6th Annual Technical Forum

GEOHAZARDS IN TRANSPORTATION IN THE APPALACHIAN REGION

August 2006

PRACTICAL CONSIDERATIONS IN WORKING WITH DATA INTERCHANGE STANDARDS

Salvatore Caronna, gINT Software, Windsor, CA

OVERVIEW

- Data interchange standards have only one purpose and that is to house data for shipment in an accurate and readily accessible manner.
- Data interchange standards are not work specifications nor are they database designs.
- Numerous special considerations must be addressed in dealing with any interchange standard.
- The DIGGS standard presents its own unique issues.

WORK SPECIFICATIONS

- The interchange standard is not a substitute for a work specification.
- Care must be taken to ensure unambiguous and efficient specification of how data are to be submitted.

WORK SPECIFICATIONS

- Specifying Code Values
  - DIGGS has lists of recommended codes for many fields.
  - These lists can be used directly, expanded, or custom lists created.
  - Whichever are used make sure they are part of the work specification.

WORK SPECIFICATIONS

- Description Classes
  - DIGGS allows multiple sets of soil and rock descriptions within each borehole.
  - The work specification needs to clearly state which type or types shall be used.

WORK SPECIFICATIONS

- Use the standard as much as possible
  - DIGGS allows expansion of the schema.
  - Only use this facility when absolutely necessary.
  - There are features in the standard that can house many types of related data:
    - MonitoringPoint: Piezometer, slope inclinometer, extensometer, settlement gauges, strain gauges, etc.
    - DrivenPenetrationTest: SPT, Dynamic Cone, Texas Cone, California Sampler, etc.
  - Make the storage location of these data explicit in the work specification.

WORK SPECIFICATIONS

- Don’t Overload Standard Fields
  - If you do have unique information not covered by the standard, expand the standard with the addition of a field, fields, or table.
  - Don’t use non-specific fields in the standard such as remarks for such information.

WORK SPECIFICATIONS

- Summary
  - The interchange standard is a vital part of the overall work specification.
  - Use the standard properly.
  - Give sufficient guidance to the data producer so that you get back the data you want in the way you want it.

DATABASE DESIGN

- Database design needs to reflect the requirements of the end user.
- Designing to mimic an interchange standard can lead to inefficient and awkward database structures.
- The design may need to be altered to ensure compatibility with an interchange standard but this is not the primary consideration.

DATABASE DESIGN

- Data Interchange Standards are not Working Databases Structures
  - Much of the DIGGS standard can be copied directly to a working database design.
  - Many parts of the standard are optimized for the purposes of data transfer and ease of expandability, not for usage in a working database.
  - The database does need to be designed to ensure basic structural compatibility with the standard.

DATABASE DESIGN

- DIGGS MonitoringPoint Object
  - This is designed to accept data from any depth-date/time monitoring such as piezometers, slope inclinometers, etc.
  - It is robust and can easily accept new types of monitoring without expanding the DIGGS schema.
  - Very difficult to deal with in a working database. Few tests require all the supplied fields and all tests are in one table.
  - Best to have dedicated tables for each monitoring type and to map to and from the MonitorPoint object.

DATABASE DESIGN

- Soil and Rock Descriptions
  - If descriptions are specified to be just single fields, can use the standard fields as-is:

Depth	Bottom	Description
0	6	Silty SAND: very loose, fine to medium, moist, green.
6	12	Silty SAND: loose, fine to medium, dry to moist, bluish red.
12	18	Silty SAND: medium dense, fine to medium, moist, brown.
18	24	Sandy CLAY: very stiff, moist, brown.
24	27	Sandy FAT CLAY: hard, wet, gray.

DATABASE DESIGN

- If you require a component description, then the interchange structure and a good working structure are vastly different:

Depth	Bottom	Main	Qualifier	Consistency	Grain Size	Moisture	Color
0	6	sand	silty	very loose	fine to medium	moist	green
6	12	sand	silty	loose	fine to medium	dry to moist	bluish red
12	18	sand	silty	medium dense	fine to medium	moist	brown
18	24	clay	sandy	very stiff		moist	brown
24	27	fat clay	sandy	hard		wet	gray

DATABASE DESIGN

- Here is how one might set up a working structure to mimic DIGGS.

Depth	Bottom	Component	Value
0	6	main	sand
0	6	qualifier	silty
0	6	consistency	very loose
0	6	grain size	fine to medium
0	6	moisture	moist
0	6	color	green
6	12	main	sand
6	12	qualifier	silty
6	12	consistency	loose
6	12	grain size	fine to medium
6	12	moisture	dry to moist
6	12	color	bluish red
12	18	main	sand
12	18	qualifier	silty
12	18	consistency	medium
12	18	grain size	fine to medium
12	18	moisture	moist
12	18	color	brown
18	24	main	clay
18	24	qualifier	sandy
18	24	consistency	very stiff
18	24	grain size	
18	24	moisture	moist
18	24	color	brown
24	27	main	fat clay
24	27	qualifier	sandy
24	27	consistency	hard
24	27	grain size	
24	27	moisture	wet
24	27	color	gray

DATABASE DESIGN

- How the Database Structure needs to change to conform to the Interchange Standard.
  - Table Key Fields
  - Table Relations
  - Data Types
  - Code Lists

DATABASE DESIGN

- Summary
  - Database design needs to reflect user requirements.
  - Database design must conform to basic structural elements of the interchange standard for compatibility.
  - Don’t let the tail (the interchange standard) wag the dog (the database design).

