A Geothermal Web Catalog Service for the Perth Basin

Thomas Poulet[1,2,3], Soazig Corbel[1,2], Michael Stegherr[1,4]
[1] CSIRO Earth Science and Resource Engineering, P.O. Box 1130, Bentley, WA 6102, Australia.
[2] Western Australian Geothermal Centre of Excellence, P.O. Box 1130, Bentley, WA 6102, Australia.
[3] The School of Earth and Environment, University of Western Australia, WA 6009, Australia.
[4] School of Science, Curtin University, GPO Box U1987, Perth, WA 6845, Australia

The Western Australian Geothermal Centre of Excellence (WAGCOE) is a science collaboration between Western Australia's leading innovative research institutions – CSIRO, The University of Western Australia and Curtin University of Technology. One of its key research initiatives is to assess geological and geophysical data from the Perth Basin to identify geothermal targets. This geothermal exploration task involves the collection and analysis of all datasets available, which can be a challenging and sometimes daunting exercise. This paper addresses the difficulties encountered in gathering data for geothermal exploration and presents our web catalog service, the application we use to manage our spatially referenced resources.

Keywords: Western Australia, Perth Basin, geothermal exploration, GeoNetwork opensource

Finding and using geo-referenced data

Geothermal exploration is a data intensive activity which requires access to all relevant information available in the region of interest, such as stratigraphic sections and horizons, downhole temperature measurements, rock properties, and hydrogeologic parameters including water quality. Because so many types of sciences are involved (geophysics, geology, geochemistry...), data may be derived from many different potential sources such as seismic surveys, petroleum well logs, mechanical property laboratories, or groundwater mineralogical assays. Data really is an indispensable element for exploration and yet it is still so difficult and expensive to discover, search, access and use; those four simple steps can sometimes represent a real obstacle course for any first-time user. Although most public datasets are nowadays accessible through companies' or governmental agencies' websites, any potential user still needs to know all key data holders in the area in order to access essential information. Once the appropriate data custodian for a particular type of data is identified, the second task is usually to search for data relevant to the location of interest, only to discover that datasets are rarely fully searchable. The third step is then to access this data, which can sometimes involve going physically to an organisation to copy several computer hard drives. The last step is generally to understand the format used, which may be non-standard and can lead to the most significant inefficiencies in a given project. Data manipulation can include for example large amounts of preliminary data conversion, deciphering any implicit assumption about units of measures and geographical references, or even trying to infer the creation process of the data in order to obtain one piece of missing information. If data is interrelated, this investigative process can loop on itself, creating workflow problems. Any user who has spent a long time going through this process can indeed wonder legitimately if newer data is now available. Our web catalog service is an attempt to eliminate the painful aspects of Perth Basin data discovery.

Using GeoNetwork opensource

GeoNetwork opensource[1] is a discovery service and catalog application, which is useful for managing spatially referenced resources and specially built to facilitate the connection between spatial information communities and their data. It is a Standardised Geospatial Information Management System based on international standards, with a large community of several hundreds of users and developers participating actively to its growth. A more complete description can be found in Ticheler & Hielkema (2007). In collaboration with AuScope Grid (Woodcock 2008) we decided to use GeoNetwork as our catalog application since it addresses most potential problems mentioned above, of data discovery, search, access and usage. GeoNetwork is accessible to users through a web interface and therefore only requires a web browser on any computer platform. Its search functionalities make data easily discoverable and

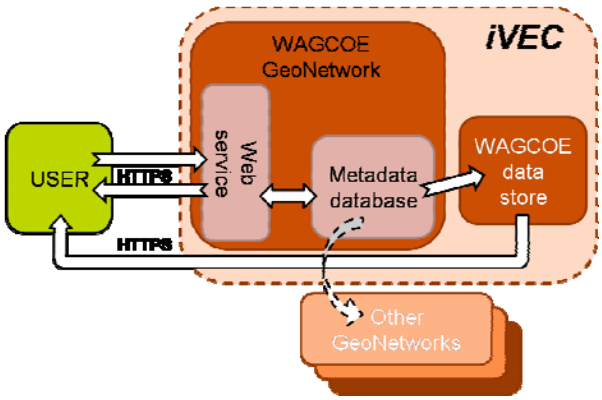

Figure 1: Workflow diagram showing the architecture of WAGCOE's geothermal catalog, using GeoNetwork as the catalog application linking to a petabyte datastore holding the corresponding data. Both components use iVEC's infrastructure and provide an efficient and secure system that can connect to other web services

