Supporting the UN 2050 Net Zero goals by reading the earth better

Nina Marie Hernandez[1*], Kim Gunn Maver[1] and Charmyne Mamador[1] present the ED2K initiative, providing multivariate earth data to support UN climate change goals.

Introduction

Against the backdrop of the current global pandemic, the UN 2050 net zero goals call for global greenhouse emissions to be cut by half by 2030 and reach net zero no later than 2050 to achieve the goal of limiting global warming to 1.5 Celsius above pre-industrial levels. The EU has pledged to become the first carbon neutral continent by 2050, and more than 110 other countries have pledged carbon neutrality by this time. Several energy companies have laid out their medium-to-long-term plans towards this objective, which includes acquisition of renewables assets, and developing competitive technologies for carbon capture and storage (CCS) and hydrogen production. Among the companies that have set zero-emissions targets are BP, Shell, Total, Repsol, Equinor and Petronas. This energy pivot will require significant capital spending to reach the goals.

To determine the technical and economic feasibility of these new energy technologies and at the same time achieve sustainable development goals, multivariate earth data, both existing and new, are required and are key to making the right management and investment decisions.

EarthDoc pointing to the future

To this end, more than 39 years of conference proceedings and publications are already available from EAGE through EarthDoc, which aggregates a wealth of subsurface information from research institutions, energy companies, service companies and dedicated professionals. The 70,000 scientific publications focus on conventional topics within geoscience and engineering especially in relation to oil and gas extraction. This subsurface information can be upcycled to provide highly valuable insights for new energy technologies. The reuse of existing oil and gas data reduces data acquisition costs, which translate into a reduction in research and development costs.

Through a collaboration between Iraya Energies and EAGE, a new database initiative has been launched with EarthDoc's repository of 70,000 scientific publications being processed using the latest in machine learning and artificial intelligence techniques and initially available to institutional and corporate subscribers. The whole data corpus is made instantly accessible and provides new tools to search and retrieve the diversity of information that one is looking for across any technical discipline.

With Big Data analytics applied to the entire data corpus of 70,000 scientific publications, additional in-depth and advanced navigation options are now available.

To extract information on a large scale, the ElasticDocs AI pipeline is applied to the EarthDoc corpus. This pipeline consists of a set of algorithms which are used to identify blocks/segments within the document corpus. Optical Character Recognition (OCR) is applied to the text segments to convert them into processable text. A Deep Convolutional Neural Network (DCNN) algorithm pipeline classifies the images into various generic and geological image classes, including tables, seismic, map, well plots, stratigraphic charts, core, thin sections, image logs, and rose diagrams. Untagged images are generically categorized as figures and remain accessible to the user.

The ingested documents are now available as structured information for analysis, which is possible in different ways:

- Metadata extraction of relevant information such as locations, names etc.
- Efficient inter and intra-search of the global text corpus for quantitative textual analysis of document contents to create automated and standardized geoscience or engineering content summary.
- Automated heatmaps to visualize density of information within a basin or country.
- Image extraction of similar image classes for efficient identification of analogues, duplicates and clusters.

Some key functionality that has a significant impact on utilizing the scientific publications is that not only are images classified according to type, but it is possible to do a search on image embedded text making the search capabilities far more advanced than just a normal search-and-find option. The search results of this database are exportable as .csv files making statistical work and analysis instantly possible across data points.

Proprietary corporate data meets published scientific knowledge.

Recognizing that energy companies hold valuable information in their repositories accumulated throughout many decades of operations, the data kept internally can now be easily integrated with published geoscience and engineering data. The combination of

[1] Iraya Energies
[*] Corresponding author, E-mail: nmh@irayaenergies.com
DOI: 10.3997/1365-2397.fb2021045

