Field Data Acquisition Using Ganfeld-NSDNR

D. F. Keppie

Abstract

Ganfeld software, primarily developed at the Geological Survey of Canada, has been refactored and revised for greater flexibility and default logic by the Geological Services Division, Nova Scotia Department of Natural Resources (NSDNR). This new version of the Ganfeld codebase is called Ganfeld-NSDNR to distinguish it from Ganfeld. Ganfeld-NSDNR has been designed so that default functionality, data tables, collected data files, and graphical interfaces to collected data files may be easily configured, changed, updated or replaced for specific uses. Key features of Ganfeld-NSDNR include: (1) the ability to dynamically create material unit maps, outcrop maps, and lineament maps, (2) the ability to create and manage station notes, diagrams, and photos, (3) the ability to collect bedrock, surficial, and coastal geology data in a single project, (4) the ability to link two structural measurements, (5) a quicker interface for adding new data either within a station collection phase or when revisiting a previous station. Ideally, with the new design, Ganfeld-NSDNR can fulfill two roles. First, it can be deployed as the software for field data acquisition across most of the geological projects at NSDNR. Second, it can be developed and refined in order to be a prototype for the next generation of field data acquisition software which will likely be developed in the coming years.

Introduction

Field data acquisition using handheld computing devices has many potential advantages over traditional methods (Broderic, 2004). One possible advantage is that maps, field books, references, global positioning systems (GPS) and cameras could be integrated into a small portable environment for use by field geologists. A second advantage is that the data collected in the field are entered directly into an electronic form within a geographic information system (GIS) potentially reducing the time from initial capture of data to publication.

The potential advantages are sufficiently attractive that many geological surveys, research groups and individuals have attempted to transition to electronic field data acquisition systems with various degrees of success (Buller, 2004). The primary challenge that must be addressed by any software implementation is to balance the flexibility needed during the data acquisition phase by different geologists, projects, and research groups, with the consistency and structure needed for data to be easily incorporated into multiproject databases and geographic information systems for eventual publication or external use. These issues have been reviewed extensively (Broderic, 2004; Shimamura *et al.*, 2008), and so are not repeated here.

The purpose of this document is to describe a recent attempt by the Geological Services Division, (NSDNR) to continue the development of the Ganfeld software for field data acquisition to further improve its balance of flexibility and consistency in active mapping projects in Nova Scotia. At NSDNR, three distinct versions of Ganfeld had emerged to support bedrock, surficial and coastal mapping,. In order to streamline the technical support required to collect multiple categories of geological data, as well as to facilitate continued updates and changes to the graphical interface, it was decided to refactor the Ganfeld codebase to explicitly support these issues. Other changes were also made as discussed below.

Keppie, D. F. 2011: *in* Mineral Resources Branch, Report of Activities 2010; Nova Scotia Department of Natural Resources, Report ME 2011-1, p. 53-57.

Development Context

In Canada, the Geological Survey of Canada (GSC) has spearheaded the development and deployment of field data acquisition software within government surveys at both the federal and provincial levels (Broderic, 2004; Buller, 2004). In the 1980s and 1990s, the primary software supported, developed and deployed was FieldLog, which was based on object-relational programming and implementing a fully-conceived ontology for field data collection (Broderic, 2004). The types of data realized in the FieldLog ontology were cartographic, geospatial, geological, and metadata (Broderic, 2004). FieldLog was highly flexible and gave the individual geologist significant control for designing their own data structures and data fields. FieldLog appears to have been used quite widely, but has since fallen out of favour

In the 2000s, the primary software developed, supported and deployed at the GSC was Ganfeld (Buller, 2004; Shimamura *et al.*, 2008), for which principal developers have been K. Shimamura, S. P. Williams, and G. Buller, among others. Ganfeld is an ArcPad® 6/7 applet implemented in VBScript® (Raper, 2009). Basic GIS functionality is provided by the ArcPad® platform, and the Ganfeld scripts principally focus on data collection (Buller, 2004). Some of the data structures and data fields implemented in Ganfeld are inherited from FieldLog (K. Shimamura, personal communication). Explicit support for a fully-conceived ontology of field data collection was not retained, however. The accessibility of VBScript® to the general user is how users can customize what data are collected and how.

Ganfeld facilitates the construction of geological field data in a small set of files in Shapefile format. Shapefiles for station data, material data, sample data, structural data, photo data, and so forth, represent the main feature sets or layers that can be collected with most versions of Ganfeld. These feature sets can then be viewed directly in the wide selection of GIS software that supports the Shapefile format. Alternatively, importing shapefiles into a modern spatial database, such as the Geodatabase format, is generally straightforward with the appropriate software and support scripts. The collection of data in Ganfeld supports a few features not directly enabled by the Shapefile format itself. One-to-many relationships between the Ganfeld feature sets are maintained using ID fields and an implied tree structure between root and branch feature types. Constraints on data entry are maintained by using user supplied lookup tables in DBF format for populating various data entry menus.