[1] http://geonetwork-opensource.org

searchable. Since we are using GeoNetwork as a cataloguing application, the granularity of the search is limited to the metadata and not the data content itself. This library function is different from a web portal connected to a full database which would make the entire data fully searchable. However, the search facility allows users to find information easily and access the corresponding metadata, including a link to download the associated data itself from a Petabyte store. Figure presents the workflow of any user accessing WAGCOE's web catalog service. Both the web service and Petabyte store are hosted at iVEC[2], the hub of advanced computing in Western Australia, and all connections are secure. The web catalog service points directly to the data store, and can also be connected to other web catalog services from other institutions.

Using that data is also greatly facilitated as GeoNetwork organises, documents and publishes data collections in a standardised and consistent way, both at the metadata and data levels, following international standards from the Open Geospatial Consortium[3]. This fully descriptive metadata in XML format improves accessibility and removes any possibility of data misinterpretation based on assumptions.

As Ticheler & Hielkema (2007) point out, GeoNetwork has also been developed using Web 2.0 techniques to allow for more interactive and intuitive use of the system and to offer building blocks for future web services. This interaction with other web services is a central feature to help reduce duplication and working towards a single point of truth paradigm. It is an essential component of WAGCOE's data strategy to link to all data custodians' databases as defined by the Australian Geothermal Energy Group (AGEG) Technical Interest Group focusing on data management. Such a framework can definitely improve geo-referenced data sharing within and between organisations, hence developing better collaboration. It can also open opportunities for data comparisons which can empower consumers

Figure 2: screenshot of WAGCOE's GeoNetwork web interface. All data are searchable through metadata, including keywords and geographical location. Results are displayed with all corresponding metadata including links for download of the underlying real data.

[2] http://www.ivec.org

[3] http://www.opengeospatial.org

(Rezabakhsh et al., 2006). In our case the first external web service we are planning to interconnect is the web interface to CSIRO's PressurePlot[4], a query tool linked to the PressureDB database. That database contains rock properties such as porosity, permeability and thermal conductivity measurements, as well as stratigraphic and downhole pressure data for more than 1700 wells across Australia.

WAGCOE's web catalog service

Figure shows a screenshot of WAGCOE's GeoNetwork web interface[5] with three entries visible. Each entry is represented by its title and abstract, as well as two small pictures. A thumbnail on the right hand side represents the corresponding data and a logo on the left hand side indicates visually the origin of the data. All current entries appear with WAGCOE's logo as GeoNetwork is not connected to any other external service yet. Part of the data we store was produced by WAGCOE and includes measurements (see Figure) as well as results from numerical models. Most of the underlying data however was not originally created by WAGCOE and comes from other organisations such as the Western Australian Department of Mines and Petroleum, Geoscience Australia, or other private companies, research and governmental organisations. This local hosting demonstrates the temporary nature of those entries as we are working toward the single point of truth principle enunciated previously. We are only hosting other organisations' data with their permission as they agree for us to do while building their own web services. We are also planning to host data for geothermal companies who choose to use our system to manage part of their data.

As shown on Figure , our infrastructure is deployed at iVEC and benefits therefore from their professional services. All connections to GeoNetwork are encrypted and secure. A custom group management function then provides the granularity required to open some data to the public while restricting the access to protected company data, for example.

The data stored in our catalog covers all major scientific areas relevant to geothermal exploration and GeoNetwork's advanced options allow users to search for entries using the following two levels of categories and sub-categories:

- *remote sensing*: thermal infra-red, landsat, spots, light detection and ranging (LIDAR),

