International data standards for geotechnical engineering

Les normes internationales de données pour l'ingénierie géotechnique

D.G. Toll
School of Engineering, Durham University, Durham, UK Email: d.g.toll@durham.ac.uk

ABSTRACT
The World Wide Web has revolutionised data access for all aspects of our daily lives. However, to allow full use of geo-engineering data from the web, as part of our professional activities, it is necessary for the data to be available in a structured and standardised form. This will allow the World Wide Web to become an international repository for geotechnical information, available to the whole community. A further advantage of developing data standards is to allow transfer of data between computer systems providing a data exchange format between different organisations or an interchange format for linking different software packages. The development of data standards for geo-engineering is now an important activity for the three international societies (ISSMGE, IAEG and ISRM) through Joint Technical Committee, JTC2 (http://www.dur.ac.uk/geo-engineering/jtc2). JTC2's role is to oversee the development of internationally agreed forms of representation of geo-engineering data. The paper discusses standardised XML schemes that are in development for geo-engineering and presents examples for borehole records and slopes. The DIGGS data format (Data Interchange for Geotechnical and Geo-environmental Specialists) is discussed and comparisons are drawn with data standards in the geosciences (eg GeoSciML). The paper also considers how web-based data could be used, such as the use of a case-based reasoning system for slope design using a global database of slope case histories.

RÉSUMÉ
Le World Wide Web a révolutionné l'accès aux données sur tous les aspects de notre vie quotidienne. Toutefois, pour permettre la complète utilisation des données de la géo-ingénierie sur le web, dans le cadre de nos activités professionnelles, il est nécessaire que les données soient disponibles sous une forme structurée et normalisée. Cela permettra au World Wide Web à devenir un référentiel pour l'information géotechnique, à la disposition de l'ensemble de la communauté. Un autre avantage de l'élaboration de normes de données est de permettre le transfert de données entre des systèmes informatiques offrant un format pour l'échange de données entre les différentes organisations ou d'un format permettant de relier différents logiciels. Le développement de normes de données pour la géo-ingénierie est désormais une activité importante pour les trois sociétés internationales (ISSMGE, IAEG et ISRM) par le Joint Technical Committee, JTC2 (http://www.dur.ac.uk/geo-engineering/jtc2). JTC2 a le rôle de superviser le développement international des formes de représentation des données de géo-ingénierie. Cet article traite les schémas standardisé XML qui sont en cours de développement pour la géo-ingénierie et présente des exemples de forages et les pentes. Le format des données DIGGS (Data Interchange for Geotechnical and Geo-environmental Specialists) est examiné et des comparaisons avec les données sont tirées des normes employées dans les géosciences (par exemple GeoSciML). Le document examine également comment les données sur le Web pourrait être utilisé, comme l'utilisation d'un raisonnement par cas (Case Based Reasoning) pour le système de la pente, en utilisant une base de données globale de cas de la pente.

Keywords : data standards, world wide web, XML, borehole, slope

1 INTRODUCTION

The World Wide Web provides us with easy access to a huge amount of information. However, at present, the data we can access exists in many formats. To allow routine use of geo-engineering data from the web, as part of our professional activities, it is necessary for the data to be stored in a structured and standardised form. The way to achieve this is by adopting XML (eXtensible Markup Language) and developing internationally agreed data standards for geo-engineering.

XML is a simple and highly extensible way to represent data, which is sufficiently flexible to allow data standards to continue to evolve to meet the needs of geo-engineering professionals. The concept of creating a geotechnical version of XML was first proposed by Mete Oner and the World Wide Web of Geotechnical Engineers (http://www.ejge.com/Gml/) in 1998. There are now a number of initiatives to develop representation schemes, both for geo-engineering and for geo-science data. The three international geo-engineering societies (ISSMGE, IAEG and ISRM) have a Joint Technical Committee, JTC2 (http://www.dur.ac.uk/geo-engineering/jtc2) to oversee the development of internationally agreed forms of representation of geo-engineering data. The data standards developed can be used to store such data on the World Wide Web and will ensure that geo-engineering data is stored in the same format anywhere on the web.

