TECHNICAL REPORT STANDARD PAGE

1. Report No. FHWA/LA.12/498	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle **Geotechnical Information Database - Phase II**	**5. Report Date** May 2013 **6. Performing Organization Code** LTRC Project Number: 10-2GT State Project Number: 30000201	
7. Author(s) Scott L. Deaton, Ph.D.	**8. Performing Organization Report No.**	
9. Performing Organization Name and Address Department of Civil and Environmental Engineering Louisiana State University Baton Rouge, LA 70803	**10. Work Unit No.** **11. Contract or Grant No.**	
12. Sponsoring Agency Name and Address Louisiana Department of Transportation and Development P.O. Box 94245 Baton Rouge, LA 70804-9245	**13. Type of Report and Period Covered** **Final Project Report** **June 8, 2012** **14. Sponsoring Agency Code**	
15. Supplementary Notes		

16. Abstract

The Louisiana Department of Transportation (LADOTD) has been collecting geotechnical data for many years in a variety of different formats. Accessing this data and combining it with new data for the purpose of design, analysis, visualization, and reporting is difficult because the data has been generated by disparate systems and stored as hard copies, scanned images, various digital formats, or other non-digital formats such as Microfilm. Essentially, there is no single system or repository nor an integrated, systematic approach for collecting, managing, reporting, archiving, and retrieving the vast amount of geotechnical data is collected or generated by LADOTD each year.

With advances in computing capabilities, software tools are now available that streamline the entire data management process from data collection through reporting, archiving, and map-based retrieval/reporting. Dataforensics created a plan to integrate and implement a customized data management system to fulfill the needs of LADOTD. This Enterprise Global Information System (GIS)-based Geotechnical Data Management System is comprised of various off -the-shelf software packages including PLog Enterprise, RAPID CPT, gINT, ArcGIS, and ArcGIS Server. It enables LADOTD to store geotechnical data in a consistent database format while improving the reliability and accessibility to key stakeholders.

17. Key Words Geotechnical Data, Data Management, Enterprise GIS,		**18. Distribution Statement** **Unrestricted. This document is available through the National Technical Information Service, Springfield, VA 21161.**	
19. Security Classif. (of this report)	**20. Security Classif. (of this page)**	21. No. of Pages 31	22. Price

Project Review Committee

Each research project will have an advisory committee appointed by the LTRC Director. The Project Review Committee is responsible for assisting the LTRC Administrator or Manager in the development of acceptable research problem statements, requests for proposals, review of research proposals, oversight of approved research projects, and implementation of findings.

LTRC appreciates the dedication of the following Project Review Committee Members in guiding this research study to fruition.

LTRC Manager

Gavin Gautreau

Research Manager

Members

Sarah Collins

Ben Fernandez

Kim Garlington

Hossein Ghara

Clay Gottschalck

Jeff Lambert

Jim Mitchell

Steve Meunier

David Ringuette

Beth Roberts

Ching Tsai

Hollis Ward

Bert Wintz

Directorate Implementation Sponsor

Richard Savoie, P.E.

LADOTD Chief Engineer

Geotechnical Information Database—Phase II

by

Scott L. Deaton, Ph.D.

Dataforensics, LLC.
6825 Jimmy Carter Blvd
Norcross, GA 30071

LTRC Project No. 10-2GT
State Project No. 30000201

conducted for

Louisiana Department of Transportation and Development
Louisiana Transportation Research Center

May 2013

ABSTRACT

The Louisiana Department of Transportation (LADOTD) has been collecting geotechnical data for many years in a variety of different formats. Accessing this data and combining it with new data for the purpose of design, analysis, visualization, and reporting is difficult because the data have been generated by disparate systems and stored as hard copies, scanned images, various digital formats, or other non-digital formats such as microfilm. There is neither a single system or repository nor an integrated, systematic approach for collecting, managing, archiving, and retrieving the vast amount of geotechnical data collected or generated by LADOTD each year.

With advances in computing capabilities, software tools are available that streamline the entire data management process from data collection through reporting, archiving, and map-based retrieval/reporting. Dataforensics created a plan to integrate and implement a customized geotechnical data management system to fulfill the needs of LADOTD. This Enterprise Global Information System (GIS) -based Geotechnical Data Management System is comprised of various off-the-shelf software packages including PLog Enterprise, RAPID CPT, gINT, ArcGIS, and ArcGIS Server. It enables LADOTD to store geotechnical data in a consistent database format while improving the reliability and accessibility to key stakeholders.

The proposed data management system will have significant implications on LADOTD's ability to respond to what might be termed "extreme events." This refers to the engineering consequences of natural events such as hurricanes and man-induced events, such as bridge and other waterfront structure impacts by ships and barges. In many cases, organizations such as LADOTD are required to be able to return the impacted structure to some level of functionality in a very short period or else demolish and rebuild a replacement structure in a "fast-track" mode. Having this GIS-based geotechnical data management system improves their ability to find, utilize, and share the appropriate geotechnical data to meet these challenging project demands.

Finally, as part of this project, historical site investigation data has been loaded into the LADOTD GIS system. This involved significant data cleaning and conversion from its native format (LADOTD gINT projects, raw CPT data files, consultant data files, etc.) into the database structure developed for LADOTD as part of this project. This includes incorporating scanned documents of boreholes, assuming they are geo-referenced.

ACKNOWLEDGEMENTS

The financial support for this project was provided by the Louisiana Department of Transportation and Development/Louisiana Transportation Research Center (LTRC) under State Project No. 30000201 and LTRC Research Project No. 10-2GT.

The authors would like to express thanks to all those who provided guidance and assistance on this project. Specifically, the following LTRC and LADOTD staff contributed to the project: Sarah Collins, Ben Fernandez, Kim Garlington, Gavin Gautreau, Hossein Ghara, Clay Gottschalck, Jeff Lambert, Jim Mitchell, Steve Meunier, David Ringuette, Beth Roberts, Ching Tsai, Hollis Ward, Bert Wintz, Pallavi Bhandari and Murad Abu-Farsakh. In particular, the guidance and assistance from Ben Fernandez and Gavin Gautreau were instrumental in the success of this project.