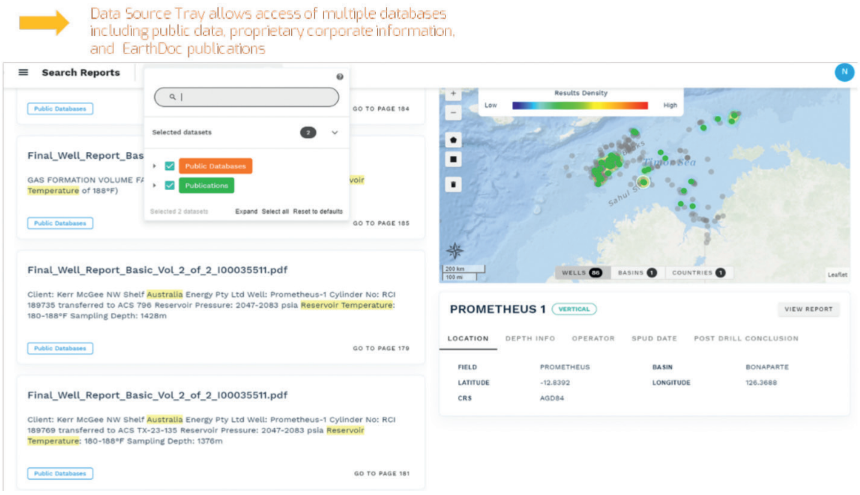

Figure 1 Data source tray in the new database allows access to multiple databases, including, but not limited to, public data, proprietary data, and EarthDoc corpus. Reference: Geoscience Australia, NOPIMS.

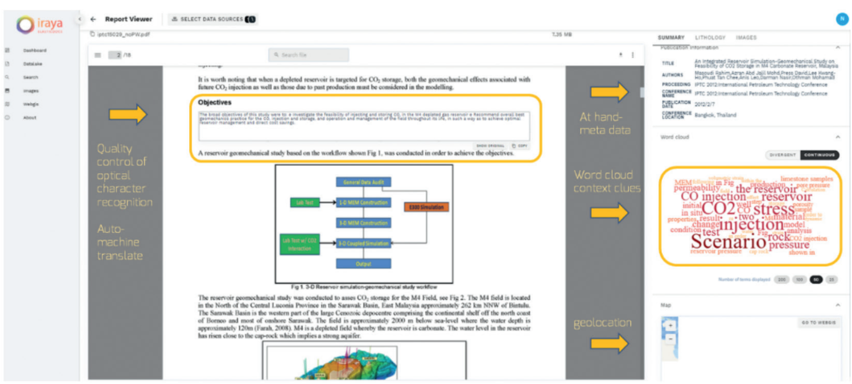

Figure 2 Earth readability tools are available in the new database dashboard (Rahim et al., 2012).

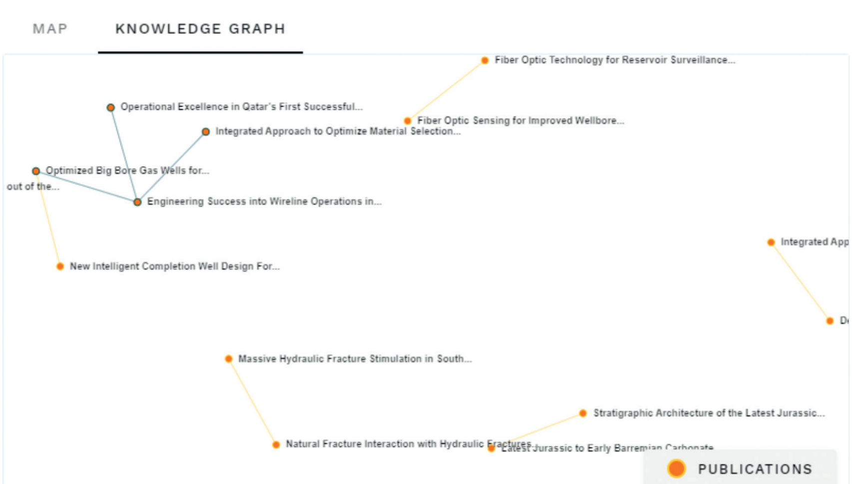

Figure 3 Visualization of knowledge graph illustrates related geological and engineering concepts.

database results in a two-way enrichment process between public and private information. Scientific studies from peers in related fields offer additional information that either validates internal company studies, or offer alternative technical perspectives from industry experts.

In the Figure 1 example above, we show how exploring temperature information available in a Final Well Reports repository that is commonly considered as proprietary corporate data, is combined with published EarthDoc contents. (Data source: NOPIMS)

Several Earth readability tools are deployed in ED2K to amplify EarthDoc capabilities. This includes access to digital textual content, the ability to quality control OCR results and read in multiple languages via machine-translate. Word clouds are simple context clues which provide a quick summary of content of the articles.