The major advantage of having data standards for web-based data is that it will make it possible to search all geotechnical data available on the web using structured search options (XQuery). For instance, it would be possible to locate XML files that contain projects within a particular geographical location or having particular soil/rock types or where a particular type of test has been performed. In this way the World Wide Web will become an international repository for geotechnical information, available to the whole community. This avoids the necessity to establish national or international geotechnical databases; each data owner can make their data directly available on their own web server.

A further advantage of developing data standards is to allow transfer of data between computer systems. XML uses very simple text files that can be easily accessed and read. Therefore, it can be used as a data exchange format between different

Proceedings of the 17th International Conference on Soil Mechanics and Geotechnical Engineering
M. Hamza et al. (Eds.)

doi:10.3233/978-1-60750-031-5-2690

organisations or an interchange format for linking different software packages. In this way XML could become the integrator between the different types of geotechnical tools (databases, knowledge-based systems, visualisation packages, conventional calculation software, numerical modelling packages etc.).

The paper discusses standardised XML schemes that are in development, with examples for borehole records and slopes. The DIGGS data format (Data Interchange for Geotechnical and Geo-environmental Specialists) is discussed and comparisons are drawn with data standards in the geosciences (eg GeoSciML). The paper also considers how web-based data could be used, such as the use of a case-based reasoning system for slope design using a global database of slope case histories.

2 DATA REPRESENTATION

The initial impetus for the development for data standards was for data exchange in geotechnical engineering. Greenwood (1988) and Threadgold and Hutchison (1992) identified the need to have a standard interchange format that was independent of particular software packages. This led to the development of the Association of Geotechnical and Geoenvironmental Specialists (AGS) data format in the UK in 1992 (the current version is available as AGS, 2004). A similar scheme for CPT data was developed in the Netherlands (CUR, 2000).

Such data exchange formats have been highly successful. Toll (1996) reported strong support for the AGS format from all involved, from data producer to data receiver/user, and also software developers. The format brought greater reliability and consistency to data transfer. However, such schemes were prone to problems when data sets were generated that did not fully conform to the standard.

The success of these data exchange formats for ground investigation data led to the development of XML data formats. Toll and Shields (2003), Chandler et al (2006), Weaver et al (2008) have described XML schemes for representing ground investigation data. The use of XML has the advantage that it can be self validating; the data standard can have a “schema” definition that defines the required structure for the data and the naming of the tags. It can even contain validation rules that specify the required format of the data within the tags, or even limiting ranges for the values. A data file that does not conform to the standard will show errors when validated against the schema.

Toll and Cubbitt (2003) also recognised that XML could do more than just provide a data exchange format. XML data formats allow storage and display of geotechnical data on the World Wide Web. In addition to “borehole” data from ground investigations, Toll and Cubbitt discussed how geotechnical entities (e.g. foundations, retaining walls and dams) could be represented in XML. Hatipoglu (2003) outlined an XML schema for storing case histories of slopes and this was extended by Toll (2007b). Styler et al (2007) described a scheme for deep foundations.

3 EXTENSIBLE MARKUP LANGUAGE (XML)

XML is a more generic form of mark-up language than HTML (Hyper-Text Markup Language), which has been the main language used on the World Wide Web. XML allows simple text files to be 'marked up' by including 'tags' that can be used to give meaning to the contents of a file; for instance data can be marked up using <slope> ... </slope> tags to indicate that all data between these tags relates to slope information. These tags can be recognised by an XML compliant web browser. XML is being widely adopted by web developers for producing web-based materials (http://www.w3.org/XML/).

The advantage of using XML for representation is that the data (stored in a simple text-based .xml file) is separated from the formatting information. Formatting can be provided by the use of a Stylesheet (.xsl) file. This means that the data can be formatted in different ways for presentation without having to make changes to the data file. Examples of how XML data can be formatted using stylesheets or represented graphically were given by Toll and Shields (2003) and Toll and Cubbitt (2003).

4 GEO-ENGINEERING DATA STANDARDS

Toll (2007a) discusses some of the current initiatives underway to develop data standards for geo-engineering. The most significant development is the DIGGS project (Data Interchange for Geotechnical and Geo-environmental Specialists) (http://www.diggsml.com/) which is discussed in more detail later in the paper.

