FAILURE:
Mountains of data that provide no actionable information*

Peter Buxton
Program Coordinator - Geoscientific Information Management and Delivery

*© Matt Flynn CC BY-SA 3.0
GSSA Geoscientific Data

Over 135 years of geoscientific data from:
- government surveys and investigations;
- industry exploration and mining activities;
- academic research; and
- groundwater investigations.

- 300,000+ drillholes
- ~350,000 surface samples
- 1.5 Million downhole samples
- Over 9000 mineral occurrences
- 90,000+ field observations
- HyLogger data
- > 150 Gb of geophysical data
- ~20,000 Report Books and Envelopes
Innovation

Machine learning:
• Predicting outcrop (with Australian Institute of Machine Learning)
• Automated extraction of data from documents
• Predicting lithology from geochemistry

Geochemical Data Quality
• Understand the variability of our geochemical data (Uni SA - School of IT and Maths Sciences)

Web service delivery of 3D drill hole
• For more info come to Discovery Day
New Directions

Data completeness:
- Understand what we have, what we could have, and what users want
- Improve workflows for ingesting data

Data quality is understood and communicated
- Better metadata

Data is accessible and usable
- Web services
- Where possible, data is structured around national and international standards
- Wide table geochemistry
- Spatially enabled
Examples of geoscientific data audit: opportunities and challenges

Liliana Stoian
Senior Geologist – Geoscience Databases
20 Year Data Growth Statewide

Drillholes

Analytical Results

140,634 drillholes captured on database

2,774,598 rock samples with geochemistry results
Case study 1: Delamerian Project Data

Audit objectives and scope

Issue: How much data we hold but it is not captured into SA Geodata

Audit criteria: searching the data captured into SA Geodata using tenement lists over the area, the associated envelopes, and compare and document the results from the data captured vs raw data submitted

- Historical technical reports and envelopes
- Pre-digital data
- Digital data from active tenements

Results of the audit as for 1st July 2019
# Case study 1: Delamerian Project Data

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>OF data captured</th>
<th>OF to be captured digital files</th>
<th>OF Handwritten to be captured</th>
<th>OF typed non digital</th>
<th>Confidential data to be captured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drillholes</td>
<td>6958</td>
<td>15697</td>
<td>668</td>
<td>4384</td>
<td>439</td>
</tr>
<tr>
<td>Auger</td>
<td>3385</td>
<td>6613</td>
<td>1879</td>
<td>1531</td>
<td>0</td>
</tr>
<tr>
<td>Downhole Geochem</td>
<td>60944</td>
<td>149627</td>
<td>26538</td>
<td>28183</td>
<td>5973</td>
</tr>
<tr>
<td>Soil</td>
<td>3847</td>
<td>28148</td>
<td>10856</td>
<td>12628</td>
<td>1847</td>
</tr>
<tr>
<td>Rock/Outcrop</td>
<td>485</td>
<td>5547</td>
<td>5576</td>
<td>2119</td>
<td>323</td>
</tr>
<tr>
<td>Calcrete</td>
<td>1051</td>
<td>3896</td>
<td>0</td>
<td>168</td>
<td>0</td>
</tr>
<tr>
<td>Stream sediments</td>
<td>339</td>
<td>780</td>
<td>2518</td>
<td>8100</td>
<td>18</td>
</tr>
<tr>
<td>Niton XRF Surface</td>
<td>0</td>
<td>42203</td>
<td>0</td>
<td>0</td>
<td>5663</td>
</tr>
<tr>
<td>Innovex XRF Surface</td>
<td>0</td>
<td>17167</td>
<td>0</td>
<td>0</td>
<td>220</td>
</tr>
<tr>
<td>Delta XRF Surface</td>
<td>0</td>
<td>6059</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Case study 1: Delamerian Project Data

Mineral Drillholes 1st July 2019

- Pre 2003: manual data entry only
- Example 1 of 500 RAB drillholes captured
- Only bottom samples with geochemistry results were captured
- Surface samples were not captured

- After 2003-today: digital data
- Large volume of data becoming recently open file
- Data checking and validation: standard data submission, files with missing data & information
- Area was on low priority
- Resources
Case study 1: Delamerian Project Data

Surface Samples

- 65% OF captured
- 14% OF to be captured digital files
- 12% OF Handwritten to be captured
- 5% OF non digital
- 4% Confidential data to be captured
Case study 1: Delamerian Project Data

Outcome

- Identifying the areas for improvement: handwritten data, lithology, stratigraphy, petrology, petrophysics logs, surface samples
- Role of the government: big data – custodian of legacy and company data, reporting, data checking and validation
- Identifying the tenements with no work and no data
- What actions to implement: prioritizing data entry & filling the gaps
- Benefits for other areas of the State to do the audit: e.g. Central Gawler
Case study 2: Core Library Inspection and Sampling

Audit objectives and scope
Issue: How much data and rock samples with analytical results are not received from core inspections

Audit criteria: searching SA Geodata state wide for core inspections on mineral drillholes where samples were taken and results are required but not submitted

- Mineral Exploration Companies
- Tertiary Institutions (Australia and overseas)
- Other Government Agencies
Case study 2: Core Library Inspection and Sampling

<table>
<thead>
<tr>
<th>Total samples</th>
<th>Results received</th>
<th>Samples with no results</th>
<th>1/2 Core TS/Geochem</th>
<th>1/4 Core Geochem/TS/Geochron</th>
<th>Cuttings geochemistry</th>
<th>Shell/microfossils</th>
<th>Sedimentary analyses</th>
<th>Non-destructive</th>
</tr>
</thead>
<tbody>
<tr>
<td>32781</td>
<td>2477</td>
<td>30304</td>
<td>2453</td>
<td>15128</td>
<td>9838</td>
<td>618</td>
<td>1223</td>
<td>1044</td>
</tr>
</tbody>
</table>

