Standards for Data and Model Reporting in the Nebraska GeoCloud

Version 1.0
July 1, 2020

Jesse T. Korus[1], Kathleen Cameron[1]

[1]Conservation and Survey Division–School of Natural Resources
Institute of Agriculture and Natural Resources
University of Nebraska-Lincoln

Contents

INTRODUCTION

This document contains standards for hydrogeological data and models in Nebraska GeoCloud (NGC). NGC is a web-based digital platform for geophysical, geological, and groundwater data and models. The purpose of NGC is to archive Nebraska's vast volume of data and make it available and accessible to both model builders and model users. The NGC consists of databases, web servers, and web interfaces designed for data storage, sharing, and distribution.

This document accompanies the document titled **Guidelines for Airborne Electromagnetic (AEM) Surveys, Data Integration, and Hydrostratigraphic Modeling in Nebraska**. Model builders contributing data to NGC should review the Guidelines to ensure data and models meet expectations of project partners.

NGC consists of one interface for Projects and another interface for Data. The Projects interface is designed to store and share project files (Fig. 1). Projects may include software files, reports, and other information related to a project. The Projects interface also contains GeoScene3D startup templates that can be used to start a new project anywhere in a given area without the need to assemble, format, and import data to a software project. After a project has been started, the interface can be used to store and share project files, or it can be used as the final repository for completed projects. It hosts a generic database for storing nearly any type of file, including .dat, .dbf, .doc, .docx, .gsmod, .id, .map, .mdb, .pdf, .prj, .rar, .sbn, .sbx, .shp, .shx, .tab, and .zip (unsupported file types can be compressed to a zip file and uploaded to the database).

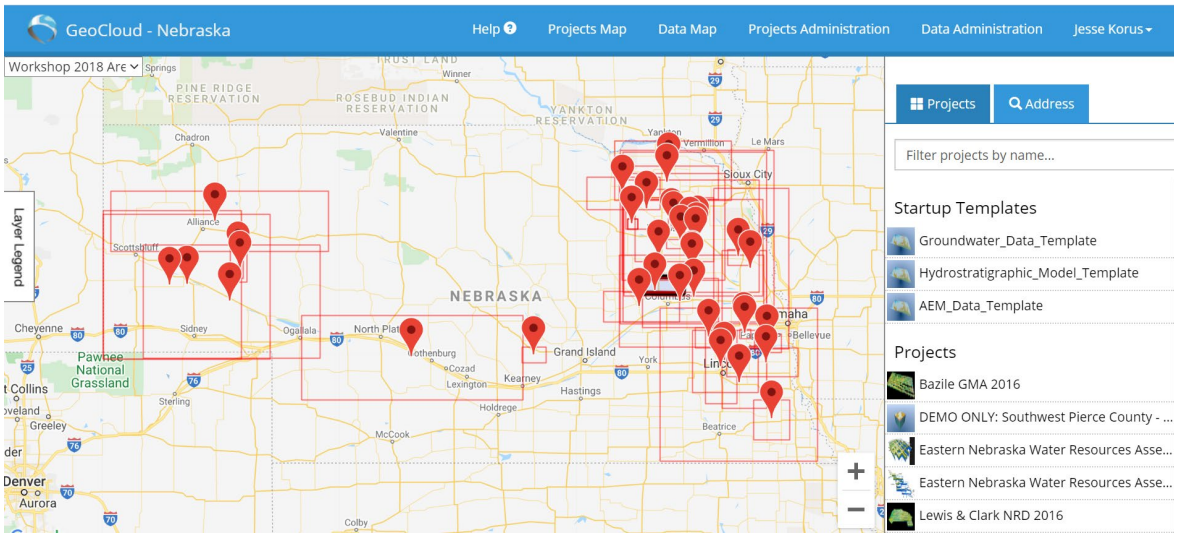

Figure 1. User interface for Nebraska GeoCloud Projects Map. The red boxes with central pins show the bounding boxes of various projects in Nebraska.