IMPLEMENTATION STATEMENT

The comprehensive Enterprise GIS-based Geotechnical Data Management system will allow LADOTD to collect, manage, and report geotechnical data significantly more efficiently and consistently. This data management system incorporates borehole data, cone penetration test (CPT) data, lab testing data, and pile load test data.

Additionally, the website developed as part of this project will allow LADOTD to locate existing boreholes, CPT data, and pile load test data on a GIS interface while providing the ability to utilize this data in a variety ways even if it was not originally reported in that format. Users can generate reports dynamically based on the dataset selected on the map-based interface. Additionally, having the data stored in a database allows users to create new reports in the future to utilize this data in different formats as project requirements evolve over time.

Users can combine information with other relevant information such as soil survey maps, geologic maps, quadrangle maps, etc. so that appropriate engineering decisions can be made with minimal time spent searching for data in the various locations it resides today.

This data management system helps meet the needs of the Geotechnical Design section by allowing them to automate various reports they manually compile today. Additionally, it builds on the work completed in project 03-1GT by aiding personnel in the Geotechnical Design section to select proper boring depths, pile depths, etc. by providing easy access to historical data near their project sites. Additionally, the comprehensive data that is stored in the database, including borehole, CPT, vane shear, and laboratory testing data combined with pile load test data, have the potential to improve the pile resistance prediction, and may ultimately reduce the cost of foundation construction.

TABLE OF CONTENTS

INTRODUCTION

The traditional process for managing geotechnical data that LADOTD currently utilizes is similar to the approach utilized throughout a large portion of the geotechnical industry. An example of the inefficient, disconnected, and discontinuous nature of this traditional process is shown in Figure 1, where in each task for the project, a subset of the data already input elsewhere in the process is repeatedly input in separate and disconnected systems. In order to ensure adequate data quality, quality assurance/quality control (QA/QC) should be performed for each of the data entry processes performed; otherwise, errors associated with repeated transcription, as well as typographical, spelling, or even calculation errors may also be introduced in the process.

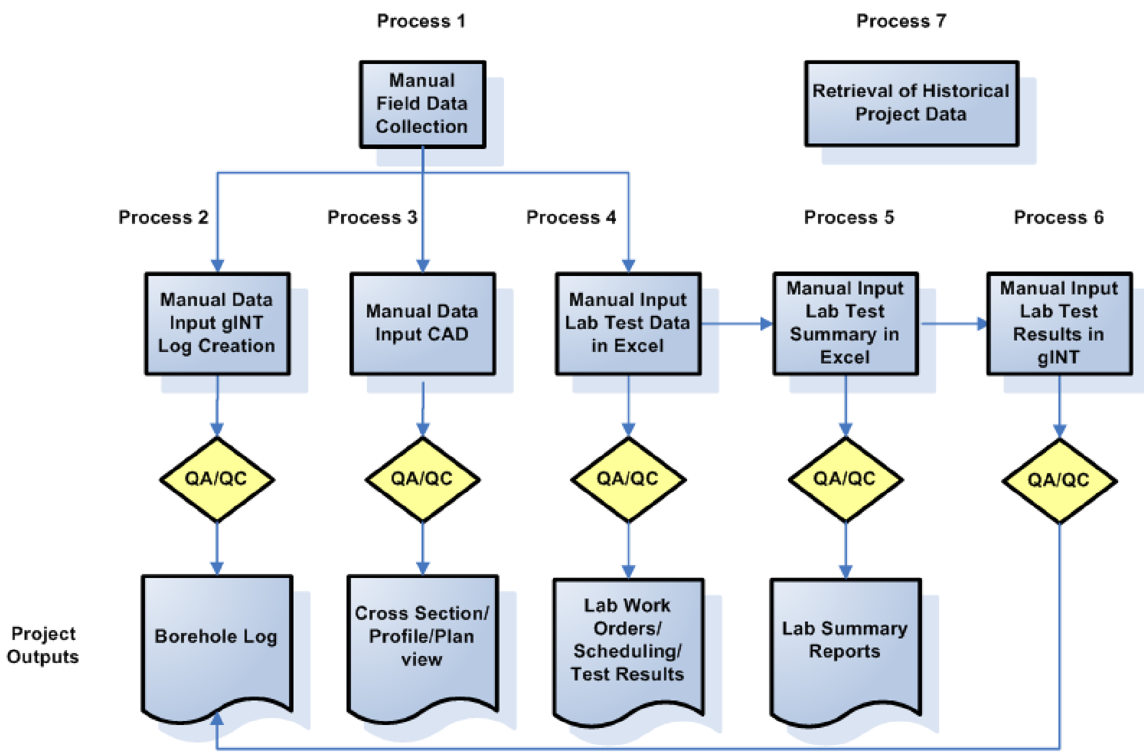

Figure 1
Typical traditional data management process

All of these undesirable inefficiencies and error sources can be overcome by a comprehensive enterprise geotechnical data management system. Within software, a process can be implemented that mimics and facilitates LADOTD business process. This enables significant streamlining of the data management process used on future projects and ensures the long term availability and usability of this data via an Enterprise GIS-based web portal.

With advances in computing capabilities, software tools are available that streamline the entire data management process from data collection through reporting, archiving, and map-based retrieval/reporting. Dataforensics created a plan to integrate and customize a geotechnical data management system to fulfill the needs of LADOTD. This Enterprise GIS-based Geotechnical Data Management System is comprised of various off-the-shelf software packages including PLog Enterprise, RAPID CPT, gINT, ArcGIS, and ArcGIS Server integrated with critical process and workflow components designed and developed as part of this project.

