

PESGB Data Management Conference December 2003



Figure 1 Data Management – before!¹

Around 140 attended the Petroleum Exploration Society of Great Britain's Data Management 2003 conference. About 25 oil company personnel attended – a rather high percentage compared to the major trade shows – suggesting that data management remains a concern. Unfortunately, no oil company folks presented papers – leaving the field wide open to rather too many vendor 'infomercials'. Many of these centered on the use of geographical information systems (GIS) to manage and use upstream data. Records management was also highlighted – with presentations on the need for systematic records management policies. Conwy Valley presented a new standards initiative in petrography. Technology-wise, insightful new presentations came from Landmark on seismic transcription, and on well deviation data acquisition and processing QC from Hydro Systems. On the governmental front it was '[Groundhog Day](#)' all over again – with yet more initiatives to preserve the Nation's data and keep the oil flowing.

Highlights

[National Hydrocarbon Data Archive](#)

[Seismic Processing Business Study](#)

[‘Open’ Seismic Storage](#)

[Petrographic Nomenclature Standards](#)

[Deviation Survey QC](#)

[Seismic Data Pitfalls](#)

¹ Image courtesy [Instant Library](#).

Contents

The PILOT Data Initiative – Jackie Banner, UK DTI.....	2
Deal Data Registry – Robert Gatcliffe, British Geological Survey	3
E&P Information Value – Paul Maton, POSC	3
Geological data management – Jenny Walsby, British Geological Survey	3
Spatial Layers – Gavin Critchley, IHS Energy.....	4
Seismic Processing Business – Adam Mitchell, Paras Ltd.....	4
Records Management – Trudy Curtis, PPDM Association	5
Petrographic nomenclature standards – Barry Wells, Conwy Valley Systems	5
E&P catalogue implementation – Glen Mansfield, Flare Consultants	6
2D seismic data rationalization - Christine McKay, Landmark	7
Seismic Data Management - Jill Lewis, Troika.	7
GIS in upstream – Nick Blake, Lynx Information Systems.....	7
GIS in upstream – Karen Blohm, Robertson Research	7
GIS implementation – Steve Ashley, Venture Information Management.....	8
Making maps from diverse data sources – Nathan Balls, Petrosys.....	8
Well deviation data QC – John Harris, John Kelley, Hydro Projects.....	8
Retention Schedules in Records Management – Jamie Burton, Instant Library	8
GIS in site restoration – Gavin Coppins, UKAEA and Mike Ayres, ESit.....	9
Services are the secret for information management – Wendy Kitson – Kadme	9
Pitfalls in seismic data retrieval – Eleanor Jack, Landmark.....	9
eSearch - Ben Trewin, IronMountain.....	9
Disaster Recovery – Paul Duller, Instant Library	10
Exhibitors and Sponsors.....	10
<i>Records Management Society.....</i>	<i>10</i>
<i>Instant Library.....</i>	<i>10</i>
<i>DeepStore</i>	<i>10</i>
Technology Watch Service	11

[The PILOT Data Initiative – Jackie Banner, UK DTI](#)

The DTI's objective is to provide easy access to UK data. Currently licensees own and have a perpetual obligation to keep data on behalf of the government. Legacy media and data have proved a stumbling for digital data sets. The '[Digital Doomsday Book](#)' project only lasted 15 years before the BBC Micro-based data was unreadable. After 1000 years or so, its paper based predecessor is still in use today!

Numerous workgroups in PILOT and the UKOOA have looked at these issues and there is now one solution, a single data catalogue, endorsed by all. Simon Toole (DTI) manages the Catalogue a.k.a. the **National Hydrocarbon Data Archive** – which is linked to the Deal Data Registry www.deal.co.uk. Data can be published by the DTI but is still owned by licensees. For well release, 76% of the DTI's effort is spent chasing missing data. There are inefficiencies due to non-compliance (and non enforcement!). The current focus is on data completeness and quality. Licensees have now agreed to reduce the

confidentiality period to 4 years for pre-20th round seismic data and there are also guidelines for spec survey data 'release'. See [IAGC website](#) and upcoming PESGB newsletter.