8% of the samples with analytical data have been submitted to the Core Library and results are captured on database.
Case study 2: Core Library Inspection and Sampling

Samples with no results submitted:

- 1/2 Core TS/Geochemistry (50%)
- 1/4 Core Geochemistry/TS/Geo chron (33%)
- Cuttings geochemistry (8%)外
- Shell/microfossils (2%)外
- Sedimentary analyses (4%)外

注: "外"表示未提供结果的样本类别。
Summary of the audits

- Opportunity to identify the gaps in the data and looking into options to fix them
- Resources and planning (short term and long term)
- Looking into options to get non-digital data into database
- We need your help to return the data from core inspections, all parties need to take responsibility regarding data return to the Core Library
- What we learned and how we apply the results of the audits for other projects
- Are new mineral deposits hidden in legacy data? Where is the next target?
Next generation Digital SA Geology: Seeking your input

Rian Dutch
Program Coordinator – 4D Geodynamic and Metallognic Evolution of SA
The current state of SA Geology:
What we deliver

- All GSSA field notes and cartography is now captured digitally
- A workflow and product designed for print maps
- Most detailed – 100k not seamless
- Best available – not readily available
- Complicated and redundant stratigraphy
- Patchy explanatory notes availability and few up-to-date
- Limited attribution and important datasets not captured/delivered

Attributes
- 100K Surface Geology
- GL code
- Main Unit
- Stratigraphic name
- Stratigraphic description
- GIS code
- Parent name
- Parent symbol
- Province
- Age
- Mineage
- Minmod
- Maxage
- Maxmod
- Maxmeth
- 100K Linear Structure
- Description
- Symbol
Next generation Digital SA Geology:
The proposal

Deliver a digital by default new SA Geology

- Time constrained layer set
- Statewide, seamless dataset
- Integrated interpreted and exposed geological map layers
- Highest available resolution data
- The development of a complete and internally consistent stratigraphic database of units, which will form the basis for a digital ‘explanatory notes’ system.

- New geological map attributes and layers;
  - Spatial structural and kinematic data,
  - Metamorphic grade data,
  - Regolith and landscape,
  - Isotopic maps,
  - Tectonic element maps,
  - Major crustal boundaries,
  - Metallogenic maps,

Proposed geology layer set

- Rock Units (polygon)
- Structure Points (point)
- Geological Boundaries (polyline)
- Linear Structure (polyline)
- Trendline (polyline)

Proposed attribute types (e.g.)

- Dominant Lithology
- Alteration Modifier
- Regolith Material Unit
- Landform
- Igneous_Type
- Redox_State
- Fractionation_State
- Depositional_Environment
- BoundaryType
- StrucName
- FaultOrder
- FaultDipAngle
- Strike
- FaultCrustalDepth
- GravityVisibility
- Event1
- EventKin1
Next generation Digital SA Geology: The benefits

Geared towards digital delivery, not paper maps
- Scaleless, best resolution mapping
- Seamless, no need for map sheet boundaries

More robust data and metadata behind it (quality/consistency/quantity)
- More attributes and internally consistent datasets
- Data formatted/international standards and machine readable – AI and Machine Learning
- Readily updateable and flexible

New and important datasets included
- Not just traditional surface geology
- Integrated time constrained geology
- Structure and kinematics
- Non-traditional maps e.g. Isotope, crustal evolution, metamorphic, age, metallogenic

Review data and delivery of Stratigraphy
- Consistent stratigraphy across regions
- Update and fully attribute/define stratigraphic units
- Digital explanatory notes system to replace published notes/green book
- Easily updateable
Next generation Digital SA Geology:
A long term commitment

- Project commencement: 31/1/2019
- Begin sandbox implementation development and test case: 1/10/2019
- Complete systems review: 1/6/2020
- Approval to deploy and resource: 28/2/2021
- Finalise business case and project plan: 31/1/2021

- Project scoping and benchmarking: 31/1/2019 - 31/1/2020
- Internal consultation: 16/5/2019 - 31/10/2019
- External consultation: 20/9/2019 - 31/1/2020
- Develop sandbox architecture: 1/10/2019 - 1/4/2020
- Central Gawler test case development: 1/10/2019 - 3/2/2020 - 30/9/2020
- Consultation and review of test system: 30/9/2020 - 31/12/2020
- Review of processes and architecture requirements: 1/8/2019 - 1/6/2020
- Development of full business case to scale to enterprise: 1/6/2020 - 31/1/2021
- Commencement of Statewide compilation and deployment: ongoing
Next generation Digital SA Geology: What we need from you!

We want this new dataset to meet your needs now and into the future

So we need your ideas and feedback;
• What spatial geological attributes do you think we should capture and deliver
• What types of digital thematic maps would you like to see
• Do you have data you’d like to contribute, point us to open file data
• Anything else you can think of.....

Consultation via the DEM website. We will email you a link
DISCLAIMER

The information contained in this presentation has been compiled by the Department for Energy and Mining (DEM) and originates from a variety of sources. Although all reasonable care has been taken in the preparation and compilation of the information, it has been provided in good faith for general information only and does not purport to be professional advice. No warranty, express or implied, is given as to the completeness, correctness, accuracy, reliability or currency of the materials. DEM and the Crown in the right of the State of South Australia does not accept responsibility for and will not be held liable to any recipient of the information for any loss or damage however caused (including negligence) which may be directly or indirectly suffered as a consequence of use of these materials. DEM reserves the right to update, amend or supplement the information from time to time at its discretion.