs. Again, is it worth the added complexity?

I don't believe these facilities have been tested successfully. If they remain, it has to proven that they do indeed work.

Units

Another theoretical advantage is the ability for Web-based tools to convert unit systems. This requires units be associated with each item of data, instead of with each field.

Except for contaminant data (which needs a separate units field), we have never seen the need for units of a field to vary with each data item. Nor do I know of any program that supports this. If this is to be left in the schema, a rule should be implemented that requires that units cannot change for a field in a project. Alternatively, units can be associated with fields with a few exceptions (like contaminant data). My understanding is then we lose the ability to use automated units conversion utilities. Again, how necessary is this facility, does it actually work, and is it worth the added complexity?

Invalid Validation

One of the main selling points to me for XML was the self-validating nature of the format. Both the AGS4 subcommittee and we have found numerous instances of invalid files passing validation.

Language

Shouldn't language be associated with the project, not with individual records? This is just one case of more information than is called for throughout the schema.

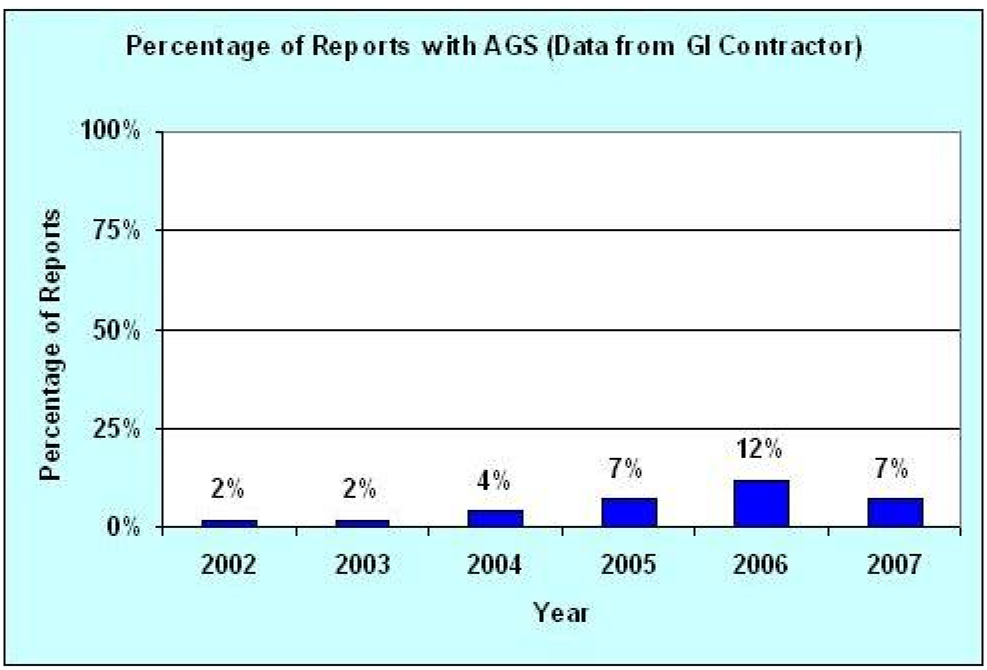
Reality Check

It is hard work to turn an industry in a new direction. It requires that the new direction be significantly better than the current situation, strong support by governing bodies, strict specifications, and that the switch must be simple.

An example in the US of a new direction that failed was the change to metric units in road projects in the '90s. Although our scope is much smaller, we are proposing something much more complex.

It required years before the AGS format gained traction in the UK. This is a very simple format based on comma separated values files with a much smaller data dictionary than DIGGS and a simple hierarchal relational structure. Granted that those were the early days and data interchange was a brave new world. However, making the jump from AGS 3.0 to AGS 3.1, which just added more groups and variables, required years as well.

To this day, 17 years after the initial introduction of the AGS format, a small percentage of geotechnical projects use the format in the UK (from David Patterson, AGS Meeting presentation; June 2008):



There still remains a large effort of education in the UK for AGS adoption. Add orders of magnitude more complexity to the interchange format and the job of increasing the format usage becomes that much more difficult.

Now we are looking at a monumental change to jump to XML. For the US market, and other markets that are not used to working with interchange standards, this is a bigger change than in the UK.

