TASK ORDER 20 (PR #NB860000-A-06887 & A-07893)
GROUND-MOTION AND BUILDING-PERFORMANCE DATABASE FOR THE
2010 CHILE EARTHQUAKE

Task 20.4 System Design
June 25, 2011

Database Design and Development Group
Ann Christine Catlin
Sudheera Fernando
Nabeel Yousef
Ruwan Gamage

Civil Engineering Group
Santiago Pujol
Brian Richter
Yan Linwood

Note: Links are only available to group "atc93" members https://nees.org/groups/atc93/members who are logged in to nees.org.

The Chile Earthquake Database is under development using the NEEShub cyber infrastructure. A MySQL database at stage.neeshub.org houses the data collected in Chile (by Chilean Investigators, and ASCE, ATC, EERI, LATBSDC, Penn State University, and Purdue University) after the 2010 Earthquake and, for reference, data from the 1985 Earthquake published by University of Illinois. The principal operational components of the database are

(1) a spreadsheet parser for data ingestion,
(2) a photo processor for creation of metadata and repository collections,
(3) a data viewer for database browse, search, explore and download, and
(4) a photo gallery viewer with keyword search

1. The Chile Earthquake Database and Data Contribution

    - Attributes of the earthquake performance data: The list of fields (or attributes) that have been identified as relevant parameters to organize and classify field observations on seismic performance has been approved and is available at
      http://nees.org/site/wiki/282/termsv.6.pdf

    - Attribute spreadsheets: Spreadsheet templates containing the attributes are used to define the schema for the database. There are 5 spreadsheet templates for the Chile earthquake database:
      Buildings Data Template:
      http://nees.org/site/wiki/282/buildings-template.csv
      Earthquake Damage Data Template:
      http://nees.org/site/wiki/282/earthquakedamage-template.csv
      Earthquake Information template:
      http://nees.org/site/wiki/282/earthquakeinfo-template.csv

Station Data Template:
http://nees.org/site/wiki/282/stations-template.csv
Data Sources Template:
http://nees.org/site/wiki/282/datasources-template.csv

- <u>Parser to convert attribute spreadsheet to database schema</u>: The spreadsheets define the database schema and a customized parser processes the spreadsheets into database tables. The ER diagram is available as a Visio file http://nees.org/site/wiki/282/chile-nisterdiagramvsd.vsd. The schema design is extensible, allowing data from other earthquakes to be added to the database. Parser code is located at nees.org in /data/database/parser.

.

- <u>Earthquake performance data</u>: Data are contributed to the database using the attribute spreadsheets templates, and the spreadsheets containing Chile building, earthquake damage, earthquake info and stations data are parsed into the database. Example spreadsheets containing data for the database are here:
Buildings Data:
http://nees.org/site/wiki/282/chile_nist_building_06_23_2011.csv
Earthquake Damage Data:
http://nees.org/site/wiki/282/chile_nist_earthquake_06_23_2011.csv
 Earthquake Information:
http://nees.org/site/wiki/282/chile_nist_earthquake_info_06_23_2011.csv
Station Data:
http://nees.org/site/wiki/282/chile_nist_stations_06_24_2011.csv
Data Sources:
http://nees.org/site/wiki/282/chile_nist_sources_06_24_2011.csv

- <u>Database ingestion over time</u>: The database will continue to be updated using spreadsheets -- this is the standard mechanism for updating HUB databases. The Civil Engineering Group submits spreadsheets to the database, and the new data rows are processed into the database by the customized parser. The parser will manage the assignment of unique identifiers to buildings, earthquakes, stations, and photographers – and will ensure the accurate ingestion of data and linkage to existing data.

- <u>Photographs and photo tagging</u>: Photos (JPG format only) are tagged by inserting the photographer name and keywords to the photo header. Photos are then uploaded using the NEEShub data transfer program SynchroNEES to the "group space" https://nees.org/groups/atc93/dropbox. A directory hierarchy has been established in the dropbox consisting of
    earthquake name-timestamp => building id => photographer id
where the ids for building and photographer are assigned by the spreadsheet parser.

