DIGITAL CAPTURE OF 1:250 000 GEOLOGICAL MAPS OF THE NORTHERN TERRITORY

NTData

Project Summary 1 August 1998 **Last Update 1 July 2000**

Background

The *Digital Capture of 1:250 000 Geological Maps of the Northern Territory* - '**NTData**' - commenced as a collaborative venture between the **Australian Geological Survey Organisation** (AGSO) and the **Central Land Council** (CLC), in cooperation with the **Northern Territory Geological Survey** (NTGS).

Significant contributions to NTData have also been made by AGSO's **North Australian Basins Resource Evaluation** (NABRE) project and the **Northern Territory Geological Survey** through the National Geoscience Mapping Accord (NGMA).

NTData was been an evolving project, with products, specifications and goals undergoing considerable change since the commencement of work (see 'NOTES' below for information on data limitations).

NTData on the World Wide Web

Information on **NTData** can be found at: http://www.agso.gov.au/information/publications/ntdata/ntmain.html .
The NTData Web pages include the information in these notes plus clickable tile indexes and an **on-line order form**. For further information please contact the NTData custodian: Mr Jon Stirzaker, Australian Geological Survey Organisation, Ph (02) 6249 9135, Fax (02) 6249 9984, email: jon.stirzaker@agso.gov.au

NTData Project Summary

Objective: Acquire and disseminate a digital dataset of geology to assist government, industry, land owners, and the community to make informed decisions on resource, land management, and environmental issues.

Dataset Name: NTData

Dataset Theme: Geology

Coverage: 90 standard 1:250 000 map tiles in the Northern Territory

Spatial Extent: 11°00' - 26°00' South, 129°00' - 138°00' East

Projection: Universal Transverse Mercator, Australian National Spheroid, Australian Geodetic Datum 1966

Data Source: Stable-base repromat from maps of the Australian 1:250 000 Geological Series published by AGSO and NTGS. Currency 1963 - 1997

Data Formats: MicroStation DGN, Arc/Info, MapInfo (others eg DXF, subject to negotiation). See 'NOTES'

Data Units: MicroStation DGN: Kilometres; Arc/Info, MapInfo: Metres

Conditions of Use: Use is subject to the AGSO Licensing Agreement

Custodian: Australian Geological Survey Organisation

Contact: AGSO welcomes inquiries and feedback on any aspect of this dataset. Please contact:
Mr Jon Stirzaker, Australian Geological Survey Organisation, GPO Box 378 Canberra ACT 2601;
Phone (02) 6249 9135, Facsimile (02) 6249 9984, Email: jon.stirzaker@agso.gov.au

Data Limitations: see 'NOTES'

Digital Capture Methodology: Data have been captured from stable-base repromat used in the production of the 1:250 000 1st and 2nd Edition printed geological maps. Acquisition was by high precision, high resolution scanning and vectorisation in conjunction with heads-up digitising over the scanned raster image. Point features (dips etc) were placed interactively using cells from AGSO's ***agsosymb.cel*** MicroStation cell library. Data were then converted into GIS (Arc/Info) format and built according to the AGSO Geological Data Dictionary.

Software: I/Scan, I/RasB - scanning, raster manipulation; ProVec - vectorisation; MicroStation V4 - interactive acquisition; Arc/Info Rev 7 - topological structuring (polygonisation etc), feature attribution

Online pricing (available July 2000)
As of July 2000 NTData tiles in ArcInfo format are available for purchase online. The price of tiles has been reduced (where possible) to reflect the significantly lower handling costs associated with online delivery. Tiles cost $110 (including GST), except for the most recent 'NGMA' data which remains at $535 (including GST). NGMA tiles include: Milingimbi, Arnhem Bay, Gove, Mt Marumba, Blue Mud Bay, Port Langdon, Cape Beatrice, Katherine, Mt Doreen, Hermannsburg. At this stage bulk order discounts do not apply for online NTData purchases.

