Western Australian
Geothermal
CENTRE OF EXCELLENCE

Perth Basin Assessment Program

Project 1:
WAGCoE DATA CATALOGUE

Authors:
T. Poulet and S. Corbel

February 2012

CSIRO

Curtin University

The University of Western Australia
Achieving International Excellence

Government of Western Australia

Enquiries should be addressed to:

Thomas Poulet
CSIRO Earth Science and Resource Engineering

Australian Resources Research Centre
26 Dick Perry Avenue, Kensington WA 6151
thomas.poulet@csiro.au

ISBN 978-0-643-10903-2

Report 2

Project 1: WAGCoE Data Catalogue,
T. Poulet and S. Corbel - PRINT

Copyright and Disclaimer

© 2012 WAGCoE To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of WAGCoE.

Important Disclaimer

WAGCoE advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, WAGCoE (including its partners, employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

CSIRO

Curtin University

The University of Western Australia
Achieving International Excellence

Government of Western Australia

CONTENTS

EXECUTIVE SUMMARY I

PREFACE II

ACKNOWLEDGEMENTS III

1. INTRODUCTION 1

2. WEB SERVICES DELIVERING GEOTHERMAL DATA 3

2.1 Department of Mines and Petroleum 3

2.1.1 Western Australian Petroleum and Geothermal Information Management System (WAPIMS) 3

2.1.2 GeoVIEW.WA - Interactive GIS-based mapping system 4

2.1.3 Geoscience Publications 4

2.1.4 Western Australian Mineral Exploration Index (WAMEX) and GeoChem Extract 4

2.2 Department of Water 5

2.2.1 Water Resource Information Catalogue (WRIC) 5

2.2.2 Water Resource Data (WRDATA) 5

2.2.3 Other DoW data 5

2.3 Geoscience Australia 6

2.3.1 Geophysical Archive Data Delivery System (GADDS) 6

2.3.2 Geoscience Australia Petroleum Wells Database 6

2.3.3 Free Data Downloads system 6

3. WAGCOE DATA CATALOGUE 7

3.1 GeoNetwork opensource 7

3.2 Structure of the catalogue 7

3.3 Overview of the web interface 8

3.4 Management tools 9

3.4.1 Metadata Exchange Format 9

3.4.2 WAGCoE metadata format 9

3.4.3 Metadata management workflow 10

3.4.4 Metadata creation in GeoNetwork 11

3.4.5 Checking for broken links 11

3.4.6 Access protection 11

4. APPLICATION TO PERTH BASIN GEOTHERMAL EXPLORATION 13

4.1 Locating temperature data in the Perth Metropolitan area 13

4.2 Extracting temperature logs from well data 15

4.3 Sharing new temperature logs for certain wells 16

5. DISCUSSION 17

5.1 Content overview 17

5.2 Use of the data catalogue 17

5.3 Future directions 18

6. CONCLUSION 21

REFERENCES 23

APPENDIX A: LISTING OF EXTERNAL PARTICIPANTS 25

APPENDIX B: CONFERENCE AND WORKSHOP CONTRIBUTIONS 25

APPENDIX C: SCIENTIFIC PUBLICATIONS 25

APPENDIX D: EXAMPLE OF CATALOGUE ENTRY 26

EXECUTIVE SUMMARY

Most companies or governmental agencies in the earth science industry advocate a web interface to facilitate user access to their valuable data. However those interfaces are all different and make it difficult for any user willing to discover, search, access and use this data in an easy and standardised manner. Those needs are particularly evident in the recent field of geothermal exploration in Australia where no such workflow is currently available.

This report presents the Western Australian Geothermal Centre of Excellence (WAGCoE) data catalogue, an application of recent geo-referenced web service technologies to address the challenges of locating, querying and accessing data in standard formats for geothermal exploration. Practical uses of this catalogue for knowledge sharing and decision making within WAGCoE and the geothermal industry are also presented.

Two particularly important aspects of this catalogue are the presence of keywords in the metadata, facilitating the search for older reports and the programmatic management toolbox. This suite of tools is built on an internal structure used to represent any entry and helps the catalogue administrators to account for internal model upgrades and external technology changes, as all entries can be automatically generated and re-imported in the web catalogue.

PREFACE

The Western Australian Geothermal Centre of Excellence (WAGCoE) was established in 2009 for an initial three year term with a $2.3M grant from the Western Australian Department of Commerce. WAGCoE operated as an unincorporated joint venture between CSIRO, The University of Western Australia and Curtin University, with CSIRO as the Centre agent. The objective of WAGCoE was to assist the Western Australian Government provide a foundation for a sustainable geothermal industry by conducting advanced scientific and engineering research into WA's geothermal resources, principally hot sedimentary aquifer resources in the Perth Basin, and to develop and transfer to industry innovative new technologies for direct heat use.

In order to deliver on this task, WAGCoE was structured into three mutually supportive research Programs:

- Perth Basin Assessments (Program 1)
- Above-Ground Engineering (Program 2)
- Deep Heat and Future Resources (Program 3)

The Perth Basin Assessments Program contained the following four research Projects:

- WAGCoE Data Catalogue (Project 1)
- Perth Basin Geomodel (Project 2)
- Hydrothermal Simulations (Project 3)
- Reservoir Productivity and Sustainability (Project 4)

This document is the final research report of Project 1 within the Perth Basin Assessments Program of WAGCoE. The document is a summary report; detailed tabulations of technical data and results are to be found in the supporting research reports, papers and thesis cited herein.

ACKNOWLEDGEMENTS

We thank the AuScope Grid development team and the GeoNetwork opensource developer community for their support and involvement, especially Ryan Fraser, Terry Rankine and Michael Stegherr.

This work has been supported by iVEC through the use of advanced computing resources provided by the Australian Resources Research Centre (ARRC) facility located at the ARRC Building, 26 Dick Perry Avenue, Technology Park, Kensington WA 6151.

This work has been funded by the Western Australian Geothermal Centre of Excellence, with some in-kind contributions from the CSIRO Minerals Down Under Flagship.

1. INTRODUCTION

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, including geophysics, geology, and 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 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.

This report presents the WAGCoE catalogue, a web service to help reduce some of the difficult aspects of data discovery mentioned above, which are valid in general and specifically for geothermal exploration within the Perth Basin. The Western Australian geothermal industry is indeed relatively new and a large amount of data already exists which is essential for their need but has been collected for other purposes initially. The WAGCoE catalogue facilitates the tasks of locating, querying, and accessing some of this data, which includes the spatially referenced resources created within the centre.

