digital energy journal

Doing more with satellite data

Value from unstructured documents

A better document search system

How to be successful with analytics

Using simulation to make better license bids

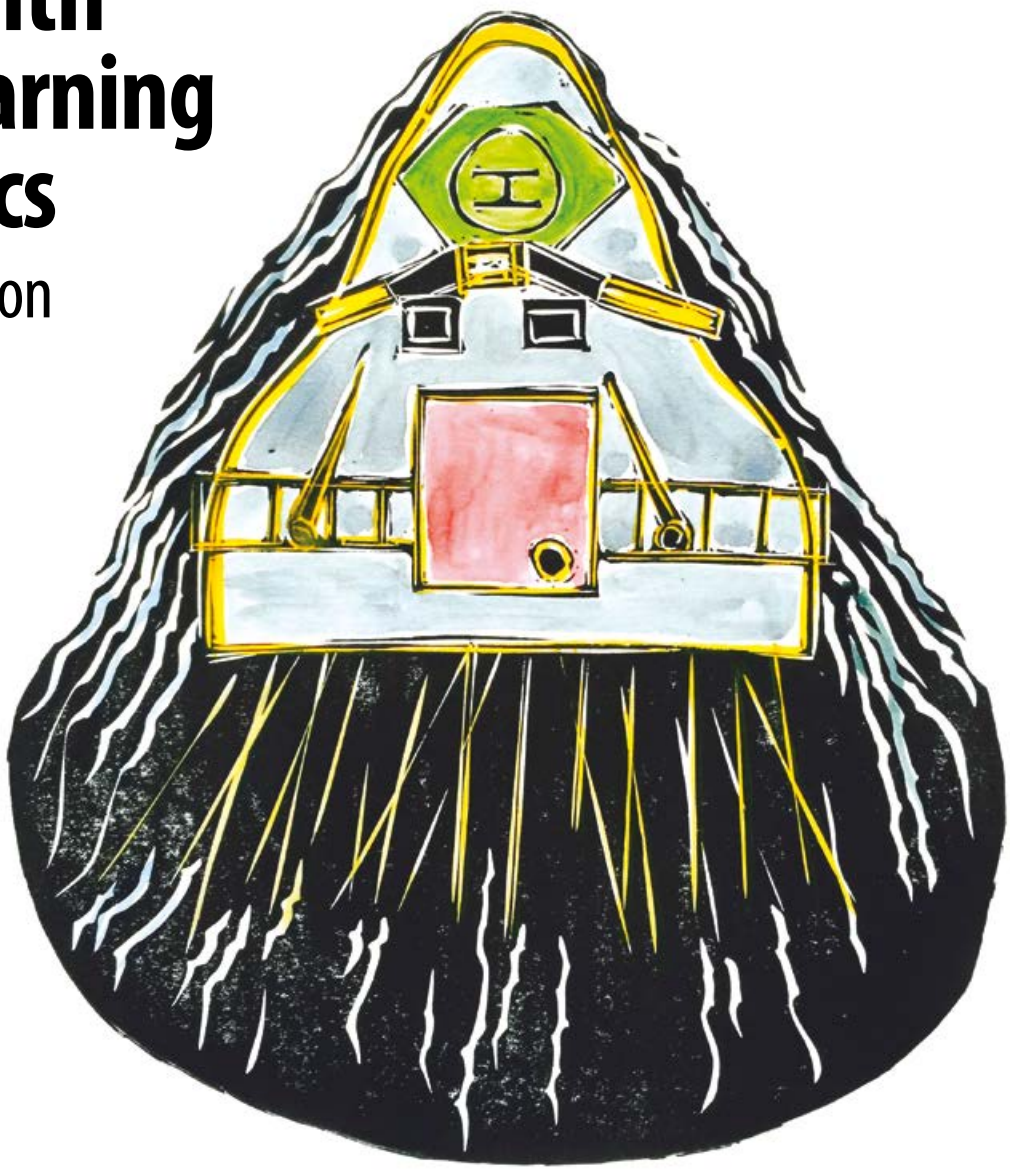
Helping your company board make decisions

Event Report, Solving E&P problems with machine learning and analytics, Nov 21, 2016, London

Special report

Solving E&P problems with machine learning and analytics

Nov 21, 2016, London



FLARE

Petromall

TERADATA

Petromall

Truthful, professional consulting; no repeating what you already know, and no spin.

Petromall is a unique oil and gas advisory service which prides itself on technical excellence in selected fields and supplementing business management and leadership in the face of uncertainty.

- Exploration
- Subsurface
- Project Capital
- Field Operations
- Security of Supply
- Decommissioning

Petromall actively participates in the facilitation and conduct of seminars worldwide, and has been a proud sponsor of this event with Finding Petroleum.

Our presence is across the globe in the UK, Malta, North America, with partners in the Middle East, North and Sub- Saharan Africa.

We help you to put science and engineering back into every decision made.
Visit us at

www.petromall.org

info@petromall.org

+44 020 3286 2556

Finding Petroleum

Solving E&P problems with machine learning and analytics

This is a report from the Finding Petroleum conference "New Exploration Ideas" held at the Geological Society, London, on Sept 20, 2016

Event website

www.findingpetroleum.com/event/f9cfb.aspx

Some presentations and videos from the conference can be downloaded from the event website.