A study of archiving costs is underway on Kerr McGee's Hutton field. The operator is responsible for data and the catalogue. Paras is acting as a 'confidential cost advocate'. The aim is to relieve the operator of storage obligations. Problems arise from inconsistent catalogue naming; it is proving very hard to aggregate the information. "There are as many E&P catalogue standards as there are operators". There is a need for an agreed set of catalogue attributes - which requires input from those who are familiar with the data. Currently it's decision time for companies. Negotiations are underway as to what/how to supply the catalogue and where to store the data. The DTI has developed a model 'to aim for'. Feedback and input from operators and service companies is now required.

[Deal Data Registry – Robert Gatcliffe, British Geological Survey](#)

Work on Deal began in 1999 – the project has subsequently been re-branded as the Deal Data Registry (DDR). The project was sold on the promise of 'up to 100,000 more jobs' resulting from 'boosting the efficiency' of offshore activity. The system should be of particular benefit to new entrants. DDR is a record of data – a web-based GIS with definitions, accurate metadata and links to vendor data stores. Gatcliffe believes the project has been 'quite successful'. Common Data Access has provided a certain amount of data free to the DDR. The Sea Fish Authority database and BGS' core data are linked to the DDR. There are currently some 1200 monthly logins. Pipelines are included in the database which allows a warning bell to ring in a trawler when it approaches within 5km of a pipe. There are still significant issues with data cleanup. Data entitlements will be addressed in 2004 and pipe and other data will be available through a new 'Infrastructure Portal' – which should help companies determine spare capacity.

[E&P Information Value – Paul Maton, POSC](#)

Maton gave a recap on WITSML, WellLogML and other Practical Well Log Standards work. Maton also called for more work homogenizing the 'huge number' of curve types and complex curve and tool naming schemas claiming that accurate data is hard to find in the public domain. A thesaurus of logging nomenclature is required – and should be kept up to date. The plan is to map service company catalogues to the Practical Well Log Thesaurus. Work on WellHeaderML has been customized to the UK DEAL environment. Another 'profile' will be 'international' and will derive from work done by PPDM². Maton described the ongoing E&P data catalogue as 'wide ranging – possibly too wide ranging'. A project to 'align' the catalogue with W3C recommendations on RDF and Dublin Core metadata standards is to be started in the next few months. Maton concluded by announcing that POSC was 'alive and well' and contemplating the opening of a Russian unit.

Q&A

What take-up is there for the catalogue outside of Shell?

It is being used – by Anadarko, Nexen – but is slow to catch on.

[Geological data management – Jenny Walsby, British Geological Survey](#)

The British Geological Survey's goal is to create a 'secure digital dataset' of borehole and geological information. BGS' million borehole paper records were scanned in 2000-2002 along with a million maps and coal borehole records, photographs, reports, memoirs etc. 2 million monochrome paper records and 180,000 color maps scanned into a 17 TB dataset - all bar coded with a new GUID. Metadata is stored (using the 'ISO metadata standard') in Oracle, including many years of changes and updates. Scanning produced about 100MB/map and 16GB/week in production. 'QC has a hard time keeping up'. Hardware includes FireWire connection to network attached storage (NAS). Data is compressed to JPEG 2000 and archived to LTO 2 tapes. User expectations were high – of instant access and current metadata. But index

² Actually Paul Maton first asked Trudy Curtis if the PPDM profile was to be for 'Alberta' or for 'Canada'. Curtis explained that the profile was for 'international'.

issues and network bandwidth make access to maps slow. Need to manage and educate users as to acceptable use. 'Digital data still needs managing'.

Q&A

Who did the scanning?

[Océ](#).

What compression/resolution was used?

Compression was 1:10 and resolution 300/400 dpi – to preserve map detail and annotations.

[Spatial Layers – Gavin Critchley, IHS Energy](#)

Critchley introduced the concept of spatial layers – cartographic representations of database features. Despite the fact that some 90% of oil and gas data has a spatial component – oil companies are 'not as digital as they think'. But end user skills are growing – and availability of map layer information increasing. GIS is the silent revolution – which has 'snuck up on us' – along with the ESRI desktop. To judge the popularity of spatial layers Critchley suggests Googling for them³. For Critchley, spatial layers equate to ESRI with its three tiers of software functionality: ArcIMS for E&P end users, ArcView for GIS interpretation and ArcEditor for the GeoMatics department. ESRI's popularity is making Intergraph the 'Wordstar' of GIS⁴. IHS Energy is leveraging ESRI technology, building a range of layers. Others are available from a developing third party market. Critchley is a firm believer in the widespread transformation of physical to digital assets – 'it has to happen'.