The full ED2K corpus is geotagged to more than 800 geological basins around the world. The geotagging is one of the most complex machine learning tasks in this implementation, because the scientific articles contain both location of the geological area of interest, as well as the location of conferences where the publications are presented. Often, these two reference locations are different and introduce ambiguity for the system.

Another feature that is implemented is the knowledge graph visualization. It illustrates the connectivity of 'related concepts' based on publication references. In the example above, it shows 'operational excellence', 'optimization' and 'engineering success' within the same network (Figure 3). Similarly, 'hydraulic fracture stimulation', 'natural fracture interaction', and 'Jurassic carbonates' show connectivity. In theory, a knowledge graph can be built in multiple ways by the user.

Utilizing ED2K to reach the zero-emission goal

The advanced access in the new database makes possible a new use of the highly valuable subsurface information and facilitates cross-discipline usage.

Many of the new cross-function usages of the existing scientific publication repository are geothermal energy, hydrogen energy, CCS and windmill foundation derisking.

A query for 'carbon capture and storage' generates an information heat map captured in Figure 4. The map shows the geographical distribution of the resulting publications either based on country or basin.

Figure 5 maps out the major carbon capture projects around the world vis-à-vis needs requirement. Europe and US show high activity in CCS initiatives, which are driven by government policy. Comparing Figure 4 and Figure 5 (left), it may be incidental, although not entirely surprising that where there is a significant amount of data to aid technical and management decision making, CCS implementations are also active.

The new database contains climate change and greenhouse gases industry discussion materials spanning over three decades. Already between 1990 and 2000, the possibility of disposing carbon dioxide (CO_2) is discussed in papers such as J. Leeb. W, (1993), and Wildenborg, F.B., et. al., (1996), in conjunction with the use of CO_2 for improved oil recovery methods (IOR).

Between 2001 and 2010, the discussions tackled issues in establishing a geological storage hub (Espie, 2000) and various potential site feasibility studies (Gregersen et. al, 2000), costs (Wildenborg et. al, 2000), use of seismic monitoring (Benson, 2003), (Gosselet et. al., 2006), pilot and numerical simulations (Domitrovic et. al., 2005), (Battistelli et.al, 2005), reservoir performance (Broad et.al, 2007) and improving facilities performance to reduce operating costs in CO_2 and H_2S contaminated fields (Swatton et. al, 2009).

From 2011 up to the present, with the advancements in seismic methods, reservoir modelling techniques and laboratory experiments (Bolourinejad, 2013) many more complex analyses on the subject of carbon storage were performed. Combined with enhanced oil recovery experiments (EOR), the amount of data modelling CO_2 behavior underground has multiplied ten-fold.

In areas where it is not possible to implement subsurface carbon capture, utilization strategies are discussed by Harsh, A. et. al (2014) on the industrial usage of CO_2 including, but not limited to, polymer processing and chemicals production.

For a deeper dive in the corpus, we draw an arbitrary areal polygon, indicated by the yellow box in Figure 5, around South East Asia. Our new database reveals some of the strategies that have been identified by an operator to manage greenhouse emissions in its operations in Malaysia (Mehta et.al.,2008). This includes, among others, ending continuous gas flaring

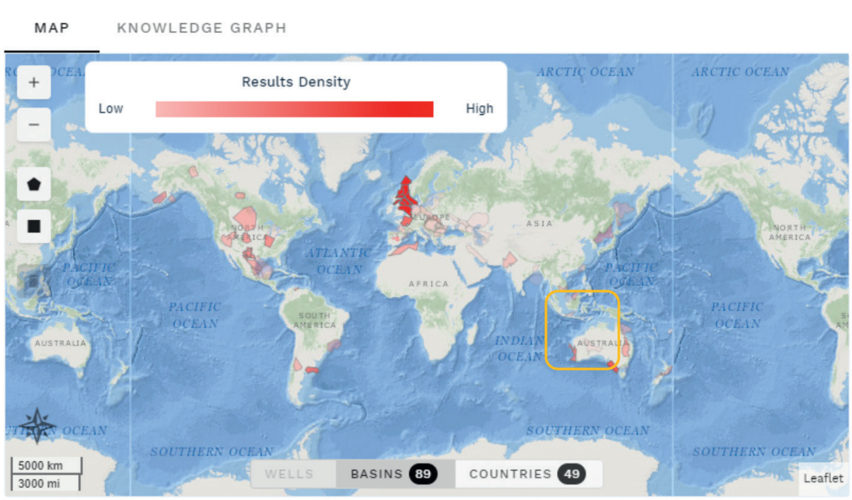

Figure 4 Indicative geolocation of global knowledge about 'carbon capture and storage' between 2008 to 2020.