Price Structure (when ordered through the AGSO Sales Centre)
The price of NTData varies according to currency of the source maps, the data format, the number of tiles ordered, and the number of users/licences (a multiplier % is applied). Discounts are available for bulk purchases and also where multiple formats are ordered at the one time, irrespective of the number of tiles ordered. This is summarised in the table below:

Single Tiles/Single Licence	***Bulk Orders***	***Multi-User/Licence***
1st Ed'n MicroStation **$215** Arc/Info, MapInfo **$375** **2nd Ed'n** All formats **$430** **NGMA** All formats **$535** (1995-98)	A sliding scale **discount** will apply for single orders of 20 or more tiles: **Less 25% Whole dataset** **Less 20% 50 or more tiles** **Less 15% 20-49 tiles**	2 users - **160%** 3-5 users - **200%** 6-20 users - **300%** 21+ users - **500%**

- Whole dataset pricing (discount applied, single user):
  MicroStation: $17 175
  Arc/Info $24 937
  MapInfo: $24 937
- Additional formats: where a customer wishes to purchase NTData in two or more formats prices will be negotiated but will be a *maximum* of 100% of the most expensive format plus 30% of each of the other formats.
- Educational institutions may be eligible for additional discounts. Please contact the custodian.

NOTES

• DATA QUALITY

Due to project resource constraints only limited quality assurance procedures have been applied. Data have been checked by visual comparison of plotted data (lines, labels) with printed maps at the capture stage (MicroStation DGN) and again after polygons have been built in Arc/Info, when full colour plots could be generated. Because not all 'polygons' are symbolised on the printed maps, some interpretation (chiefly of colours) and extrapolation (of labels) has been required at the data acquisition stage. It is possible therefore that a small number of polygons may be incorrectly labelled or missing. At this time we have done no statistical analysis of the accuracy or completeness of the data. Users finding errors or omissions are encouraged to fill out the Data Correction Report or contact the project manager directly.

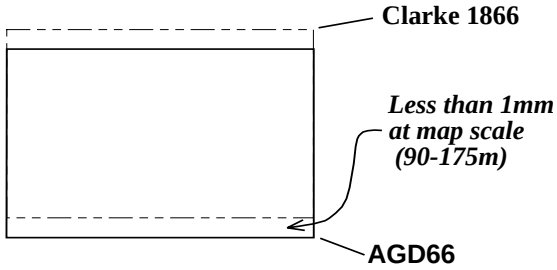
• LIMITATIONS

During the acquisition and processing of the NTData dataset we have encountered a number of issues which users should be aware of. Chief among these are interpretation inconsistencies (dealt with broadly under **Edge Matching**), and the positional accuracy of features on early maps (see **Positional Accuracy**).

Edge Matching

Although these data are not seamless coverages, where possible features crossing tile borders (map frames) have been edgematched. Geological unit labels however may differ from one tile to the next even where boundaries coincide. Differences may range from merely 'superficial' (eg Czl on one sheet may be labelled Tl on another but the rock unit descriptions are essentially the same), to unresolvable interpretation inconsistencies caused by relative ages of the maps, improvements in techniques etc (eg Clb abuts Pmt). Rock unit attributes such as *'Unitname'* are provided so that users can group related rocks beyond tile boundaries even when labels differ.

The accuracy and completeness of data along tile edges is particularly questionable where old (pre early 1970's) and new (1980's and 90's) maps abut. This is because historically spheroid differences were ignored for paper map production, the object being to ensure that where possible the units 'lined up' on the paper prints.



Positional Accuracy

It is inevitable that some positional inaccuracy will be present in digital data derived from archived map material. In the case of NTData we believe that there are three major causes for any such inaccuracy. Listed in order of their potential impact, these are: (i) *inaccurate topographic bases,* (ii) *spheroid/datum issues,* (iii) *map production processes*.

(i) Inaccurate topographic bases: A comparison of recent topographic information with the geological source maps indicates that the positional accuracy of the drainage on some early geological maps is, in places, very poor. This necessarily means that the geological features (both 'point' and line) carry a similar inaccuracy, as the geology is 'tied' to the drainage during the compilation processes. **[**Improvements in the geodetic control network and airphoto rectification techniques have contributed to the 're-positioning' of drainage detail on newer topographic map products.**]** The only current solution to this problem is re-compilation of the geology using modern bases and techniques and this is beyond the scope of the NTData project. Positional errors may vary locally within any single tile. Robinson River SE5304 contains examples of this type of problem.