The report is articulated in four parts. Section 2 presents existing services from various public agencies which represent a major source of information. In Section 3 we introduce the WAGCoE catalogue, its structure and management tools. In Section 4 we illustrate an application of the catalogue to geothermal exploration in the Perth Basin, and Section 5 discusses the usage of the catalogue and future directions.

2. WEB SERVICES DELIVERING GEOTHERMAL DATA

Western Australia hosts important mineral and petroleum resources and as a result offers many sources of geological, geophysical, hydrogeological and geochemical data. These data can be obtained from the literature, various public institutions, or private companies. This section presents the main public sources used within WAGCoE. They include, in no specific order, the Western Australian (WA) Department of Mines and Petroleum, the WA Department of Water and the federal Geoscience Australia agency.

2.1 Department of Mines and Petroleum

The Department of Mines and Petroleum (DMP) provides geoscientific information on mineral and energy resources and manages title systems for the mining, petroleum and geothermal industries. It also gathers data acquired by those industries and make them available on the DMP website (1). DMP offers several web interfaces to discover various types of data including geochemistry data, geological maps or petroleum and geothermal information. The online systems relevant to geothermal and used by WAGCoE are summarised in the following subsections.

2.1.1 Western Australian Petroleum and Geothermal Information Management System (WAPIMS)

This service is mainly providing data from:

- petroleum wells, including any report related to the well and geophysical logs
- seismic surveys
- other geophysical surveys

WAPIMS offers a Data Browser, including a search functionality (eSearch) to query the data and an Exploration Map to query the data by location. All data can be downloaded from the website, but a login is required. WAPIMS is updated almost every six months and represents an important source of information for WAGCoE.

The eSearch option enables to query the database on different criteria. However it is also possible to download spreadsheets of petroleum wells analysis per basin, in Data Browser Public > Information Links > Download Well Analysis Data. The spreadsheets contain data such as porosity and permeability, Drill Stem Tests data, formation tops, or water analysis. Queries can be easily performed from those spreadsheets for users who prefer the convenience and flexibility of working with their favourite spreadsheet software rather than the eSearch option.

2.1.2 GeoVIEW.WA - Interactive GIS-based mapping system

This service makes the following data accessible:

- topography
- geochronology data location
- geochemistry data location
- geology
- mineral and petroleum information
- geothermal titles
- geothermal lease applications and others.

GoeVIEW is a mapping system that enables users to download layers to make geoscientific maps and to extract data from the screen view, up to a scale of 1:1000000. State-wide datasets in GIS formats are also available for download in the Data and Software Centre system.

The catalogue Search option is linked to the Geoscience Publications system and it is possible to search for publications in a selected area of GeoVIEW. GeoVIEW also contains some seismic line locations that are not available in WAPIMS.

2.1.3 Geoscience Publications

This service mainly provides data from publications, maps or reports related to geosciences in Western Australia.

Geoscience Publications is a document delivery system that enables to view, print and download any publication, map or report of the Geological Survey of Western Australia (GSWA). It is useful to gain access to geochemical, hydrogeological and geological reports. The system provides a Search for Documents tool to query the publications using many different keywords.

2.1.4 Western Australian Mineral Exploration Index (WAMEX) and GeoChem Extract

WAMEX and GeoChem were only used a few times within WAGCoE. However they are cited here as they may be of interest for more geothermal studies.

WAMEX gives access to open-file reports on exploration for minerals. To extract data such as formation tops or well location, it is necessary to go inside the reports. No data were extracted for the mineral reports in the form of a database like WAPIMS.

GeoChem Extract gives access to geochemistry data over WA. The only way to query the data is using the Interactive Map Sheet Selection or by geographical location. The data can be downloaded as spreadsheets, one spreadsheet per dataset.

2.2 Department of Water

The Department of Water (DoW) is the best source for shallow hydrogeology data. DoW manages the State's groundwater and surface water resources and ensures adequate water services. Their website (2) also provides publications, maps, atlas and groundwater monitoring data throughout Western Australia. Like DMP, DoW currently provides a few different web interfaces, depending on the type of data served. Those useful for geothermal data are:

2.2.1 Water Resource Information Catalogue (WRIC)

This service provides:

- Flow data
- Water quality sampling data
- Water level data
- Rainfall data

The WRIC online system enables to search water data by map or by criteria. However, once the data have been identified, users cannot download it from the website and need to make a request for information by email.

2.2.2 Water Resource Data (WRDATA)

This service provides:

- Summary data for Rainfall sites
- Summary data for Streamflow sites
- Summary data for Groundwater sites

All datasets are accessible from their Site ID or Site name, so a preliminary usage of WRIC (see subsection 2.2.1) is required. It is also possible to access WRDATA from WRIC directly. All datasets are available in text file format or zipped CSV format for water levels.

2.2.3 Other DoW data

Not all of DoW's data is available through their web services and must be obtained directly. WAGCoE contacted DoW to locate hydraulic conductivity data for the Perth Metropolitan area for example. As DoW requires all operators to submit a bore completion report after drilling a well, many such reports have been collected which contain precious hydrogeologic data such as estimated hydraulic conductivities or storage coefficients when pumping tests were performed. Due to legal limitations, these reports cannot be released to the public but can be accessed at the DoW offices.

DoW does not currently have the funding to search these reports and compile hydrogeologic data in point locations. After discussions between WAGCoE and DoW, DoW agreed that WAGCoE could access the reports and compile the measurements. The data would then be returned to the DoW for their dissemination to the public. Initial investigations in the Perth Metropolitan Area identified at least 25 reports which should contain quantitative hydrogeologic information. Future funding should be targeted towards collation of this essential data.

2.3 Geoscience Australia

Geoscience Australia represents Australia's national geological survey and as such is the custodian of a huge amount of data including geological and geophysical maps, 3D models, geochemistry and geochronology data. Their website (3) gives access to several web interfaces delivering those different data.

2.3.1 Geophysical Archive Data Delivery System (GADDS)

This service provides:

- Magnetic data
- Radiometric data
- Gravity data
- Digital Elevation data

The data is searchable by location, either by entering the extents, using a map query tool or using 1:250 000 map sheets areas. Once the dataset is identified, it is necessary to request it.

