



Re-printed from Oilfield Technology
September 2011

Felix Balderas, Geophysical Insights, USA, discusses how geoscientists can utilise the newest technology to solve today's E&P challenges.

ADVANCING SEISMIC RESEARCH WITH MODULAR FRAMEWORKS

In many disciplines a greenfield project is one that lacks any constraints imposed by prior work. The analogy is to that of construction on 'greenfield' land where there is no need to remodel or demolish an existing structure. However, pure greenfield projects are rare in today's interconnected world. More often one must interface with existing environments to squeeze more value from existing data assets or add components to a process, manage new data, etc. Adding new technologies to legacy platforms can lead to a patchwork of increasingly brittle interfaces and a burgeoning suite of features that may not be needed by all users. Today's challenge is to define the correct 'endpoints', which can join producer and consumer components in a configurable environment.

This article highlights a strategy used to develop new seismic interpretation technology and the extensible platform that will host the application. The platform, which is codenamed Paradise, includes an industry standard database, scientific visualisation and reporting tools on a service-based architecture. It is the result of extensive research and technology evaluations and development.

Geophysical Insights develops and applies advanced analytic technology for seismic interpretation. This new technology, based on unsupervised neural networks (UNN), offers dramatic improvements in transforming and analysing very large data sets. Over the past several years, growth in seismic data volumes has multiplied in terms of geographical area, depth of interest, and multiple attributes. Often, a prospect is evaluated with a primary 3D survey along with five to 25 attributes

servicing general and unique purposes. Self-organising maps (SOM) are a type of UNN applied to interpret seismic reflection data. The SOM, as shown in Figure 1, is a powerful cluster analysis and pattern recognition method developed by Professor Teuvo Kohonen of Finland.¹

UNN technology is unique in that it can be used to identify seismic anomalies through the use of multiple

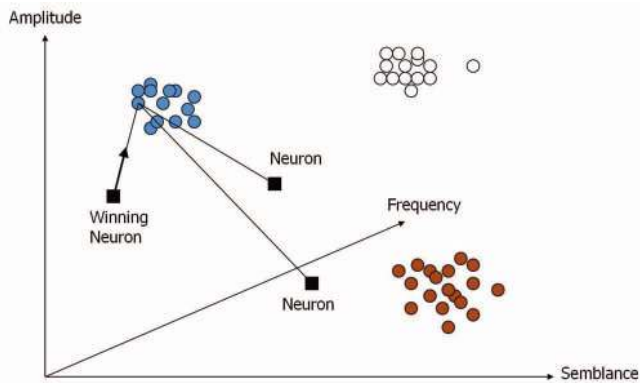


Figure 1. Self-organising map (SOM) attribute clusters.

seismic attributes. Supervised neural networks operate on data that has been classified so the answer is known in specific locations, providing reference points for calibration purposes. With seismic data, a portion of a seismic survey at each logged well is known. UNN however, do not require the answer to be 'known' in advance and therefore are unbiased. Through the identification of these anomalies, the presence of hydrocarbons may be revealed. This new disruptive technology has the potential to lower the risk and time associated with finding hydrocarbons and increases the accuracy of estimating reserves.

The company decided to build new components separately, then, with loosely coupled interfaces, add back the legacy components as services. In its efforts to build a new application for the neural network analysis of seismic data, Geophysical Insights struggled to find a suitable platform that met the goals of modularity, adaptability, price and performance. While in the process of building new technologies to dramatically change seismic interpretation workflow, an opportunity arose for a new approach in advancing automation, data management, interpretation and collaboration using a modular scientific research platform

with an accessible programming interface.

With this new technology concept underway, an infrastructure was needed to support a core technology and make that infrastructure available for others. To deploy their own core technology, they would potentially need databases, schemas, data integration tools, data loaders, visualisation tools, licensing, installers, hard copy, and much more. While not everyone would need the numerous lower level components, they would find that there is more work to be done on the supporting infrastructure than on the core technology itself. Without a platform, each vendor would have to undergo the long process of gathering requirements, developing, testing, and evaluating numerous frameworks, all for something that is not their core product. This is not only a major distraction from developing the core technology, but also an expensive endeavour that most are not ready