Report written by Karl Jeffery,
editor of Digital Energy Journal
jeffery@d-e-j.com
Tel 44 208 150 5292

Sales manager
Richard McIntyre
rmcintyre@d-e-j.com
Tel 44 208 150 5291

Conference chairman and producer:
David Bamford

Layout by Laura Jones, Very Vermilion Ltd

Cover art by Alexandra Mckenzie

Digital Energy Journal

www.d-e-j.com

Future Energy Publishing,
39-41 North Road,
London, N7 9DP, UK
www.fuenp.com

digital energy journal

Finding Petroleum held a forum in London on November 21, 2016, "Solving E&P problems with Machine Learning & Analytics".

Topics covered included how the industry can do more with satellite data, ways to get value from unstructured reports, how to set about doing data analytics, using simulation and network modelling, and other areas analytics can add value to the industry.

The big promise is that these technologies might lead to an impact on the oil and gas industry which is as big as the impact of 3D seismic or horizontal drilling, said conference chairman David Bamford in his introduction.

"Not a day goes by, including weekends, where I don't get an e-mail or a link talking about some analytics technique which will absolutely transform the economics of the Permian or Bakken or Barnet or some other shale enterprise," Mr Bamford said.

Terrabotics – doing more with satellite data

UK company Terrabotics is helping oil and gas companies make more use out of satellite data. The company is a spin-off from research at the founders' doctoral research at Imperial College, London, launched 2 years ago.

Gareth Morgan - Founder and director, Terrabotics

There is a massive growth in satellite data available. There are more and more satellites being launched, and they can take more high resolution photographs. "We're developing and deploying smart algorithms to handle all this data and make sense of it," said Gareth Morgan, founder and director of Terrabotics.

He was speaking at the Finding Petroleum forum in London in November, "Solving E&P Problems with Machine Learning and Analytics".

Companies are talking about 'earth observation 2.0', combining a number of different technologies, with machine learning and computing power, he said.

Satellite imagery has never routinely established itself as part of exploration processes thus far. It may be used here and there, but not throughout the whole organisation, Dr Morgan said.

Oil and gas activities cover a large area of the earth's surface – an estimated 26m square kilometres of active oil and gas license areas – so you might expect satellite data to offer a great deal of value, he said.

The capabilities of satellites may have been oversold to the oil and gas industry in the 1980s and 1990s, he said. But in the past, satellite imagery was much more expensive than it is now – around $4400 per image in the 1990s – and there wasn't computing ability able to store, manage and analyse them. The images were not reliably available.

Today, the computing power is here, the reliability is here and the costs have dropped, with a great deal of imagery available free, he said. Prices are as low as $5 for a photograph covering a square kilometre, with pixel size on the ground of between 30cm and 1m.

Many different fleets of observation satellites are being launched, including a number of companies with constellations of satellites. We are getting video recorded by satellite.

It would be possible for oil companies to buy their own satellites if they wanted to, and have even more control over what was available.

"We are awash with earth observation data, the supply problem is increasingly going away," he said.

Image stacking

Because so much imagery is being recorded, it becomes possible to work with multiple images of the same part of the world.

Images can be "stacked", the way seismic data is stacked. The more stacking you do, the more signal you have and less noise. Other "atmosphere perturbations", such as a cloud, can also be filtered out in this way.

"This gives us super enhanced signal boosted images; much better to then perform analysis on"

It is also possible to see how images are changing over time, for example to track a spill or construction activity.

In the past, analysts would usually only have one satellite image available, and apply a number of manual or semi-automatic techniques on it to try to get some understanding of it, in a process which was "not very scientific," he said.

Taking 374 images of the same part of the world would have cost $1.6bn in 1990, but now it is free.

Data from drones can play a supporting role to the satellite images, he said. "We work a lot with drone/UAS data as well," he said. "It's still early days for processing drone data. There's lots of data not processed correctly."

Using the data

The raw images have been used in the oil and gas industry to see where oil wells are, perhaps for a company which has lost its records.

They can also be used to see what drilling activities other companies are doing. The computer can be trained to look for well pads in satellite photos (area cleared for a drilling rig to work), by comparing one picture with the previous one of the same spot.

The satellite images have proved very useful for companies who need evidence for authorities.

There was one example from a small island country, where the government was blaming a chemical plant operator for a paraffin spill into the ocean. Analysis of satellite images, showing the path of the plume on the water, showed it was actually coming from a government facility, Mr Morgan said.

With something like this, it would previously have been basically luck of whether it was possible to obtain the right satellite images, he said. "Today, you have data every week, and soon every day, and there's radar data that gets through the clouds," he said.

Terrabotics' aim is to enable scalable machine learning on the images, for example so a computer could spot a certain geological feature.

To do this, the machines need to be trained by humans. This is similar to the way that social media sites ask people to train the computers to identify images. There are probably costs involved to this – people won't expect to be paid for tagging their own faces, but a geologist might expect to be paid for tagging basalt cones.

Terrabotics also works with people with all kinds of expertise in machine learning in imagery but not just satellite imagery – for example self-driving cars and robotics.

The analysis work can be easily scaled, to cover a whole basin.

Teradata – how to get more insights from data

Finally, oil companies are showing more interest in data analytics – but making it work is still far from easy. Duncan Irving of Teradata provided some tips.