2.3.2 Geoscience Australia Petroleum Wells Database

This service provides data from petroleum wells over Australia, including geochemistry data, reservoir/facies data and biostratigraphy data.

The well database is searchable by different categories. It is possible to extract graphs and Excel spreadsheets for each well. The Geoscience Australia Petroleum Wells Database includes all that is in WAPIMS plus additional wells.

2.3.3 Free Data Downloads system

This service provides mainly:

- geology data
- geophysics data
- geochemistry data
- geochronology data
- digital elevation data

The data can be searched for an area of interest. We mainly used this system to access well temperature data from OZTemp database (4).

3. WAGCoE DATA CATALOGUE

Due to the difficulties encountered (see Section 1) to access geothermal data from the wide range of sources presented in Section 2, WAGCoE decided to implement a new web application dedicated to geothermal exploration, based on the above listed requirements, using GeoNetwork opensource (5). The WAGCoE catalogue (6) can be accessed at the following address: http://www.geothermal.org.au/Catalogue.htm

3.1 GeoNetwork opensource

GeoNetwork opensource is a free and open source catalogue-repository solution, designed to manage spatially referenced resources through its web interface (7). It is specially built to facilitate the connection between spatial information communities and provides a suitable technical solution that addresses all WAGCoE data requirements. GeoNetwork opensource is widely supported by the geoscience community and uses common standards specifically designed for geographic data (ISO19139 and Federal Geographic Data Committee) as well as the international standard for general documents (Dublin Core). As free and open source software, it also benefits from a large community of geospatial users and developers, who contribute actively to further advance the software and its increasing popularity.

GeoNetwork opensource is also a service catalogue supported by the AuScope Grid project (8), a part of the AuScope initiative for an Australian collaborative research infrastructure strategy for geosciences. One of the major aims of AuScope Grid is to deliver a software infrastructure to provide the geoscience community with interoperable access to data. This project is currently working with several Australian geological surveys and government agencies, helping build web services to make their data more easily available and exchangeable using interoperable formats. AuScope Grid is supporting the development of GeoNetwork opensource for the needs of WAGCoE.

3.2 Structure of the catalogue

GeoNetwork opensource represents a managed catalogue-repository solution with a web service interface. A catalogue is a formal mechanism for allowing interested parties to announce or discover the availability and state of any resource. A repository is the physical or virtual facility where those resources are stored. As shown in Figure 1, the GeoNetwork instance of the WAGCoE data catalogue does not make use of the GeoNetwork

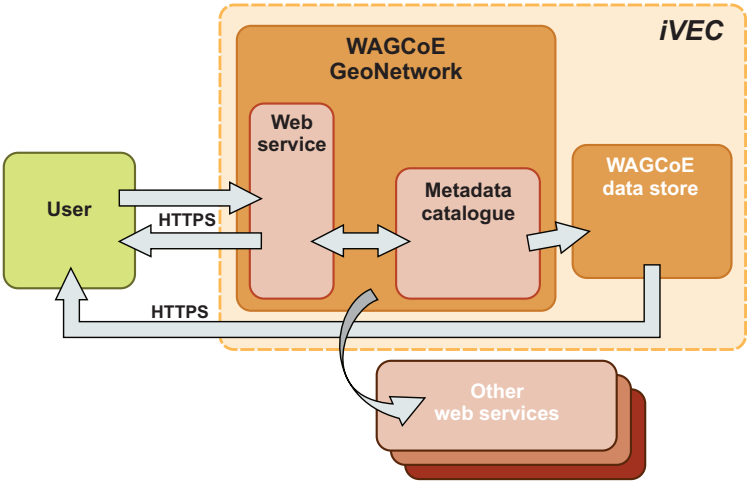

Figure 1: Structure diagram of WAGCoE data catalogue. Both GeoNetwork and its associated data store are hosted at iVEC, the hub of advanced computing in Western Australia.

repository component and uses instead a file based system in a separate data store. The choice to use a separate file system repository in a Petabyte Data Store allows WAGCoE members to access this data easily while the whole system was being developed. Both the GeoNetwork web service and the Petabyte Data Store are hosted at iVEC, the hub of advanced computing in Western Australia (9), which provides services including the backup of all data for catastrophe recovery.

Users can access the WAGCoE data catalogue via a secured connection using HTTPS protocol. All metadata entries are subject to GeoNetwork groups and users management mechanisms which allow specification of access rights within the application. Users with appropriate credentials can access metadata entries and download the associated physical data through secured connections.

3.3 Overview of the web interface

GeoNetwork opensource offers a user friendly interface divided in two parts (Figure 2). In the upper part of the interface, users can query the data catalogue using free text, geothermal oriented keywords and categories. They can also search by location, date of publication in the catalogue or date of acquisition of the data. The results of the query are displayed in the lower part of the interface. Users can click through the different metadata entries and download all associated data. An example of a full entry with all its associated metadata is presented in Appendix D: Example of Catalogue entry.

Figure 2: Screenshot of the GeoNetwork instance developed for the WAGCoE data catalogue. This example shows a result of the query for all entries associated with the keyword "Thermal Infra-Red" and located in the south-western part of Western Australia (10).

3.4 Management tools

Some time was invested at the start of this project to build appropriate management tools, which ensured a scalable development of the catalogue. This section describes such existing or developed tools and techniques used.

3.4.1 Metadata Exchange Format

GeoNetwork opensource provides additional services which make it an easy-to-use tool. For example, generating catalogue entries can be automated with a batch import using Metadata Exchange Format (MEF). MEF files enable to define the metadata associated with a resource, its category, its privilege and associated thumbnails(5). Using the MEF batch import ensures that all data are registered properly and following the metadata standards.

3.4.2 WAGCoE metadata format

The ISO19139 profile used by GeoNetwork opensource to describe geographic data based on the ISO19135 standard is however not specifically designed for geothermal data and a large proportion of the metadata is therefore irrelevant or inappropriate for our specific purposes. To address this issue WAGCoE is involved within a Technical Interest Group (TIG8)(11) from the Australian Geothermal Energy Group (AGEG) to identify and if necessary create a specific geothermal GeoNetwork profile. While a community-agreed vocabulary is being defined, WAGCoE is using a vocabulary and controlled concepts defined internally. We also designed different metadata templates for geothermal data by adapting the standard to fit our needs. Five data categories were identified for all geothermal exploration data: interpolated maps, interpreted maps, well data, seismic data and measurements.

