Natural Resources Canada | Ressources naturelles Canada

GEOLOGICAL SURVEY OF CANADA
OPEN FILE 8188

Supplementing public geoscience knowledge with archived industry data: An example from northwest Canada

T.D. Finley, K.M. Fallas, and R.B. MacNaughton

2017

Celebrating 175 yrs

Canada

GEOLOGICAL SURVEY OF CANADA
OPEN FILE 8188

Supplementing public geoscience knowledge with archived industry data: An example from northwest Canada

T.D. Finley[1,2], K.M. Fallas[1], and R.B. MacNaughton[1]

[1] Geological Survey of Canada (Calgary)
[2] Current affiliation: University of Victoria

2017

© Her Majesty the Queen in Right of Canada, as represented by the Minister of Natural Resources, 2017

Information contained in this publication or product may be reproduced, in part or in whole, and by any means, for personal or public non-commercial purposes, without charge or further permission, unless otherwise specified.
You are asked to:
• exercise due diligence in ensuring the accuracy of the materials reproduced;
• indicate the complete title of the materials reproduced, and the name of the author organization; and
• indicate that the reproduction is a copy of an official work that is published by Natural Resources Canada (NRCan) and that the reproduction has not been produced in affiliation with, or with the endorsement of, NRCan.
Commercial reproduction and distribution is prohibited except with written permission from NRCan. For more information, contact NRCan at nrcan.copyrightdroitdauteur.rncan@canada.ca.

doi:10.4095/299605

This publication is available for free download through GEOSCAN (http://geoscan.nrcan.gc.ca/).

Recommended citation
Finley, T.D., Fallas, K.M., and MacNaughton, R.B., 2017. Supplementing public geoscience knowledge with archived industry data: An example from northwest Canada; Geological Survey of Canada, Open File 8188, 18 p. doi:10.4095/299605

Publications in this series have not been edited; they are released as submitted by the author.

Abstract

Mineral industry exploration data archived in the public domain are an underutilized source of geoscience knowledge in Canada. Efforts to rescue archived data can contribute to the repository of public geoscience knowledge, thereby meeting the objectives of the Government of Canada's Open Data Initiative and increasing the cost effectiveness of modern mapping projects. This report documents a data rescue exercise dealing with a high quality set of maps produced by a 1975 Rio Tinto exploration program in the Mackenzie Mountains, NWT, (RT claim group, northern NTS 106B and southernmost NTS 106G) and now residing in the public domain. These maps were digitized in preparation for Geological Survey of Canada (GSC) field work in the Mackenzie Mountains. Line-work, lithological information, and structural data were captured in a GIS database. These archival maps cover a small area in much greater detail than was possible for the regional maps published by the GSC. As a result, the digitized dataset will allow for strategic selection of sites requiring a revisit, potentially streamlining the GSC's present and future mapping campaigns in the region. The high density of structural measurements contributed to clearer understanding of map unit distribution during GSC field work in 2016, and may contribute to a detailed structural analysis of the region. The 1975 Rio Tinto exploration program cost over $450,000 CAD ($2 million in 2016 dollars). The cost of rescuing the data was minor by comparison, consisting mainly of the salary required to pay a junior scientist to carry out the digitizing.

1. Introduction

The Geo-mapping for Energy and Minerals (GEM) Program of the Geological Survey of Canada (GSC) aims to generate new regional-scale geological maps and data sets for northern Canada, in support of increased natural resource exploration, informed land-use decisions, and responsible resource development. A key driver behind the GEM Program is that the geology of many areas in northern Canada had previously been mapped at reconnaissance scale only. While published maps for these areas still act as industry standards, they are of varying quality, often lack fine detail, and warrant further study. Given the logistical challenges and costs of research in Canada's remote northern regions, it is desirable to make use of good-quality archival data during field planning and data compilation, to reduce costs and augment data coverage.

In addition to the mapping undertaken by the GSC and by provincial and territorial surveys, numerous studies by geologists from industry and academia have been undertaken in the remote regions of Canada. While such studies do not generally result in regionally comprehensive maps like those produced by government geoscience surveys, individual localities or regions may have been investigated in greater detail than is possible during government geoscience projects. If the archival data were preserved with adequate care, they can be candidates for data rescue. Efforts to rescue archival geoscience data from government (e.g., Fallas *et al.*, 2015; Riganti *et al.*, 2015) or industry sources (e.g., Smith, 2012, 2015) commonly are driven by the twin desires of ensuring ease of access and machine-readability, and thus are in keeping with the principles of the Government of Canada's Open Data Initiative (see http://open.canada.ca/en/open-data).

Although contemporary industry data generally are proprietary and unavailable to government geological surveys, there exists a trove of information in legacy maps, reports, and primary data sets that have been archived in the public domain by federal, provincial, and territorial agencies. These records can be a useful supplement public geoscience knowledge. The present report documents the procedures used to rescue archival, public-domain mineral-exploration data in the context of a GEM Program bedrock-mapping activity. These data were gathered during a mineral-exploration campaign in the Mackenzie Mountains, NWT, by Rio Tinto Canadian Exploration Ltd. (Sanguinetti *et al.,* 1976), and subsequently were placed in the public domain. The rescued data are a worthwhile supplement to the GSC's digital records and will contribute to the repository of public geoscience data in Canada, meeting the objectives of the Open Data Initiative. Having access to these maps in digital format will increase the effectiveness of current mapping projects, taking advantage of money already spent, and making more efficient use of public funds.