[Seismic Processing Business – Adam Mitchell, Paras Ltd.](#)

Paras have recently completed a survey of the geophysical marketplace for a client. Seismic data management, processing and interpretation are increasingly integrated. Bandwidth and media are 'struggling to keep up' with field data volumes. Few other industries are as demanding of their data. Existing software was not built to handle multiple attributes and depth domain – again – struggling to meet user requirements. Oil and gas data management may be an order of magnitude greater than that of the genomics industry (c.f. the Wellcome Trust [Sanger Institute](#)). Oil and gas companies tend to work in business units – 'filials' – few have strict corporate data management policies.

³ We did with interesting results. On Google's first page on 'spatial layer' all references point to Oracle Spatial or Mapinfo. Maybe ESRI isn't paying Google enough!

⁴ This is probably truer of upstream oil and gas than other industry sectors.

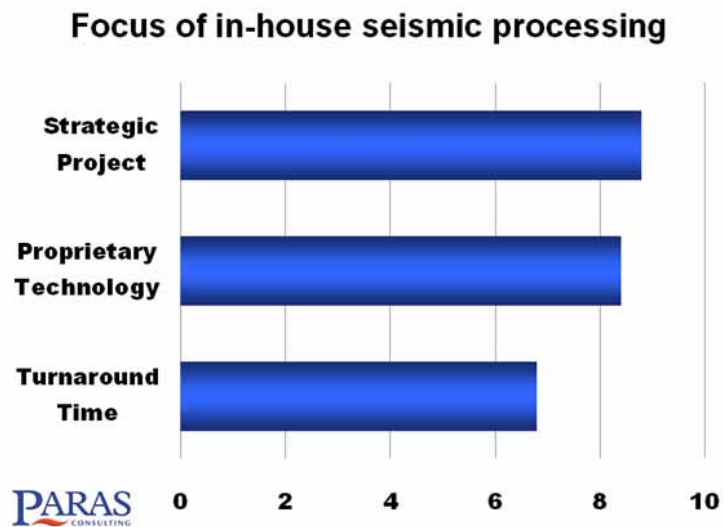


Figure 2 In house processing focus⁵

Few oils have their own tape libraries – most storage is outsourced. Almost always 3590 tapes are used – but disk storage is now being considered. SEG-Y/SEG-D formats are almost universal – few companies use encapsulation or compression. There is a high onus not to take legal or operational risks with the outsource provider. Cost, flexibility and security are key issues. Outsourcing solutions are a lot more available and efficient than five years ago.

Seismic processing is moving to Linux. Scale is the key because of the issues of moving data around the world. More (up to 50%) processing is being done in-house – a big change from five years ago. Processing effort breaks down thus: 56% PSDM, 28% PSTM, 11% 4D, 3% 4C. All companies surveyed forecast that in-house processing would grow. All use in-house proprietary algorithms – but bring in key external providers. Companies believe that in-house processing provided a major differentiator. Some companies reported as much as 50% of their seismic activity was production related.

Records Management – Trudy Curtis, [PPDM Association](#)

Today, one person does the work done by two or three a few years back. Records management (as defined by ISO 15489) is increasingly important. Structured information is relatively well managed. Unstructured less so. Data should be managed or else ‘you may go to jail!’ In Canada, PIPEDA can audit information management systems. In the UK, the data protection act places similar constraints on what can be done with corporate data. A major problem is that today’s users are all untrained records managers. Taxonomies and metadata are the way to go. PPDM is working on [Dublin Core](#), and the [Web Ontology Language](#) (OWL). These are being ‘used to integrate taxonomy into PPDM along with Flare Consultants and POSC’. Also ran – ISO, EPSG, IEEE.

Petrographic nomenclature standards – Barry Wells, [Conwy Valley Systems](#)

Wells gave something of a treatise on what standards are about: what is a well? Sometimes nomenclature and definitions can be tortuous – e.g. ‘monocrystalline’ means ‘less than four crystals’, a ‘turbidite’ may be just a ‘sandstone’ – mixing genesis and description. The [IUGS commission](#) for systematics in petrology has just started working on sediments. The IUGS will meet in Florence Aug 2004 and there will

⁵ Courtesy [Paras](#).

be a re-run at the 2005 AAPG in Calgary. Work from the IUGS commission is built into [Conwy Valley's Petrog](#) product.