A suite of administrative tools was developed, using the Python programming language, to build and manage a sustainable workflow for populating and maintaining the catalogue. The main purpose of this programmatic toolbox is to address the common problem of consistency of all entries, both across people creating them and over time. The format of each entry in GeoNetwork is indeed flexible enough to accommodate some variability, which can easily lead to a wide range of discrepancies though the catalogue. Imposing the creation of all catalogue entries through a script then allows some consistency checks to be performed.

A specific list of keywords was defined for each of the five data categories. A configuration setup then defines which keywords are compulsory or optional, which data types are valid for each keyword, and other validity rules based on the existence and values of various combinations. A trivial illustration is that numerical values for example are checked to be within bounds. For all keywords, associated values other than free text are enforced to be members of predefined allowed sets prescribed in the configuration setup. This set of keywords constitutes our internal format, for which we defined an internal WAGCoE metadata XML file format to save the data on disk.

3.4.3 Metadata management workflow

Figure 3 presents a workflow diagram involving WAGCoE metadata XML files, the internal Python programming structure (based mainly on Python dictionaries) and GeoNetwork metadata files. In terms of format conversions, the management toolbox allows administrators to perform the following actions: (i) reading an XML file to get the corresponding internal structure, (ii) writing an internal structure to an XML file and (iii) writing an internal structure to a GeoNetwork metadata file, ready to be imported directly in GeoNetwork. This workflow allows us to keep a backed up copy of all our catalogue entries in WAGCoE metadata XML format, from which we can easily recreate all GeoNetwork entries.

This is a major feature of our workflow to ensure consistency across the whole catalogue. Since most metadata is represented by keywords (except for free text entries like titles, abstract, etc.) we have then separated as much as possible the content from its display in GeoNetwork. This mechanism allows us to accommodate changes which invariably occur over time, where the displays of specific fields need to be changed for all entries.

For example, the provenance for all entries is stored as a keyword in a list of allowed values. The keyword for WAGCoE data is then "WAGCoE", which is displayed in the web interface as "WA Geothermal Centre of Excellence (WAGCoE)", and we can change this label for all entries easily. This mechanism proved extremely useful with time as labels were changed and also when the GeoNetwork file format changed slightly with different versions of the software.

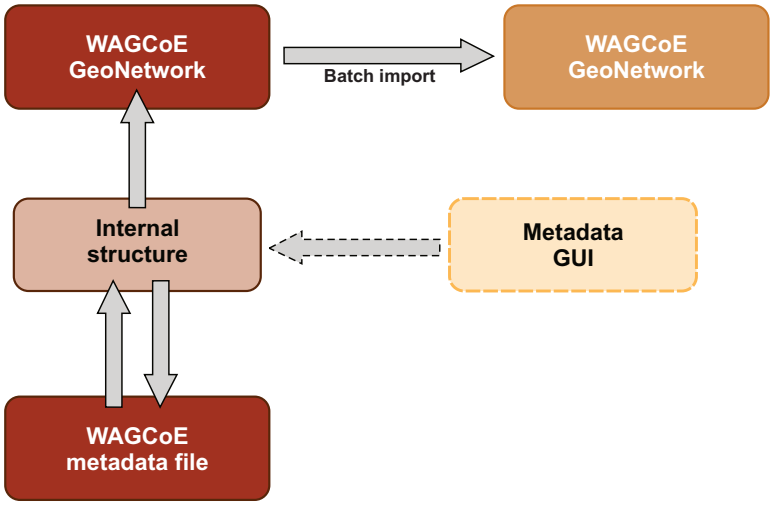

Figure 3: Workflow diagram describing WAGCoE's metadata management

The system is constantly changing as the tools and formats used are in continuous development, and accounting for this evolution is an important part of the catalogue management. For example, as the geothermal community is progressively defining its needs and corresponding standards to represent them in GeoNetwork, we are currently using some WAGCoE metadata keywords which do not have any appropriate place of their own in GeoNetwork. This information is therefore stored with well-defined keywords in our internal format, but simply appears as text under the "Supplemental Information" tag in the web interface. As standards evolve and new keywords are created, the amount of information displayed as supplemental information is reduced, and we can accommodate every step along this process by updating all entries when such steps occur.

3.4.4 Metadata creation in GeoNetwork

The administrative toolbox also handles file transfers and helps respect rigorously the procedure put in place. As mentioned above, all metadata are kept in WAGCoE XML format on iVEC's petabyte store as the single point of truth for the whole catalogue, in a separate directory from the real data it points to, and all this data is backed up automatically to another physical location by iVEC. At any time, we can regenerate the whole catalogue. This has been done a number of times over the duration of the Centre, following the evolution of our internal format, GeoNetwork format, or simply to improve the display. Another important advantage of our methodology is that all entries are now limited to a WAGCoE metadata XML file, whose size and complexity are one or two orders of magnitude smaller than for the complex GeoNetwork XML file.

Anyone within the Centre can create new entries for the catalogue by editing one of the existing templates for the data type selected and providing this file with the corresponding data to one of the administrators. The administrator can then use a programmatic tool to check the validity of the entry, generate the corresponding GeoNetwork MEF file (archive containing the metadata and permission files), upload that file on the server running the web service, and upload the WAGCoE metadata to iVEC's petabyte store for storage. All those steps were thoroughly documented on WAGCoE's collaborative web site (using Twiki (12) technology), along with all administrative procedures, but they were also automated to remove any human error. All scripts were also developed following best practices in software management, using revision control and implementing appropriate unit tests to automate defect prevention (13).

3.4.5 Checking for broken links

Storing all metadata in our internal structure also provides full flexibility to use the Python programming language for various tasks, including checking the validity of all links. The catalogue is not a full database, and the data is stored separately from the metadata. This choice was made to allow researchers in the Centre to access the data outside from the web interface, simply by browsing the directory structure on iVEC's petabyte store (with read access permissions only). There is therefore no automatic mechanism to update the metadata of existing entries if the corresponding data changes location. The administrative tools can check if links are broken automatically so that any administrator can fix the problem. This functionality minimises the negative experience of users discovering broken links when browsing the catalogue.