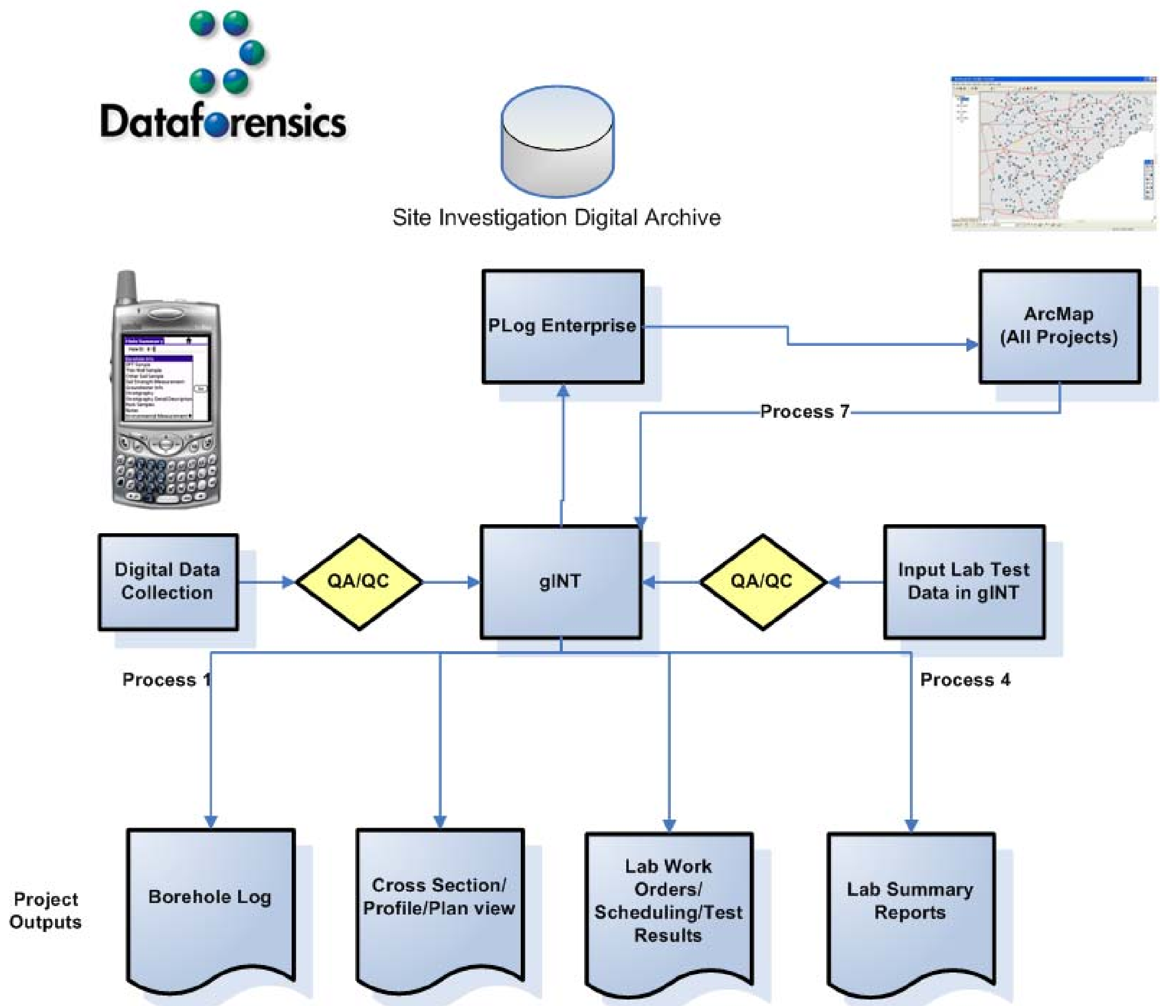

Figure 2
Enterprise GIS-based geotechnical data management system

Figure 2 depicts the data flow within the proposed Enterprise GIS-based Data Management system. Contrasting the data flow for this system with the traditional data management approach shown in Figure 1 highlights numerous achievable efficiencies. Specifically, significant efficiencies result from single source data entry in the field and lab testing data entry in gINT. Furthermore, QA/QC must only be performed once for each data entry point, much of which can be automated via validation rules programmed into the data management system. Finally, having a single integrated system then allows the reports to be generated for all of the various needs without any data duplication. Even if the data is not input in the field digitally, the streamlining is quite significant from that point forward in the system. All of the data can be archived into the PLog Enterprise GIS system, such that it can be re-used and combined with data from future exploration projects.

OBJECTIVE

A typical engineering project is comprised of a number of phases including a review of historical information, a field investigation, a laboratory testing program, office-based analysis and design studies, full-scale test programs of selected components, construction supervision, and post-construction performance monitoring. With the exception of the analysis and design studies phase, all other phases typically result in the generation of information and associated reports, which are used in an integrated fashion to complete the project. Too often, however, these phases are regarded as relatively discrete activities, and this leads to the type of inefficiencies and error sources previously noted. The scope of this project is to allow LADOTD to move to the forefront of research and practice by creating an enterprise GIS-based geotechnical data management system that allows for the integration of historically acquired (likely in non-digital format), recently acquired (likely in digital format), and future acquired data (in digital format) so that not only does the composite database generated for a particular project benefit that project, but it becomes part of a larger knowledge base available for use on other projects being undertaken by LADOTD in the future. In order to realize this goal, Dataforensics identified a scope of work that incorporated tasks and strategies that require expertise in geotechnical engineering, database systems, GIS technologies, process flow, as well as software development and integration.

SCOPE

The scope of the project focused on the creation of comprehensive database structure and reports that facilitate the data management process for geotechnical data including borehole data, CPT data, vane shear data, test pile load test data, and laboratory testing data.

The type of data included in this database structure is limited to:

- Borehole data and laboratory data that is often generated as a result of samples obtained from the borehole.
- Data types for which LADOTD has access to in a nonproprietary, digital format (i.e., CPT data or Test Pile Load Test Data) such that it can be easily imported.
- Data from proprietary data formats such as Geocomp lab testing files and binary data associated with test pile information CAPWAP, WEEP, etc. are not imported into the database as they require the proprietary software to read these files.