E&P catalogue implementation – Glen Mansfield, [Flare Consultants](#)

Mansfield reviewed the POSC E&P Catalogue 'Standard' V 0.5, originally developed for Shell Expro's Discovery project. EpiCat consists of a set of attributes and valid values for E&P objects. Catalogues are used in document management for control, retention, bibliographic studies etc. Catalogues bring context to a query. A petroleum engineer's query will return items pertaining to a petroleum engineering related process. In this context the POSC Business Process model (ex Shell⁶) will provide the contextually valid value list in EpiCat.

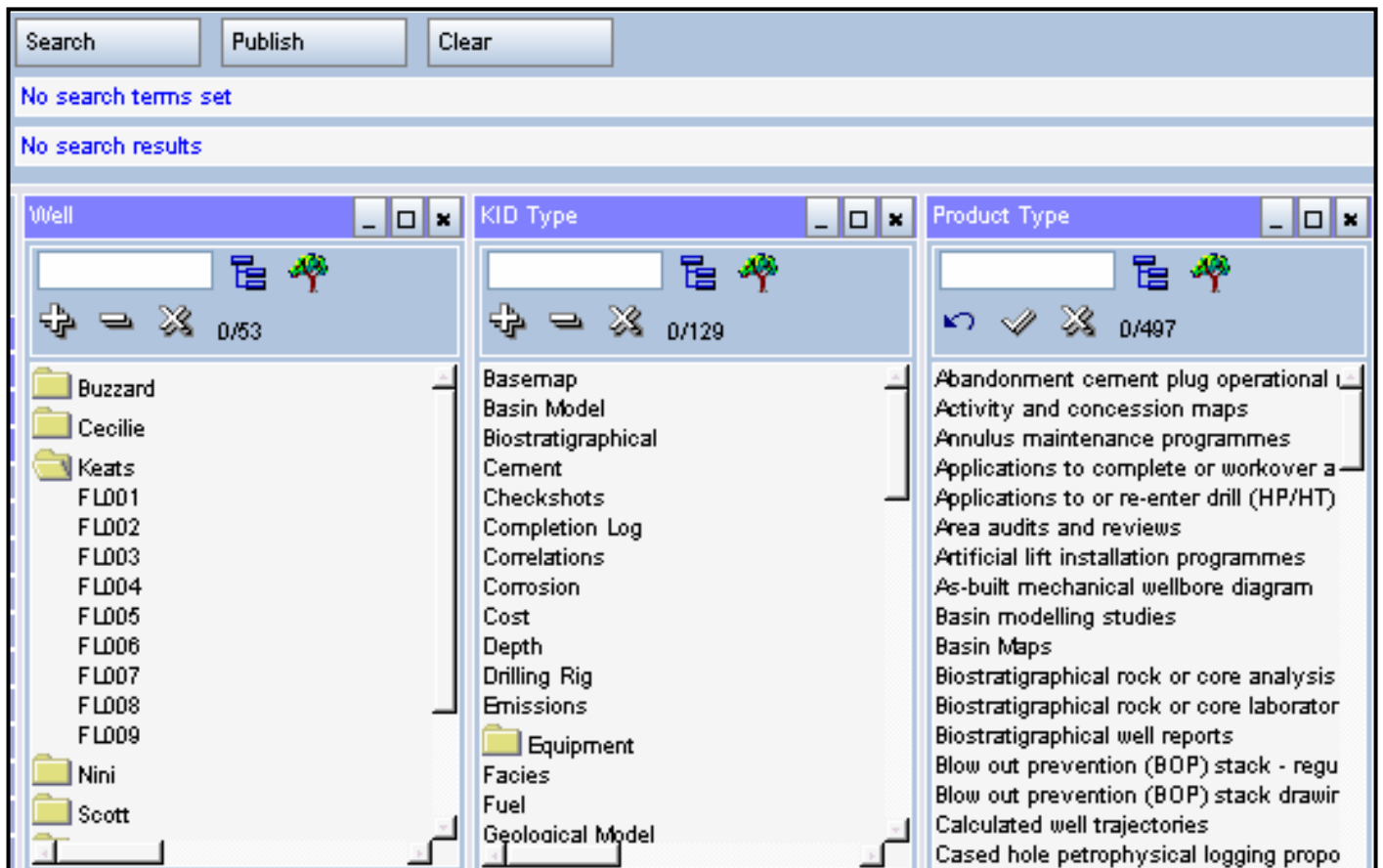


Figure 3 - Catalogue 'Product Types'⁷

The 'product type' - e.g. well completion - will set off other attributes and relationships - ownership etc. The catalogue does not replace existing data sources but rather provides a common indexing method, augmenting existing data sources with 'common context'. As delivered to POSC, the catalogue is too complex with 52 attributes (as opposed to 3 in a typical EDMS). Some implementations have huge valid value lists - of disciplines and processes with up to 500 product types. The key to successful integration therefore is to present the catalogue in a meaningful way to the end user. Just putting EpiCat into an EDMS is too complex. But stripping down the valid value list breaks the link to the standard. What is required is a good GUI with appropriate integration⁸. Change management is also an issue. Coping with the well file information explosion across multiple repositories requires a strategy of reducing pick lists

⁶ According to Mansfield's handout, 'In 2001, Flare worked with Shell on the Discovery, building on Flare's previous experiences and work. [...] A derivative of this project was published by POSC in March 2002 as the E&P Cataloguing Standard V0.5 (EpiCat).'

⁷ Image Courtesy Flare Consultants

⁸ Guess what Flare sells!

down as product groups are selected. Flare's product brings up native format document and physical item icons. Most recently, Flare is working on an XML standard to support cataloguing - CAML. Flare feeding back findings to the standards effort. Developing connections to Microsoft Excel and industry application software.

[2D seismic data rationalization - Christine McKay, Landmark](#)

