

# digital energy journal

The future of BP's  
Field of the Future

Shell and HP's  
wireless seismic  
technology

March 2010

Issue 23



Finding hydrocarbons with dielectric resonance  
How good are geologists?  
Finding oil in the university library  
Dangers of "workflow"

**Finding  
Petroleum**

Silver sponsor



**ENIGMA DATA  
SOLUTIONS**  
World Leaders in Oil & Gas Archiving



Associate Member

**energistics™**  
The Energy Standards  
Resource Centre

# Finding oil in university archives

As an industry we have focused on the management of in-house and proprietary data. Is this just the tip of the information iceberg? What about the thousands of other sources of information that are not easily accessible? And assuming we can find all the data sources, are they of any value?

**By Robert Winsloe, president, Digital Earth**

There is a vast amount of useful exploration data all over the world, stored in university archives, attics, computer systems, and people's heads.

Of course the value of (and the need for) external data sources varies through the oil field life cycle. In new ventures for example, as we screen thousands of opportunities looking for the one or two that best match our investment criteria, we rely heavily on external data sources and external expertise. And at each major decision point through the life cycle we again look for the information and experience that will enable us to optimise performance and maximise profits.

The larger external data providers such as IHS, Woodmac, Fugro, TGS-Nopec are well known to us. And there are many medium sized data companies who fill a niche such as C&C Reservoirs' field analogues database. But below this there are literally thousands of smaller obscure data providers, from geologists with data stashed away in their garage to university archives in developing countries and of course the largely unindexed knowledge held by hundreds of thousands of individuals around the world. It is these people and these smaller, often hidden, data providers that are the subject of this paper.

Three major oil companies decided to find out how valuable these smaller sources might be by commissioning studies across three separate areas: The Arctic (2007); Iraq - finding data that had been removed from the country (2008); and China with a focus on Chinese language sources (2008).

The Arctic study identified 231 of which were not previously known and were considered to be of significant value to their exploration programme.

For Iraq, of the 72 sources identified more than half were previously unknown and considered to be of value. Many individuals (both inside and outside Iraq) have data in attics and garages, and some data has been sold on an exclusive basis to a major oil company.

The China study added the foreign language dimension, and relied heavily on native speakers with knowledge of the local oil and gas sector. There is a wealth of infor-

mation online and in the public domain.

### Sources

In undertaking their research at university it is common for PhD students to have better access to hard data than oil companies. The resulting interpretations can contain valuable insights into potential plays and prospects but in many countries these reports are archived and rarely see the light of day. Several oil companies have "found oil" in these rarely seen documents, often engaging the services of local students (who understand the language) to carry out the research.

Other little known sources of data include the obscure Russian Institute of Foreign Geology ([www.vzg.ru](http://www.vzg.ru)) which collects data gathered by Russians abroad, a similar institute in China (World Data Center for Geology, China), Woodenski.com which sells out of date geology books, and The University of Tulsa which has 1 million abstracts online - [pa.utulsa.edu](http://pa.utulsa.edu).

Besides finding valuable sources of information many lessons were learned about the challenges associated with finding, indexing and integrating external data.

### Finding the sources

This step was considered to be a great value when going into a new area. And better sources were found by engaging local in-country resources (who often knew where to find the data outside their own country).

In the absence of a global database of people, the recruitment companies offer a stop gap alternative. Rather than searching for skills ask the recruitment company to search for the basin and field names you are seeking to explore.



*People's attics - where lots of oil and gas data is stored*

And of course one of the most valuable sources will be the oil companies who have previously relinquished acreage in your target area - assuming of course they can still find the data (no guarantee there!).

### Indexing and integrating

Data was often poorly and inconsistently formatted. Metadata structures were inconsistent, even on header information for authored documents coming from the same source. Where large document databases exist they often come with poor search interfaces. It is no surprise that the customer lists of these providers is restricted to the larger companies who can afford librarians to make sense of the complex search interfaces.

Even with cultural data dictionaries and comprehensive taxonomies, natural language processing code still had to be developed to scan and rewrite the metadata so that it could be integrated with each company's in-house systems.