Oil companies are starting to show more interest in data analytics, and there is a lot of capability for data analytics to add value to the industry, said Duncan Irving, Practice Partner, oil and gas with Teradata.

He was speaking at the Finding Petroleum forum in London in November, "Solving E&P Problems with Machine Learning and Analytics".

We have a lot of challenges of getting this out beyond 3-6 month projects and some point solutions, he said. "It's not going to be easy bedding this in to our current way of doing things."

The top level picture is sometimes simply that companies have been told by management consultants that they should be doing "big data" and then they bring in a company like Teradata to show them how to do it, Mr Irving said.

'Big data' techniques can help companies work better with the large amounts of data they generate which don't fit into any existing schemas or workflows, he said.

Through these techniques, many companies have got a better understanding about how people interact, how they connect in social media, how they connect with organisations when we buy their products and services.

Companies using data science are starting to bring in a geographical context, so people can understand the surrounding environment. "It is starting to help machines put themselves in a context," he said. "It will help us understand how we fit into a population."

"We've always had this desire, ever since we started creating data, to get more insights from it, something more sophisticated than extrapolating," he said.

The big challenges are understanding where analytics can add value, getting access to data sets, making sure the data is good quality and the data has business value, he said.

For analytics to work, the systems need to have a close coupling with whoever in the business will actually make decisions.

Another major challenge is getting analytics from a one-off project to something you use every day, he said.

There is an opportunity for the industry to get started at the moment though - much more than there was a year ago, when most industry people were mainly concerned with whether they would keep their jobs.

The adage that people in the oil and gas industry are "good at going second" is still true (i.e. no-one wants to be the first). But people are now "going second a lot faster," he said.

Sometimes the results of the analytics only show ways to change your results by small amounts, such as 4 to 6 per cent. This can mean that analytics is used more widely in industry sectors which run on tight margins, such as retailing. The oil and gas industry might be used to thicker margins, but it is having to get used to margins like this now.

Success stories

There are not many success stories about how analytics is used in the oil and gas industry – because people who are successful tend not to talk about it, he said.

"There are two sorts of companies that don't talk about analytics - the ones that aren't doing it and the ones that have been doing it for a long time," he said.

When companies have been doing it for a long time, "it is such a competitive enabler they're not able to tell anyone that they are doing it let alone what they are doing," he said.

However, Shell has told some success stories. "They are very mature in the way they do this, everything in Shell is measured and often analysed and often, I guess, acted upon," he said. "Every now and then they'll do an industry presentation on it. They are analytically driven. They used to be a trading company back in the day so that's in their life blood."

"I'm sure there are other ones out there of that size that do it. BP are doing a massive analytical transformation internally. They are excelling in some ways in the way they are applying analytics. They quietly get on with it."

"It is the next tier [of oil companies] that aren't culturally ready for it. How do you sell it to them? You can sell it until you are blue in the face. It will polarise the industry into leaders and laggards."

"Once someone realises they are not benchmarking with their peers anymore [it changes]. That typically is what's happened in other industries."

Defining 'data science'

Artificial intelligence has been in discussion since the 1950s, he said, coming out of efforts to crack codes in the Second World War.

In the 1970s, people started talking about "data mining," the idea of getting data into one place and looking for relationships in it. What you could do was limited for decades by computational capability.

In 2016, the term used is usually 'data science', although it means roughly the same as data mining, he said. "If I'm talking to someone 10 years older than me, I say 'data mining', because that's [what they] want to hear about," he said.

You should not get too worried about the terms, but try to understand the "simple and fundamental" concepts behind them, he said.

"The whole thing is still moving so quickly, it is difficult to nail down what it is a data scientist should be," he said.

"There are many components you have to bring to data science. You have to have domain and computing expertise, the right maths and statistical insight, and the right blend for any particular project."

You can say that data science is still at very early stages, like the aviation industry at the biplane stage, he said.

Getting started

To get started on data analytics, you need to put the right team together, have a suitable budget, and deliverable expectations from the budget holder, he said.

Recruiting people with the right data science skills is becoming difficult, with banks and retailers also on the hunt for the same talent.

With about two thirds of the project time about to be spent on collecting data, it is important that everyone understands how long the project will take, he said.

You need a feel for what it might be possible to deliver for a certain budget.

Many different skills are involved. "You can't just put the data scientist, the domain expert and python coder in a team, and say 'go do this for 3 weeks' and see what they come out with," he said.

"You need a geophysicist who has an understanding of statistics, a statistician who has an understanding of the algorithms and how they are coded."

"You need to have someone who understands the architectural aspects of how to scale this up to 100,000 wells, when you started with 1,000 wells."

You can gradually develop a sense of what sort of analytics might work in your domain.

For example in geology, you might use analytics to try to work out how similar the formation you are looking at is to another formation, or why it is different. There are data science projects to do automated boundary and feature classification.

You can do analytics on passive seismic data recording. Rather than store all of the data (which can be an enormous amount of data, if you record continuously with a number of receivers and channels), you can do analytics to find out if the data shows anything useful, and dispose of the rest of the data. You can see what the dominant frequency is, or if it is changing as the water pressure around the well changes.

Duncan Irving, Practice Partner, oil and gas with Teradata

As you develop your ideas, there can be lots of loops to go around. You might have an idea for data analytics but then discover that the person who owns the data is not willing to share it because you are from the wrong company department, and have internal company negotiations to make.