For Shell, '2D data is not dead'. Shell wants to be able to display 2D navigation data for all of the UKCS, to see its own data, other companies' data and speculative surveys. The advent of GIS makes data more available – but also reveals flaws in the data. There was a lack of 3D data in the 21st round. Old 2D lines from the 1980's proved vital. One 1973 survey was used and even some paper sections. But a single set of consistent survey names was lacking. Shell started with the CDA/DEAL data stored in PetroBank. Issues arose with multiple versions of the same line, 'unexpected' line groupings and inconsistent names. Landmark was asked to 'streamline' the Shell dataset of 5,000 surveys and over 120k lines. The process was automated 'as far as possible'. QC plotting –was reduced to a 'minimal but acceptable risk level'. Lines loaded to Oracle Spatial with spatial analysis scripts run and compare lines identifiers. This has allowed 80% of conflicts to be resolved. Otherwise a map is plotted for further QC and investigation.

[Seismic Data Management - Jill Lewis, Troika.](#)

Troika has 10 years as a 'transcription enabler' and can 'put a system in, anywhere in the world'. 14 projects are currently running using SeisDB, RODE, PetroBank, IBM Tivoli and FileTek's new GeoCube formats. Many issues have been encountered with lack of agreement on metadata storage – like where to put the survey name. Lewis warns that Qualstar – the only remaining manufacturer – has just announced the end of the line for 9 track tape support. IBM has announced a 3592 drive with 300GB capacity – and has a roadmap for up to a 1TB tape. Metadata management on these super high density supports will become increasingly important. 4D/4C will further complicate things – especially for smaller companies. Lewis introduced Troika's new 'Open Architecture Scalable Information Store' - OASIS. OASIS is a binary object store built on Open Systems – Linux, mySQL/PostGres – offering seismic metadata management, positioning data and random access to trace data.

Q&A

If OASIS is 'open' what Open Source license will apply?

Not sure if OASIS will be in public domain – may go for some sort of collective development.

[GIS in upstream – Nick Blake, Lynx Information Systems](#)

For Blake, GIS is now firmly in the mainstream of upstream interpretation. For Lynx, GIS is synonymous with ESRI – with a '90% share of the upstream market'. Blake states that ESRI has some 400,000 sites (all industries), 1 million licenses and a \$500bn turnover. Blake walked through various uses of ESRI's software – presumably for the benefit of the 10% of oil and gas users who are not already believers. A plug for Lynx's (very pretty) GIS-based geological map of Iran. Also ran – gridded [ETOPO](#) data and a 'DIY PetroBank' – Seisview.