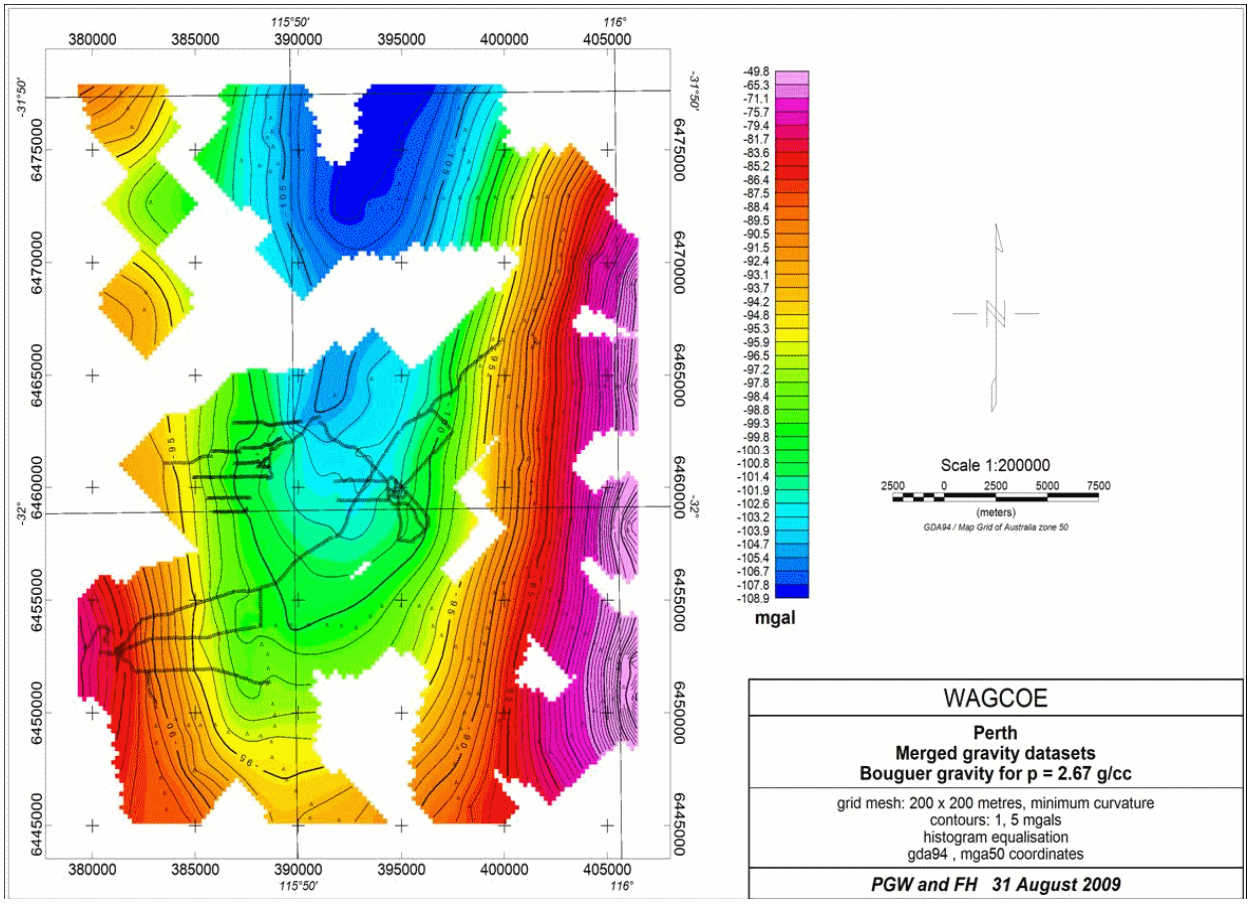

Figure 3: Example of data accessible through our catalog web service. This map shows Bouguer gravity from regional Geoscience Australia data combined with new survey data acquired by WA Geothermal Centre of Excellence.

[4] http://www.pressureplot.com

[5] https://wagcoe.ivec.org/geonetwork

• *geophysics*: seismic, velocities, magnetics, gravity, electomagnetics, radiometrics, geophysical logs, stress measurements, seismology data, magnetotellurics,

• *geology*: geological maps, palynology data, sedimentology data, thin sections, grainsize analysis,

• *hydrogeology*: hydraulic heads, water levels, water flows, rainfall, evapotranspiration data, recharge data, specific yield, storativity,

• *geochemistry*: water quality, rock geochemistry, thermochronology,

• *rock physics*: permeability, porosity, thermal conductivities, heat flux.

While all mechanisms have been implemented for us to manage these data types, populating the catalog with data references is a work in progress; not all sub-categories have data available yet.

WAGCOE's catalog management tools

We have developed a suite of administrative tools in order to build and manage our system as shown in Figure . We can therefore easily enforce consistency between the metadata on GeoNetwork and the corresponding data on the petabyte store. The user's group management in both places is also greatly simplified, although not completely automatic. The main feature of our administrative toolbox is a converter between GeoNetworks' metadata XML file format and an in-house extension of it allowing us to handle specific metadata information which is not supported in GeoNetwork by default and is not part of our customised version yet.