- <u>Keywords</u>: The list of keywords for rapid tagging of photos has been approved and is available at http://nees.org/site/wiki/282/keywords.v.3.pdf

- <u>Photo processing to create metadata and build the photo repository</u>: A customized JPG processor operates on the dropbox hierarchy. A spreadsheet with photographer information is present in the dropbox. The processor operates as follows:

  For each photo in the dropbox
      Copy photo to repository collection location with photographer name pre-pended
      Validate and extract photo header information
      Identify building ID from the database; insert building name and coordinates in header
      Create metadata entry in database to identify photographer, keywords, building and
          repository location of the file
      Create thumbnails for photo gallery viewing

- <u>Other earthquake documentation</u>: Diagrams, reports, drawings, and other building documentation are processed from the upload location (the dropbox) into the document repository and the repository path is inserted into the database. Here is the current list of considered documentation types and their relationship to buildings, earthquakes and sources:

  general_drawings
      <building ID>
          <source ID>
              Varying file types
  typ_floor_drawing
      <building ID>
          <source ID>
              Varying file types
  general_files
      <building ID>
          <source ID>
              Varying file types
  analysis_files
      <building ID>
          <source ID>
              Varying file types (accompanied by a readme file)
  ground_motion_records
      Maule.Chile.2010
          <station ID>
              <source ID>
                  Varying file types
      Valparaiso.Chile.1985
          <station ID>
              <source ID>
                  Varying file types

2. Data Viewing, Exploration and Export

- <u>Data Views:</u> A collection of customized "spreadsheet" data views is available for browsing the database. The data views display the building, earthquake and photograph data that has been processed into the database.

Data in the views can be

1) Searched by column or across the spreadsheet,
2) Sorted by column for both numeric and textual data,
3) Filtered from one spreadsheet 'data view' to another 'data view' through column links, e.g., column selections for a building has a link to an earthquake damage view or a pop-up earthquake damage comparison,
4) Filtered by pop-up dialog box selections,
5) Linked to Google maps for building locations and metadata identification
6) Customized for user specified data columns

These views are now available:
Essential Earthquake Data
https://stage.neeshub.org/dataview_dev/spreadsheet/chile_nist_essential_items/

Common Earthquake Data
https://stage.neeshub.org/dataview_dev/spreadsheet/chile_nist_items_we_will_most_likely_have/

Full Data View of All Fields Available in Earthquake Database
https://stage.neeshub.org/dataview_dev/spreadsheet/chile_nist_all/

Stations Information
https://stage.neeshub.org/dataview_dev/spreadsheet/chile_nist_stations/

Customizable Data View
https://stage.neeshub.org/dataview_dev/spreadsheet/chile_nist_custom/

Data Views are used to verify the correctness of the data and data linkage, and also for assessing the completeness of the database entries. Data views for database validation are here:

Building Data:
https://stage.neeshub.org/dataview_dev/spreadsheet/chile_nist_building_default/

Earthquake Damage Data:
https://stage.neeshub.org/dataview_dev/spreadsheet/chile_nist_earthquake_default/

Station Data:
https://stage.neeshub.org/dataview_dev/spreadsheet/chile_nist_station_default/

Sources Data:
http://stage.neeshub.org/dataview_dev/spreadsheet/chile_nist_source_default/

Photo Keywords
https://stage.neeshub.org/dataview_dev/spreadsheet/chile_nist_photo_keywords_default/

Photo Metadata:
http://stage.neeshub.org/dataview_dev/spreadsheet/chile_nist_photos_default/

- Documents, reports, diagrams, images and photo galleries are associated with buildings through the 'data view', and have thumbnail representations with links to their repository locations.

- Google maps are associated with building coordinates, building damage, and links to information in the database about the building.

- Data can be exported from each data view as a spreadsheet. Additional data views can be created and tailored to export requirements. Export in ESRI format is being developed.

- The component for customization of data views is under development.

3. Photo Gallery Search and Exploration

A customized photo gallery viewer is under development for search and discovery across the entire repository of photos from Chile.

- Searches are performed for keywords, photographers and other metadata (such as building name)

- Maps are included with the photographs and geospatial queries will be supported, users can switch between map and photo view. Map view automatically zooms and pans to show all markers for buildings, current selection is animated.

- Keywords are displayed as a keyword cloud.

- Photo metadata is displayed along with the photo, including the photographer/source name, building name/year, description, keywords.

- Thumbnails to select photo, with start/stop filmstrip

A initial view of the interface is here http://stage.neeshub.org/photo_gallery/

4. Data Security

Security is managed by the NEEShub cyber infrastructure.

5. Delivery of the Chile Database System to NIST

HUB database systems are transferred from one hub to another as follows 1) transfer the Joomla components, 2) transfer the MySQL database and configuration files, 3) create the HUB resource page for accessing the database.