You might discover that the data is in the wrong format, or much of it is missing, or there isn't data for enough wells, or it has an error because the drill depth was measured wrongly. "All of these things take you around this loop several times," he said.

Modelling

Eventually you get good enough data to "break into the world of modelling," he said.

You build models and analyse them to see if they show anything useful, for example can you get some understanding about what usually happens when wells are drilled into this sort of formation, or what amount of logistics capability you usually need for this rate of field development, based on what happened last time.

Some companies are getting faster and faster at updating their models, including in the US unconventional oil and gas industry. Some companies in the refinery industry are rebuilding their analytics models every 12 hours, he said.

Spurious relationships

The first problem you run into is that the relationships shown up by the analytics might not actually be a useful relationship. Instead of one parameter driving another parameter, actually there's a hidden parameter driving both of the ones you are looking at, so they seem to change at the same time.

An example of this is when some scientists claimed there was a correlation between autism and subscriptions to cable television, he said.

However people with statistical skills can analyse this, and show whether one factor is causing another factor, or if both factors are caused by another (perhaps unknown) factor.

But do not use the 'spurious relationships problem' to argue that all analytics is useless, he warned.

An example is the story about "Google Flu Trends," when an analytics team at Google thought they could predict where flu was about to break out by correlating flu searches with people's locations, but it was later shown that they were just showing which parts of the world were in Winter.

But the same analytics team went on to develop a useful tool which could predict the price a house would sell for.

"If you get the domain expertise right it really does deliver value," he said.

Crossing to production

If the data analytics starts to show something useful, it is important to make sure that the business starts using it, and people in the company trust it. "You have to be able to cross that void into production."

The industry needs to keep records of what works, and how it is built, and share code around internally.

They should manage the work in-house, rather than with external consultants, and have company employees in charge of it.

IT infrastructure

To make this stuff work, you need an IT infrastructure which can make the data available, computer platforms to run the analytics on, and people who understand the mathematics of data science, he said.

One obstacle is that oil and gas industry has got in the habit of expecting to have software tools to do everything it wants, or 'push

button functionality', he said, expecting either a special software application or a plug-in.

The industry now usually ends up with software plugged together in a very ad-hoc fashion.

This can all make running analytics projects "really difficult," he said.

Data quality

To get started with data science "you have to have really good data," he said. "That's got us stuck already as an industry."

Bad data means that most of the time is spent sorting out data, rather than analysing it.

"I don't think there's a single [analytics] project where we have spent less than two thirds of the time preparing the data for analysis," he said. "That's a surprise the first time you do it."

"If you think about the time and cost of that, whether a consultancy doing it, or [it is done] within the organisation, that is a significant overhead before you get to the most simple of reporting on what the data has in it."

"Our data has often been collected from instrumentation that has been out there for decades. The sensors were deployed and data was collected under a different paradigm," he said.

Once you have enough clean, high quality data, "you can perform some pretty fantastic analysis on it, give yourself some very robust insights," he said.

The best way to achieve good data quality is caring – if people care about the data they are entering themselves, and they also want the data to be useful in the company, which means they are happy to let other people access it.

"Someone has to be in charge of it, responsible for it, own it," he said.

Teradata is involved with one project in seismic acquisition, where it gets all the companies involved in the data chain to take more responsibility and ownership over the data, so that at the end you get something people feel they can trust more.

Data packages

One challenge is that the industry typically keeps its data in packages, for example a log file which contains a header and different log files all as one package.

This is equivalent to keeping all of your Lego in the small boxes it came in, rather than tipping it all out into one bin, he said.

"Instead of having all these silos you centralise the data," he said.

"You should be able to let all of the different analytical workflows loose across all of your different data types, without being constrained by data type, business unit, data type, data age, any of those things," he said.

"That sounds like crazy talk, because we're a long way away from this place in the oil industry."

It is technically easy to move data into a 'sandbox' so there is nothing lost from messing around with it, except perhaps a week or two of someone's time.

Where it can go wrong

Data projects can go wrong if the data you begin with is not good enough - it isn't big enough, wide enough, or doesn't go long enough back in time.

Projects can go wrong if you have the wrong data scientist. The data analysis can be quite domain specific, so someone who has done data analytics for hedge funds might not be good at seismic data analysis.

It can go wrong if the aspirations are wrong. "If someone says, 'I want a fully predictive oil refinery that works like Waze does and tells me where the traffic is going to go' and there's full stop or no pause for breath in that sentence, and you go 'yeah', you're on a hiding to nothing," he said.

"They are expecting their oil refinery to become a sentient being within 6 months."

"You have to stage and articulate what you are doing in a very granular level."

Projects can go wrong if the people driving it are just doing it because they want something with Hadoop on their CV, rather than actually wanting to add value or remove cost to a particular business function, he said.

Some people suggest that the oil and gas industry should develop domain specific data scientists, such as a geologist data scientist.

But a risk with that is approach is that the data scientists spend so much time with a particular team that they lose their objectivity and ability to find new answers, he said. Some companies, including Facebook, have a policy of moving data scientists through different business units, so they maintain an objective approach into the best way to use data to find useful answers.

Who should be in charge?

There is a big question of who should be in charge of the project - IT department, the business sponsors or the data scientists. Having tried all three, there is no clear answer, Mr Irving said.