The Data interface of NGC is built upon structured databases that support the upload and download of certain data and models used in typical hydrogeological studies. Users can access the data contained in this part of NGC via the GeoScene3D data portal. These data can also be viewed on an interactive web map (Fig. 2) and they are accessible via a web map service in GIS programs

(e.g. ArcGIS, QGIS). The databases support a variety of data types: AEM resistivity-depth models (.xyz), 2D grids (.grd, .asc, .tif), 3D grids (.grd3), borehole lithology logs and well information (.mdb/.accdb), borehole geophysical logs (.mdb/.accdb), point data (.csv), ESRI shapefiles (.shp, .shx, .dbf), and color scale files (.clr).

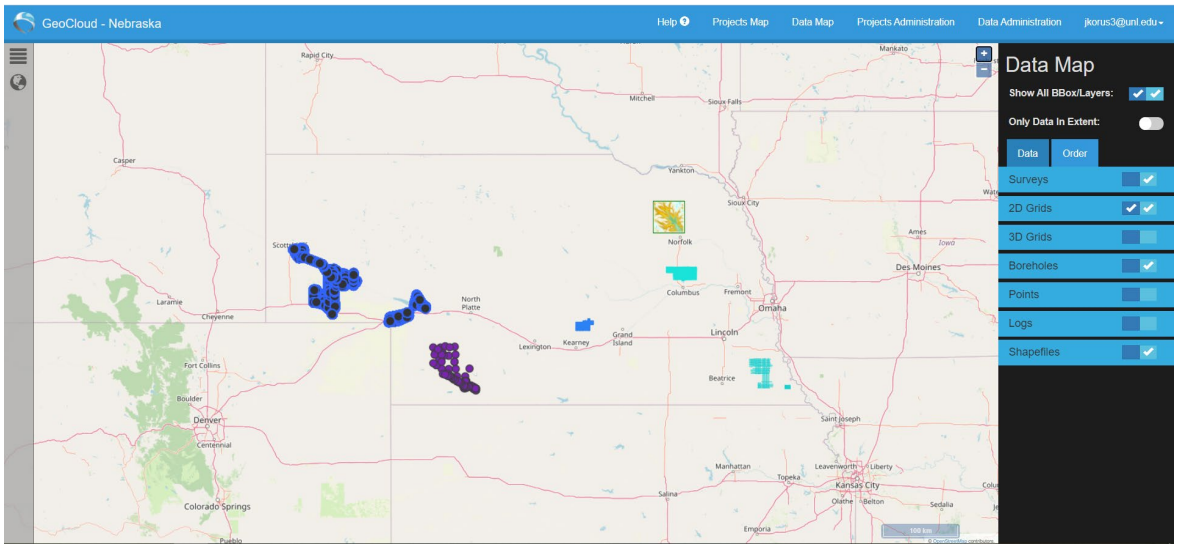

Figure 2. User interface for Nebraska GeoCloud Data Map. A selection of various data types from across Nebraska are shown on the map.

The standards in this document are a set of strict requirements for uploading data, reports, models, and metadata to NGC. Model builders must follow these requirements when submitting files to NGC for permanent archiving and sharing.

COORDINATE SYSTEM AND UNITS OF MEASUREMENT

All data served on the Nebraska GeoCloud are projected to the Nebraska State Plane meters coordinate system referenced to the North American Datum of 1983 [European Petroleum Survey Group (EPSG) 32104]. The Nebraska State Plane coordinate system was chosen for all data served through the Nebraska GeoCloud because it is a single system for the entire state, unlike Universal Transverse Mercator (UTM), which has three zones in Nebraska. The State Plane system is the standard for Nebraska public agencies as required by the Nebraska Information Technology Commission (https://nitc.nebraska.gov/standards/3-Chapter.pdf). Furthermore, the State Plane system is a projected coordinate system, so it minimizes scale distortions (ESRI 2019). Prior to inclusion to NGC, spatial coordinates of supporting data sets must be re-projected to the Nebraska State Plane meters (EPSG 32104) coordinate system.