The current implementation of Ganfeld has a finite lifetime. One reason is because the technology on which it is based will soon become obsolete and will only be actively supported by Microsoft®, in the case of Windows Mobile/XP®, and ESRI®, in the case of ArcPad® 6/7, for a few more years. Windows Mobile® support is scheduled to end in 2014, for example (K. Shimamura, personal communication). ESRI® is currently deploying ArcPad® 10 and has abandoned VBScript® as one of the scripting options in the more recent versions of ArcPad®. A second reason is because funding for Ganfeld development at the GSC may soon end. Ganfeld development was previously and is currently funded through the TGI-3 and GEMS funding vehicles at the GSC, respectively (K. Shimamura and L. Nolan, personal communication). It is almost certain that surveys, such as the GSC and NSDNR, will have to move to and/or develop a new implementation of field data acquisition software in the next decade. In the meantime, however, Ganfeld is the main system in use at NSDNR.

With the goal of deploying field data acquisition software in current mapping projects at NSDNR, NSDNR decided to adapt the Ganfeld software to its needs. In so doing, it was decided to re-factor the codebase originally shared by the GSC, in order to enable more flexibility and the addition of new features. The hope is that Ganfeld can continue to be actively maintained at NSDNR. The design and features implemented during this period will still be of value after Ganfeld is retired because they can serve as a reference or prototype design for the next generation of data acquisition software. Formally, this version of Ganfeld is dubbed Ganfeld-NSDNR, as long as it remains distinct from the original Ganfeld codebase. The potential for engaged collaboration with the GSC and other provincial surveys will be explored. One possibility is that any parallel development efforts that have occurred in the past few years may be able to be merged back into a single Ganfeld codebase. This possibility requires discussion amongst the various interest groups.

Design of Ganfeld-NSDNR

The main goal in re-factoring the Ganfeld codebase was to substantially increase the flexibility for both the user and the developer to enable, support, and change the basic behaviour of the software.

First, simple support for configuring Ganfeld-NSDNR from a file was implemented. The main advantage of using a configuration file is that users and developers have a single location, not embedded in the VBScript® code, where they can identify or specify some of the basic functionality. The configuration file support is minimal because I was not aware of a configuration file library for use in VBScript®, but the main benefits are realized simply through the extraction of the many items that might otherwise be hard-coded at various places in the scripts.

Second, the files and folders distributed with the Ganfeld applet were restructured to facilitate a cycle of active and custom development. In Ganfeld-NSDNR, there is now one system script, one applet script, and one applet folder that contain most of a Ganfeld distribution. A few icons and fonts need to be placed elsewhere. In the Ganfeld Applet folder, various subfolders clearly separate different parts of the program. Specifically, Ganfeld-NSDNR is configured to support different project types. Also, for each project type, Ganfeld-NSDNR is configured to support different shapefiles to build, different formats for the shapefiles to build, different lookup tables to use for entering data, and different GUIs to use for entering data. All of these parts of Ganfeld-NSDNR are managed through specific relationships between the configuration file, the folder structure and file content distributed with a deployment of Ganfeld. Moreover, all of the code is implemented to be independent of the project type, so it can be copied and pasted into new GUI interface scripts for new shapefiles and new project types with minimal editing. This facilitates the rapid development of new functionality. Also, UTM projections that can be selected for use with a Ganfeld Project are those supplied in a corresponding folder. This eliminates the possibility a user might try to build a project with a UTM projection for which support is not available.

Finally, a new implementation of menus in Ganfeld-NSDNR now enables the displayed entries to be configured on a per menu per user basis. This helps to mitigate the impact of too many choices. There is also support for configuring default selections in various menus to minimize the typing and clicking needed during data entry.

Key Features of Ganfeld-NSDNR

Along with changes to the design of the Ganfeld codebase, a number of features were modified or added to the version of Ganfeld on which Ganfeld-NSDNR was based. It is my understanding that some of these features have also been added to the current version of Ganfeld at the GSC, but others are unique to Ganfeld-NSDNR. The main priorities in changing or adding features to Ganfeld-NSDNR is to facilitate more efficient entry of data in the field and to enable interactive interpretation of collected data on various map layers for map units and lineaments, for example. Here, more efficient entry of data means faster and more flexible data entry options. Construction of map layers is supported by adding in appropriate symbology in the Shapefile interface specification and linking this to data fields in the Shapefiles. A description of some of the key features in Ganfeld-NSDNR follows.