Within lab testing, basic laboratory testing of soils including natural moisture/natural density, Atterberg Limits, sieve, hydrometer, specific gravity, Proctor, unconfined compression, and chemical testing are all available in a data format. More advanced strength testing such as consolidation, direct shear, and triaxial have key test results stored in the database but utilize the laboratory testing equipment software for the processing of the data in their proprietary format. The proprietary raw data files for these tests are stored in the database for future usage, as well.

Additionally, data cleansing and data migration tools were developed to clean existing geotechnical data so it can be archived within the PLog Enterprise system. The scope was limited to data compiled in the 03-1GT project including scanned documents for historical boreholes, existing LADOTD gINT projects, gINT projects obtained from consultants, and the LADOTD CPT raw data format. Other older boring logs, pile driving records, load test reports, and other geotechnical data were not archived and would need to be included as a separate project.

A custom web-based workflow management/tracking system was developed to facilitate tracking the status of projects by LADOTD personnel.

Lastly, a customized GIS-based website was created to allow LADOTD personnel to easily visualize and report the data in the geotechnical data management system.

METHODOLOGY

The methodology utilized on this project can be divided into four main tasks with a deliverable associated with each task:

- gINT Database and Report Development
- Workflow Tracking System
- PLog Enterprise Web Portal
- Historical Data Cleansing and Archiving

gINT Database and Report Development

The initial task of the project involved researching the current data management processes and software utilized for processing and storing borehole logs and other geotechnical data within LADOTD. The investigation included interviewing LADOTD personnel using questionnaires and teleconferences, as well as meetings with the PRC committee members; Headquarters Geotechnical Design Group; the Material Laboratory; Information Technology (GIS, Content Manager, etc.); and others. During this process, the types of data to be stored within the system as well as representative examples of the various reporting formats were obtained.

Since streamlining the geotechnical data management process was the primary focus of this study, the next task focused on developing a comprehensive database structure in gINT and PLog Enterprise that encompassed borehole, CPT, vane shear, test pile load test data, and laboratory testing data. Once the database structure was designed, validation rules were added to ensure the quality and reliability of the data was acceptable. A comprehensive description of the validation rules for each table in the database is included in the interim report for the project. Examples of validation rules include:

- Verifying that specific data is input for a particular test or record
- Verifying coordinates are reasonable (i.e., longitude must be negative in Louisiana)
- Verifying a user has the necessary permission to add, edit, or delete a record
- Automating various calculations (N-values, CPT correlations, etc.)
- Verifying samples, layers, and other data types are less than or equal to the depth of the hole

Once the database structure and validation rules were finalized, importers for the LADOTD CPT format and Test Pile Load Test Data were developed to facilitate automatically, loading data types from their native format into the geotechnical data management systems. Additionally, custom correlations for the Zhang & Tumay Fuzzy and Probabilistic Soil Classification system were incorporated into the RAPID CPT software based on research by M. Abu-Farsakh for LTRC/LADOTD *[2]*.

Once data could readily be compiled within gINT, reports were developed based on the requirements determined in the initial task to mimic existing processes whether already in gINT, other software, or manually generated.

Lastly, permission for various groups of users (roles) were assigned in gINT to ensure the reliability of the data and eliminate the potential for data being edited or deleted after it has been reviewed and approved by the appropriate personnel.

Workflow Tracking System

The original intent of this task involved designing and developing a workflow tracking system similar to the one utilized by the Kentucky Transportation Cabinet (KYTC). However, as researchers initiated the design process, it became clear that the tracking system utilized by KYTC was significantly different both in scope and functionality compared with what LADOTD requested.

Accordingly, Dataforensics created a draft design document that was presented to the PRC as part of the interim report including flow charts depicting the business process today and with the proposed system as well as wireframes of the various user interface forms. Feedback was provided at that time and the design was refined further. At that time, specific design details were guided by a member of the PRC, Ben Fernandez, who would be ultimately managing the system. Additionally, the data structure of the ERP system influenced design considerations.

Six weeks after Dataforensics informed LADOTD that the workflow tracking system was available to be reviewed as a beta version with significant functionality above and beyond the KYTC system was added, Dataforensics was informed that a SharePoint site would be a preferred alternative to the workflow system. However, utilizing SharePoint technology had not been indicated to Dataforensics in any of the design phases of the system and would require a substantial amount of effort to modify. Since Dataforensics had already exhausted the budget for the development of the site in its current format, Dataforensics ceased development on the workflow tracking system. Dataforensics believes the system is approximately 90-95% complete and will serve as an excellent design for adapting it to work within SharePoint.

PLog Enterprise Web-Portal

The PLog Enterprise system already contains a PC-based toolbar that runs in ArcGIS (9.2, 9.3 and 10.0) that allows users to visualize their geotechnical data within a simple, map-based user interface. From the map, any report configured in the gINT Database and Report Design task can be generated based on the live data in the PLog Enterprise database. Accordingly, this functionality was duplicated within a web-based environment to facilitate broader access and utilization of the geotechnical data across various LADOTD sections. Essentially, the goal is to make the data in the system accessible via the preconfigured reports developed to facilitate the LADOTD workflow process, but not requiring users to have ArcGIS installed.

Accordingly, a flex-based user interface was developed and combined with a newly developed ASP.NET middleware application that integrates with existing PLog Enterprise functionality. The ASP.NET middleware application and PLog Enterprise are based on .NET framework 4.0. The middleware application utilizes pooling technology to allow multiple simultaneous users to generate gINT reports served in a PDF format directly to the Flex based user interface. A schematic of the system architecture is shown in Figure 3.