I believe this is a necessary change but it must be as simple as possible. This will require going back to fundamentals and solving 80% of the issues, not 99%.

A Way Forward

With any project, the fundamental steps are:

1. Make it work
2. Make it right
3. Make it complete, efficient, and elegant

I believe our fundamental problem was that we tried jumping directly to step 3.

The advantage to our current situation is that there is a large body of work that provides an excellent foundation to move forward. We know a lot more now because of the process we have gone through than when we started.

Possible steps:

- Data Dictionary
    - Ensure that the new AGS4 data dictionary can be mapped to DIGGS.
    - Put the DIGGS data dictionary in an Excel, CSV, and/or ACCESS format so that users can map their own database structures to and from DIGGS. This will allow an excellent means for users to better understand DIGGS and will uncover problems early.
- Schema
    - Make decisions on the high level issues such as GML, unique identifiers, key fields, etc.
    - Rework the schema so that it can be flattened and a proper, empty XML file can be generated. Both these process must be able to be easily handled by generally available, inexpensive XML tools. The Visual Studio.NET programming tools, in particular must be supported.

AGS Data Format

Development and Governance
1991 to 2009

Stephen Walthall

Chair AGS Data Management Committee

Who are the AGS?

AGS is Limited Company with a board of directors.
It is a trade organisation for the UK Geotechnical and geo environmental industry

The Data Management Committee (formally known as the Data Format committee) is one of the constituted sub-committees of the AGS.

The Data Management Committee has responsibility for the Data Transfer format

Key dates for the Data Format Committee

1991 -convened in after a conference to discuss electronic data transfer
1992 AGS v1
1994 AGS v2
1999 AGS v3
2002 AGS-M
2004 AGS v3.1
2004 Launch of the web site
2008 renamed **Data Management Committee**

Committee format has been :-

2 Representatives from each of

GI contractors

Consulting engineers

Client organisation

Software houses

Laboratory testing facilities

(approx 10 members plus co-opted people as required)

The committee is on voluntary basis,
meets about 4 times a year with small groups carrying out specific tasks as required

Basic Transfer rules

Only transfer data required by standards
Engineering units only (not millivolts)
No data which can derived from other data

Documents published:-

Version 1 (1992) written in 9 months Very little data was formally submitted in Version 1 although time was spent developing software.

Version 2 (1994) 2 years later after a number of errors and omissions were realised.

Version 3 (1999) included user defined fields and paired tables

AGS-M (2002) for monitoring data was prepared under contract to CIRIA and is fully V3 compliant

AGS3.1 (2004) combined V3 with AGS-M into one document

Versions 1 and 2 were sold but v3 onwards are available freely on the internet

Publicity

Each version has been launched by a seminar, held at Birmingham Motorcycle Museum

Numbers have been typically 80 but reached 100 in 2008.

User group meting following similar format have been held about every 2 years.

Technical papers and articles appear frequently written by members of the committee and others.

The logo below was made available for inclusion on printed logs to indicate that an electronic version was available

Support

Web Pages on the AGS web site www.ags.org.uk

Discussion Board

Seminars / user group meetings

Commercial training courses

Usage 1

Most UK contractors can produce the data

Most UK major clients ask for the data in their contracts

Most major consultants ask for and receive the data.

Users are requested to register with the AGS after downloading the Document   Presently there are about 108 registered users listed

A number of high profile projects have used
AGS data transfer to populate their data bases and been very successful

Used successfully for monitoring data

Usage 2

Very few major contracts issue the data at construction stage, even though piling contractors, in particular, would like to see it.

About half the contactors use propriety software to produce the data, others use in house software, including Excel scripts, to produce the transfer file.

A significant part of the UK geotechnical work load is for small projects and electronic data is not usually used.

Usage 3

It is used extensively in Hong Kong where the GCO implemented its use in the mid 1990’s

It is being used in other countries including Australia, Singapore, New Zealand by major clients

and on many specific contracts where UK based consultants are involved throughout the world.

AGS

Reported problems and perceptions

For many years the quality of the data was of concern as it often contained errors of two distinct sorts; format errors and content

Format errors have become less common due to improved software

Content has become subject to interpretation,
often of the form *"where do I put XXXX"* and the problems created by 'orphans'.