Another initiative is a collaboration between JTC2 and the ISRM’s Commission on Testing Methods to develop standardised forms of representation for rock test data. This is being achieved through the establishment of a joint Working Group on for *Representing ISRM Suggested Methods in Electronic Form* (RISMEF).

JTC2 will also work with JTC1 (Joint Technical Commission for Landslides and Engineered Slopes) to develop forms of representation for slopes (Toll, 2007b).

5 OTHER DATA STANDARDS

In addition to Geo-Engineering schemes there have been parallel developments of XML schemes for geological sciences and mining. These include eEarth, XMML (eXploration and Mining Markup Langauge) and GeoSciML. eEarth (http://www.eearth.eu) was a European funded project that links the Geological Surveys of six European countries. The project aimed to increase availability, use and distribution of the digital subsurface data across Europe. The project is now completed and the website provides access to borehole information from several countries and in multiple languages. XMML (https://www.seegrid.csiro.au/twiki/bin/view/Xmml) was aimed at geoscience and exploration information. It has now been subsumed by GeoSciML (http://www.geosciml.org/) which aims to represent geoscience information associated with geologic maps and observations, as well as being extensible in the long-term to other geoscience data.

There have also been developments in generic representations schemes, using XML, that have wide application, such as Geography Mark-up Language (GML) and SensorML. GML (http://www.opengis.net/gml/) is a widely used and well-developed standard for representing geo-spatial data. It has constructs for representing coordinate schemes, locations and features that are geo-referenced. SensorML (2005) allows the representation of data from sensors (http://vast.uah.edu/SensorML/). Although initially developed to represent sensors for earth observation, much of the work is generic enough to be applicable to other types of sensors (see Toll, 2008a for applications to geotechnical engineering).

6 DIGGS

DIGGS represents the most advanced development of data structures in geo-engineering. This is a collaboration between the Federal Highway Administration (FHWA), United States Environmental Protection Agency (US EPA), US Army Corps of Engineers, US Geological Survey (USGS), Eastern Federal Lands Highway Division (EFLHD) and a number of Departments of Transport in USA, funded through the Transportation Pooled Fund. The UK Highways Agency is also a collaborator. DIGGS brings together existing standards