We can cross out IT. "I'm pretty adamant that data science, data mining, and the analytics agenda should not be driven by the IT organisation in any large company," he said. "They are there to enable but they are not there to lead on this one."

Perhaps not business people either. "I think that you can talk to very smart people on the business side, but they have a day job to get on with. They may have some opinions, which are valid because they are the ones who have to make the decisions. But they need to be sold the credibility of this," he said.

"I'm also pretty certain you shouldn't let the data scientist drive. You can think of reasons why someone with a stats background may have the right answer and you can think of reasons why they would have the wrong answer."

"I think that someone who has an understanding of all three of those is probably the right answer (IT, business and analytics). "But then where do they live in an oil company? Should it be an asset team, should it be central facility?

Flare – value from unstructured documents

UK consultancy Flare Solutions is developing a range of methods to get value from unstructured documents – including working out when one document is similar to another, and helping people to find and manage the documents they need.

Dave Camden, IM Consultant with Flare Solutions

UK consultancy Flare Solutions is helping oil and gas companies get more understanding from unstructured text based documents – and find new ways to find what they are looking for.

The work involves using computers analyse the raw text, looking for and recognising document analogues (where the content of a report seems similar to a 'standard' reference one).
The work is structured around probabilities, not absolutes, said Dave Camden, information management consultant with Flare Solutions. It shows you the most likely best answers, not just a 'yes' or 'no'.

This should make the system more robust than a system based on rules, like the 'expert systems' of the 1980s, which will easily collapse if the data behind them does not fit with one of the rules for some reason, he said.

As far as possible, the system aims to reduce cognitive bias, which can influence human decisions through pre-conceived ideas that are not always based on real facts or knowledge.

But when building up a knowledge base, for example by learning from existing text, you cannot entirely avoid cognitive bias since the text is written by people expressing their ideas and concepts. Choosing peer reviewed articles, looking across texts from multiple authors or choosing very large data sets will minimise the problem.

It is better if decisions can be made based on 'known facts', something people have collectively, rather than individually, shown to be true, he said.

"We have tried to strike a balance between using standard 'reference' values and relationships (known facts) and using information extracted from the target text (new facts that we can learn from)," Mr Camden said.

Analogues

Much of Flare's work is based around trying to help identify 'analogues', or where something is similar to something else.

As human beings, much of our understanding of how the world works is based around spotting patterns, where we have seen something similar to the event which is happening now.

Similarly, in geoscience, it can be helpful to find examples of wells, or rock type, from our past experience, which are similar to the wells we are working on now, he said.

The same approach could be used in reducing downtime on operational equipment. As (non-digital) human beings, we build up expertise about how things fail and indications that something might be about to fail. For example we learn that a small leak in a water pipe could indicate that the whole pipe is about to fail.

In the digital world, we can try to compare what is happening now with what has happened in the past, because equipment might fail by going through a similar sequence of events every time.

Inverse analogues can also be helpful, when you find something which is very different to what you would expect, based on what happened in the past. So maybe a computer analytics system can deliver these too, he said.

The human computer challenge

Computers are a long way from being able to understand much in the human world. As an example consider how hard it would be to explain a joke to a computer, he said.

Although you can program the computer to classify something as a 'joke', it could struggle with identifying similar jokes and would certainly not experience the joke as a human does.

Similarly, things in the real world can be very case specific – you could train a computer to understand when a certain component is about to fail, but the computer may not use that knowledge anywhere else.

This means that humans with domain expertise, or deep knowledge about a specific domain, will always be extremely important, because they will have to train the computer systems how to understand that domain, he said.

Explicit training or training based on human-written texts are currently a major component on making computers seem smart.

"Humans are passing knowledge onto systems and the systems then are learning from that and making inferences," he said.

The CDA challenge

Common Data Access, an organisation which helps manage the UK's national oil and gas data, recently issued a technology

challenge for companies to come and try to generate some value from the organisation's archive of exploration and production data and reports.

Flare participated in the challenge, working through about 25,000 unstructured data reports, mostly well documents for the North Sea.

Flare decided to try to look for 'analogues', parts of the North Sea which are very similar, but the similarity was not previously known.

OCR

The first step of working with the 25,000 exploration and production reports was to scan them with Optical Character Recognition (OCR), so the computer could 'read' them.

There was a continuous learning process – getting to know alternative spellings, and spotting errors from the optical character recognition, he said.

OCR doesn't get everything completely correct, but "many mistakes tend to be somewhat predictable, such as confusing zeros with noughts, ones and the letter l," he said. Spell checkers, or comparing the words with a reference list of words, can pick out many OCR errors. "We'll build a big knowledge base to enable us to do that."

OCR failures may add more 'noise' to the system, but are unlikely to create false matches, he said.

Another task was to remove so-called 'stop' words, like 'if' and 'but', which are commonly used in English but don't add anything to the technical understanding.

Formation Analogues

This project was based on characterising geological formations and finding analogues based on the words (terms) that occurred around the formation names in the 25,000 documents.

This particular project did not start with a list of known formations, rather the formation name were extracted from the text (although a previous similar project did use an existing field list).

It set out to characterise each formation by a number of factors – the type of lithology, the age and depositional environment, taking this data automatically from the text.

Text analytics

The next stage was to analyse the text to see which words (terms) occur in proximity to each other (co-occurrence). There were hundreds of thousands of terms in the CDA text set.