Conflict of fixed format with freedom

Availability of 'cheap' software to use the data (*Why cannot I load it into Excel?)*

Confusion over electronic copies, pdf, Autocad logs, database/transfer format

Drivers for implementation

Interested users

Specified by clients/engineers

Need for data in a useable format

Cost saving (not realised)

Still considered as a ‘bolt on’ not part of the investigation process

Future

AGS 4 (?)

Working party with other AGS sub committees

- Business Practices
- Laboratory
- Contaminated Land
- Data management

DIGGS Strategic Planning & Roadmap

March 25, 2009

Suggested Rules…

- Cell phone/Laptops need to be off… *REALLY off.*
- No egos.
- Be passionate about getting the best solution… for DIGGS.

Orientation

- Be introspective
  - Why are you here?
  - Why is this important to you?
  - What do you absolutely need?
  - What do you want?
- If it's about ***winning*** and ***losing***, then we've already lost

Steps for effective consensus building…

- LISTEN
  - *Hear* what others have to say.
- Identify how this fits with your understanding?
- Identify if there are any problems with this approach/idea.
- Identify possible improvements if any
- Present your understanding and proposal
  - It should include items two and three
- Use “I” Statements

We only win if we end up with a solution.

- If you are not willing to help reach understanding and consensus, but are here to defend territory, then you have already doomed this meeting.

First Two “Straw Men”

- Close down the project…
  - Shut everything down
  - Refund money that is left to donor organizations.
- Keep Status Quo…
  - Do Nothing…
  - No consensus… project basically dies.

Is this what we want?

- Both throw away a lot of work and leave a vacuum that still needs to be filled…

Basic Business

- Need a recorder…
- Break at …3:15?

- **Deliverable:** Standard - Consensus
- Hire GML/XML expert (How many?)
  - Task 1 – solve quick issues (date below w/ Data dev group)
  - Task 2 – do through review
- Work with Data development group (core development team – SIG)
- Prioritize key issues to solve V1.0
- Release date chosen by June, Release V1.0 when we can demonstrate that following issues have been fixed:
  - Flatten files (recursion, inheritance, remarks, etc) – Be able to map data
  - Produce blank XML file (using free and purchased tools)
  - List working XML mapping tools that allows DIGGS -> XSLT app (mapforce, stylus Studio, etc)
  - Fix key simple things to get release
    - Consistent terminology
    - Standard code list
    - Etc
  - Attached (scanned) document
  - MUST: pass a data file among 4 software vendors and read/export – in order to release

- Task 3 – OK!
  - Contractor vs Volunteer
    - Task & time factor based
      - Review and Advise – volunteer
      - Shorter timeframe – needs contractor
      - Tangible product (code, software, documentation) – contractor
      - In-kind efforts (funded by others) – leveraging of development
      - Critical milestone, on critical path – then contract
      - If the group does not have the expertise or timeframe is critical - contract

- Scope: V1.0 (4~6 months)
  - Geotech (Borehole, Lab, Insitu)
    - Documentation, schema change tracking, data dictionary, Schema
  - Deep foundations, Geophysics (tools by individuals/groups)
  - GeoEnv (schema only)
- Scope V1.x or 2.0 (24 months)
  - Improved schemas
  - Tools and pilots (Borehole, Insitu, Lab) – funded (need minimum set) (Style sheets,
  - Nice: Documentation and Pilots for Deep Found, Geophysics, GeoEnv
  - Agreed!!

- Task 2:
  - Develop 2 year project plan
  - Fund project manager to run project
    - Issues task orders, contracts, etc

- Review Tools
- Documentation

- Pilot Apps

- Community Development Env

Governance

Day 2

DIGGS Governance

- DIGGS so far is a pooled fund project
  - Develop the Schema
- Need a governance structure to:
  - Market and encourage adoption
  - Maintain schema (updates, new versions)
  - Support schema and users (answer questions, provide fixes)
  - Consider new object types
- Need a funding mechanism to ensure long term stability
  - Adoption will take years (AGS experience)

Governance

- AGS is an Industry Group
  - Funding is minimal
    - Pay small membership fee (£200/year)
  - 2 members from each constituency on Board
  - Work is volunteered