### Conclusions and recommendations

There is a gap in the market. We have powerful generic search engines such as Google at one extreme, and industry specific but proprietary data vendors such as IHS at the other. What is missing is a search tool with the power of Google but with a structured energy index that makes sense of oil industry da-

ta and knowledge. There are some partial solutions in the market already.

Metacarta integrated with IHS basins, fields, wells does enable you to attach thousands of publicly available documents to wells, fields, basins etc. The drawback is that you need a license for both Metacarta and IHS to make this work.

There are a growing number of information portals including Onepetro, Petroleum Abstracts, search4oil.com.

To bridge the gap a number of elements are needed:

An “opensource” cultural data dictionary (wells, fields, basins, blocks, pipelines)

An “opensource” industry subject taxonomy (Digital Earth is seeking partners to build this element)

Against this, all data, documents, news AND people can be indexed in an industry specific and freely available search site, including a collaboration platform for automatic uploading and indexing of data, documents and CVs.

As the industry specific search technology improves the number of sources will

grow exponentially.

And the same tools will allow for a viral indexing of people through an online collaboration platform.

Internally I would recommend a similar approach. It would be easier to achieve (although not without major expenditure) by licensing the IHS cultural data dictionary and the Petroleum Abstracts taxonomy. Software tools that have been employed internally with success include Autonomy, Metacarta.

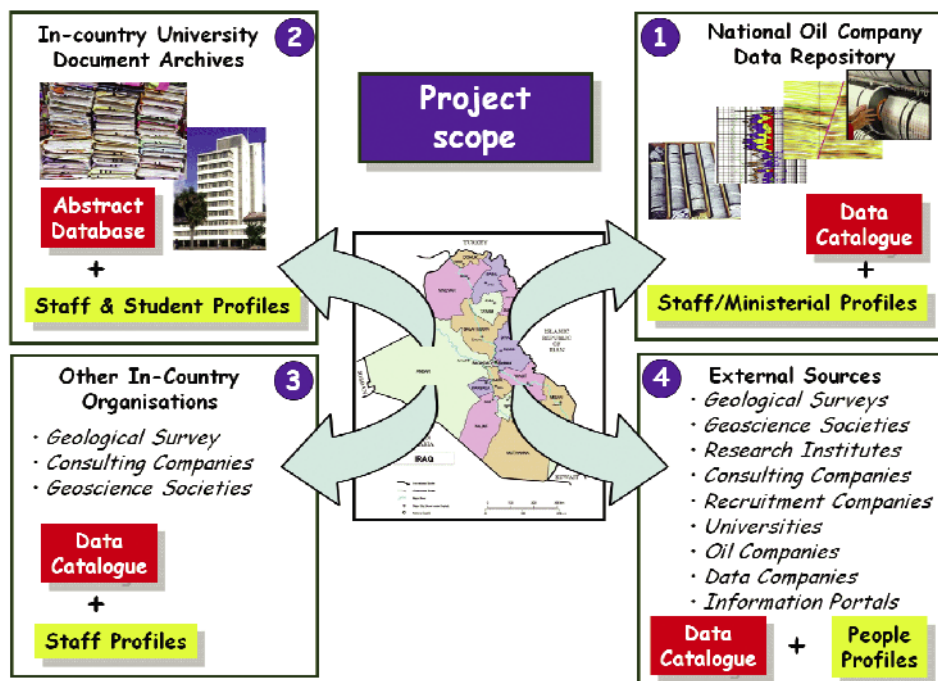
Today, the major oil companies have the resources to play this game, to find the obscure and yet valuable resources, to track down that one document in the University archive that will lead them to a major discovery, and to find the people who will help them make the most of that opportunity.

And like e-Bay and Google, the new industry specific search and collaboration technologies will enable the thousands of smaller companies to compete on an equal footing with the major companies.

Our internal data is not so much the tip of the iceberg but the head of a very long tail.



*Finding more places you can find oil and gas data - Robert Winsloe, president, Digital Earth*



*Places you can find oil and gas data*