"Words that occur together generally share a similar concept," he said. "The idea is that you can tell a word by the company it keeps. Words that are infrequent probably contain more information than words that are frequent."

The outcome of this analysis is, for every term, a 'co-occurrence fingerprint' of 300 values. By comparing these fingerprints we can measure how similar terms are.

Some of the terms are the formation names we have extracted from the text, others are the lithology values that we 'know' from our reference knowledge base that also occur in the text. For each formation, we now compare its co-occurrence fingerprint with those of each of 200+ lithology terms, thereby creating a 'lithology fingerprint' of 200+ values for each formation.

To match formation based on lithology, we just look for the most similar 'lithology fingerprints' between a user-chosen formation and other formations.

A similar process is carried out for other aspects of formations, like geological age, depositional environment, production, problems.

The system can work in any (human) language, although if it is based on a certain knowledge base (such as oil and gas documents), they are likely to be mainly in one language (English in this case).

Search

This text analysis can then be used as a basis for more sophisticated search tools.

Flare showed a prototype search based on the same similarity methods used in the Formation Analogue system. The used can input one or more terms (for example, 'turbidite, shale, and tuff') and the system will respond with a ranked list of formations that best match those terms.

For the future, Flare is currently developing systems around graph databases (which shows which terms are related to which other terms). "That will give us a lot of capabilities in this space to do a lot of this kind of work," he said.

Information management

Mr Camden's company, Flare, sees itself as an information management company, but "we're looking to blur the boundaries between managing information and exploiting it," he said, "trying to link the information management world with people who consume the information."

"Analytics is about trying to glean insight from information that's already out there," he said.

"As information people, we've struggled for years with trying to make the thing relevant as far as the business end user is concerned. We see analytics as one way to achieve this."

You can watch Mr Camden's talk on video at **www.findingpetroleum.com/event/f9cfb.aspx**

Simudyne – how simulation and modelling can help predict

Simudyne is helping oil and gas companies use simulation and network modelling, along with machine learning, to make better predictions and improve decision making.

Justin Lyon, founder and director of Simudyne.

UK company Simudyne is bringing together machine learning, simulation and network modelling to help oil and gas company executives make better decisions.

Together, these three elements provide "incredibly powerful techniques for predictive analytics," said Justin Lyon, founder and director of Simudyne.

"One of them is really powerful, but when you combine all three, you combine statistical techniques with techniques for understanding organised complexity."

Decision makers in businesses today are trying to forecast, manage risk and achieve results, all at the same time. "It requires brilliant decision making at pace and at scale," he said.
To help them to do this, decision makers should have sophisticated models available to them about how the real world works, he says. These models should also run on real data.

Decision makers also need tools to provide them with better answers built very quickly. "If it takes longer than three months, it's an R+D project, and no-one is going to pay for it except R+D," he said.

The models need to be sophisticated, because the real world is very complex. In the past, many decision makers, including banks, have made decisions based on standard linear regressions (an understanding of how one parameter will change if another parameter changes).

But these tools do not provide any understanding of the deeper cause and effect relationships, he said. As a result, they do not have a clear idea of what is driving what. "When the crisis hits, they wonder what happened," he said.

Mr Lyon sees executives in two types, the gut instinct driven and data driven.
"We don't work with the [gut instinct driven] people," he said. "We're interested in people who want to have a fact based data driven approach to making decisions. They [are happy for their] assumptions to be scrutinised and made explicit by computer models."

One challenge with trying to build interest in these sorts of tools is that no-one wants to talk about their successes, in case other people copy them.

But, "we need to talk about successes so people can know there's gold in the hills and get the investors to go and dig it out," he said.

Also, no-one wants to talk about their failures either. "But we need to, because only thorough our failures can we learn to make successes possible," he said.

For building the simulation models, typically you have a multi-disciplinary team to build the solutions. This can include a number of data scientists, for example one specialising in machine learning, one specialising in computational simulation, one on network modelling.

Sometimes you can re-use some of the work, for example if you build a tool to help make bidding decisions that could be used in other scenarios, the same user interface but different data. "We're also finding that some of the models are surprisingly also re-usable," he said.

Mr Lyon is a former information and physical security contractor with the Bank of England, developing policies for information and physical security (2011-2015). He has also worked on a simulation for US Health Care (2012), and a project on retail banking technology (2012).

Simulating bidding in Mexico

Simudyne was involved in a project for a (name undislosed) oil and gas company which was bidding for oil and gas rights in Mexico.

The company wanted to use simulations to try to work out the best bid strategy, based on a prediction of the bid strategies other companies might follow.

Simudyne put together a large scale computational simulation. The simulation includes 'agent based simulation' (simulating how individual 'agents' might behave), as well as traditional data and calculation models.

It basically builds "a virtual world of the whole landscape," he said.

The simulation took into consideration the business environment and where the hydrocarbons were thought to be.

Simudyne acquired a range of data, and built a range of simulation models, including geotechnical issues, price outlook, returns expectations, capital expenditure.

By using the tool, the project team "predicted the formation of a consortium which they [otherwise] thought was never going to happen," he said. "In the real world it did happen."

Simudyne wanted to provide the company decision makers with tools to try out different decisions well in advance of the deadline, he said.

The decision makers can take the simulated results, analyse them and use them to adjust decision making, before they make decisions in the real world, he said.