Governance

- GMS/GDC/Core Team is an excellent project structure – needs to control TPF project
- AGS has a long standing and excellent structure for UK – Local Implementation Group (LIG)
- US needs an LIG
- Need to consider International Coordination
  - So far, pooled fund has accomplished this through the project implementation (joint meetings)

General Governance Issues

- Do not charge for schema use (charging will kill the adoption)
- Need some mechanism to keep "derivatives" consistent (stop proliferation)
- Must be independent of software vendors
- Come under umbrella of AGS (could be temporary)
- Limit what can be localized in the schema (make LIG work easier)
- Some mechanism to handle maintenance & support issues (pay, community, etc)
- Examine a variety of models (COSMOS, ASTM, ASCE, AASHTO, AGS, etc)
- Will "control" be lost If handed to another group/org (e.g. ISO)
- Geoenv – National groundwater assoc. could be involved
- ASCE – Geoinstitute possible for US LIG
- Include rules in documentation for extension, changes, additions, etc – and trademark/copyright

General Governance Issues

- Action Item: Task a subcommittee to evaluate ideas and report
  - GDC and GMS approve
- Need transition plan to hand off (once V1 released, then body takes it over)
- Need existing governance body to had off to (AGS-M worked this way)
- Need agreement to say free standard (if handed off)
- Need some type of funds to maintain
- Consider selling paper documentation
- Conferences generate funds
- How to ensure unified standard (international)

Governance Task XX

- Task a subcommittee to evaluate ideas and report on how to form a US LIG
  - GDC and GMS approve
  - Develop a few options to present to GDC & GMS
    - Include International and/or US group
    - Address funding options
    - Include details on structure, organization
    - How to maintain and support schema long term
  - Must keep international cooperation and single standard
    - Must address local (country) needs
    - Suggest how AGS and US LIG cooperate
  - Recommend best option
  - Include any funding required from TPF
  - Recommendation ready by June date (when technical team reports)
  - Hoit will chair committee

Governance task

- Mohamad Mullah (NCDOT) wants to help
- Get name from Tom for GeoInstitute
  - Renaldo Luna (computational geo-mechanics)
- Tom Lefchik wants to be member

- AGS requirements
  - US based group in place – for AGS partnership
  - International structure (can be handled later)
    - Probable US & UK for first 10 years
    - AGS and US LIG form cooperative management group
  - Funding ideas
  - DIGGS MUST work! (in order to be adopted)
    - DIGGS V2 is probably earliest that can be adopted
    - Would adopt as AGS 5 (not 4)

International Governance

- AGS & US-LIG organize joint annual meeting
  - Groups sign MOU on process for change
  - Good for at least five years

Tasks

- Task 1 – review & fix
  - Quick review for V1
    - Punch list of problems to fix
    - Track on forum (use issue tracking)
  - Longer review
  - Authorize Loren to spend XX to accomplish fixes found in Task 1
- Task X – Governance
- Task Z – Maintenance & support
  - Message on web – new corrected version coming
  - 2 year transition solution
  - Need paid position to handle – with authority to make changes
  - Post on blog and documentation
- Task W – Community development forum
  - Be sure assignments clear
  - Give guidance on how to get messages when posted
  - Establish tools to support community, transparency, maintenance
  - Wiki for documentation
- Task Y – Deployment Team (Core team?)
- Task V – Fill out other Schemas (Deep Found, geophysics & Geo-env)
- Task U – Develop Tools

TPF Tasks – Needed Tool & Pilots

- AGS -> DIGGS mapping (AGS responsible)
- Simple Tools to help people not in the group
  - web app/style sheet that displays DIGGS file (data browser)
    - Put repository of style sheets by community on website (how to vet quality)
    - Very simple and VERY low cost
  - Checker to validate non-schema validated info
    - Schematron, etc (Continuous layers, etc)
  - Simple example spreadsheet that reads e.g. SPT and plots (to show how you can program)
- Develop business cases for adoption
  - AGS doing similar task – work together
  - Document peoples experience at adopting DIGGS
    - Find X number of groups adopting and use DIGGS person to document
  - Establish a “benefit log” – listing of hard and soft benefits of adoption
- Planned Pilots with dates and milestones
  - Test of interoperability between FL, CA, Ohio – once complete

Additional Ideas for Sustainability

- Community Development Model
- Outsource support for schema