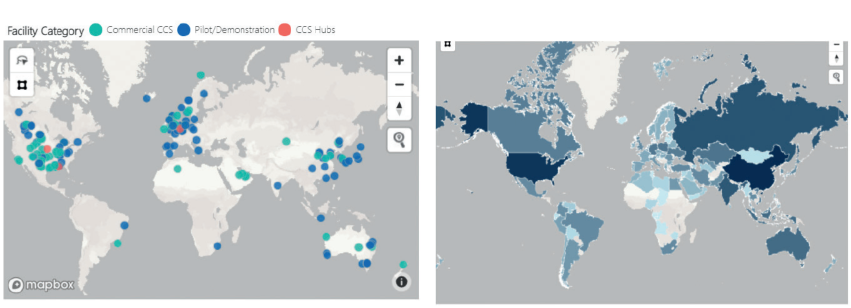

Figure 5 Location of major carbon capture projects around the world (left) and the requirement index based on fossil fuel production and consumption (right). Reference: Global CCS Institute.

and minimizing gas venting, improving energy efficiency in the design of assets and production operations, accounting for the cost of emitting greenhouse gases in investment decisions, supporting development of CCS infrastructures, and policy advocacy.

The rich diversity of data available in the new database are illustrated in Figures 6 and 7. These include graphical information of PVT analyses, miscibility, flow rates, and time-lapse pressure profiles, which are useful for reservoir simulation studies focused on the interaction of reservoir rocks with CO_2 during injection. Also available are petrography data that makes it easier to interpret reservoir modelling results by being able to look at the structural fabric of the storage rocks down to microscopic levels.

New energy from old data

With this data-driven strategy it will be possible to facilitate the pivot to new energy from valuable existing data. No part of the data is left unprocessed. It may be that not all relevant information will exist within the 70,000 scientific publications, but this can be confirmed instantly saving valuable time and resources doing data exploration. On the other hand, if the information is available, it will be immediately accessible, trackable and put into context with other relevant information and geographically.

For example, the geothermal gradient is a key parameter of interest in relation to geothermal energy. The temperature data in the new database has been acquired by energy companies mostly for the purpose of understanding the hydrocarbon generation window, petrophysical interpretation and reservoir modelling analyses. They can be relooked at for further exploratory geothermal applications. It currently excludes temperature data from geothermal companies.

We are barely scratching the surface on the data and insights available – multiple data stories are waiting to be reimagined, reconnected and retold in the context of the future of energy.

Accelerating internal digitalization initiatives

All the elements of the new database are stored and structured in a digital data warehouse. They will be optionally available as an API link to be used for additional geological analysis, data analytics, or machine learning experimentations. Already in structured format, they can be fed into additional natural language processing or image segmentation processing for in-house experimentation.

Opportunity for the energy geoscientists and engineers of the future

While the energy industry has faced significant headwinds, it is now moving faster than ever towards new, cleaner energy production. It is possible to see that multiple opportunities remain, as already pointed out by Raistrick (2008), and remain relevant in 2021, for the geoscientists and engineers who are looking at the