The data-rescue exercise documented in this report reflects the ongoing transition from analog to digital mapping technologies. All stages of a geological mapping program, from field work to final compilation, can make use of Geographic Information System (GIS) software. Although hardcopy maps are a portable and useful depiction of a geological interpretation at a particular scale, digital maps have the advantage of offering the opportunity to investigate and analyse the underlying data, as well as generate new interpretations or representations at different scales for different purposes. Archived hard-copy maps drawn at different scales, using topographic base-maps of variable quality and/or employing unfamiliar unit nomenclature, are often difficult to relate to contemporary maps. Poor scan quality can also discourage the use of archived maps viewed on a digital platform. As such, there has been an apparent data loss in the shift to digital methods. The effort required to utilize both mediums when preparing for field work and compiling a map post-field work tends to dissuade researchers from making use of older data, despite its potentially high quality. The data-rescue efforts documented herein are presented with the goal of helping to improve this situation.

2. Historical Background

2.1. Operation Norman

In 1952, the GSC began using helicopters to access remote areas; a rapid increase in the rate of mapping of Canada's bedrock geology followed (Officers of the Geological Survey of Canada, 1959). By the late 1960s, one of the final regions that remained unmapped was a large area extending west from the margin of the Precambrian Canadian Shield to 132° W, and south from the Arctic Ocean to 64° N. Several physiographic regions are found in this area, including the Mackenzie Mountains, Mackenzie Plain, Peel Plain, Peel Plateau, and the Franklin Mountains. A campaign called Operation Norman was mounted between 1968 and 1970 to fill the gap. Bedrock was mapped at a reconnaissance scale and observations were recorded on standardized note cards. Locations were marked on air photos and topographic maps. Sparse vegetation in some regions allowed for much of the mapping to be done from the air, but most lithological and structural contacts on the final map lacked precise spatial control. Following the completion of fieldwork, maps and cross-sections from Operation Norman were published at a variety of scales and formats (see Fallas *et al.,* for a full summary). Bedrock maps for the Mackenzie Mountains were published either as multicoloured, professionally drafted "A series" maps (Aitken *et al.,* 1974; Aitken and Cook, 1979a,b) or as Open File reports, which at that time consisted of map compilations drawn and lettered by hand by the authors upon a cartographically prepared map base (Aitken and Cook, 1974). Hard-copy maps were eventually scanned to a digital Portable Document

Format (PDF). The primary observations from the field (e.g., note-cards and annotated air photos) were carefully preserved at the GSC's Calgary office, and therefore available to be digitized in a process outlined by Fallas *et al.,* (2015a).

2.2. Rio Tinto exploration program: RT claim group

Operation Norman laid the groundwork for further industrial exploration in the Mackenzie Mountains. In the summer of 1974, working on behalf of Rio Tinto Canadian Exploration Ltd., geologists of Cordilleran Engineering Ltd. discovered Mississippi Valley-type zinc-lead mineralization in Neoproterozoic (then believed to be Mesoproterozoic) dolostone and Cambro-Ordovician rocks near the upper Gayna River (64°57'N; 130°43'W). This campaign was referred to as the "Gayna River Project". Following preliminary geological and geochemical investigations, a total of 878 claims were staked in northern NTS 106B and southernmost NTS 106G (Fig. 1). These claims are referred to herein as the "RT claim group". More thorough field mapping and prospecting took place in the summer of 1975; two-person teams mapped the majority of the property at a 1:31680 scale, the core of the claim group at 1:4800, and specific mineral showings at 1:1200. In total, approximately 1000 structural measurements were recorded. Surficial bedrock geochemical samples were collected at significant showings for spectrographic analysis and soil samples were collected on a grid spacing of 60 to 120 meters where topography would permit. Induced polarization resistivity and magnetic anomaly were measured along 12 transects near the primary mineral showings. Seventeen holes were drilled in one showing to better constrain the form of the mineralized zone. The field work of 1975 prompted the addition of 20 claims in the northwest corner of the claim group and the concession of 133 others. In total, the 1975 field season was reported to have cost $453,568.24 (Sanguinetti *et al.,* 1976), equivalent to more than $2 million today according to the Bank of Canada (2016). As per the requirements of the Mining Regulations of the Northwest Territories Land Act, maps, cross sections, and summary reports pertaining to the RT claim group (Sanguinetti *et al.,* 1976) were submitted to the N.W.T. Mining Recorder's Office. Historical hard-copy documents submitted to that office were eventually scanned and made available to the public in an online database maintained by the Northwest Territories Geological Survey (see http://www.nwtgeoscience.ca/mineral-exploration-assessment-reports).

The records in the online database comprise an invaluable resource that is easy to access and has the advantages of being searchable and downloadable. However, the utility of the data can be limited by their digital format, which is most commonly scanned copies of maps and reports archived as PDF or TIFF files. Such files are not machine-readable and, depending on the reproduction quality of the original report, the resolution of the scans may not be optimal. For example, many older reports were archived as carbon copies, and drafting standards varied widely from report to report. Depending on the quality of the scan and the size of the original font and linework, fine details can be obscured or lost entirely, decreasing their value as a research tool and as a public dataset. The inability to read, query, or manipulate data points individually prevents their integration with other data sets for scientific analysis, and use of the reports in a modern GIS-based mapping project is not straightforward.