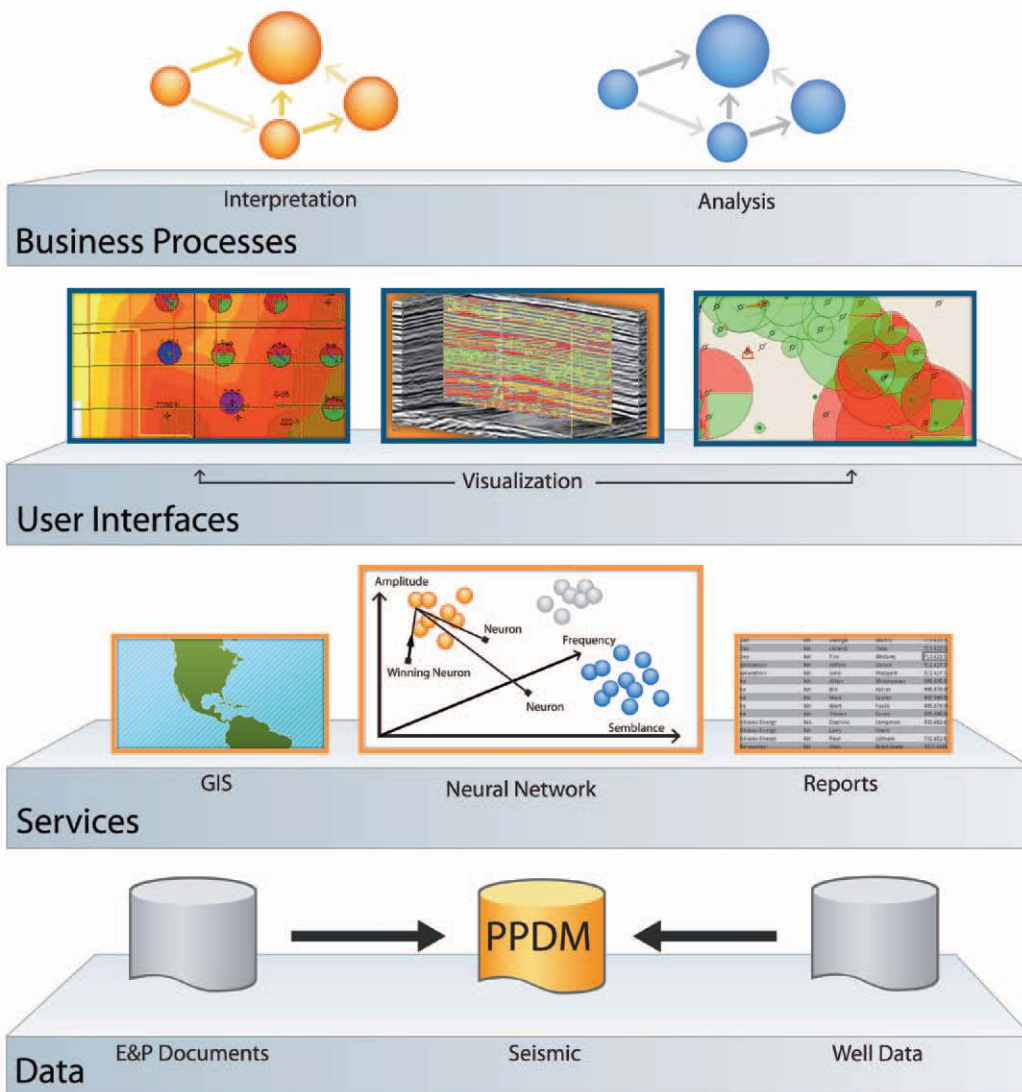


Figure 2. A software framework for hosting oil and gas software applications.

to make, and in some cases perhaps a deal breaker for the project.

The company decided to move forward, developing a platform for itself that would be useful for others. The basic concepts around the chosen architecture are depicted in Figure 2. The goal was to build an affordable, yet powerful platform that could be used by small and large organisations alike, for building and testing new software technologies and shortening the time between design and deployment of new components. Developing a platform separate from the core component meant that it was possible to overlap the development activities for the core component and platform. This minimised the impact that changes in the platform had on the science component and vice versa, thus reducing delivery time. Similar platforms already existed but due to their price, these were out of reach for many smaller vendors and potential end users. Any vendor wanting to promote a simple tool integrated on pricey platforms would find a limited audience based on who could afford the overall platform. End users would probably pay for extra but perhaps unused features. One of the company's goals was for a modular, affordable overall platform. A vendor of a new component can choose to license portions of the Geophysical Insights' platform as needed. A good software design practice is to include end users early in the process, making them part of the team. One thing that they made clear was their sensitivity to price, particularly maintenance costs.

The new generations are more accustomed to working with social collaboration and mobility tools. No longer can a scientist bury himself behind a pile of literature in a dark office to formulate a solution. With the changing demographics of geoscientists entering the workforce and declining research funds, the lag time between drawing a solution on the white board and when it can be visualised remotely across many workstations must be reduced.² These are some of the challenges this platform tackles.

Design and architecture are all about trade-offs. One of the earliest decision points was the fundamental question of whether to go with open source or proprietary technology. This decision had to be made at various levels of the architecture, starting with the operating system, i.e. Linux versus Microsoft or both. Arguments abound regarding the pros and cons of open source technologies such as security, licensing, accountability, etc. In the end, although it was felt that Linux dominated server applications, when looking at the potential users, the majority would be using some version of Windows OS. This one early decision shaped much of the future direction, such as programming languages and development tools.

Evaluations were conducted at various architecture levels, taking time to try out the tools. The company designed data models and evaluated databases. For programming languages and development platforms, C++, C# with MS Studio IDE and Java with Eclipse IDE were evaluated to search for mixed-language interoperability, reliability and security. Java/Eclipse IDE did not meet all the set goals, instead better mixed language programming support was discovered between managed C#/ .NET code and unmanaged Fortran for some scenarios. Other scenarios required multiple simultaneous processes.