The metric system is used for the official units of measurement for all data served on the NGC. The use of metric units makes processing and computation easier for geophysicists and geologists,

improving the overall quality and reliability of results. It also makes the NGC compliant with international scientific standards.

Other coordinate systems [i.e. State Plane (feet), UTM (meter)] or standard units (feet, miles, etc.) can be uploaded to NGC as part of the AEM dataset if requested by the client. However, these coordinate systems and units will be included only as additional data fields: State Plane meters and metric units will still be required.

STANDARDS FOR PROJECT REPORTING

Projects are hydrogeologic investigations, research, or consulting work that involve a specific area or site, typically completed within a specific timeframe. Projects may involve any variety of data, software, and deliverables. Hydrogeologic Project files are uploaded to Projects Administration and are available for download in the Projects Map. On-going projects can be uploaded to NGC for file sharing during the project. User access can be defined to restrict access until the project is complete. After completion, the project files and final reports can be upload to NGC for long-term archiving and public access.

Metadata for Hydrogeologic Project Files

Hydrogeologic data integration and visualization projects should be uploaded to the NGC via the Projects Administration interface at https://geocloud.live. Projects can contain a variety of data and information, including software files and supporting datasets. GeoScene3D project files are supported. However, if other software was used, the files can be compressed to a .zip file and uploaded to NGC. It is not necessary that a project contains software files. The specific contents of each project will depend on client-consultant agreements and deliverables.

Metadata are required for projects in NGC so that the data sources, methods, and results can be identified. Projects involving AEM surveys through 2019 have been uploaded to NGC and have been given a project name. Future projects shall be given names as follows:

- A descriptive place name or client(s) name;
- The year the project was completed

Example: Lewis & Clark NRD 2016

The description below the name shall identify the consulting company or institution responsible for the project, the type of project (AEM survey, hydrogeologic assessment, water supply study, water quality study, etc.), a short description of the project contents, and any related identifying information.

The project shall be given geotags, including a x, y coordinate for pin position and rectangular extent. Projection system of coordinates should be Nebraska State Plane meters coordinate system referenced to the North American Datum of 1983 (EPSG 32104).

The access levels of users shall be defined to specify which users can read and download the project. The access levels will depend on client-consultant agreements. Incomplete projects and working files can be uploaded to NGC to allow file sharing. In such cases, it may be necessary to restrict user access to only those parties involved in the study. After project completion, user access can be changed to broaden the audience if desired.

File attachments should include software project files, raw data suites for AEM surveys, and a report or project documentation. The report or documentation should describe the methods, sources of data, and processing procedures used in the project. The **Project Documentation Template** can be used in lieu of a report.

If a project involves an AEM survey, the suite of raw AEM data files (.gex, .skb, .lin, sps etc.) will be compressed to a .zip file and uploaded to Projects Administration at http://geocloud.live. If an AEM system other than SkyTEM is used, the ASCII files with an associated Readme file should be uploaded to the project. The raw data files shall be accompanied by a report that details the AEM survey specifications. This report should include information about the AEM system and data acquisition, as outlined in the guidelines above. Example reports from previous surveys can be downloaded from NGC.

Standards for Airborne Electromagnetic (AEM) Data Reporting

Both raw data files and inversion results from AEM surveys shall be delivered to NGC following the completion of each project. Adhering to these standards will ensure that data is accessible and usable by the broad community of groundwater professionals beyond the life of the project. Unless otherwise agreed upon, it shall be the responsibility of the consultant responsible for overseeing the survey to upload raw data, inversion results, and metadata to NGC.