First, Ganfeld-NSDNR now clearly supports the collection of three types of information: data layers, interpretation/product layers, and supplementary files. Data layers are the point feature sets that store station, material, sample and structural data. Interpretation/product layers are polygon and line feature sets that reflect interpretation of the underlying point data and may be preliminary versions of map products that will be published. Currently, Ganfeld-NSDNR supports building structural linework maps, map unit polygon maps, and outcrop outline maps. Supplementary files are text or image files that accompany various aspects of the program. Affiliated with individual stations are a text file for general notes, diagrams for drawing sketches, and photos for various station features. Affiliated with a deployment of Ganfeld-NSDNR is a text file for reporting problems or feature requests. These reports can be submitted to those responsible for active development of Ganfeld-NSDNR upon returning from the field. Also, affiliated with a deployment of Ganfeld-NSDNR are a set of files that describe the various map units encountered. These could be entered prior to starting mapping based on previous work, but these descriptions can also be built in the field for current descriptions of units. There is now for access to explicit support files from within Ganfeld-NSDNR. These currently include tide charts and arbitrary reference material that is loaded into a Ganfeld-NSDNR deployment prior to starting mapping.

Second, Ganfeld-NSDNR includes some changes and additions that support greater flexibility and speed when entering data. The fetch mode has been reworked to involve less navigation of selection forms. Now, you simply select the layer of interest, point to the location of interest and are shown the data of interest, unless there is more than one, in which case you navigate through one select form. Fetch is now viewed as another way to enter data and incorporated into the data entry section of the Ganfeld toolbar. Buttons have been provided that allow a user to finish data entry for one shapefile and proceed directly to the next without returning to the Ganfeld toolbar. This facilitates the rapid entry of one station, one material, and one structure measurement, for example. Or it facilitates the rapid entry of many photos or structural measurements in sequence. Using this functionality, linking two related structural measurements together has been added as a minimal way to support the more complex nature of structures. Entering data to the material shapefile has been made modular so that a large number of fields can be presented in related subsets, smart defaults can be used, yet the application does not become too slow on Windows Mobile devices such as the HP IPaq devices. This allows bedrock, surficial and coastal data to be collected in the same project. For material and structural data, the concept of auxiliary data has been added explicitly. This allows users to collect up to three more types of data for each entry that is not already supported elsewhere in Ganfeld, and allows them to do so in a structured and consistent manner because Ganfeld remembers the auxiliary data fields the user has previously entered in the specific project. The ability to delete data with appropriate cascading deletes for related subdata is also included.

Finally, Ganfeld-NSDNR has incorporated support for users changing how map units, lineaments and structures are displayed on a per Ganfeld deployment basis. Specifically, users can maintain three different palettes or styles of symbols at once and switch between them. In practice, it is envisaged that one symbolization scheme for each of these layers would be supplied by the corporate standards of the given research group, whereas the other two could be used interactively by particular users. It turns out that symbology is one of the more tricky aspects to support in Ganfeld, so some preliminary support for preparing different symbol sets for Ganfeld-NSDNR has been supplied as well. This should facilitate Ganfeld-NSDNR being adapted for other research groups.

Deployment Schedule

A preliminary version of Ganfeld-NSDNR was sent for testing on January 21, 2011, by NSDNR staff, the GSC and the Department of Natural Resources, Newfoundland. It is hoped that feedback provided by late February to early March will be addressed and incorporated into a version to be distributed in late April for use in the mapping projects at NSDNR during the 2011 field season.

A key feature that has been identified to be part of this April release is the ability to manage what symbols are actually to be displayed or not on a summary map view. Another outstanding issue is how best to handle the drawing of diagrams on Windows Mobile devices. The lack of a free paint program with a file browser that can explore the full directory structure of a Windows Mobile device means that loading and saving diagrams on Windows Mobile devices ends up only working outside the Ganfeld project directories and the user has to be vigilant to maintain file names correctly. A good solution to this issue is probably needed for the diagram feature to work seamlessly on Windows Mobile devices. Windows XP devices, such as the UMA handhelds deployed at NSDNR, do not have this issue because Paint for Windows XP provides the necessary functionality.

Conclusions

Ganfeld-NSDNR is a re-factored and extended version of the Ganfeld software. The new design and new features are an attempt to make Ganfeld suitable for use in all mapping projects at NSDNR. It is also hoped that making it easier to add new functionality will mean the software can be developed during use to be a reference implementation or design for the next generation of field data acquisition software. A version of Ganfeld-NSDNR will be distributed to staff in late April for use during the 2011 mapping season. Collaboration with other surveys and research groups may allow Ganfeld-NSDNR to become part of a broader initiative.

References

Broderic, B. 2004: The design of GSC FieldLog: ontology-based software for computer aided geological field mapping; Computers & Geosciences, v. 30, p. 5-20.

Buller, G. 2004: Ganfeld: geological field data capture, *in* Digital Mapping Techniques ’04-Workshop Proceedings, ed. D. R. Soller; United States Geological Survey, Open-File Report 2004 – 1451, p. 49-53.

Raper, J. editor 2009: ArcPad®: a field user’s guide; ESRI® Press, 380 New York Street, Redlands, California, 92373-8100, 148 p.

Shimamura, K., Williams, S. P. and Buller, G. 2008: Ganfeld user guide: a map-based field data capture system for geoscientists; Geological Survey of Canada, Open File 5912, 90 p.