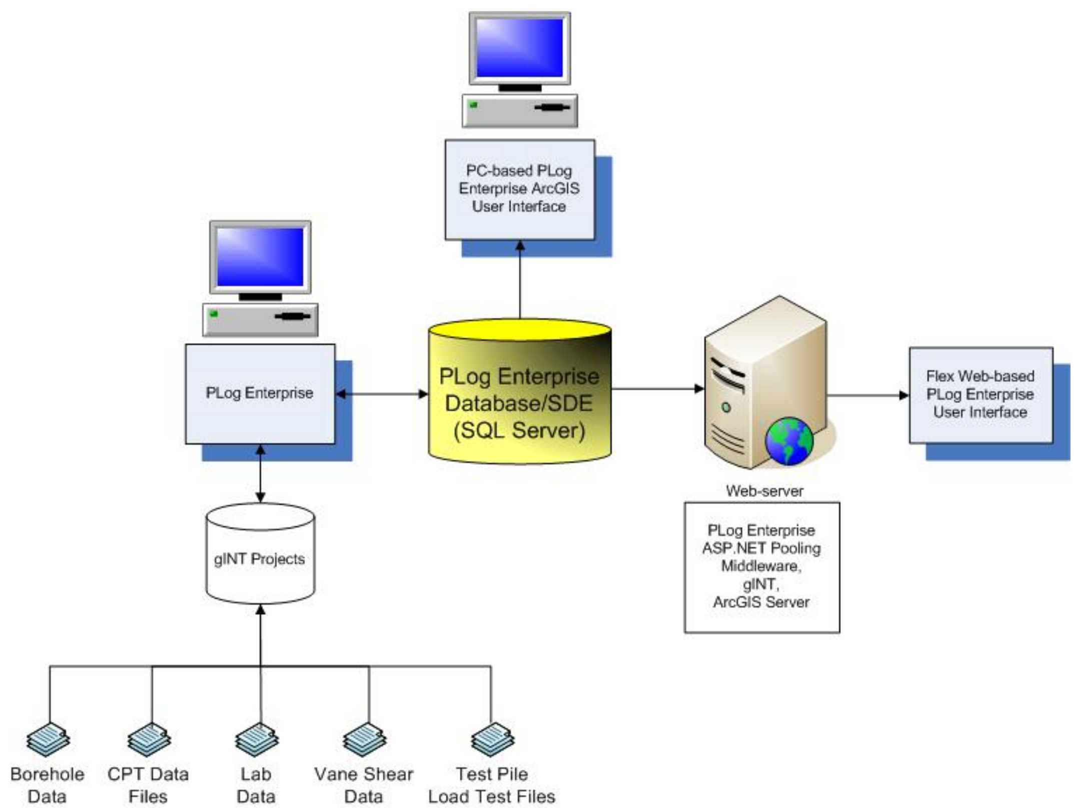

Figure 3
Enterprise GIS-based geotechnical data management system

Historical Data Cleansing/Archiving

A significant amount of effort was required to compile, cleanse, standardize, and archive existing historical data utilized in the 03-1GT project. Data cleansing was the most time consuming of these processes but was key to the future usage of the data. The predominant hurdles that had to be overcome with the various data formats were:

- Raw CPT data files with latitude and longitude in various formats
- Existing borehole data in gINT with latitude and longitude in various formats
- Overloading data fields
- Existing borehole data without coordinates
- Existing CPT data and scanned boreholes without coordinates or with coordinates in another file format
- Categorizing each type of hole
- Incorporating scanned borehole logs into PLog Enterprise
- Nonsensical coordinates
- Conflicts between coordinates determined in project 03-1GT and the current project
- Projects with varying unit systems
- Duplicate Project IDs
- Duplicate data
- Text data that should be numeric

DISCUSSION OF RESULTS

The results of this project can be found in the gINT files provided to LADOTD and will be located on the LADOTD Intranet. A discussion of the results follows.

Raw CPT Data Files Latitude and Longitude

The CPT data files from the 03-1GT project were attempted to be imported into gINT using the LADOTD CPT importer developed as part of this project. It quickly became apparent that the majority of the CPT files could not be imported properly because the latitude and longitude were in a variety of different formats. Accordingly, a VB.NET tool was created to attempt to convert and log the various formats for the latitude and longitude. In this process, six regular expressions were developed for deciphering the latitude and longitude that comprised approximately 18 various formats. The latitude and longitude were predominantly in either a degrees-minutes-seconds format or a decimal-degrees format. However, it is easy to imagine the combinations that can occur: spaces may not exist between each component (degrees, minutes, seconds), the minutes (‘) or seconds (“) symbol may not exist at the end of each component, the “N” or “W” indication in the string may not exist, and/or a “-” may not exist in front of the longitude.

Existing Borehole Data in gINT Latitude and Longitude

The gINT data files compiled as part of this project were significantly more consistent than the raw CPT data files. Only two regular expressions were developed to parse the latitude and longitude into a consistent format, one for decimal degrees and one for degrees, minutes, and seconds.

Overloading Data Fields

The existing gINT data files compiled as part of this project contained multiple data types within the same field [ex. standard penetration test (SPT)_or_ undrained unconsolidated (UU) field in the sample table]. This is a violation of the normal forms of database design because it makes it much more difficult to work with the data. One type of data was the format was the undrained strength and confining pressure for the UU test. It was indicated as 5.933@16.5 where 5.933 is the strength in tons per square foot (tsf) and 16.5 is the confining stress in pounds per square inch (psi). In the same field, there were also blow counts and N-values stored. To determine if it was an N-value, it is stored as N=*value*. To determine if it was three separate blow counts, it was indicated as 1-2-3, where each number is the blow count for a particular interval.

To clean this data, Gavin Gautreau enlisted a student worker to replace all N-values with the individual blow counts in the same format (i.e., 1-2-3). Then, the data was migrated from the old database structure to the new database structure developed as part of this project. During this process, UU data was transferred into its own table with fields for "Confining Pressure" and "Undrained Strength," while the blow counts were separated into three different fields with a fourth field that automatically calculates the N-value.