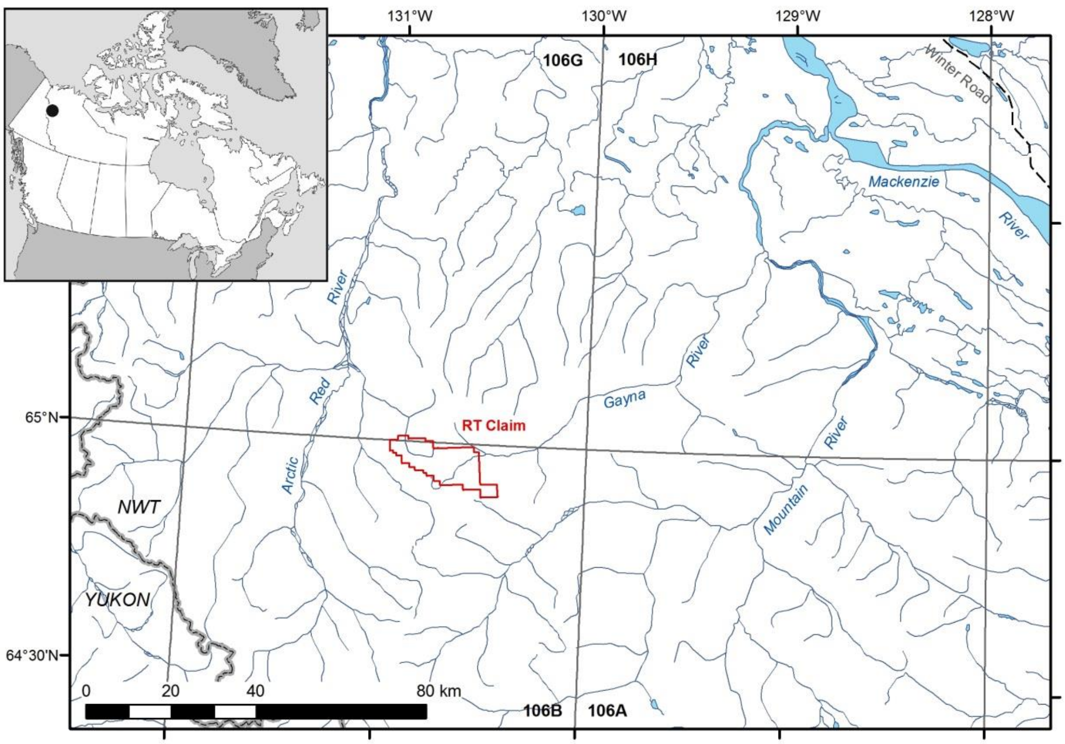

Figure 1. *Map showing the areal extent of the RT claim group in the northern Mackenzie Mountains.*

2.3. Geo-mapping for Energy and Minerals (GEM) program: Shield to Selwyn geo-transect

In 2008, the federal government announced $100 million in funding over 5 years for the Geo-mapping for Energy and Minerals (GEM) Program. The goal of the program was to modernize the geological understanding of Canada's north in order to support resource exploration and aid in responsible land use planning. In keeping with the Government of Canada's Open Data Initiative, a key outcome was the production and publication of GIS-enabled bedrock geology maps. The first phase of the GEM program was a success and funding was renewed in 2013 for an additional 7 years.

During the first phase of GEM, GSC scientists re-mapped the region around the Norman Wells oil field, updating several Operation Norman maps (Fallas *et al.,* 2012). An activity within the second phase of the GEM program, the Shield to Selwyn Geo-transect, also overlaps with the area covered by

Operation Norman. GSC researchers are now undertaking systematic re-mapping of the 1:250,000 maps first compiled in the early 1970s (e.g., Fallas *et al.,* 2015b). In the summer of 2016, field work was focussed on NTS maps 106G (Ramparts River) and 106H (Sans Sault Rapids), an area that includes the Canyon Ranges of the Mackenzie Mountains (Fallas *et al*., 2016). Published Operation Norman maps for these areas are of "A series" quality (Aitken and Cook, 1979a,b). Field work activities in 2017 will be further south in the Backbone Ranges of the Mackenzie Mountains on NTS map 106B (Bonnet Plume Lake), an area that was only partly mapped during Operation Norman and for which the resulting maps were Open Files (Aitken and Cook, 1974). The archived data from Operation Norman, and from industrial activities like the RT claim group, informed planning and execution of field work in 2016. The Rio Tinto data will be an especially significant contribution to mapping in 2017, as the majority of those stations are in the 106B map area.

3. Methods

The methods used in rescuing the geological data from the RT claim group are a modification of those described by Fallas *et al.,* (2015a). Although a wide variety of data types were gathered for the claims by Sanguinetti *et al.* (1975), only bedrock mapping data were dealt with during the present activity.

3.1. Georeferencing

Because neither recorded geographic coordinates nor annotated air photos were available to help locate the Rio Tinto field stations, it was necessary to obtain their geographic coordinates directly from the report maps. Topographic features on the original industry compilation maps were georeferenced to modern satellite imagery in ArcMap™, using river bends and intersections as the primary controls on location. Larger watercourses were given preference, except where wide braided channels involved the possibility of significant positional shifts between the time of map compilation (circa 1975) and that of the satellite imagery acquisition (2008). A minimum of 10 reference points were placed on each image. Topographically, the larger scale (smaller area) maps were very accurate, and matched the satellite imagery with a root-mean-square error of <10 m. Thus these maps were used for data rescue wherever possible. The georeferencing of generalized, smaller scale (larger area) property maps was much less precise, with a root-mean-square error of >100 m. As a result, positional accuracy errors were greater when rescuing data for areas mapped only at the smaller scale.

3.2. Line-work

The Rio Tinto compilation maps were preserved as black-and-white scanned copies of originals and this made it difficult to discern the finer details of depositional and faulted contacts. Map unit polygons were generated from digitized contacts and fault traces and then colourized to clarify the distribution of lithologies and reduce the possibility of misreading the map when digitizing station points later on. The resulting improvement in clarity is shown in Figure 2. Line tracing was generally done at a 1:10,000 scale and more emphasis was placed on speed than on node density, since the result was to be used only to provide context for field stations and not for inclusion in a new publication.