They could adjust the parameters, and see what might happen as a result of various bids they could make. They could look at the perspective from different companies.

Every time they suggest a bid, the software

would run millions of different scenarios around it.

The simulation could be run to look at what might happen on individual blocks or companies. Some of the parameters was preset, and some were adjustable.

Altogether the company could put together a chart of what each company might be likely to bid.

The project needed to be put together within a tight deadline, because the license blocks were going to be awarded in four months time.

Executive communication

The 'visualisation' part of the software, building the thin client 'apps' which company decision makers would actually work with, should be done separately to developing the actual simulation, he said.

Left to themselves, "the data scientists produce something that's elegant for data scientists and not understandable for the executives," he said.

It is very important that the decision makers can play around with the simulation easily, he said.

"We had brilliant mathematicians, but if their insights aren't communicated in a fashion that executives can get their heads around, the insights are missed and lost."

"The idea of giving this to a decision maker on a Surface Pro or iPad and letting them explore with parameters and fail safely [is important]," he said. "They can run an infinite number of scenarios themselves."

The skill of building the visualisation layer is quite similar to traditional website design, he said. You need to have someone who will have conversations with the decision makers who will use the software, about what they want to work with.

The software needs to be able to support multiple users signing in simultaneously, with different teams making decisions about different parts of the system, he said.

If you want to optimise a complex value chain, there will be many different decisions, and a decision in one place will have an impact somewhere else.

Silos

A challenge with developing simulations is that companies often keep their data in different data silos, and a lot of the value is only generated when they can be brought together, he said.

Studies show that 99 per cent of the data generated by offshore oil platforms is never used.

"McKinsey is doing lots of reports saying, as you combine the data, that's when you generate the real value," he said. ""If you don't bring the data together, you leave so much money on the table."

Other business applications

One audience member asked whether the software could have predicted the outcome of the US presidential election.

Mr Lyon replied that it might be possible if you could build a large scale model of the entire US population, and gather as much information as you could about them, including from their Facebook profiles (although the privacy issues could be an obstacle).

The software could handle 300m independent agents, he said. You would create a model of how each person is likely to behave.

So technically, it could be done, legally maybe it would not be so easy, he said.

Simudyne is also considering using the software for modelling London house prices, trying to predict when prices will burst and the reasons that might cause it.

David Bamford – where analytics can help

Machine learning and analytics could be used to help company boards make decisions (or help you make a case to company boards); it could help combine different sorts of survey data; it can help you work out which parameter is most important, and can help you.

David Bamford

Analytics would do a useful service if it could help make it clear to the boardroom that a certain sort of spending is worthwhile, said David Bamford, consultant with Petromall and a former head of exploration with BP.

He was speaking at the Finding Petroleum forum in London in November, "Solving E&P Problems with Machine Learning and Analytics".

One example (from the 1990s) was when board members of BP were persuaded that spending on large 3D seismic surveys would make a material impact to a metric which board directors did care about, the overall exploration success rate.

"[When] you plug that parameter into the very simple economic models that got traction in the boardroom, and change exploration success rate from 20 per cent to better than 50 per cent, it's a very convincing argument," he said.

Another example was an analytic study to compare seismic contractors, which could show the relative contribution each company's seismic made, in terms of ultimately helping the company to improve production.

The study showed that data from one seismic contractor was ultimately much more useful than data from another one, he said. This is also very useful information for company decision makers.

Working with satellite images

Analytics could also help in the oil and gas industry if it was used on satellite imagery, particularly if the results were compared with other sorts of surveys, including on-site work.

Companies might be able to use high resolution satellite imagery to get better at looking at structure, drainage patterns, and geomorphology (study of physical features of the surface of the earth).

In a similar way, industrial agriculture firms are using satellites and airborne data to try to better understand stresses on crops, including from water drainage, he said.

The satellite images could be used to detect subsurface hydrocarbon accumulations, if small amounts of hydrocarbon seep up to the surface, and cause a detectable change to the rocks or vegetation.

As an example, petroleum near to the surface can start depleting oxygen from rocks and soils, which can lead to rocks changing colour, or vegetation being stressed. There are examples in the US or Russia where there is a change in the colour of a forest in a certain spot due to petroleum below. This could be detected by satellite.

It may be possible to detect useful materials using magnetometers (measuring magnetism). In one example, geologists working on the ground mapped a pyrite (iron sulphide) mineralisation near an oilfield by drilling lots of boreholes. This could have been done more easily by flying over it with a magnetometer on a drone, he said.

Companies are also doing gravity surveys in much higher resolution, a technology which "has been a driving force for onshore exploration in East Africa," he said. The gravity survey can help you get a better understanding of the structure.

So by combining all this together, satellite images, seepage maps, geomorphology, aerial gravity and magnetometer surveys, and on site geology, you can get a much better understanding.

"The problem is that I have not seen anybody who is capable of delivering an integration of these different measurements," he said.

Picking the key parameter

Analytics can also help work out which parameter has the biggest impact on increasing production (or whatever your main objective is).

This is not always obvious. As an example, many unconventional drillers thought that the more proppant they pumped into the well (which forced the rock open during hydraulic fracturing), the more production they would get.

Analytics showed that this is only true to a point, after which more proppant does not lead to more production. "There's a lot more going on than just a simple correlation," he said.

Other parameters, such as rock type, age and depth, which company is doing the work, have proved to have a strong impact on the overall result.