Existing Borehole Data without Coordinates

The existing borehole gINT data files in some instances lacked coordinates in the gINT file. If boreholes were missing coordinates, Dataforensics requested the coordinates of the holes from LADOTD personnel and added them to the gINT files as latitude/longitude in decimal degrees format. All existing borehole data had coordinates and thus were categorized as "Legacy Boreholes" to allow for differentiation of this Hole Type on the GIS map.

Existing CPT Data without Coordinates

The existing CPT data files in some instances lacked coordinates in the raw data file. In these scenarios, the GIS layers developed as part of 03-1GT were utilized if the coordinates were present there. These CPT soundings were classified as a Hole Type of "CPT." Alternatively, if no coordinates were identified in the GIS coverage, the Excel spreadsheets compiled as part of the 06-6GT project were utilized for identifying the location of the soundings. If no location was known from the spreadsheets, the average location of all known CPT soundings was determined for a project and each sounding with an unknown location was assigned this latitude and longitude. In these scenarios, Hole Type was indicated as "CPT – Est. Location" so that users of the GIS application would see different symbology for these holes and understand that the location is approximate.

Scanned Boreholes without Coordinates

The scanned borehole data in some instances lacked coordinates in the scanned document. In these scenarios, the GIS layers developed as part of 03-1GT were utilized if the coordinates were present there. These boreholes were given a Hole Type of "Scanned Borehole". Alternatively, if no coordinates were identified in the GIS coverage, the Excel spreadsheets compiled as part of the 06-6GT project were utilized for identifying the location of the borehole. If no location was known from the spreadsheets, the average location of all known boreholes was determined for a project and each sounding with an unknown location was assigned this latitude and longitude. In these scenarios, Hole Type was indicated as “Scanned Borehole – Est. Location” so GIS users would see different symbology for these holes and understand that the location is approximate.

Categorizing Each Type of Hole

As indicated in the previous two sections, different hole types were identified depending on the source of the data and the reliability of the location. Other categories were included in the gINT configuration, such as Borehole and Test Pile. Borehole will be utilized by LADOTD as new data is input in the gINT format. Test Pile will be utilized by LADOTD as new test pile load test data is imported into the system. Dataforensics recommends making the “Hole Type” field required in order to ensure differentiation of the points on the GIS map in the future.

Incorporating Scanned Borehole Images

In order to make the scanned borehole logs accessible in the PLog Enterprise system, the scanned boreholes must be added into a gINT project. The process requires adding a new record for each borehole in the Point table and inputting critical metadata such as PointID, HoleDepth, Latitude, Longitude, Hole Type (Scanned Borehole or Scanned Borehole – Est. Location) and inputting the file path for the scanned image in the associated file in the Point table.

Nonsensical Coordinates

For some projects, the latitude and longitude provided in the scanned boreholes were incorrect coordinates (meaning they were incorrectly printed on the scanned log report). Dataforensics identified these coordinates as nonsensical because they contained either minutes or seconds that were greater than 60. In these scenarios, LADOTD employees were utilized to identify the correct coordinates, which were then loaded into the system.

Conflicts between Coordinates Determined in 03-1GT Project and Current Project

For some projects, the latitude and longitude indicated in the ArcGIS layer differed from what Dataforensics identified as coordinates from the CPT data files or other sources. In these scenarios, Dataforensics deferred to coordinates from the 03-1GT project.

Projects with Varying Units

For some of the projects with scanned boreholes, the scanned images were in metric units; thus, the station, hole depth, and elevation values input into the system were converted to English units to be consistent with all the other data. There were no projects in gINT that were provided in metric, so the correspondence file written to convert historical data from the previous LADOTD gINT database format to the database format created as part of this project only supports English units.

Duplicate Project IDs

Several projects could not be imported into the database, because there were already projects with the same identifier and boreholes with the same identifier in the database. The duplicate project IDs were limited to three projects where the Design Number was indicated as “General Engineering.” For each of these projects, the Design Number value was changed to the name of the project, e.g., LA 488 Slope Failure, LA 8 Slope Failure, and US 84 Slide.

Duplicate Hole IDs

Several projects that contained duplicate Hole ID’s for a Project ID had already been loaded into the PLog Enterprise system. The source of this problem seemed to be in the compilation of data from one of the previous projects; the previous researchers allowed a CPT sounding to be referenced in multiple folders/Excel files. As Dataforensics developed gINT projects, a new project was created based on the name of the folder and then the raw CPT data from the text file was imported. This resulted in the same CPT file being loaded into multiple gINT projects because the same raw data files were present in multiple folders. Table 1 below indicates the folders where duplicate data for specific Project ID and Hole ID combinations resulted.

Table 1
Duplicate CPT data file locations

Project ID	Hole ID	Original Folder Name	Duplicate Folder Name(s)
455-05-62	BT3	DIST 08\RAPIDES\455-05-62\DEE0562	DIST 08\RAPIDES\455-05-62\DEE0562\RAMPD2 DIST 08\RAPIDES\455-05-62\DEE0562\RAMPD9 DIST 08\RAPIDES\455-05-62\DEE0562\RAMPD9T
700-23-0200	CPT1	DIST 03\IBERIA\700-23-0200	DIST 03\IBERIA\700-23-0200\LOREAUVILLE
700-23-0200	CPT2	DIST 03\IBERIA\700-23-0200	DIST 03\IBERIA\700-23-0200\LOREAUVILLE
005-07-0057	1	DIST 02\LAFOURCHE\700-29-0120\005-07-00	DIST 02\ST. CHARLES\005-07-0057
005-07-0057	CPT1	DIST 02\LAFOURCHE\700-29-0120\005-07-00	DIST 02\ST. CHARLES\005-07-0057
005-07-0057	CPT2	DIST 02\LAFOURCHE\700-29-0120\005-07-00	DIST 02\ST. CHARLES\005-07-0057
362-01-0009	B1	DIST 08\NATCHITOCHES\362-01-0009	DIST 08\NATCHITOCHES\362-01-0009\BAYOU NID AIGLE
362-01-0009	B2	DIST 08\NATCHITOCHES\362-01-0009	DIST 08\NATCHITOCHES\362-01-0009\RAT BOIS BAYOU

Project 424-07-21 provided a different set of challenges that expanded on the challenges identified in the previous paragraph. There were 26 CPT soundings in the project, half of which were missing coordinates and the other half already archived in PLog Enterprise. This is a result of the problem identified above, compounded by the fact that the raw data files for the CPT soundings were renamed either in one of the previous projects that compiled the data or when the data was generated. Refer to Table 2, which indicates the duplication of the data. Note that all values are identical, except the Hole ID with the one exception of the total depth for Hole IDs 105+53/CP1. It is Dataforensics assertion that it is very difficult if not impossible to push duplicate soundings at exactly the same start date/time (to a precision of seconds) and to the same total depth to a precision of 0.01 ft. Therefore, the soundings on the right half of the table were removed.