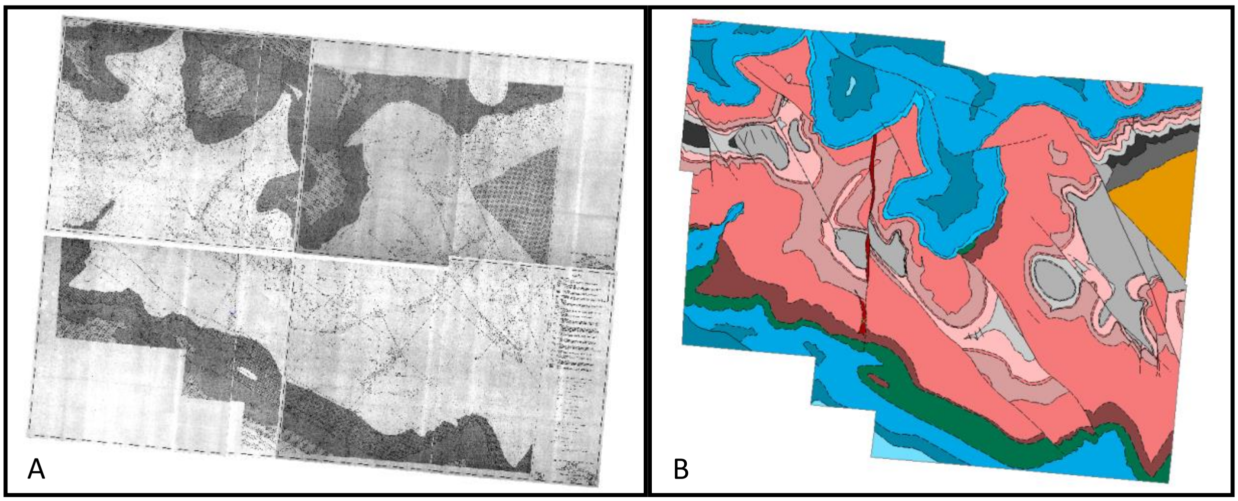

Figure 2. *A) Black-and-white scan of original map from the RT claim group. This particular mapsheet was the highest quality of all material submitted with the assessment report. Note the folds in the paper and the taped seams. B) Digitized linework and colourized unit polygons of the same map area. The finer details of lithological and structural relationships are much more clear. Unit names are contained in the GIS project attribute tables and can be turned on when needed.*

3.3. Database configuration

The GSC has standardized the organization of geological data recorded and stored in its GIS geodatabases, ensuring long-term usability of the data across multiple user groups, and allowing for simplified migration to a publicly available format. The present data rescue project used workflows and organizational structures that adhere to GIS principles outlined by MacDonald (2014) and Fallas (2014). In accordance with the workflow, station location points were placed at the best-determined location of observation (as outlined in Section 3.4), and key metadata (e.g., physical environment, observer, date) were recorded in appropriate fields. A station identifier was assigned to each point, containing information on year, collector, and a unique number. For the RT claim group, stations were assigned a unique alphanumeric code using the following format: 75RT-Gayna-####, where “75” indicates the year of collection, “RT-Gayna” indicates that the data was gathered by Rio Tinto from their RT claim group (Gayna River Project), and “####” is a unique 4-digit numeral. Lithological and structural data points were “snapped” to the station point and “one-to-many” relationship rules were established so that no lithological or structural data could exist on the map without being associated with a station point via a common station identifier.

3.4. Structural data

Station points were expressed on the original Rio Tinto compilation map only as structural measurements. It was apparent that they were not necessarily plotted at the exact location at which they were recorded, but rather at locations chosen to minimize clutter and overlap (e.g., measurements gathered in a stream cut were plotted adjacent to the stream rather than in it). The compilation map showed rough outlines of confirmed outcrop (as opposed to talus and/or felsenmeer), and it was generally assumed that the actual station locations were on those outcrops and not adjacent to them (Fig. 3A). In many cases, clusters of several structural measurements were found near each other. Where outlines of bedrock exposure and visible outcrop on satellite imagery supported this approach, it was assumed that these were all measured at the same outcrop, and so a station point was placed at their centroid on the probable outcrop, and all structural measurements were attached to it (Fig. 3B). Any positional errors introduced by this methodology are estimated at 25 m or less.

As is common on geological maps, structural measurements on the claim maps were drawn with the dip value displayed numerically and the strike value represented only by the cardinal orientation of the symbol on the map. In order to accurately capture the strike, the “Coordinate Geometry (COGO)” tool in ArcMap™ was used to measure the azimuth in accordance with the right hand rule.