[GIS in upstream – Karen Blohm, Robertson Research](#)

More pushing at the open door of 'GIS is good for you' and another commercial plug for Robertson's 'Explorationist's Workstation'. This offers support for data collection, management and analysis and also uses ESRI technology. The Workstation offers connections to Finder, Recall, OpenWorks, OpenRSO, GeoFrame and Petrosys. The key to spatial integration is in the metadata – identifying data, author, source, project, coordinate system and 'enclosure'. Robertson has developed 'rule-based' utilities for metadata population. Geoprocessing is used to convolve maps of source rock development, reservoir extent etc. for fairway analysis – illustrated with another pretty map of the Murzuq Basin, Libya. A plug for Robertson's Telus database.

[GIS implementation – Steve Ashley, Venture Information Management](#)

GIS has a lot to offer – but more focus is needed on implementation of E&P specific GIS. High quality data is the key. Venture uses an ‘Information Maturity Matrix’ after D’Angelo & Troy⁹. The ‘Matrix’ (aided by Venture) helps companies move from the ‘heroic effort’ state to ‘predictable risk’. Venture has also developed metrics for managing GIS implementations, although no examples were shown.

[Making maps from diverse data sources – Nathan Balls, Petrosys](#)

Petrosys’ ‘DirectConnect’ makes very pretty maps from diverse data sources. Tracing the history of GIS in oil and gas – Balls reflected on the growing GIS data volumes and the duplication of datasets subsequent to mergers. Many solutions fail to catch on (POSC, PPDM, GeoShare, OpenSpirit – and Petrosys’ own dbMap) because they failed to address the issue of finding and using data across multiple Oracle instances. In 2001, Petrosys introduced DirectConnect – allowing GIS users to take data directly out of multiple databases – including SDE, dbMap, Finder, OpenWorks etc. Composite GIS objects can be built e.g. well trajectories from Finder, tops from OpenWorks. DirectConnect provides instant update on change and a multitude of tabular reports leveraging XML/XSL. Uses Blue Marble technology and Schlumberger and Landmark dev kits.

Q&A

Will DirectConnect be sold separately from dbMap?

No, folks are generally not prepared to pay for infrastructure.

The GIS wars are over?

You no longer need to go the ESRI route of putting all data in the Geodatabase. Petrosys has more experience of E&P data objects than ESRI – the war is not over.

What skill set is required to manage metadata? What tools are available for physical to logical mapping – especially as data models and software are upgraded?

If you want all your metadata in Oracle – that can be arranged. For migration, SQL files can be extracted from C++ code – giving more control on migration.

[Well deviation data QC – John Harris, John Kelley, Hydro Projects](#)

Harris traced the history of offshore positioning systems from the hyperbolic systems of the 1970’s through Transit Satellites, short range radial systems (which were sometimes switched off because they interfered with drillers’ TVs!). Hydro Projects was involved early on developing software to verify well coordinates for Amoco. Kelly asked ‘do we really know where a well bore is?’ Survey errors are cumulative down the borehole – there is no redundancy to correct errors or ‘blunders’. A mistake in point n^o 1 propagates right down the well. Usually, no account is taken of the earth’s curvature in computing deviated well trajectories. Kelly showed how various errors and misapplied corrections could impact a long offset horizontal well. These could be over 100m in bottom hole location for a 11km horizontal well. HydroProjects [WellDev](#) program used to QC and recompute data.

[Retention Schedules in Records Management – Jamie Burton¹⁰, Instant Library](#)

Retention schedules are the ‘missing link’ in E&P records management. ISO 15489 defines what a ‘record’ is. Companies need to develop retention strategies – and shouldn’t keep records for longer than necessary. Regulatory compliance is a legal necessity – with the data protection act, HSE legislation and with legal ‘holds’ for pending court cases. Corporate policy has to be developed and enforced.

⁹ ‘Data Management creates value.’ D’Angelo and Troy, American Oil & Gas Reporter, July 2000.

¹⁰ Burton is editor of the Journal of Records Management.

[GIS in site restoration – Gavin Coppins, UKAEA and Mike Ayres, ESiT](#)

The UK Atomic Energy Authority (UKAEA) is working on the environmental restoration of UKAEA sites. UKAEA commissioned [ESiT](#) to develop the 'Images System' – for QC and tagging of data quality. Images leverages [ISO 19115](#) for all GIS metadata.