Table 2
Duplicate CPTs from project 424-07-21

Hole ID	Total Depth	Start Date	Hole ID	Total Depth	Start Date
105+53	144.03	11/8/1995 8:13	CP1	144.01	11/8/1995 8:13
113+89	155.31	11/8/1995 10:29	CP2	155.31	11/8/1995 10:29
120+66	154.81	11/7/1995 10:15	CP3	154.81	11/7/1995 10:15
121+42	124.91	11/7/1995 11:21	CP4	124.91	11/7/1995 11:21
122+18	125.31	11/7/1995 0:20	CP5	125.31	11/7/1995 0:20
122+95	127.31	11/7/1995 13:24	CP6	127.31	11/7/1995 13:24
123+71	144.61	11/7/1995 14:30	CP7	144.61	11/7/1995 14:30
124+47	113.51	11/7/1995 15:41	CP8	113.51	11/7/1995 15:41
125+23	135.31	11/7/1995 9:12	CP9	135.31	11/7/1995 9:12
125+98	117.71	11/7/1995 8:05	CP10	117.71	11/7/1995 8:05
126+75	123.81	11/6/1995 16:33	CP11	123.81	11/6/1995 16:33
127+51	118.11	11/6/1995 15:28	CP12	118.11	11/6/1995 15:28
113+90	95.01	11/7/1995 16:46	TESTC2	95.01	11/7/1995 16:46

Text Data that Should be Numeric

In the process of converting data from the old structure to the new structure, there were several instances where various values for elevation or hours after drilling were indicated as the combination of a number and some string describing it (i.e., approximately 0 or 18 hrs). These were manually cleaned as the problems were discovered.

Once all the data was cleaned and converted to a consistent format, it was then loaded into the PLog Enterprise system. A total of 1645 distinct locations were archived in the system. This can be subdivided into the specific Hole Types categorized herein, as shown in Table 3. Note that for the Scanned Boreholes and Scanned Boreholes with Estimated Locations, the exact number of locations is known for those that appear on the map, but in some cases the scanned images contained multiple boreholes, which were not all geo-referenced, therefore more borings may exist than the counted number on the map.

Table 3
Hole types archived in PLog enterprise

Hole Type	Number of Holes
Scanned Boreholes	153*
Scanned Boreholes with Estimated Location	35*
Legacy gINT Boreholes	297
CPT Soundings	1080
CPT Soundings with Estimate Location	11
Boreholes in new gINT Format	69

Process for Archiving Historical Data

During this project, a significant amount of data was archived into the PLog Enterprise system. The process for loading historical data depends on the format of the data. The formats of data that were archived are:

- CPT data files
- LADOTD gINT projects
- Ardaman gINT Projects
- Scanned PDF logs

A description of the process to load the aforementioned files into gINT is described below, followed by a description of archiving the data in PLog Enterprise.

CPT Data Files

CPT Data files can be imported into gINT using RAPID CPT directly.

- Choose the appropriate importer from the Add-ins menu.
- Select the correspondence file (ladot rapid cpt.gci).
- Select the file(s) to import.

Once the data is in gINT, ensure the coordinate system is defined in the Project tab as well as the Hole Type and coordinates in the Point table. Add any other metadata about the Project/Points and then proceed to archiving it in PLog Enterprise.

LADOTD gINT Projects

Dataforensics has written a correspondence file that converts the previous LADOTD gINT project database structure to the new database structure that is more comprehensive. The process to convert the data is:

- Create a gINT project based on the new LADOTD gINT files.
- From the file menu->Import/Export->Import from database.
- Choose the legacy gINT project as the source database.
- Choose the correspondence file (ladot legacy to ladot 2012.gci).
- Click ok.

Once the data is in gINT, ensure the coordinate system is defined in the Project tab as well as the Hole Type and coordinates in the Point table. Add any other metadata about the Project/Points and then proceed to archiving it in PLog Enterprise.

Ardaman gINT Projects

Dataforensics has written a correspondence file that converts the Ardaman gINT project database structure to the new LADOTD database structure. The process to convert the data is:

- Create a gINT project based on the new LADOTD gINT files.
- From the file menu->Import/Export->Import from database,
- Choose the Ardaman gINT project as the source database.
- Choose the correspondence file (ardaman to ladot 2012.gci).
- Click ok.

Once the data is in gINT, ensure the coordinate system is defined in the Project tab as well as the Hole Type and coordinates in the Point table. Add any other metadata about the Project/Points and then proceed to archiving it in PLog Enterprise.

Scanned PDF Logs

The new LADOTD gINT database structure incorporates the capability to include a document associated with each Point (Borehole). The process to load this data into gINT is as follows:

- Create a gINT project based on the new LADOTD gINT files.
- Input the Project Information.
- Input a record in the Borehole table for each separate point location (borehole, CPT, vane shear, etc.). Specify the coordinates for each point and in the Associated File field, specify the filename of the PDF that contains the scanned document.

The coordinate system must be defined in the Project tab, as well as the Hole Type and coordinates in the Point table. Add any other metadata about the point as necessary and then proceed to archiving it in PLog Enterprise.