VALIDATION

- DIGGS files are self-validating
  - the schema can be checked for accuracy
  - data are checked against the specified data types
  - values in fields associated with code list can be validated
  - minimum and maximum values can be assigned to fields

VALIDATION

- There are no automated methods inherent in the standard for performing dependent validations.
  - RQD must be greater than or equal to Total Recovery.
  - If an SPT penetration is 1.5 feet, there must be three blows.
  - Specimen depths must be in the range of the corresponding parent sample depth range.
  - Layers within the same description classification must not overlap.
  - If the liquid limit is 35, a plastic limit of 52 is unreasonable.

VALIDATION

- With time, budget, and enough people, these dependent validation rules can be written.
- In the meantime, do not assume that because DIGGS is self-validating that the data are all reasonable.

DIGGS IS NOT HUMAN

- Some data interchange standards can be edited by real people using text editors or spreadsheets.
- DIGGS is not one of those standards.

DIGGS IS NOT HUMAN

```
- <subsurface>
  - <Hole gml:id="D6DD2E0C-7BFF-4ebc-83E4-4F69BC1D71A1">
      <gml:name codeSpace="http://www.ags.org.uk/id">TS150</gml:name>
    - <geometry>
      - <gml:LineString gml:id="72A638B0-731A-4474-BA1F-BC4CF48DC052" srsName="urn:ogc:crs:epsg:6.9:27700">
          <gml:pos dimension="3">97488.580 103170.658 54.894</gml:pos>
          <gml:pos dimension="3">97488.580 103170.658 54.894</gml:pos>
        </gml:LineString>
      </geometry>
    - <gml:engineeringCRSRef>
      - <gml:EngineeringCRS gml:id="43AF3014-9E65-4faf-B93C-E78A2A938559">
          <gml:srsName>TS150 CRS</gml:srsName>
        - <gml:usesCS>
          - <gml:LinearCS>
            - <gml:usesAxis>
              - <CoordinateSystemAxis gml:id="E1A43C6C-E9B4-4593-9D10-472C12F809E4" gml:uom="units.xml#m">
                  <gml:axisName>Depth</gml:axisName>
                  <axisDirection xlink:href="72A638B0-731A-4474-BA1F-BC4CF48DC052" />
                </CoordinateSystemAxis>
              </gml:usesAxis>
            </gml:LinearCS>
          </gml:usesCS>
        - <gml:usesEngineeringDatum>
          - <EngineeringDatum gml:id="374AD02E-4CEE-4d16-ADC3-260B799DAD69">
            - <origin>
              - <gml:Point srsName="urn:ogc:def:crs:epsg:6.9:27700">
                  <gml:pos>97488.580 103170.658 54.894</gml:pos>
                </gml:Point>
              </origin>
            </EngineeringDatum>
          </gml:usesEngineeringDatum>
        </gml:EngineeringCRS>
      </gml:engineeringCRSRef>
      <type codespace="AGSHoleTypeCodeList.xml">INST</type>
    - <remark>
        <comment>Tunnel progress 0</comment>
        <dateTime>2001-05-02T12:00:00</dateTime>
      </remark>
    - <remark>
        <comment>Tunnel progress 0</comment>
        <dateTime>2001-05-07T12:00:00</dateTime>
      </remark>
    - <remark>
        <comment>Tunnel progress 0</comment>
        <dateTime>2001-05-11T12:00:00</dateTime>
      </remark>
    - <remark>
        <comment>Tunnel progress 0</comment>
```

DIGGS IS NOT HUMAN

- DIGGS is based on XML (Extensible Markup Language) and GML (Geographic Markup Language) which makes self-describing and self-validating.
- It also makes DIGGS understandable to some degree by many GIS software packages without special translation.
- Finally, a host of tools are available for validation, coordinate transforms, and unit conversions.
- The downside is that no human can hope to properly create or edit a DIGGS file in any reasonable period of time.

DIGGS IS NOT HUMAN

- Specialized software will be needed to read and write DIGGS files.
- The experience in the UK with the AGS is that this is probably a good thing.

SUMMARY

- The holy grail of easily interchangeable data is within our reach.
- This brave new world will require significant changes in the way we deal with data and software.
- New structures need to be put in place and work must be put into the exchange process to ensure that it be routine and accurate.