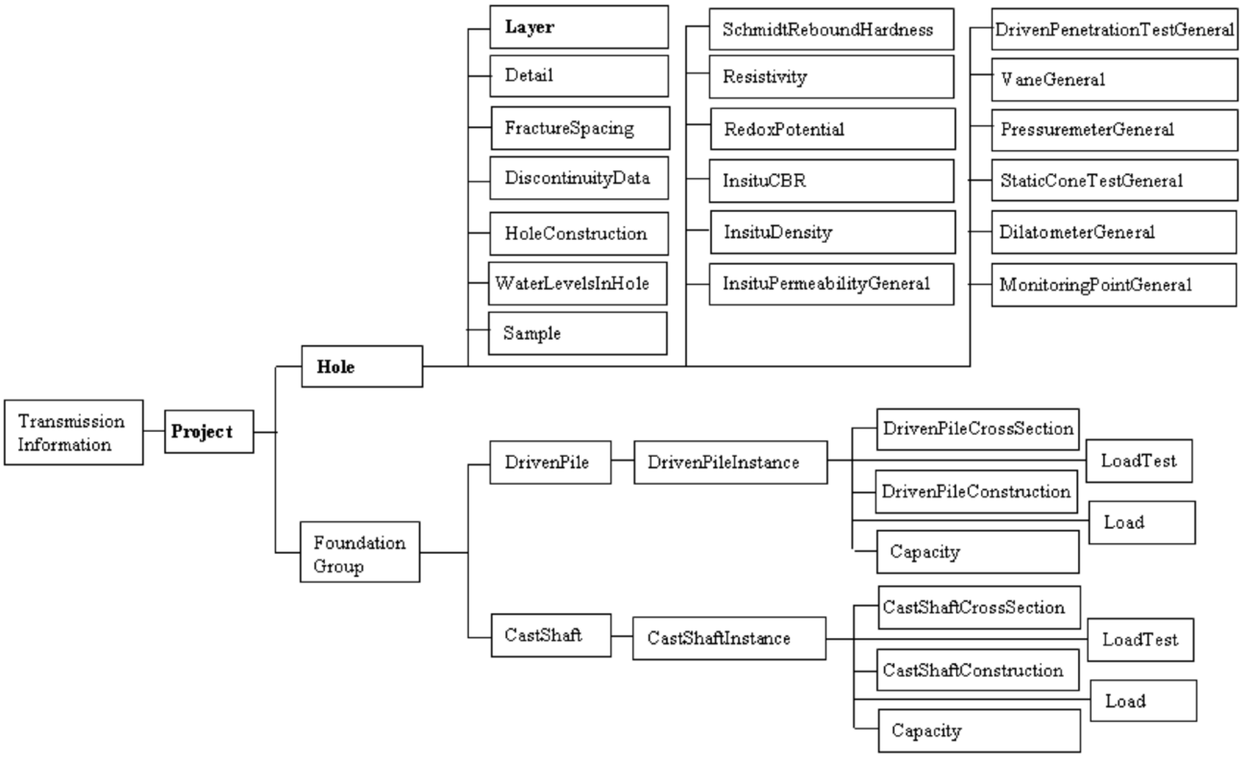

Figure 1. Data structure for DIGGS based on DIGGS, 2006 and Styler et al, 2007 (Toll, 2007a)

developed by AGS (www.ags.org.uk/agsml), Consortium of Organizations for Strong-Motion Observation Systems (COSMOS) (http://www.cosmos-eq.org/GVDC.html) and the University of Florida, Department of Civil Engineering (http://fdot.ce.ufl.edu/).

Currently the DIGGS framework defines over 300 geotechnical, geoenvironmental, monitoring and piling objects and associated properties (Bray et al, 2008). The high level data structure for the ground investigation and piled foundation objects are shown in Figure 1.

There were some concerns with the emerging DIGGS standard. Toll (2007a) identified the problem of a lack of consistency between geo-engineering schemes (such as DIGGS) and geo-science schemes (such as eEarth and GeoSciML). It is vital that these different naming systems are eliminated to avoid a lack of compatibility between common data entities being used by both the geo-engineering and geo-science communities.

DIGGS has made extensive use of GML for representing spatial data. However, Toll (2007a) suggested that the location representation used in early versions (DIGGS, 2006) could use a simpler GML construct to aid compatibility with other schemes. This seems to have been addressed in the latest release (DIGGS, 2008).

7 REPRESENTING SLOPES

Slope data is used by a range of professionals: geotechnical engineers, geomorphologists, geologists and planners. They may each need to represent information in different ways. Even if we consider only topographic information, different levels of representation are likely to be needed. Those dealing with hazard assessment (using GIS systems) will typically work with data on the slope height, slope angle and aspect (as well as lithology and land use). Geomorphologists may want to divide a slope up into segments having different land forms. Geotechnical engineers will usually produce quantitative cross-sections showing detailed topography.

It is therefore essential that a representation scheme is capable of operating at any of these different levels of representation. Toll (2007b) outlined a three level representation of topography that can be used to satisfy the varying requirements. At Level 1 the slope is defined simply by an overall slope angle and overall slope height (or Crest and Toe elevations). At Level 2 the slope is divided into segments (each referenced by height above the toe) which can be described as Slope Segments (defined by angle/height), Bench Segments (primarily defined by bench/berm width) or Wall Segments (primarily defined by wall height). Of course, each segment can have other properties attached, such as surface cover or geomophological descriptions. At Level 3 the topography is defined by coordinates allowing a very detailed topographic representation (in 2D or 3D).

Whetton (2009) has further developed the representation scheme and also provided ways to link the slope data with ground investigation data in the DIGGS format. Further tools have been developed, particularly the use of an SVG (Scalable Vector Graphics) application to display the slope data in a browser (Figure 2).

SVG is an XML application for describing images. It allows slope data stored in XML format to be transformed into an image for display within a web browser. The image is generated from the XML data describing the topography, the ground conditions and defined failure surfaces. This is greatly preferable to storing an image file (such as a .jpg or .gif file) as it means that if changes are made to the XML defining a slope, the image can be regenerated directly to reflect any changes.