Figure presents a workflow diagram of all conversions occurring, between our WAGCOE metadata XML file format and GeoNetwork's original one, through an internal Python structure which allows easy data handling and processing, including checking validity, updating time stamps or creating unique identifiers. We can then store our extended metadata along with the original data and still retain the ability to efficiently evolve

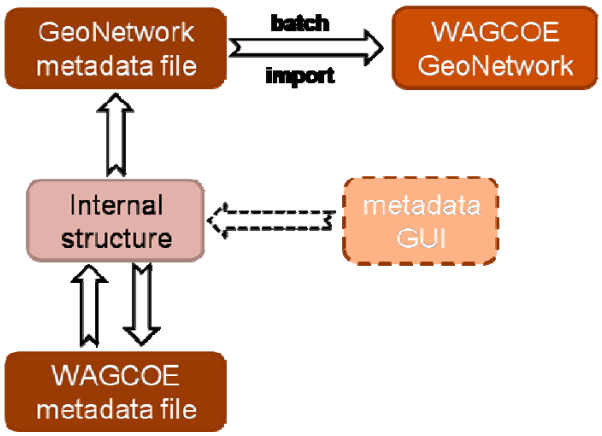

Figure 4: workflow diagram describing WAGCOE's metadata management

our GeoNetwork implementation through official updates or in-house customisations. All metadata information is stored in the most appropriate format for the particular data type, even though the additional information is displayed currently only as part of a "supplemental information" paragraph for each entry. Any future enhanced display handling in GeoNetwork can be easily accommodated with a script updating all GeoNetwork entries from our original extended metadata files.

Future work

WAGCOE's geothermal catalog is currently under development and even though nearly all major tools have been developed, some major work still remains to populate the catalog. The last major upgrade to our administrative tools remains a Graphical User Interface (GUI) to enter metadata more easily, as shown on Figure , instead of directly editing metadata files or internal Python structures. There are also several features planned for the future which will come from WAGCOE's strong collaboration with the AuScope Grid project (Woodcock 2008). The first major enhancement will be to connect our catalog to other web services, starting with CSIRO's PressurePlot web service. This linkage will provide access to a real pressure and temperature database allowing full content searches (whereas current searches on our catalog are limited to metadata). We are also planning to register our data collection to the Australian National Data Service (ANDS) Register My Data service[6] which will then automatically harvest all metadata and provide greater exposure to our public entries. Finally, we are monitoring closely the AuScope Grid portal development as it could provide an even more user friendly interface to our catalog. Van Oort (2009) pointed the cost of building such a portal, which only emphasises the importance of AuScope Grid's generic and open source tools development.

Conclusion

In the last few years, web services such as GeoNetwork have redefined the way users can discover, search, access and use geo-referenced data. This powerful tool is currently seeing a major uptake as it allows a dramatic change of practise for users who can now spend their time using data rather than searching for it. WAGCOE's GeoNetwork catalog brings this enormous advantage to the geothermal exploration industry and provides an easy, remote, and secure access to various geothermal data sources through an intuitive web interface. It helps to reduce the risk for geothermal exploration in the Perth Basin and could also help some

[6] http://ands.org.au/services/register-my-data.html

companies to store and access their data. We are hoping to connect it as soon as possible to other web services from various geological surveys, companies, research and governmental organisations, thanks to some of the technology developed by the AuScope Grid project

References

van Oort, P.; Kuyper, M.; Bregt, A. & Crompvoets, J. (2009), Geoportals: An Internet Marketing Perspective. Data Science Journal, 8, 162-181.

Rezabakhsh, B., Bornemann, D., Hansen, U., & Schrader, U. (2006), Consumer Power: A Comparison of the Old Economy and the Internet Economy. Journal of Consumer Policy V29, 3-36.

Ticheler, J. & Hielkema, J. U. (2007), GeoNetwork opensource, OSGeo Journal vol. 2, sep 2007.

Woodcock, R., (2008), AuScope: Australian Earth Science Research Information Infrastructure, Geological Society of Australia and the Australian Institute of Geoscientists. Australian Earth Sceinces Convention (AESC) 2008. New Generation Advances in Geoscience. Abstracts No 89 of the 19th Australian Geological Converntion, Perth Convention and Exhibition Centre, Perth WA, Australia. July 20-24, 276 pages.