(ii) Spheroid/datum issues: The NTData coordinate system is based on the Australian National Spheroid using the Australian Geodetic Datum (AGD66). However, in some cases the geological maps (the primary data source) were compiled on topographic bases which used an earlier spheroid, most likely either Clarke 1858 or 1866. Where we could establish that this was clearly the case, map data were assumed to be Clarke 1866 and were transformed to AGD66. This transformation results in the data being shifted about 150m to the north (the precise shift varies across the Territory). Our limited investigations - using GPS and geophysical image data - indicate that the shift places the data much closer to its true ground position. Incomplete/non-existent information about spheroids/datums on the source maps has however meant that the need for transformation is open to conjecture for some tiles. The possible error range for untransformed tiles (ie tiles which perhaps should be transformed and have not been due to lack of information) is 90-175 metres or less than 0.8mm at map scale.

(iii) Map production processes: NTData's source material (the geological map pre-plate repromat) was the culmination of many different cartographic and photo-mechanical processes. These manual procedures involved some 'best-fitting' of cartographic layers as films expanded and contracted during processing, and it is likely that they have contributed to positional inaccuracy of the digital data. The resultant error has been estimated at 125 metres or 0.5mm at map scale.

General: As a rule of thumb, data may not be suitable for use at scales larger than twice the capture scale ie 1:125 000, even where positional accuracy is high. The integrity (smoothness) of linear features may not be preserved if users choose to zoom in much beyond this threshold.

• DATA FORMATS

'CAD': *MicroStation DGN, AutoCAD DXF;* **GIS:** *Arc/Info, MapInfo*

MicroStation DGN: This is the data capture format (after vectorisation). Data are structured in layers (63 levels per DGN file, not all used) and generally separated into three files - **geo**: geology, **lin**: lines, **sym**: symbols. A fourth file, **pol**: polygons, has been constructed from translated Arc/Info data. Data are essentially 'CAD' quality with no attribution present, although the file/level structure provides excellent feature separation and facilitates extraction into other systems. Documentation of levels appears on lv63 in each DGN file. Data are projected (UTM, AGD66, ANS) and master units are kilometres (working units: km,m,1000,100). Limited checking has been applied to these data.

DXF: Currently these files are translated from the MicroStation data and retain their level separation, but contain no attribute data. Depending on customer requirements DXF may also be provided with attributes, using Arc/Info translation utilities.

Arc/Info (Rev 7.03): Data have been converted from MicroStation using *IGDSARC* and in-house AML's, and structured according to the AGSO Geological Data Dictionary (some early tiles still require upgrading to this standard). Polygons and feature attribute data have been added, and full-colour check plots created, making these data superior in quality and more complete than the 'CAD' formats. Data are projected (UTM, AGD66, ANS) and master units are metres. Arc coverages are provided as uncompressed Export files in double or single precision.

MapInfo Professional (V4.02): Data have been converted from Arc/Info using the *ArcLink (V3.1.2)* translator and retain their attributes and structure.

• POLYGON LABELLING

Initially, project specifications (as agreed with CLC) determined that only minimal letter symbolisation (polygon labelling) was to be carried out. However as the project progressed it became clear that the labelling of all units was a fundamental requirement, and the label coding system below was used in the MicroStation DGN files during the data acquisition stage.

Character	Meaning	Map Label	map_symbol	Other Examples
-P	Proterozoic 'P'	Pfb	-Pfb	Qa,Qr / Tla ➔ Qa/Tla, Qr/Tla
-C	Cambrian 'C'	COlj	-COlj	
p-C	pre-Cambrian	p Cg	p-Cg	
-R	Triassic 'T'	Rm	-Rm	Tla, Qr / Qc ➔ Tla, Qr/Qc
>	Subscript	Pg	-Pg>a	
^	Superscript	Cla	-Cla^1	
,	Combined units	Qa,Qs	Qa,Qs	Qs / Pla,Plm ➔ Qs/-Pla, Qs/-Plm
+	Combined units	Qa and Qs	Qa+Qs	
/	Overlying	Qs / Ar	Qs/Ar	

GIS *(Arc/Info, MapInfo)***:** The **GEOL** file has an item, **map_symbol**, which contains coded map unit labels (letter symbols on the map). The field width of this attribute has been extended to 30 characters to facilitate inclusion of combined units eg Pmei,Pmea,Pmp,Pmx,Pmnh (Bauhinia Downs 2nd Edition). Earlier data contained an additional attribute 'igds_text' to overcome this problem. The coding system in the table above applies to map_symbol.