3.4.6 Access protection

The user's group management is also greatly simplified with some management tools, although not completely automatic. The creation of all entries includes a definition of access permissions which are set at a group level. GeoNetwork is connected to one of CSIRO's internal user database through Lightweight Directory Access Protoco (LDAP), allowing us to use the same user definition both for GeoNetwork and the petabyte store. Current limitations of the software however do not allow us to use LDAP groups directly, and our management tools cover this limitation by ensuring that groups on both sides (for the metadata and data) are identical. The management tools also generate

automatically the appropriate server configuration files (Apache htaccess files) from all internal metadata entries. This feature is another critical justification of our choice to store all metadata in internal format as the single point of truth.

4. APPLICATION TO PERTH BASIN GEOTHERMAL EXPLORATION

An important task to be carried out by WAGCoE in the Perth Basin is to assess the temperature coverage of the Perth Metropolitan area and its surroundings. In this section we are using this work as an application of the WAGCoE data catalogue.

4.1 Locating temperature data in the Perth Metropolitan area

The WAGCoE data catalogue currently contains the five generic types of data. Each entry of the catalogue contains all metadata associated with its data type as well as links to the files corresponding to the data type of the entry (including reports or measurements).

Locating a specific geothermal resource in the WAGCoE data catalogue has been simplified due to the implementation of a multi-level keyword approach. Each entry is associated with one or more sets of at least two levels of keywords from a predefined table (Table 1).

For example, using the keyword “Rock Physics/Temperature”, the WAGCoE data catalogue returns all data associated with any kind of temperature data. Applied to the larger Perth metropolitan area, this query returns, the artesian monitoring bores from DoW containing temperature logs as well as the few petroleum wells from DMP which contain bottom hole temperature measurements in their completion report.

Geochemistry	Geochronolgy	
	Rock geochemistry	Fluid inclusion analyses Gamma-ray spectrometry Stable isotope analyses XRD analyses XRF analyses
	Thermochronology	
	Water quality	
Geology	Core photography	
	Geological maps	
	Grainsize analysis	
	Lithology description	
	Palynology data	
	Petrography data	Diagenesis Mineralogy Optical photomicrographs Provenance analysis
	Sedimentology data	Facies analysis Sedimentary log
	Seismic interpretation	
	Stratigraphy interpretation	
	Structural data	Orientation data
		Structural analysis
	Structure maps	
	Thin sections	
Geophysics	Electromagnetics	
	Geophysical logs	
	Gravity	
	Magnetics	
	Magnetotellurics	
	Seismic	2D Reflection 2D Refraction
	Seismology data	
	Stress measurements	
	Velocities	
	Vertical Seismic Profile	
Hydrogeology	Evapotranspiration data	
	Hydraulic conductivity	
	Hydraulic heads	
	Rainfall	
	Recharge data	
	Specific yield	
	Storativity	
	Water flow	
	Water levels	
Remote sensing	Bathymetry	
	Digital Terrain Model	
	Landsat	
	Light Detection and Ranging	
	Spots	
	Thermal Infra-Red	
Reservoir engineering	Pressure testing	
	Numerical simulation	
Rock physics	Density	
	Heat flux	
	Permeability	
	Porosity	
	Temperature	
	Thermal conductivity	
	Water saturation	

Table 1: List of keywords defined on three levels for WAGCoE data catalogue

4.2 Extracting temperature logs from well data

Each entry in GeoNetwork opensource displays its metadata, including a "Distribution Information" section containing links to download the associated data and their respective sources. Well metadata also display available well logs and the format of the corresponding files. Figure 4 presents the example of the petroleum well Barberton 1. It shows that users can read and directly query information associated with that well (calliper, gamma-ray, temperature and other logs in LAS format) without having to open and search the corresponding reports. In that example two reports are associated to the well and available for download.

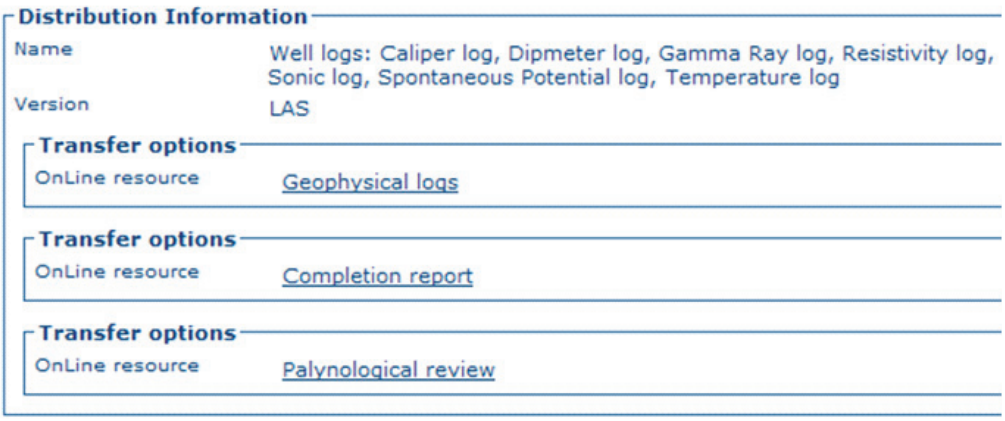

Figure 4: Screenshot of the Distribution Information section of metadata entry of petroleum well Barberton 1.

To simplify the work of users, metadata have been designed to contain all information necessary to import the data in a modelling package (Figure 5). Users can also see the provenance of the data and relevant copyrights are associated to it. They can then download data from the catalogue and use them immediately.