Helping follow rules

Another application of analytics is helping subsurface interpreters follow the rules.

For hundreds of years, geologists have been developing sets of rules about how geology works, and today "you can regard geology as a rule driven science," he said.

People can work out how something was formed from its geometry. It does not need any imagination. "We all know what those rules are, if we work in this industry," he said.

However there is a great deal of interpretation in how the rules are applied.

The UK's Oil and Gas Authority recently did an analysis of wells drilled offshore in the Central North Sea and Outer Moray Firth to try to get a better understanding of dry holes and failures. It showed that "one of the biggest reasons for failure is interpreter error," he said.

So perhaps the computer systems could help by automating the rule-based steps, and perhaps identifying more rules, through automatic analysis of company archives.

You can watch David's talk on high resolution video at
www.findingpetroleum.com/event/f9cfb.aspx

Solving E&P problems with Machine Learning & Analytics
The Geological Society, London, November 21 2016 Attendee List

Iain Poole, Head of Oil and Gas Consultancy, Barnett Waddingham

John Boucher, Director, Beagle Geoscience

Philipp Chladek, Analyst, Bloomberg Research

Argiri Patouni, BP

Bob Spence, Director, Capital Project Partners Ltd.

James Goldwater, Senior Developer, Cegal Ltd

Neal Coles, Software Architect/Product Owner, Cegal Ltd. (Formerly Blueback Reservoir)

Mustafa Elsherif, Sales Manager, CGG

John Glass, Consultant Geologist, Cloverfield Consulting Ltd.

Feargal Murphy, Consultant

Christopher Roberts, Principal, Consultant

Micky Allen, Consultant

Dan Kunkle, Director, Count Geophysics

Maria Mackey, Energy Sector, Cray UK

Hermas Amaewhule, Data Scientist, DMV Solutions

Juan Cediel, Consultant, DNV GL

David Cleverly, Geophysicist, ENGIE

Telamon McCullough, Reservoir engineer, ENGIE

Jonathan Moore, Product Manager, Evaluate Energy Ltd.

Chris Jepps, Technical Director, Exprodat Consulting Ltd.

Karl Jeffery, Editor, Finding Petroleum

Richard McIntyre, Sales Manager, Finding Petroleum

Avinga Pallangyo, Conference Producer, Finding Petroleum

John Leggate, Managing Partner, Flamant Technologies

Dave Camden, IM Consultant, Flare Solutions

Glenn Mansfield, Director, Flare Solutions Limited

David Sagi, Structural Geologist, Getech

Thomas O'Toole, R&D Geophysicist, Halliburton

Ben Saunders, Geoscientist, Halliburton Landmark

Wally Jakubowicz, Managing Director, Hampton Data Services

Norman Hempstead, Director, Hempstead Geophysical Svcs

Alastair Reid, Consultant, IHS

Glenn Jones, Research Associate, Imperial College London

Sam Fielding, Independent

Manouchehr Takin, Independent consultant

Rahul Malhotra, Associate Partner, Infosys

Ronald Doherty, Manager Field Developments, INECTSEA

Clayton Ludik, Information Systems Consultant, INTECSEA

Peter Allen, Consultant, Layla Resources

Neville Hall, Director, Llahven Ltd

Alan Smith, Director, Luchelan Limited

James Dodson, Business Development Director, NEOS

Mac Chipunza, Head Of IT, Ovationdata

Robert Parker, Consultant, Parker

Alastair Skelton, Process Consultant, Pegasystems

Vincent Sheppard, Chief Geophysicist, Petrofac

Mohit Khanna, Chief Geologist, Petrofac

Temitope Adeyinka, Energy Analyst/Economist, Petroleum Development Consultants

David Bamford, Petromall Ltd.

Kes Heffer, Director, Reservoir Dynamics Ltd.

Nick Bright, Retired

Brian Hepp, President, Rocky Mountain Limited

Robert Waterhouse, Director, Rosha Resources Ltd.

David Webber, Seismic Operations Supervisor, Sceptre Oil & Gas

Sophia Fannon-Howell, SFH Consulting Ltd.

Tom Martin, Director, Shikra Consulting

Justin Lyon, Founder and director, Simudyne

Tony Edwards, CEO, Stepchange Global

Peter Brownsmith, Head of Business Development - Energy UK&I, Tata Consultancy Services

Duncan Irving, Practice Partner, Oil & Gas, Teradata

Gareth Morgan, Managing Director, Terrabotics Ltd

Peter Roberts, Business Development Manager, Tessella

Matthew Jones, Tessella

Brian McBeth, Managing Partner, The Oxford Consultancy Group

Hugh Ebbutt, Independent, Upstream Adviser

Ugur Algan, Director, Volantice Ltd.

Katerina Gunningham, Principal Consultant, Wipro Technologies

Andrew Zolnai, Owner, zolnai.ca

What did you enjoy most about the event?

It was a concise review of the subject. Speakers were naturally enthusiastic about their subjects but were also frank in acknowledging their limitations. The final talk by Mr David Bamford was particularly appropriate.
Manouchehr Takin (Independent Consultant)

Focused content. great networking.

Interesting and relevant themes.
Dave Camden (Flare Solutions)

Simudyne's presentation.

Learned about the diversity of players in E&P who are working towards using machine learning methods.

The stats on the satellites etc was good.