8 USING XML DATA ON THE WEB

Fyson and Toll (2008) demonstrated a case-based reasoning system (SlopeSafe) for preliminary design of slopes. SlopeSafe

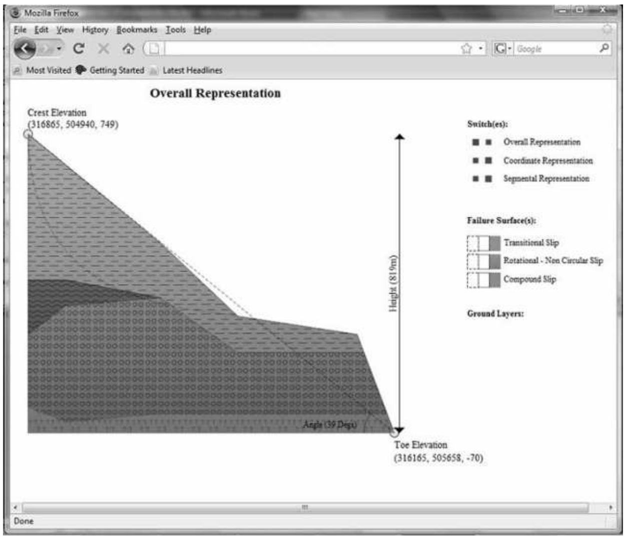

Figure 2. An SVG representation of a slope generated from XML

draws on a case-base of records of successful and failed slopes to give an indication of the likely success of a proposed slope by matching its geometry and ground conditions to the slopes held in the case-base. XML was used to create the case-base. The system, as originally designed, used a local XML file containing almost 3000 case records. However, Toll (2008b) has argued that slope case histories could be made available on the web, using a standard slope representation scheme, extending that reported by Toll (2007b). In this way, it would be possible for a case-based design system to access an international case-base of slope case records, by performing an XQuery search. Such a design system would automatically draw on new experience every time a new case history was added to the web, anywhere in the world.

9 CONCLUSIONS

It is necessary for geo-engineering data to be stored on the World Wide Web in a structured and standardised form if it is to provide a useful repository of data. The way to achieve this is by adopting XML (eXtensible Markup Language) and developing internationally agreed data standards for geo-engineering. There are now a number of representation schemes in development, with the DIGGS (Data Interchange for Geotechnical and Geo-environmental Specialists) standard being at the most advanced stage of development.

The advantage of using XML to represent data on the World Wide Web is that the data is separated from the formatting information. The XML data can be viewed in a variety of ways within a web browser, either formatted using stylesheets or represented graphically, for instance using Scalable Vector Graphics (SVG). An example is given that allows slope data and ground investigation data (in DIGGS format) to be linked and displayed in a web browser using SVG.

The use of a standardised XML data representation scheme will make the World Wide Web into an international repository for geotechnical information, available to the whole community. Such developments will make possible the use of case-based reasoning systems using global databases of case histories. Such a design system would automatically draw on new experience every time a new case history was added to the web.

REFERENCES

AGS. 2004. Electronic Transfer of Geotechnical and Geoenvironmental Data, Edition 3.1, Association of Geotechnical and Geoenvironmental Specialists, Beckenham, Kent (with addendum May 2005) (available from http://www.ags.org.uk).

Bray, C.J, Chandler, R.J., Walthall, S., Hoit, M. and Lefchik, T. 2008. Extending the Geotechnical Dictionary: Best Practice for Customizing the International Framework for Geotechnical Data, Proc. GeoCongress 2008 (eds. A.N. Alshawabkeh, K.R. Reddy & M.V. Khire), Geotechnical Special Publication No. 179, Reston: American Society of Civil Engineers, pp. 549-556.

Chandler, R.J., Quinn, P.M., Beaumont, A.J., Evans, D.J. and Toll, D.G. 2006. Combining the Power of AGS and XML: AGSML the Data Format for the Future, Proc. GeoCongress 2006: Geotechnical Engineering in the Information Technology Age, Atlanta, USA, Reston: American Society of Civil Engineers, pp. 112-117.