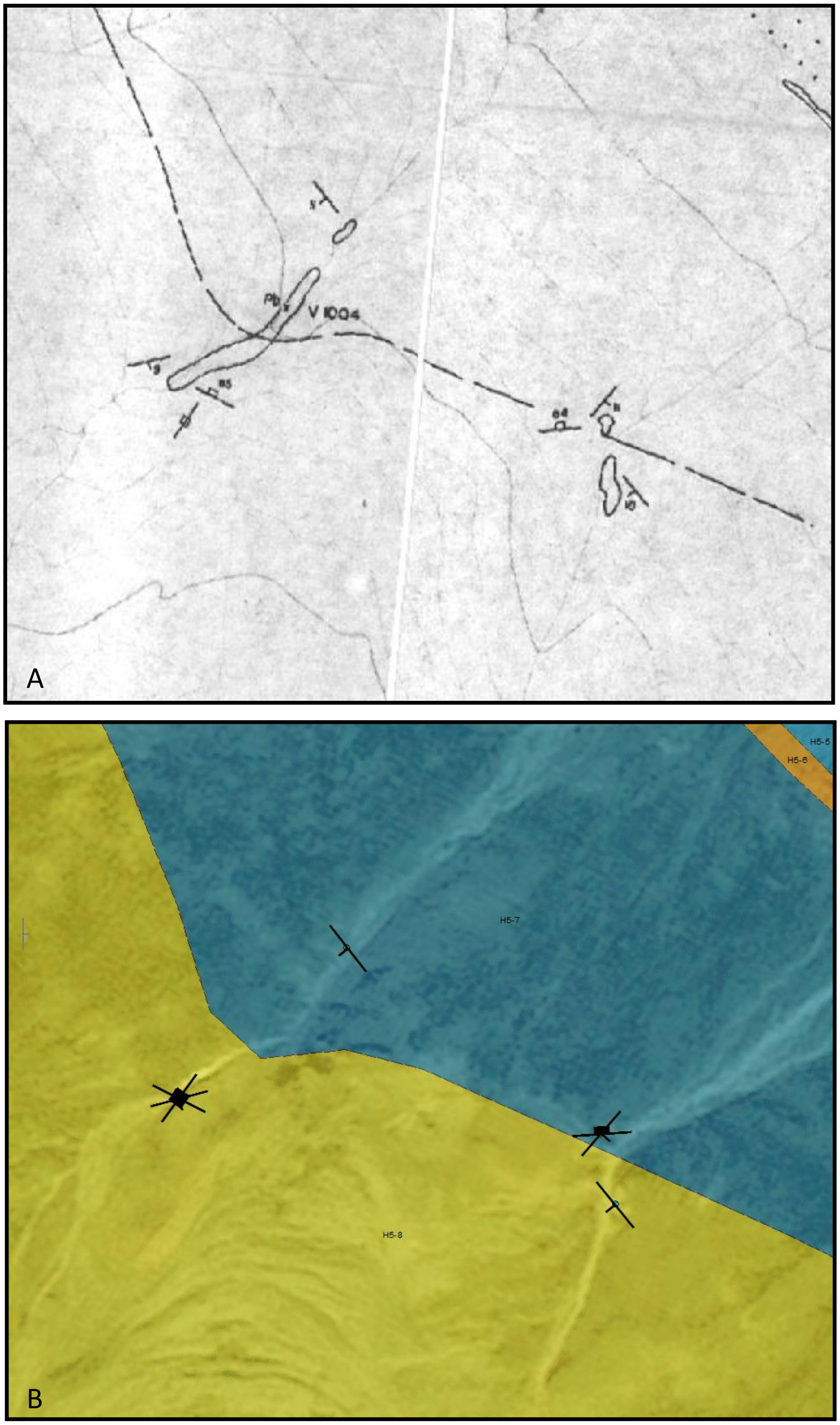

Figure 3. *A) Portion of scanned original claim map showing structural measurements recorded in two small stream cuts. Outcrop exposure is encircled by a solid line, while lithological contacts are dashed or dotted. B) Digitized version of the same map area, with structural symbols placed directly on the outcrop locations rather than adjacent to them. Mapping is underlain by satellite imagery to help identify outcrop exposure.*

3.5. Lithology

Primary observations from each station were unavailable. Lithological information was assigned based on the descriptions provided for each unit in the map legends (Fig. 4A). The digitized unit polygons (see Section 3.2) made this task simple, as it was immediately obvious which unit a station was in. These brief lithological descriptions were "decomposed" (i.e., broken down) into standard qualifiers that could be recorded in appropriate attribute fields (Fig. 4B) as outlined in the data model (Fallas, 2014). The resulting properties are useful for querying and sorting in the GIS geodatabase. Description details that could not be adequately decomposed were transposed into the "REMARKS" field.

A

UH

UPPER HOST

Thin bedded, dark grey to brownish grey, fine to medium grained dolostone. Typically cliff forming. Frequently stromatolitic; oolitic. Numerous breccias of dolostone fragments cemented by dolostone mud; some soft sediment deformation of fragments in breccia. Abundant carbonate influx, most typically in secondary breccias such as rubble, mosaic, and crackle breccia healed by white sparry calcite or dolomite. Recrystallization, vug filling, fracture filling, zebra rock and other features common. Carbonate influx frequently mineralized with sphalerite. Primary breccias rarely mineralized.

B

Field	Value
OBJECTID	19419
MAP_THEME	bedrock, at surface
FEATURE	lithology
STATION_ID	75RT-Gayna-0852
LITH_ID	75RT-Gayna-0852A
OCCURRENCE	bed
LITHGROUP	sedimentary
LITHDETAIL	dolostone
FIELD_UNIT	Upper Host Unit
MAP_UNIT	Gayna Formation
SOURCEUNIT	UH
COMP_QUAL	<null>
TEXT_QUAL	cliff forming, brecciated
STRUC_QUAL	stromatolitic, oolitic
MINERALS	sphalerite
SED_MIN	sphalerite
MIN_NOTES	carbonate influx frequently mineralized with sphalerite. Primary breccias rarely mineralized
GR_SIZE_MN	finely crystalline (16–62 um)
GR_SIZE_MX	medium crystalline (62–250 um)
FR_COLOUR	dark grey to brown-grey
W_COLOUR	<null>
BEDDING_MN	thin bedded (3–10 cm)
BEDDING_MX	thin bedded (3–10 cm)
REMARKS	Numerous breccias of dolostone fragments cemented by dolostone mud. Soft sed. defm. of fragments in breccia. Abundant carbon
SOURCE_REF	Rio Tinto, Gayna River Property, RT Claims, Mineral Assessment File 080567
MAP_ID	106B

Figure 4: *A) Original legend entry for "Upper Host" unit on a RT claim group map. B) Decomposed lithological information in appropriate attribute fields in the GIS geodatabase. The "MAP_UNIT" field reflects modernized lithostratigraphic unit nomenclature, while the "FIELD_UNIT" and "SOURCEUNIT" fields contain the name and label provided on the Rio Tinto maps. Note that detailed qualitative information is placed in the "REMARKS" field.*

The full, paragraph-form unit description was also captured in a separate table, in case the decomposed qualifiers should later require modification. Lithological units were correlated with modern lithostratigraphic nomenclature so as to maximize the value of the data as a tool while in the field and during future recompilation. In particular, the recently formalized units of the Little Dal Group (Turner and Long, 2012) were applied to the best-fitting divisions on the Rio Tinto maps (Figure 4B).