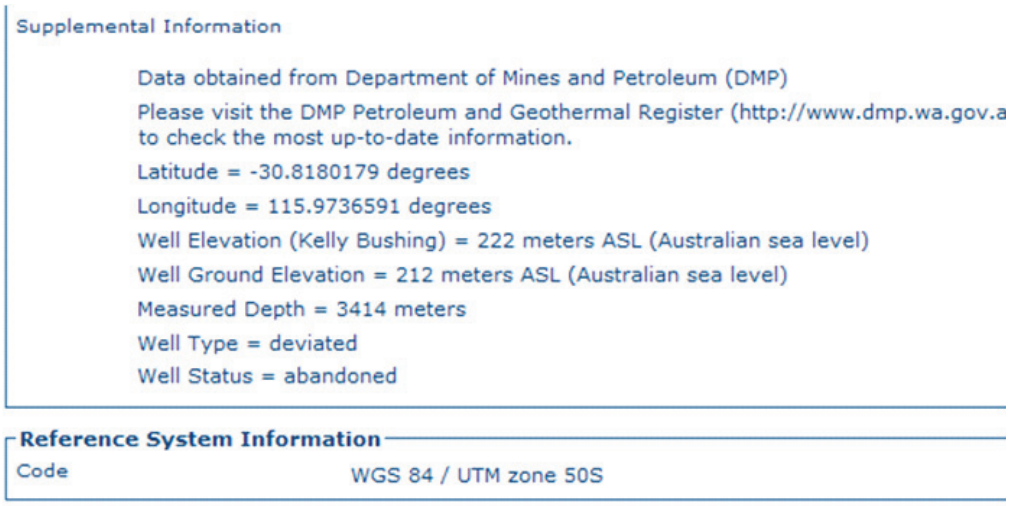

Figure 5: Screenshot of the Supplement Information section and Reference System Information section of metadata entry of petroleum well Barberton 1

4.3 Sharing new temperature logs for certain wells

Most water bores from the Department of Water have geophysical logs associated to them, providing information such as gamma-ray, resistivity or temperature logs. Those logs can be requested from DoW but can only be obtained in PDF format, including images of temperature or gamma-ray plots with depth. Those plots are extremely informative, but the image format is inappropriate to input the data directly into a modelling package. WAGCoE therefore started to digitise existing temperature logs, 55 in total, and also acquired 17 new temperature data from DoW boreholes. All those new resources are available in the data catalogue. Figure 6 shows an example of temperature data available in the data catalogue.

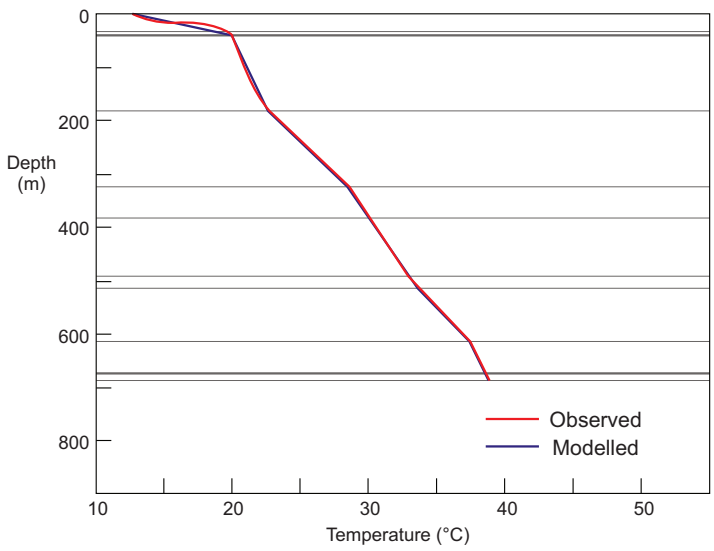

Figure 6: Some data entries come with a thumbnail giving an overview of the data. This thumbnail is associated with DoW artesian monitoring well AM40. The logged temperature (in red) is superimposed over a modelled temperature profile (in blue), from WAGCoE's report on Temperature regime in the Perth Metropolitan Area (14), also available in the data catalogue.

WAGCoE temperature logs were created in format LAS 2.0, a standard format for well logs defined by the Canadian Well Logging Society (15). Seismic data gathered by WAGCoE in the Perth Basin are also stored in SEGY rev1 standard format (16) to ensure that geophysical and seismic data are self sufficient and can be imported directly into a modelling package.

5. DISCUSSION

5.1 Content overview

A summary of the data available currently in the data catalogue is presented in Table 2. This table is only provided for information as an overview of the content and does not reflect the real flexibility of the catalogue which handles various data types from different sources for some of its entries. The WAGCoE catalogue was designed to handle different data types and contains diverse data such as 3PDF, LAS files, Excel spreadsheets, GIS files, PDFs and others.

Number of data entries	Data type	Main associated data	Main data source
194	Petroleum wells	Reports and geophysical logs	DMP
72	Water bores	Temperature logs	WAGCOE
5	Geological maps	Maps	DMP
3	Geophysical data	Raw data	GA
14	Geophysical data	Raw data	WAGCOE
1	Seismic data	Raw data and report	WAGCOE
3	Rock physics measurements	Report and raw data	WAGCOE

Table 2: Summary of data available in WAGCoE data catalogue.

5.2 Use of the data catalogue

The WAGCoE catalogue includes an access protection mechanism (see Section 3.4.6) and not all data is public. Some entries are restricted to WAGCoE researchers and collaborative companies, with different levels of protection, and require users to request credentials to log in. This can be done through the web interface which contains a “contact us” link to email the catalogue administrators (see Figure 2). There are currently 18 users in 4 groups registered in the data catalogue. The catalogue was largely used to access completion reports from petroleum wells containing specific information (through the search functionality, see Section 3.3). However, here are examples of more specific needs for different projects:

- Locating structure maps for the Perth Basin

Sebastian Gutbrodt, a WAGCoE student, used the data catalogue to gather structure maps available in the Perth Basin. Some petroleum reports contain maps with interpreted faults, folds or formation depth, but this information can be hard to locate as it often constitutes a minor part of the reports which is presented in passing under a section with a different focus. Using the data catalogue, locating those reports become easy using the keywords Geology/Structure maps, Geology/Geological maps and Geology/Structural data. Sebastian therefore compiled a list of structure maps that were used for the 3D model of the entire Perth Basin (see associated report “Re-assessment of the 3D structural and stratigraphic architecture of the Perth Basin”, Wilkes et al. 2012).

- Gathering salinity data in the Central Perth Basin

Sarah Glasson, another WAGCoE student, investigated salinity within the Yarragadee Aquifer in the Perth Area (17). Sarah used the data catalogue to identify the petroleum wells containing Spontaneous Potential or resistivity logs in the Central Perth Basin and then downloaded the geophysical logs. Querying well data on their geophysical logs is quite easy in the data catalogue. Geophysical well logs are part of the metadata and input in a standard way. Therefore a simple search using the words "Resistivity Spontaneous Potential" returns all the well data with that kind of geophysical logs and using the Search by Extent tool, Sarah limited her search to the Central Perth Basin.