Archiving Data in PLog Enterprise

Once the data has been loaded into gINT in the LADOTD format, it can then be archived into PLog Enterprise, which automatically makes the data available to anyone utilizing the PLog Enterprise system whether via the web application or ArcGIS. To archive the data in PLog Enterprise, refer to section IV in the PLog Enterprise User's Manual.

CONCLUSIONS

This project developed a comprehensive geotechnical data management system that will allow LADOTD to streamline their data management process for borehole, lab testing, CPT, in-situ vane, and test pile load test data while providing long-term availability of the data via a web-based GIS portal. By standardizing the database structure, and incorporating validation rules and custom reports LADOTD personnel in various sections can more easily access and report their geotechnical data while simultaneously improving the quality and reliability of the data.

Dataforensics provided a single system for managing all active project data and enabling it to be easily archived into the PLog Enterprise GIS based data management system accessible via a customized web-based GIS, which will allow the Geotechnical Design section access boring logs and other geotechnical documents via the Intranet. The GIS interface can access many different sources and types of data both within and outside the Department. The quick and easy access to valuable data, including the mapping applications in the GIS, will streamline and facilitate the analysis of data.

This project builds upon the work completed in project 03-1GT to create a reference resource that will continue to grow over time that will aid in the evaluation of specific geotechnical site data, which will allow for more accurate and cost effective design decisions. Additionally, this project archived data compiled in the 03-1GT project and the 06-6GT project as the first data loaded into the system. If LADOTD were to try and re-create this same data by re-investigating the same sites, the cost would be at least $16,000,000 ($10000 per borehole/CPT sounding). Accordingly, the return on investment for this task is extraordinary as approximately $20,000 was expended to turn $16,000,000 worth of data into a usable asset.

Various meetings and demonstrations of the system have been given to various members of the PRC. The technology utilized in the system described herein uses off-the-shelf software products PLog Enterprise, RAPID CPT, gINT, ArcServer, Flex and ASP.Net technology. The website is located on the Intranet and is ready to serve the Geotechnical Design section and others in the Department.

RECOMMENDATIONS

Website and Server Maintenance

Website and Server Maintenance: Perpetual licenses were included with research, meaning LADOTD can use the software indefinitely. Maintenance and periodic updates should be applied to the software in order to ensure proper function with an optimal performance. This will require combined efforts and ongoing support of both the GIS Team at LADOTD and Dataforensics. One year of support for RAPID CPT and PLog Enterprise is included in the current contract.

Personnel and Access

It is recommended that a Geotechnical Design section employee be designated to manage the PLog Enterprise system (i.e., upload the gINT projects whether developed by consultants or in-house) so that records are kept timely and accurate in digital format.

The website will be located on the LADOTD Intranet. A goal for a follow-up project should be to open the website to the public. Additional efforts regarding firewalls will need to be implemented for this to occur.

Site/Materials Manager

The incorporation of data from the Site/Materials Manager is unknown at this time. Dataforensics believes that incorporating borehole and lab test data from shallow boreholes into the PLog Enterprise system should be accomplished in order to maximize the value of the LADOTD database. The data types for the basic laboratory testing have been implemented already within the database structure, but reporting capabilities in gINT likely need to be enhanced and tailored to LADOTD's specific requirements. This was outside the scope of this project since it was focused on deep boreholes.

Loading Additional Historical Data

LADOTD now has a platform on which a tremendous amount of historical data can be loaded into the system quite easily. The difficulty involved with loading historical data is obtaining/finding the historical data and/or converting it into the standard format. For historical data that consists of scanned images, there is no conversion necessary. For this type of historical data, it simply requires creating a gINT project and filling in the appropriate project and borehole metadata to identify the project/hole and associating the scanned image with the point(s). A student worker could perform this task with minimal guidance from Dataforensics.

Dataforensics recommends a follow-up project where the budget can be explicitly allocated to pay a consultant to compile the gINT projects and provide to Dataforensics for the work they have performed for the LADOTD in the last few years. This would then provide datasets that can be cleansed and converted into the new LADOTD format.

Workflow Application

Dataforensics recommends a follow-on project to complete the workflow application using SharePoint functionality since it is an internally supported platform that is now available and suitable to manage the workflow related data and documents.

Future Modules

All current and future data should be recorded and stored digitally within PLog Enterprise so that as the data grows, it can be accessed easily via the map-based GIS interface. Enhancements to the system such as incorporating additional data types (i.e. extracting binary data from CAPWAP, WEEP, or the GeoComp systems) could prove to be useful from a data mining perspective over the long term.

Additionally, LADOTD may consider adding spatial and non-spatial data mining algorithms into the system to automatically search for and identify trends in the data that are not obvious until they are specifically examined. Data mining allows users to gain a better understanding of their data by discovering new patterns in their large datasets. PLog Enterprise provides the foundation for this to occur.

ACRONYMS, ABBREVIATIONS, AND SYMBOLS

CPT	Cone Penetration Test
GIS	Geographic Information System
KYTC	Kentucky Transportation Cabinet
LADOTD	Louisiana Department of Transportation and Development
LTRC	Louisiana Transportation Research Center
QA/QC	Quality Assurance/Quality Control
SPT	Standard Penetration Test
UU	Unconsolidated Undrained Triaxial Test

REFERENCES

1. Gautreau, G.P., and Bhandari, P., “Development of a Geotechnical Information Database.” 03-1GT Final Report, LTRC, Baton Rouge, La., Nov 2008.

2. Abu-Farsakh, M., Zhang, Z., and Tumay, M., “Development of MS-Windows CPT Soil Classification Software,” 87th Transportation Research Board (TRB), TRB 2008 Annual Meeting CD-ROM, 2008.

3. Zhang, Z., and Tumay, M., “Statistical to Fuzzy Approach Toward CPT Classification.” *Journal of Geotechnical Engineering*, Vol. 125, No 3, 1999.

4. Alshibli, K., Okeil, A., Alramahi, B., “Update of Correlations between Cone Penetration and Boring Log Data,” LTRC Report 439, FHWA/LA.08/439, 2008.