CUR. 2000. Description of the GEF Language, Civieltechnisch Centrum Uitvoering Research en Regelgeving, Gouda (available from http://www.geffiles.nl/).

DIGGS. 2006. Data Interchange for Geotechnical and Geoenvironmental Specialists, Technical Manual Draft Version 0.8 July 24, 2006.

DIGGS. 2008. Data Interchange for Geotechnical and Geoenvironmental Specialist, DIGGSML v0.10 Data Dictionary (http://www.diggsml.com).

Fyson, M. and Toll, D.G. 2008. Case-Based System for Slope Design, Computers and Geotechnics, Vol. 35. No. 3, pp. 468-478.

Greenwood, J.R. 1988. Developments in computerised ground investigation data, Ground Engineering, Sept. 1988, pp. 36-41.

Hatipoglu, B. 2003. The Development of an Integrated and Intelligent Design Environment for the Investigation of Slope Stability Problems, PhD Thesis, Istanbul Technical University.

GML. 2004. Geographic information – Geography Markup Language (GML) Version 3.1.1, Report OGC 03-105r1, Open Geospatial Consortium Inc. (http://www.opengeospatial.org/standards/gml).

Styler, M., Hoit, M., and McVay, M. 2007. Deep Foundation Data Capabilities of the Data Interchange for Geotechnical and Geoenvironmental Specialists (DIGGS) Mark-up Language, Electronic Journal of Geotechnical Engineering, http://www.ejge.com/2007/Ppr0772/Ppr0772.pdf .

SensorML. 2005. Sensor Model Language (SensorML) Implementation Specification, Report OGC 05-086, Open Geospatial Consortium Inc. (http://www.opengeospatial.org/).

Threadgold, L. and Hutchison, R.J. 1992. The Electronic Transfer of Geotechnical Data from Ground Investigations, Colloque International Géotechnique et Informatique, Paris, Presse de l'Ecole National des Ponts et Chaussées, pp. 749-756.

Toll, D.G. 1996. Data Management: Discussion Reviewer's Report for Session II, in Advances in Site Investigation Practice (ed. C. Craig), Thomas Telford, London, pp. 250-256.

Toll, D.G. 2007a. Geo-Engineering Data: Representation and Standardisation, Electronic Journal of Geotechnical Engineering, http://www.ejge.com/2007/Ppr0699/Ppr0699.htm .

Toll, D.G. 2007b. Representing Slopes in XML, Proc. Specialized Session S02 of 11th Congress of the International Society for Rock Mechanics, Lisbon, pp. 13-17 (on CD) (http://www.dur.ac.uk/geo-engineering/jtc2/ISRM2007/ISRM_Specialised_Session_S02.pdf).

Toll, D.G. 2008a. Representing Geo-Engineering Data from Instruments and Transducers, Proc. GeoCongress 2008 (eds. A.N. Alshawabkeh, K.R. Reddy & M.V. Khire), Geotechnical Special Publication No. 179, Reston: American Society of Civil Engineers, pp. 581-588.

Toll, D.G. 2008b. International Data Exchange: The Future for Geo-engineering, Proc. 12th International Conference of International Association for Computer Methods and Advances in Geomechanics (IACMAG), Goa, India, pp. 1958-1965.

Toll, D.G. and Cubitt, A.C. 2003. Representing Geotechnical Entities on the World Wide Web, Advances in Engineering Software, Vol. 34, 11-12, pp 729-736.

Toll, D.G. and Shields, R. 2003. A Web-based Data Format for Ground Investigation Data, Electronic Journal of Geotechnical Engineering, http://www.ejge.com/2003/Ppr0321/Abs0321.htm .

Weaver, S.D., Lefchik, T., Hoit, M. and Beach, K. 2008. Geoenvironmental and Geotechnical Data Exchange: Setting the Standard, Proc. GeoCongress 2008 (eds. A.N. Alshawabkeh, K.R. Reddy & M.V. Khire), Geotechnical Special Publication No. 179, Reston: American Society of Civil Engineers, pp. 557-564.

Whetton, J. 2009. Slope Representation on the World Wide Web. MEng Final Year Project Report, School of Engineering, Durham University.