At the GUI level, the company looked at Qt, WinForm and WPF. It was decided to use WPF because it allows for

a richer set of UI customisation including integration of third party GUI controls, which was also evaluated. Licensing tools, visualisation tools and installers were also examined. (All of this is a bit too much to discuss here in detail, but Geophysical Insights advocates taking the time to evaluate the suitability of the technology to the application domain.)

The company also considered standards at different levels of the architecture. There is usually some tension between standards and innovation, so caution was needed about where to standardise. One component that appeared as a good candidate for standardisation was the data model. Data assets such as seismic and well data were among the data that needed to be worked with, but the information architecture also required new business objectives that were not common to the industry. For example, the analytical data resulting from the neural net processes. A data model was required, which was simultaneously standard, yet customisable. It also needed to have the potential to be used as a master data store.

Professional Petroleum Data Model (PPDM) is a great, fully documented, and supported master data store, which shares a lot of common constructs with several other proprietary data stores, and has a growing list of companies using it. PPDM builds a platform and vendor independent, flexible place to put all the E&P data. The company actively participates in that community, helping to define best practices for the existing tables while proposing changes to the model.

Research, including attendance and participation at industry conferences and discussions with people tackling data management issues, made it clear that the amount of data, data types and storage requirements are growing exponentially. The 'high-water' marks for all metrics are moving targets. It will be a continuing challenge to architect for the big data used by the oil and gas business. 'Big data' refers to datasets that are so large they become awkward to work with using typical data management and analysis techniques.³ Today's projects may include working with petabytes of data. Anyone building a boutique solution today will have to be prepared for rising high-water marks, and if they depend on a platform, they should expect the platform to be scalable for big data and extensible for new data types.

Neural networks in general, when properly applied, are adept at handling the big data issues through multi-dimensional analysis and parallelisation. They also provide new analytical views on the data while automated processing eliminates human-induced bias, enabling the scientist to work at a higher level. Using these techniques, the scientist can arrive at an objective decision at a fraction of the time. In the face of a data deluge and a predicted shortage of highly skilled professionals, automated tools can assist in achieving the increasing productivity demands placed on people today.

The usefulness of a platform depends heavily on the architecture. Geophysical Insights has witnessed how rigid architectures in other software projects can become brittle over time, causing severe delays for new enhancements or modifications. However, business cannot wait for delayed improvements. Rigid architectures limit growth to small incremental steps and stifle the deployment of innovations. Today's technology change rates call for a stream of new solutions, with high-level workflows including the fusion of multi-dimensional information.

A well designed architecture allows for interoperability with other software tools. It encompasses the exploration, capture, storage, search, integration, and sharing of data and analytical tools to comprehend that data, combined with modern interfaces and visualisation in a seamless environment.

Good guidelines for a robust architecture include Microsoft's *Oil and Gas Upstream IT Reference Architecture*. Another is IBM's *Smarter Petroleum Reference Architecture*.

It was decided to implement the platform on Microsoft frameworks that support a service-oriented architecture. A framework is a body of existing source code and libraries that can be used to avoid having to write an application from scratch. There are numerous framework and design pattern choices for different levels of the architecture and too many for a review here. Object Relational Mapping is a good bridge between the data model and the application logic and the company also recommends N-Tiered frameworks.

Personnel who understand the business domain and the technology must carry out the implementation – otherwise one must plan to spend extra time discussing ontology and taxonomies. They must adhere to efficient source code development practices. The changing work environment will require tools and practices to deal with virtual teams, virtual machines and remote access. The company is using a test driven development (TDD) approach. This approach increases a developer's speed and accuracy. It keeps requirements focused and in front of them, eliminating time spent on unnecessary features. It also enables parallel development of

interdependent systems. In the long run, it yields dividends by reducing maintenance and decreasing risk. Using TDD, a developer can deliver high quality code with certainty.

Ultimately, the science has to come down to business. A good licensing strategy is one that will maximise revenues and allow users to buy products a la carte as opposed to a one-size-fits-all approach. Some vendors attempt to be creative with bundles for different levels of upgrades, but a configurable platform allows maximum user choice among available, even perhaps competing technologies. The market will favour vendors that innovate and manage data and licenses well.

Geophysical Insights' neural network application presents an opportunity to examine seismic data in ways and means orthogonal to those of the legacy systems today. The research platform enabled the company to use this application as a configurable service. Making the right choices in information and application architectures and frameworks was the key to achieving the business objectives of modular services. The company can now move forward with additional science modules and tangential neural network processes, servicing a rapidly changing landscape, licensed to fit specific needs. **01**

References

1. Kohonen, T., *Self-Organising Maps*, 3rd ed. (2001).
2. American Geological Institute, 'Geoscience Currents', www.agiweb.org/workforce/currents.html (Accessed on 29 July 2011).
3. The Professional Petroleum Data Management Association, www.ppdm.org (Accessed on 29 July 2011).