4. Discussion

4.1. Strategic mapping

Proper use of archived data can greatly reduce financial expenditures in a project, particularly in remote regions where sites can only be visited via helicopter. Recent efforts by Fallas *et al.* (2015a) have demonstrated the value of rescuing old GSC hardcopy data for reference in modern field campaigns. If archival mapping data are deemed to be of satisfactory quality, the need to revisit sites can be reduced, thereby allowing for a greater number of new stations to be visited, increasing the total data coverage.

Similarly, the systematic digitization of the archived Rio Tinto data allowed for the identification of strategic sites in need of re-examination during the current mapping campaign. For example, since the time of the 1975 Rio Tinto exploration program, the Katherine Group quartzite has been formally divided into seven formations as outlined by Long and Turner (2014). The Katherine Group is undivided on Rio Tinto's Gayna River map sheet. This may reflect the fact that Rio Tinto's work predated any detailed subdivision of the Katherine Group (Aitken *et al*., 1978) but it also appears that the Katherine Group was not of economic interest to the Rio Tinto geologists as they recorded only one field station within it. It would be worthwhile to revisit this particular stratigraphic interval to better constrain the formations within it.

In contrast to the treatment of the Katherine Group, the maps produced by Rio Tinto provide a significantly higher level of detail within units now assigned to the Little Dal Group. Where the published GSC maps (Aitken and Cook, 1974; 1979a) show the Little Dal Formation and 3 units assigned to "H5", the Rio Tinto maps have subdivided H5 into 9 units representing 5 of the formations formalized within the Little Dal Group by Turner and Long (2012). Similarly, recent stratigraphic studies by MacNaughton and Fallas (2014) have defined a new unit, the Nainlin Formation, at the base of the Franklin Mountain Formation. The RT claim maps include a thin red-bed unit that matches the type description of the Nainlin Formation at the base of the Franklin Mountain Formation. Being that the thickness and lateral continuity of the Nainlin Formation and the subdivisions of the Little Dal Group are already well constrained in the map area, fewer revisits will be required for these units, and focus can be placed on sites most likely to yield key stratigraphic information.

4.2. Structural Analysis

Faults, bedding angles, and joints in the RT claim group were mapped on a much finer scale than was possible during Operation Norman. The high density of recorded data points may prove useful in structural analyses of the region. With more than 1000 structural measurements, it may be possible to see statistically significant changes in bedding angles or joint sets across significant unconformities. For example, the sub-Cambrian unconformity in this region is thought to have been enhanced by uplift of the Mackenzie Arch (MacNaughton and Fallas, 2014). The RT claim group contains rocks that bound this horizon, and it may be possible to observe different structural trends that can be associated with either the uplift of the arch or with the younger, Cretaceous Laramide orogeny. Additionally, numerous minor faults in the RT claim group were mapped as cross-cutting the basal Cambrian red-beds (now Nainlin Formation after MacNaughton and Fallas, 2014), but not the overlying Franklin Mountain Formation. Further examination of these structures may aid in understanding Cambrian tectonics. These structural analyses will be the subject of further research both in the office and in the field.

4.3. Scale and level of detail

Regionally, maps produced by Operation Norman have proven to be broadly reliable, albeit subject to improvements in light of more detailed data (see, e.g., discussion in Fallas *et al*., 2015). However, the level of detail on the RT claim group maps is much higher than what was achieved by Operation Norman due to the significantly higher spatial concentration of stations. Minor map units were identified and displayed on the map, whereas many were grouped together by Operation Norman geologists. This is a reflection of differences in both scale and coverage. Topographic maps used in the compilation of Operation Norman geological maps were of 1:250,000 scale and lacked fine detail. The Rio Tinto exploration project made use of much higher resolution topographic maps and the lithological boundaries reflect that improvement. For example, the boundaries of horizontal units perfectly match topographic contours at 1:50,000 scale, whereas the generalized contours on older 1:250,000 scale maps do not allow for the illustration of this relationship. Rio Tinto visited >1000 stations per 250 km^2 whereas Operation Norman visited <3 stations over the same areal extent. Even if Operation Norman geologists had been able to identify minor units in their reconnaissance mapping, many of the smaller divisions would have been too small to display properly on 1:250,000 maps without adding clutter.

The practical value of the archival Rio Tinto data was illustrated during the 2016 field season when it proved difficult to delineate the Little Dal Group in the southeast corner of NTS 106G due to issues in identifying stratigraphic units and relationships. The Rio Tinto geologists had mapped the region immediately southeast of the problem area in sufficient detail to delineate formation boundaries. These could be extrapolated into the problem region, supporting a working hypothesis as to which formations were present, and ruling out unlikely interpretations. The Rio Tinto data are expected to be similarly helpful during the 2017 mapping campaign.

4.4. Data source

Unlike the archived data from Operation Norman, there are no primary observations (i.e. field notes) preserved from the exploration field work in the RT claim group. It is not known whether such records exist and no attempts were made to obtain them from Rio Tinto. The information available for rescue included the compilation maps, cross sections, and summary reports that were submitted to the NWT government following the completion of Rio Tinto's field work. These documents mainly record secondary interpretations and should not necessarily be regarded as ground truth. During the digitization process, care was taken to document the source of the data to ensure that future workers know that they are dealing with rescued industry data rather than primary GSC data. In the absence of primary observations, it will be especially important to perform spot-checks to ground-truth the interpretations of the Rio Tinto geologists. Because the unit lithological descriptions provided on the Rio Tinto maps formed the basis for assigning strata to modern formalized units, spot-checks will also serve to validate the unit assignments made during the present project.