Label inconsistencies

There is no complete reference list /look-up-table for NTData polygon labels due to inconsistencies between maps and the lack of resources to look at what is a very complex task. Therefore, users wishing to combine tiles to make seamless coverages need to be aware of three issues with regard to labels:

(1) Different labels may have been used for the same units eg Hermannsburg Sandstone is Dr on the Hermannsburg map and Pzr on Lake Amadeus.

(2) The same label may have been used for completely different units eg Pua on Hermannsburg is Areyonga Formation, while on Hay River Pua is used for the Grant Bluff Formation.

(3) Lithological descriptions can change significantly from tile to tile for similar geological units (this is often the rule rather than the exception). This may not be a significant issue for users unless they are using polylabels to search for specific rock types, in which case it would be better to use lith_desc as the search field. For seamless coverages, lith_desc would have to be amended to take into account the lithologies of each of the combined tiles.

Treatment of distinct lithological bands and 'marker' units

On some maps particular rock types ('lithological bands') within a geological unit have been depicted by patterns eg the basal conglomerate of Jasper Gorge Sandstone on the Auvergne map sheet. These will appear digitally as separate polygons, be labelled with subscripts eg Paj>b (although this is not how they appear on the map), and have distinct lithological descriptions, but have the same name as the 'parent' unit.

Geological units which were too thin to show as coloured outcrops on the printed maps (often referred to as 'marker' units and depicted by a single line on the maps, eg Julie Formation on the Hermannsburg map sheet) have been made into narrow polygons (10 metres wide) by creation of a buffer zone around the line. These 'artificial' polygons have the same attribute items as other polygons.

• AGSO CODE

agso_code is the main feature code for geological elements in the GIS data. It is an integer derived from the Symbol No. in the AGSO publication ***'Symbols Used On Geological Maps'*** (1989)**.** For example 5.5.1, the code for an overturned anticline(p15 of the book), becomes 551 as a feature attribute.

• COLOUR TABLES

Due to resource limitations there are currently no useful colour tables available for either the whole of the dataset, or for individual tiles.

• METADATA

A polygon coverage (**nt_index**) comprised of 1:250 000 map tiles covering the Northern Territory is available. Refer to **NTDATSTR.DOC** for information on this coverage.

Mandatory metadata as described by the Australia New Zealand Land Information Council's (ANZLIC) *Core Metadata Elements for Land and Geographic Data Directories* is provided in **NTMETA.DOC** (Word6 file)

• FILE NAMING CONVENTIONS

Coverages conform to an early version of the AGSO data dictionary standard with additonal coverages as required. Refer to **NTDATSTR.DOC** for a summary of the data structure for each coverage. For exporting these data we have adopted separate naming conventions for **UNIX export** (double precision uncompressed), **PC export** (single precision uncompressed), and **MapInfo**. Following are examples of coverages containing line, annotation, and polygon/point features from **Bauhinia Downs (e5303)** showing the variations:

Arc Coverage	**UNIX Export**	**PC Export**	**MapInfo tables** (.dat, .id, .map, .tab)		
basline	e5303bas.e00	e5303ba.e00	e5303bal	e5303baa	e5303bap
deposit	e5303dep.e00	e5303de.e00	e5303del	e5303dea	e5303dep
fossil	e5303fos.e00	e5303fo.e00	e5303fol	e5303foa	e5303fop
geol	e5303geo.e00	e5303ge.e00	e5303gel	e5303gea	e5303gep
geophys	e5303gph.e00	e5303gp.e00	e5303gpl	e5303gpa	e5303gpp
sites	e5303sit.e00	e5303si.e00	e5303sil	e5303sia	e5303sip
strlines	e5303lin.e00	e5303li.e00	e5303lil	e5303lia	e5303lip
struc	e5303str.e00	e5303st.e00	e5303stl	e5303sta	e5303stp
volcs	e5303vol.e00	e5303vo.e00	e5303vol	e5303voa	e5303vop
polygon	e5303pol.e00	**Available by prior arrangement only**			