- Accessing Australia Mean Annual Surface Temperature (MAST) for the Perth Basin

Florian Wellmann, from the hydrothermal modelling group of WAGCoE, used the data catalogue to directly access MAST dataset compiled by Frank Horowitz (10). The raw data is available for download in a format compatible for modelling packages. Therefore Florian was able to easily input it into the hydrothermal modelling software (Figure 7) for the hydrothermal modelling work on the entire Perth Basin (see associated report by Reid et al. 2012).

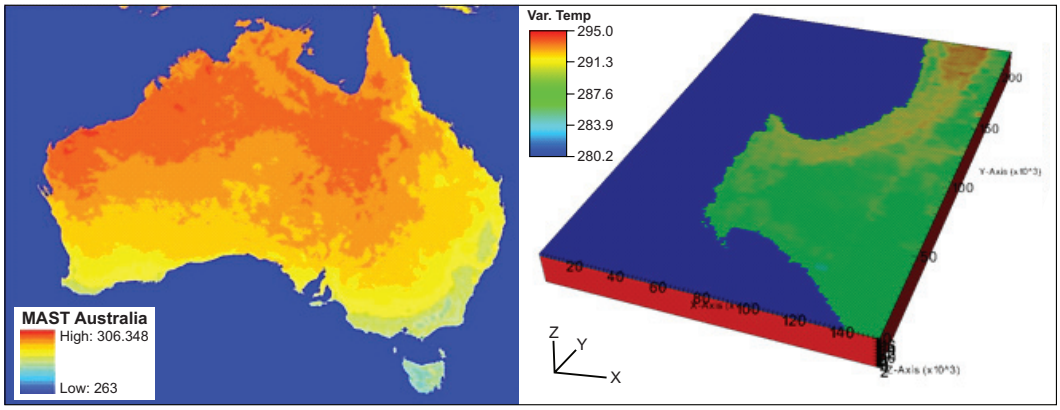

Figure 7: Example of data available in the data catalogue (left) and its easy direct use into a hydrothermal modelling package (right) due to the use of standard data format.

5.3 Future directions

This section lists some future directions to continue the work done within this project past the end date of WAGCoE, depending on future funding.

WAGCoE's geothermal catalogue has been in constant development through the duration of the Centre. The GeoNetwork software has been evolving at a rapid pace and several version upgrades were performed over the last three years. This tool is still not fully mature yet and will certainly keep on evolving for the foreseeable future. The management tools have also been continuously improved and documented to facilitate our workflow, and the next major upgrade to our administrative toolbox would probably consist in a Graphical User Interface (GUI) to enter metadata more easily, as shown in Figure 3, instead of editing directly metadata files or internal Python structures. The collaboration

with the AuScope Grid development team could also be extended further and the WAGCoE catalogue could make use of their portal development to provide an even more user friendly interface to our catalogue. Van Oort (18) pointed the cost of building such a portal, which only emphasises the importance of building on AuScope Grid's generic and open source tools.

The main work that still needs to be performed though consists of gathering members from the geothermal community to build some community-agreed standards. As mentioned in Section 3.4.2, WAGCoE is involved in a Technical Interest Group which has been created in Australia to address this issue but this movement is still in its very early stage and preliminary discussions about international collaborations on the subject are only beginning to occur. One of the most exciting aspects of this work is that most technical issues have been solved to allow interoperability between the various data custodians, allowing anyone to create a portal to discover, search, access and use for those various sources. The human side of this collaborative adventure however represents a challenge which should not be underestimated. The AuScope Grid project (8) delivered some of those major technical developments and demonstrated the road to follow to build such a community for the more general field of geosciences. This knowledge could be applied further to the more specific area of geothermal science and exploration.

6. 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 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 catalogue brings this enormous advantage to the geothermal exploration industry and provides easy, remote, and secure access to various geothermal data sources through an intuitive web interface. It helps to reduce the risk for geothermal explorations in the Perth Basin and could also help some 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.

REFERENCES

1. *Department of Mines and Petroleum website.* [Online] http://www.dmp.wa.gov.au/.

2. *Department of Water website.* [Online] http://www.water.wa.gov.au/.

3. *Geoscience Australia website.* [Online] http://www.ga.gov.au/.

4. *OZTemp.* [Online] http://www.ga.gov.au/meta/ANZCW0703013802.html.

5. *GeoNetwork opensource website.* [Online] http://geonetwork-opensource.org/.

6. *WAGCoE catalogue.* [Online] http://www.geothermal.org.au/Catalogue.htm.

7. *GeoNetwork opensource.* **Ticheler, J. & Hielkema, J. U**. 2007, OSGeo Journal, Vol. 2.

8. *AuScope: Australian Earth Science Research Information Infrastructure.* **Woodcock, R.** Perth : s.n., 2008. Australian Earth Sciences Convention (AESC). Vol. New Generation Advances in Geoscience.

9. *iVec website.* [Online] http://www.ivec.org.

10. *Mean annual surface temperature (MAST) and other thermal estimates for Australia and New Zealand from 6 years of remote sensing observations.* **Horowitz, F. G., & Regenauer-Lieb, K.** NZ Geothermal Workshop, The University of Auckland : s.n., 2009. pp. 65-68.

11. *TIG8 website.* [Online] http://www.pir.sa.gov.au/geothermal/ageg/technical_interest_groups/group_9_-_data_management.

12. *Twiki - the Open Source Enterprise Wiki and Web 2.0 Application Platform.* [Online] http://twiki.org/.

13. **Huizinga, D. and Kolawa, A.** Automated Defect Prevention: Best Practices in Software Management. s.l. : Wiley-IEEE Computer Society Press, 2007.

14. **Reid, L.B., Bloomfield, G., Botman, C., Ricard, L.P., and Wilkes, P.** *Temperature Regime in the Perth Metropolitan Area. Results of temperature and gamma logging and analysis, June/July 2010.* 2011. p. 197, CSIRO Report EP111422.

15. *The Canadian Well Logging Society. Log ASCII Standard (LAS).* [Online] http://cwls.org/las_info.php.

16. *SEG Y rev1 Data Exchange format1.* **Norris, M.W. and Faichney, A.K., [ed.].** Tulsa, OK : s.n., 2002, Society of Exploration Geophysicists.

17. **Glasson, S.** *Investigation of salinity within the Yarragadee Aquifer in the Perth area. 2011.*

18. *Geoportals: An Internet Marketing.* **van Oort, P., et al.** 2009, Data Science Journal, Vol. 8, pp. 162-181.