4.5. Workflow challenges

This data rescue project was faced with several challenges, all surmountable, but noteworthy for future projects.

4.5.1. Typographical errors

As with any digitization effort, the work was repetitive, emphasizing the value of ensuring machine-readability of future datasets. Because it was common for small typographical errors to go unnoticed during data entry, procedures were put in place to check the data for accuracy. The data entries were periodically audited by visually scanning the attribute tables in the ArcMap™ project, looking for inconsistencies. Some common issues were duplicate station numbers, attributes in the wrong field, spelling mistakes, missing digits, and extra digits. The low variability between entries made errors easy to spot while scrolling through the table as they would often break the pattern. Lithological and structural data were checked using the “Validate” tool in ArcMap™ to ensure that composite relationship class rules were obeyed (i.e. all structural and lithological information was related to an existing station point, and no duplicate station numbers existed). The validation tool also found attribute values that did not match the required format of the attribute fields they were placed within, such as text entries in number-only fields.

4.5.2. Quality of archived data

The archived Gayna River property maps, from the RT claims group, were in some cases difficult to read due to the resolution of the scanned copies. The structural data in particular were very small and it was difficult to discern the numerical dip angles. This problem was worsened by the distortion introduced by georeferencing the images in ArcMap™ and the map imagery often needed to be examined in its original format before dip angles were recorded in the GIS environment. Stations that could not be deciphered were omitted from the digitized map. It was also difficult to interpret line-work on the colourless maps. For example, the line symbology for lithological contacts and bedrock outcrop boundaries looked very similar and lines commonly crossed one another, so care had to be taken to ensure the correct lines were traced.

While some map sheets, like the one shown in Figure 2A, were of high topographical accuracy, others were less accurate. In some cases georeferencing yielded an RMS (root-mean-square) error of >100 m. Some distortion of the georeferenced map image was noted. This may have been due to errors in the original topographic base, or warping of the original paper copies caused by folding or moisture damage. As such, the locations of stations on those map sheets lacked the same spatial accuracy as those found in the central core of the claim group. A caveat noting this uncertainty was included in the attributes of each of these stations.

4.5.3. Data loss

Not all components of the original Gayna River property maps have been captured in GSC geodatabases as of yet. Mineral showings and drill holes did not have enough accompanying information to permit ready digitization, but it may be worth expending effort to add them in the future. Outcrop outlines on the archived maps were found to add clutter to the digitized versions and thus were not included. Although lithological information from the map legends was captured in the geodatabase, some of the nuanced detail contained in the original paragraph format was lost, a common problem with modern pick-list methods of capturing geological information. To address this, a table containing the full paragraph-form descriptions of each unit was added to the geodatabase (see Section 3.5).

Now that the Gayna River property maps have been digitized, it is less likely that researchers with access to the digital data will feel compelled to examine the originals. Therefore data not included in the digitized map is more likely to be overlooked in the future. This is important to consider when choosing which data to capture and which to ignore, as any loss of data is undesirable.

4.6. Cost comparison

The 1975 Rio Tinto exploration project was reported by Sanguinetti *et al.,* (1976) to have cost $453,568.24, roughly equivalent to 2 million dollars in 2016 (Bank of Canada). Since the results of this major financial expenditure now reside in the public domain, it is desirable to see them used to greatest effect, particularly considering that budgets for current GSC summer field campaigns in the region can run to several hundreds of thousands of dollars. Within these budgetary constraints, it is not feasible to remap the Shield-to-Selwyn project area to the same level of detail as the Rio Tinto work. However, the Rio Tinto data can be utilized during the present project following a comparably inexpensive data-rescue exercise that amounted to approximately 100 hours of digitization (roughly 10 hours for line-work, 75 hours for entry of structural and lithological data points, and 15 hours for miscellaneous database management) performed by a junior scientist at an approximate cost of $1600. Scanned copies of the original RT claim maps were available due to a major initiative of the NWT government at an unknown cost. The cost of scanning the relevant material for this particular area however, is undoubtedly low when compared to field work. Digitization of records can involve significant time and effort rather than cost, but if done correctly, the data increase immensely in present and future value.

4.7. The interplay between industry and government geoscience data

The flow of information between government and industry geologists can be largely unidirectional. Maps produced by the GSC may not include industry data, but industry maps may be based in large part on published GSC maps. This is evident, for example, in industry's use of formalized unit nomenclature that, in many cases, was established by GSC investigators. On the other hand, when federal scientists look to update the regional maps, they are likely to look first to archived GSC records, while public-domain industry data may not be consistently utilized. The procedures outlined in this report demonstrate that there is significant value in looking beyond the primary records of the GSC and engaging in a more comprehensive approach to natural resource exploration in Canada. The public stands to benefit not only from tax-funded primary research, but also from industry data released into the public domain.

5. Conclusion

The 2016 summer field season provided an initial opportunity to test the utility of the rescued industry data. Field work planned for NTS 106B in 2017 will probably make greater use of the data, given that the largest part of the RT claims were in that map area. Having access to the data in a machine-readable format aids the field party in making strategic decisions about which sites to visit in the limited time available. There is no need to revisit previously visited locations unless there is an apparent issue requiring resolution, allowing for more time spent collecting new data in areas with few or no previous observations. The maps that are produced from field work as part of the Shield to Selwyn geo-transect will be made available to the public; the Rio Tinto data will have played a role in the compilation of those maps. More than 1000 structural measurements and unit identifications recorded by Rio Tinto geologists will also be made publicly available in the same format as new observations from the GSC. These data will be suitable for ongoing scientific analysis and integration with other datasets, contributing to the objectives of the Government of Canada's Open Data Initiative. Future GSC project leaders should endeavour to digitize any useful archived, public domain data that they come across in their research, both for the interests of their own projects, and for the benefit of the public. Digitization of legacy data is an obstacle faced by many scientific organizations today, and opportunities to modernize relevant datasets should be treated as a priority, in order to preserve their value into the future. When capital investments made in historical exploration campaigns can be leveraged in this way, new scientific investigations can build upon a more comprehensive base of archived research, while saving costs in modern field campaigns.