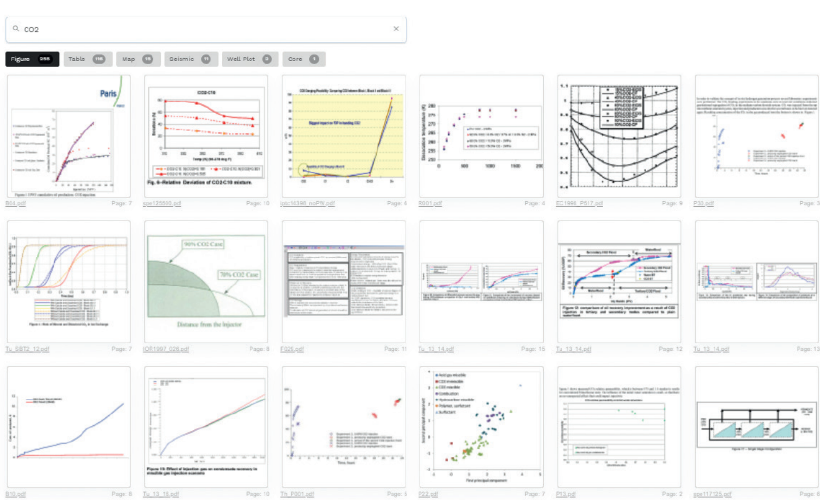

Figure 6 Experimental engineering data of CO_2 behaviours in enhanced oil recovery operations of mature fields can be transferable to carbon storage design and monitoring.

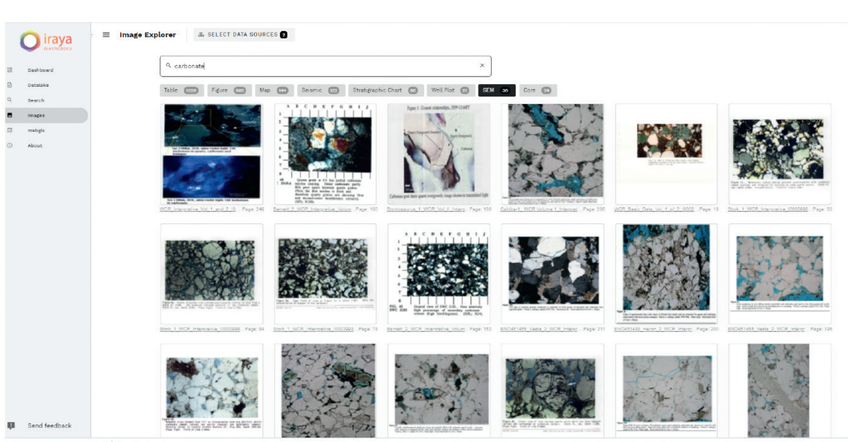

Figure 7 Combination of carbonate petrography data and information extracted from well reports.

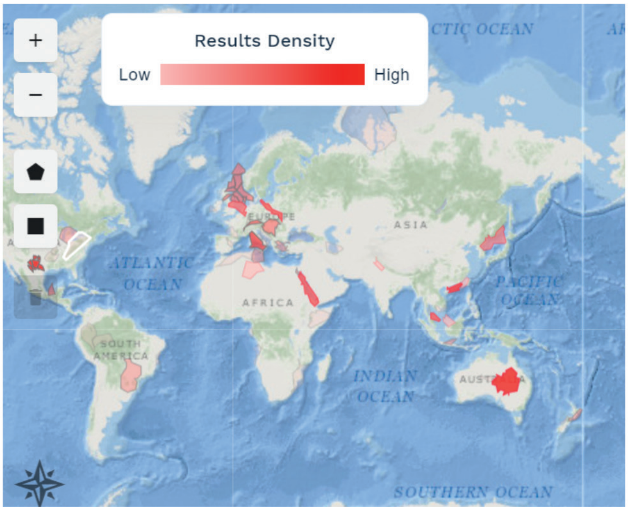

Figure 8 Locations of relatively 'high geothermal gradient areas' based on existing corpus, data acquired by oil and gas companies. Data excludes information from geothermal companies.

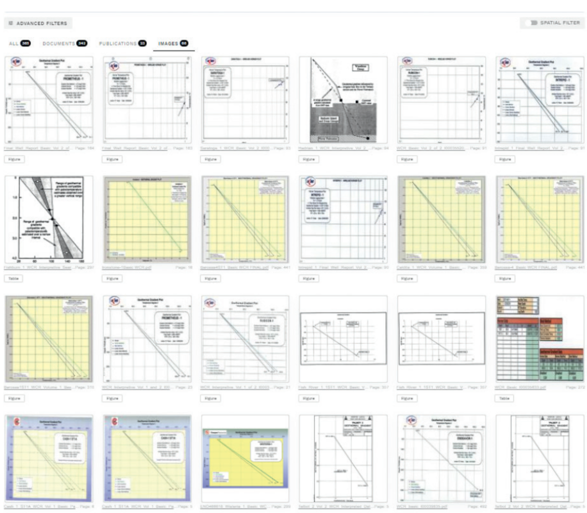

Figure 9 Compiled graphical temperature information filtered by country, basin or an arbitrary polygon location. Reference: Geoscience Australia, NOPIMS.

future of energy. Strong, flexible technical skills will be needed to explore for suitable carbon capture facilities, assess their storage, containment and injectivity capacities. Meanwhile, the new energy industry will continue to gather, integrate and analyse empirical data, whether it is on the reservoir, sub-surface or at hydrocarbon or future hydrogen production facilities.