Metadata is required for both the AEM survey *and* any resistivity-depth models that are created from the survey data. Although most AEM surveys in Nebraska currently have only one inversion product, it is possible that some AEM surveys will have several models as new inversion results are created in future studies.

Metadata for AEM Surveys

Each AEM survey shall also be added to Data Administration on NGC at https://test.geocloud.live. The AEM survey shall be accompanied by metadata as specified in the **NGC AEM Metadata Standard**. The AEM survey will serve as a holding spot for all resistivity-depth (inversion) models created for that survey.

Metadata for Resistivity-Depth (Inversion) Models

The inversion results shall be uploaded to Data Administration at https://test.geocloud.live/ under the appropriate AEM survey. The inversion results will be uploaded as 1D resistivity-depth models with accompanying metadata. The NGC supports upload of .xyz (Geosoft) files for inversion results as well as ASCII data. The inversion file shall be accompanied by metadata specified in the **NGC AEM Metadata Standard.**

Versioning

It is important to keep in mind that each AEM survey will have only one entry in NGC, but each survey may contain many inversion files. Because of this one-to-many relationship, it is necessary that each inversion file uploaded to NGC be associated with the correct AEM survey and that each be given a date and version number. As part of the **NGC AEM Metadata Standard**, each new inversion dataset associated with a single AEM survey will be given a separate, consecutive integer. An inversion log shall be provided if more than one inversion is delivered. The log will have a separate line for each inversion and will list the inversion parameters used for that particular inversion.

STANDARDS FOR HYDROGEOLOGIC DATA REPORTING

Many of the standards for data reporting, including test hole logs, driller's logs, water levels, and water quality, are implemented through other programs administered by various agencies, as described above. Because these standards are already in place, the standards dealt with here are for documentation of data sources and any modifications to the original datasets. In some cases, such as the addition of driller's logs to NGC, certain modifications of the data are necessary in order to make it useful for hydrogeologic studies. These modifications should follow the standards established herein such that data is consistent and seamless across the state.

The following sections detail metadata requirements for data files. Data files for boreholes, point data, and shapefiles will be uploaded to Data Administration and are available for download in the Data Map or via a web map service (WMS).

Metadata for Boreholes

CSD test hole data, including lithology, stratigraphy, and geophysical logs, shall be uploaded to NGC and maintained by designated CSD personnel.

Driller's logs from DNR registered water wells may be added to NGC by model builders for specific projects. The driller's descriptions shall be aggregated into the simplified terms contained in the **Standardized Lithology Keywords** (SLK) file. To facilitate a standard approach to keyword aggregation, the CSD has developed the **Lithology Keyword Automation Tool**, or LithoKAT, a word-pattern-recognition tool that quickly converts unstandardized lithologic descriptions obtained from driller's logs into a uniform set of lithology keywords (Korus et al. 2018). Keywords can be mapped to a lithology keyword file, allowing the logs to be imported into various software programs for visualization. This tool is contained within a Microsoft Excel™ spreadsheet as a series of logical statements and lookup, search, and reference functions. The

spreadsheet operates on the premise that the first lithologic term in a string of terms likely represents the primary lithology. Unstandardized terms like “sticky grey mud” can be replaced with “clay” or another standard term. Sand and gravel, being aquifer materials, can be subdivided into additional classes on the basis of grain size descriptions, if available. Search terms and keywords are user-defined, allowing the program to be customized to any hydrogeologic setting or site-specific AEM resistivity-lithology relationship. Please refer to the LithoKAT program and documentation for more information.

LithoKAT is not intended as a replacement for a geologist’s judgement. The tool should be accompanied by a program of quality control. Manual edits to the keywords may be necessary if problems are encountered.