19. *GeoSciML Cookbook, How To Map Data to GeoSciML Version 2.* [Online] http://www.geosciml.org/geosciml/2.0/cookbook/ .

20. *PressurePlot website.* [Online] http://www.pressureplot.com/.

APPENDIX A: LISTING OF EXTERNAL PARTICIPANTS

Three external persons were funded through this project to help populate the database.

- Aspasia Kouhsen - 498 hours
- Cate Walker - 79.5 hours
- Andrew Minto - 12 hours

Their contributions consisted mainly in

- parsing existing reports to identify all type of data available and tag each reports with appropriate keywords,
- generating the corresponding catalogue entries, and
- helping the administrators identify the workflow, test and improve the administrative tools.

APPENDIX B: CONFERENCE AND WORKSHOP CONTRIBUTIONS

This project was presented at three conferences, including one international.

- National Groundwater Conference, GroundWater 2010, 31 Oct – 4 Nov 2010, Canberra.
  Corbel S., Poulet T. & Stegherr M. "Implementation of a multi-disciplinary data repository to aid geothermal exploration"
- Australian Geothermal Energy Conference 2010, 16 – 19 Nov 2010, Adelaide.
  Poulet T., Corbel S. & Stegherr M. "A Geothermal Web Catalog Service for the Perth Basin"
- IEEE e-Science Conference 2010, 7 – 10 Dec 2010, Brisbane.
  Corbel S., Poulet T., "WAGCoE data catalogue for geothermal exploration"

APPENDIX C: SCIENTIFIC PUBLICATIONS

This project resulted in two publications:

- Poulet T., Corbel S. & Stegherr M. A geothermal web catalog service for the Perth Basin. *Proceedings - Australian Geothermal Energy Conference 2010,* **2010**.
- Corbel, S. & Poulet, T. WAGCoE data catalogue for geothermal exploration *Proceedings - 2010 6th IEEE International Conference on e-Science, eScience 2010,* **2010**, 89-94

APPENDIX D: EXAMPLE OF CATALOGUE ENTRY

This appendix displays one example from the WAGCoE catalogue in full to illustrate the information presented by the web application.

WELCOME TO THE WAGCOE DATA CATALOGUE

Aggregate Results matching search criteria : 1-1/1 (page 1/1), 0 selected Select : all, none actions on selection

FAULT-RELATED STRAIN ACCOMMODATION AND DIAGENESIS IN THE NORTHERN PERTH BASIN: IMPLICATIONS FOR FLUID PATHWAYS.

Abstract Honours thesis report 2011, department of Applied Geology, Curtin University.

Keywords Geology, Core photography, Geology, Grainsize analysis, Geology, Lithology description, Geology, Petrography data, Geology, Sedimentology data, Geology, Structural data, Geology, Thin sections, Rock physics, Permeability, Rock physics, Porosity

Metadata Edit Delete Other actions Owner: admin

Send a reference to this record by email Permanent link to this description Bookmark on Delicious Bookmark on Digg Bookmark on Facebook Bookmark on StumbleUpon

Identification info

Title	Fault-related strain accommodation and diagenesis in the Northern Perth Basin: Implications for fluid pathways.
Date	2011-07-30T08:46:07
Date type	**Publication**: Date identifies when the resource was issued
Other citation details	Olierook, H., Timms, N.E., Hamilton, J. (2011). Fault-related strain accommodation and diagenesis in the Northern Perth Basin: Implications for fluid pathways, May 2011, Honours thesis, 273 pp.
Abstract	Honours thesis report 2011, department of Applied Geology, Curtin University.

Point of contact

Individual name	Nicholas E. Timms
Organisation name	Department of Applied Geology, (WA Geothermal Centre of Excellence), Curtin University, Perth, Western Australia
Role	**Author**: Party who authored the resource

Descriptive keywords	Geology , Core photography .
Descriptive keywords	Geology , Grainsize analysis .
Descriptive keywords	Geology , Lithology description .
Descriptive keywords	Geology , Petrography data .
Descriptive keywords	Geology , Sedimentology data .
Descriptive keywords	Geology , Structural data .
Descriptive keywords	Geology , Thin sections .
Descriptive keywords	Rock physics , Permeability .
Descriptive keywords	Rock physics , Porosity .

Extent

Temporal Extent

Begin date	2011-01-01T00:00:00
End date	2011-05-31T00:00:00

Figure 8: Example of catalogue entry with all metadata displayed.

Extent

Geographic bounding box

WGS 84

North bound
-29.31590

West bound
115.07100

East bound
115.07110

South bound
-29.31600

Supplemental Information

Data obtained from WA Geothermal Centre of Excellence (WAGCoE)

More descriptive keywords: Diagenesis, Facies analysis, Mineralogy, Optical photomicrographs, Orientation data, Provenance analysis, Sedimentary log, Structural analysis

See also: Olierook, H., N.E. Timms, J. Hamilton and S.M. Reddy. 2011. Characterization of fault damage zones in the Northern Perth Basin. West Australian Geothermal Energy Symposium 2011, Perth Convention and Exhibition Centre, Perth, Western Australia, 21-22 March, 2011.

Distribution Information

Transfer options
OnLine resource: Fault related strain accommodation and diagenesis in the Northern Perth Basin:Implications for fluid pathways report

Transfer options
OnLine resource: Appendix A - Sample catalogue

Transfer options
OnLine resource: Appendix B - Core photointerpretation

Transfer options
OnLine resource: Appendix C - Core logging

Transfer options
OnLine resource: Appendix D - Optical microscopy

Transfer options
OnLine resource: Appendix E - Scanning Electron Microscopy (SEM)

Spatial representation info

Transformation parameter availability: true

Checkpoint Availability: false

Figure 8 continued: Example of catalogue entry with all metadata displayed.

Reference System Information

Code	WGS 84 / UTM zone 50S

Metadata

File identifier	fa0c574f-e774-46bb-87c5-ff6f94f8e177
Character set	**UTF8:** 8-bit variable size UCS Transfer Format, based on ISO/IEC 10646
Date stamp	2010-06-18T17:21:41

Contact

Organisation name	WA Geothermal Centre of Excellence
Role	**Point of contact:** Party who can be contacted for acquiring knowledge about or acquisition of the resource

Figure 8 continued: Example of catalogue entry with all metadata displayed.