We have already seen a lot of data. It is up to us to use the right tools to read the earth better, and get a head start towards new energy.

References

Global CCS Institute (https://www.globalccsinstitute.com). CCS Facilities Database (https://co2re.co/) Geoscience Australia, NOPIMS (http://www.ga.gov.au/nopims).

Bolourinejad, P. and Herber, R. [2013]. Experimental and Modeling Study of Salt Precipitation during Injection of CO2 Contaminated with H2S into Depleted Gas Fields in Northeast Netherlands - (SPE-164932), 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013, London, UK.

Battistelli, A., Giorgis, T. and Marzorati, D. [2005]. Modeling Halite Precipitation around CO2 Injection Wells in Depleted Gas Reservoirs, 67th EAGE Conference & Exhibition, Madrid, Spain.

Broad, J., Ab Majid, M.N., Ariffin, T., Hussain, A. and Basher, A.B. [2007]. Deposition of "Asphaltenes" during CO2 Injection and Implications for EOS Description and Reservoir Performance, IPTC 2007: International Petroleum Technology Conference, Dubai, United Arab Emirates.

Domitrovic, D., Tuschl, M. and Šunjerga, S. [2005]. CO2 Pilot Injection at Ivanic Oil Field – Numerical Simulation, IOR 2005 - 13th European Symposium on Improved Oil Recovery, Budapest, Hungary.

Gosselet, A.C. and Singh, S. [2006]. Elastic Full Waveform Inversion for CO2 Sequestration monitoring - 1D Synthetic Data Investigations,68th EAGE Conference and Exhibition incorporating SPE EUROPEC 2006

Gregersen, U.N., Johannessen, P., Kirby, G., Chadwick, A. and Holloway, S. [2000]. Regional Study of the Neogene Deposits in the Southern Viking Graben Area - a Site for Potential CO2 Storage,62nd EAGE Conference and Exhibition - Special Session on CO2, Glasgow, UK.

Harsh, A.H. and Anne, V.A. [2014]. Carbon Dioxide Capture, Utilization and Storage (CCUS),76th EAGE Conference and Exhibition 2014, Amsterdam, Netherlands.

Leeb, W. [1993]. Case study of CO2 disposal in aquifers - A solution to reduce the greenhouse effect, 55th EAEG Meeting, Stavanger, Norway.

Mehta, A., Hj-Kip, S. and Foo, J. [2008]. Managing Greenhouse Gas Emissions in Upstream Operations in a Carbon-Constrained World, IPTC 2008: International Petroleum Technology Conference, Kuala Lumpur, Malaysia

Rahim, M., Azran, A., Press, D., Lee, K.H., Phuat, C.T., Anis, L., Darman, N. and Othman, M. [2012]. An Integrated Reservoir Simulation-Geomechanical Study on Feasibility of CO2 Storage in M4 Carbonate Reservoir, Malaysia, IPTC 2012: International Petroleum Technology Conference, Bangkok, Thailand.

Raistrick, M. [2008]. Carbon capture and storage projects to challenge governments, scientists, and engineers, *First Break*.

Wildenborg, A.N., Breunese, J.G.H.and van der Meer L. [1996]. Potential of CO2-disposal in deep reservoirs and aquifers of the Netherlands,58th EAGE Conference and Exhibition, Netherlands.

Wildenborg, A., Floris, F.D., van Wees, J. and Hendriks, C. [2000]. Costs of CO2 Sequestration by Underground Storage,62nd EAGE Conference and Exhibition - Special Session on CO2, Glasgow, UK.

Swatton, M.J.R., van Soest-Vercammen, E. and Klein Nagelvoort, R. [2009]. Innovation and Integration in LNG Technology Solutions, IPTC 2009: International Petroleum Technology Conference, Doha, Qatar.