The source of location information for registered water wells is a critical piece of information that must be supplied in each file. Coordinates in the DNR database are either from GPS readings supplied by the driller or calculated from location information on the registration form. For the latter, this location information is either measurements made on the ground in reference to section boundaries or legal description (township, range, section, subsection). Therefore, the borehole collar file uploaded to NGC shall contain a field specifying the source of x, y coordinates:

- GPS
- Measured
- Legal description

Bear in mind that coordinates derived from legal descriptions may be highly inaccurate. Errors in the elevation and location of a borehole can lead to erroneous interpretations of geologic units, screened intervals, and water levels.

At this time, NGC supports the upload of Microsoft Access databases for borehole data. The database must have a collar table specifying borehole identification (ID) and location information, and a related table containing the matching borehole ID and depth-dependent data (either lithology keywords or geophysical measurements). Borehole data added to NGC will follow these standards:

- **NGC Borehole Metadata Standards**;
- **NGC Geophysical Logs Metadata Standards**.

Metadata for Point Data

Point data can be added to NGC for a variety of data types. The most common types will be groundwater levels and water chemistry data. Normally, these data will be derived from the USGS or CSD databases described above. Metadata should clearly specify the source of the data and any modifications that were made. However, in some circumstances where data that are not from the state or federal databases are used for a study, it may be necessary to upload this data to NGC. In these cases, metadata should clearly state the source of the data and the quality control measures used to verify accuracy.

Point data for multiple other types of point data (samples, quantities, etc.) can be added to NGC. Regardless of the data type, all uploads shall be accompanied by metadata following the **NGC Point Data Metadata Standards.**

Metadata for Shapefiles

A wide variety of vector data can be displayed in GIS programs or 3D visualization software using ESRI shapefiles. Shapefiles can display interpretive products such as geologic maps, aquifer boundaries, and other important hydrogeological information. Shapefiles from CSD or USGS studies and published maps will be uploaded by those agencies. However, model builders will likely generate new products and upload them to NGC. In order for the end user to understand the product, the data and methods used to produce the shapefile shall be described in a detailed report. This report, along with other identifying information, shall be referenced in the **NGC Shapefile Metadata Standards.**

STANDARDS FOR HYDROSTRATIGRAPHIC MODEL REPORTING

As new projects and hydrogeological investigations continue into the future, it is expected that model builders will contribute to a growing population of 2D and 3D grids in the NGC. In order to facilitate comparison between projects and integration of multiple results, it is necessary that data uploaded to NGC follow a strict set of standards.

Metadata for 2D Grids

Geological surfaces are represented numerically by 2D grids. Model builders shall provide appropriate metadata following the **NGC 2DGrid Metadata Standard**. In addition, every 2D grid should reference a hydrogeological study report or model documentation file so that the data, methods, and results can be understood beyond the life of the project.

Metadata for 3D Grids

Physical properties within geologic volumes are represented numerically by 3D grids. Model builders shall provide appropriate metadata following the **NGC 3DGrid Metadata Standard**. In addition, every 3D grid should reference a hydrogeological study report or model documentation file so that the data, methods, and results can be understood beyond the life of the project.

ACKNOWLEDGEMENTS

The authors wish to thank the Nebraska GeoCloud project sponsors: Central Platte, Lewis and Clark, Lower Elkhorn, Lower Loup, Lower Platte North, Lower Platte South, Nemaha, Papio-Missouri River, Twin Platte, and Upper Elkhorn Natural Resources Districts. Technical review of this document was provided by Jared Abraham, James Cannia, and Ted Asch of AquaGeoFrameworks (AGF), LLC, under contract with the project sponsors.

REFERENCES

ESRI, 2019, ArcMap version 10.5 software documentation, https://desktop.arcgis.com/en/documentation/, Accessed June 22, 2020

KORUS, J.T., CAMERON, K., HOBZA, C.M., JENSEN, N.-P., RICO, D., AND MUNOZ-ARRIOLA, F., 2018, Integrating AEM and borehole data for regional hydrogeologic synthesis: Tools and examples from Nebraska, USA: 7th International Workshop on Airborne Electromagnetics.