Acknowledgements

The work documented in this report was carried out during tenure of a Co-op student position at the Geological Survey of Canada by the first author. An earlier version was submitted to the University of Victoria in partial fulfillment of the requirements of its Co-op student program. The authors thank Christine Deblonde for her helpful review of an earlier draft.

References

Aitken, J.D. and Cook, D.G., 1974. Geological maps showing bedrock geology of the northern parts of Mount Eduni and Bonnet Plume Lake map areas, District of Mackenzie, Northwest Territories; Geological Survey of Canada, Open File 221.

Aitken, J.D. and Cook, D.G., 1979a. Geology, Upper Ramparts River, District of Mackenzie; Geological Survey of Canada, Map 1452A, 1:250 000 scale.

Aitken, J.D. and Cook, D.G., 1979b. Geology, Sans Sault Rapids, District of Mackenzie; Geological Survey of Canada, Map 1453A, 1:250 000 scale.

Aitken, J.D., Cook, D.G., and Yorath, C.J., 1974. Geology, Carcajou Canyon, District of Mackenzie; Geological Survey of Canada, Map 1390A, 1:250 000 scale.

Aitken, J.D., Cook, D.G., and Yorath, C.J., 1982. Upper Ramparts River (106G) and Sans Sault Rapids (106H) map areas, District of Mackenzie; Geological Survey of Canada, Memoir 388, 48p.

Aitken, J.D., Long, D.G.F., and Semikhatov, M.A., 1978. Progress in Helikian stratigraphy, Mackenzie Mountains; in Current Research, Part A; Geological Survey of Canada, Paper 78-1A; 1978. p. 481–484.

Bank of Canada, 2016. Consumer Price Index Inflation Calculator: http://www.bankofcanada.ca/rates/related/inflation-calculator/ (accessed April 2016)

Fallas, K.M., McLean, B.C., MacNaughton, R.B., and Hadlari, T., 2012. New bedrock map compilations for the central Mackenzie corridor, Northwest Territories; Geological Survey of Canada, Scientific Presentation 11, poster.

Fallas, K.M., 2014. Canadian Geoscience Map Bedrock Data Model v 4.0.0; Earth Sciences Sector of Natural Resources Canada, Geological Map Flow Project Internal Report, 69p.

Fallas, K.M., MacNaughton, R.B., and Sommers, M.J., 2015a. Maximizing the value of historical bedrock field observations: An example from northwest Canada; GeoResJ, v. 6, p. 30-43.

Fallas, K.M., MacNaughton, R.B., Turner, E.C., and Sommers, M.J., 2015b. GEM 2 Mackenzie Project: Colville Hills bedrock mapping, stratigraphy, and related studies; Geological Survey of Canada, Open File 7963, 19p.

Fallas, K.M., MacNaughton, R.B., Finley, T.D., and Gouwy, S.A., 2016. Report of activities for the GEM 2 Mackenzie Project: Northern Mackenzie Mountains bedrock mapping, stratigraphy, and related studies; Geological Survey of Canada, Open File 8132, 15p.

Long, D.G.F., and Turner, E.C., 2014. Formal definition of the Neoproterozoic Mackenzie Mountains Supergroup (Northwest Territories), and formal stratigraphic nomenclature for terrigenous clastic units of the Katherine Group; Geological Survey of Canada, Open File 7113 (revised), 40p.

Macdonald, L.E., 2014. Bedrock mapping work flow reference manual using ArcGIS™ Desktop version 10.2; Earth Sciences Sector of Natural Resources Canada, Geological Map Flow Project Internal Report, 158p.

MacNaughton, R.B., and Fallas, K.M., 2014. Nainlin Formation, a new Middle Cambrian map unit from the Mackenzie Mountains, Northwest Territories; Bulletin of Canadian Petroleum Geology, v. 62, n. 2, p. 37-67.

Officers of the Geological Survey of Canada, 1959. Helicopter Operations of the Geological Survey of Canada; Geological Survey of Canada, Bulletin 54, 60p.

Riganti, A., Farrell, T.R., Ellis, M.J., Irimies, F., Strickland, C.D., Martin, S.K., and Wallace, D.J., 2015. 125 years of legacy data at the Geological Survey of Western Australia: capture and delivery. GeoResJ, v. 6, p. 175-194.

Sanguinetti, M.H., Verley, C.G., and Webb, G.G., 1976. Summary report on the RT Claim Group, Gayna River area, Mackenzie Mining District, Northwest Territories; Cordilleran Engineering Limited, v. 1.

Smith, I.R., 2012. The application of seismic shothole drillers' log data to petroleum exploration and development. Bulletin of Canadian Petroleum Geology, v. 60, p. 59-68.

Smith, I.R., 2015. Seismic shothole drillers' lithostratigraphic logs: unearthing a wealth of regional geoscience information in northwestern Canada. GeoResJ, v. 6, p. 21-29.

Turner, E.C., and Long, D.G.F., 2012. Formal definition of the Neoproterozoic Mackenzie Mountains Supergroup, NWT, and formal stratigraphic nomenclature for its carbonate and evaporate formations; Geological Survey of Canada, Open File 7112, 57p.