[Services are the secret for information management – Wendy Kitson, Kadme](#)

IM (not DM) – unstructured data – in oil industry. Services, not products are the powerhouse of IM in the oil industry. Decentralization has created havoc for the data manager – with data everywhere and no checks and balances on use. Bell's law – 'there is a new computer platform every 10 years'¹¹, further complicates matters. In 2000 a study showed that 80% of corporate IT budgets went on database administration – with five times as much spent on maintenance over cost of software. Business consultancy and services are where the money will be made in the next few years.

Q&A

Going for the IT service market with the acquisition of SEMA actually lost Schlumberger billions in market capitalization.

Yes, but IBM managed a big turn-around with its acquisition of PricewaterhouseCoopers.

How do you build the skill set needed for data management?

Look outside your company.

[Pitfalls in seismic data retrieval – Eleanor Jack, Landmark](#)

Data loss has many origins – tapes can get lost and be omitted from a job, files may be only partially transcribed or corrupted in transcription or demultiplexing. The problems stem from the particular blocked nature of SEG-Y data. Samples per trace appear in two places – in binary header and in the trace header – there may be conflict. Recording techniques differ – traces may not be of constant size – that indicated in the binary header may be just a maximum. When copied to disk, physical inter-record gaps on tape disappear. TIF is used to encapsulate blocked data to disk, the headers verified and then the TIF encapsulation removed.

It is actually easy to lose the complete data set! Jack showed some examples using RWEA data. In one survey, the format code said '2 byte fixed point' – 'it was lying!' Should have been 4 byte – the whole data set had been scrambled by the transcription contractor. Using the wrong sample rate or number of traces per shot can cause data loss. This is very hard to QC for uncorrelated Vibroseis data. Transcription programs are 'very robust' – they keep going even with the wrong input parameters. A major cause of data corruption is human error. Some data is priceless and cannot be re-shot.

[eSearch - Ben Trewin, IronMountain](#)

IronMountain (IM) acquired Hays Information Management in July 2003. IM recently signed a deal with Schlumberger for the merger of Schlumberger's AssetDB product with IM's eSearch. The product merge will complete in 2004. AssetDB and OpenRSO clients will be migrated 'real soon now'.

Q&A

Will eSearch have GIS functionality?

Probably

Will the data structure be configurable – with plug-in catalogs?

Yes

¹¹ c.f. McNaughton's law – 'there is a new law about something or other every now and then'

Disaster Recovery – Paul Duller, [Instant Library](#)

How quickly will your company be up and running after a disaster? A survey¹² showed that a major IT disaster was almost invariably bad for a company's health – and often proved fatal. Companies must plan for disasters – singling out 'vital records' as priority targets for the protection effort. Need to meet BS5454 storage standard. Some practical suggestions – move stuff way from pipes, keep things dry with roofs that deflect water from shelving, install smoke alarms etc.

Exhibitors and Sponsors

Records Management Society

For information contact [Jude Awdry](#) and <http://www.rms-gb.org.uk>

Instant Library

Various library, document management and consulting services. [Enquiries](#) and www.instant-library.com.

DeepStore

[DeepStore](#) secure document storage in Cheshire salt mine – www.deepstore.com.

¹² [Comdisco](#) Continuity Services Study 1997.

[Technology Watch Service](#)

This report was produced as part of The Data Room's [Technology Watch Reporting Service](#). For more information on this subscription-based service please visit the above link or contact

The Data Room
7 Rue des Verrières
92310 Sevres
France

Email info@oilit.com

Tel +331 4623 9596

Fax +331 4623 0652

2003 Technology Watch Program		
Date	Conference	Location
10-11/02	SMi Data and Information Management	London
11-12/02	IBC EU SCADA Conference	London
10-12/03	ESRI Petroleum User Group	Houston
12-14/05	AAPG Annual Conference & Exhibition	Salt Lake City
12-13/05	Datamation Lifecycle Process Management	London
20-21/05	IQPC Oil & Gas Knowledge Management	London
22/05	POSC Member Meeting	London
28-30/05	PNEC Data Integration Conference	Houston
10-11/06	Landmark European Forum	Aberdeen
29/09-1/10	Schlumberger information solutions Forum	Miami
6-8/10	SPE ACTE	Denver
20-22/10	AspenTech European user conference	Paris
27-29/10	SEG Conference and Exhibition	Dallas
3-5/11	PPDM AGM 📄	Calgary
8-9/12	PESGB Data Management 🗣️	London

N McNaughton acted as 📄 summarizer, 🗣️ Session Chairman.