

Big Data Is a Big Deal

Pradeep Anand, President, Seeta Resources



Pradeep Anand is president of Houston, Texas-based Seeta Resources (www.seeta.com), a business consulting

firm he founded in 1994. He also holds an adjunct faculty position at Rice University's Jones Graduate School of Business, where he teaches "Marketing Management in the Energy Industry," in its MBA program. Previously, Anand was vice president, marketing, at Landmark Graphics; manager, North American Operations, at a division of Baker Hughes; and the first marketing and business development manager for LWD/MWD at NL Sperry Sun. In 2009, Anand was co-chairman of the SPE Emerging Technology Workshop on "Delivering and Using New Technology to Make Money in E&P."

Anand received a BS degree in metallurgical engineering from the Indian Institute of Technology, Bombay, where he received a Distinguished Service Award in 2001, and an MBA from the University of Houston. He serves on the advisory boards of the University of Houston's College of Technology and India Studies program. He is the author of the novel *An Indian in Cowboy Country*.

The global population is forecast to grow to about 10 billion people by the middle of this century. While this population growth will generate a growth in global GDP, it will also create significant stresses on resources needed to feed the population and fuel its economic growth, with the demand for food-water-energy creating a "stress nexus."

The Food and Agriculture Organization of the United Nations predicts that by 2030 demand for food will increase by 50%. The International Food Policy Research Institute expects demand for water to increase by 30%. And the International Energy Agency forecasts that energy demand will surge by 50%, despite projected progress in improving consumption efficiencies.

Alternative fuels are forecast to find ready markets and increase their percentage of the energy mix, yet oil and gas are still expected to deliver about 60% of the energy needs of the future. For example, in 2040 the exploration and production (E&P) sector is projected to be faced with delivering about 30% more oil production liquids, about 110 million b/d of oil equivalent, than it does today. The E&P industry is expected to find and develop new types of resources through innovations in technology used in deep water, the Arctic, oil sands, tight oil, unconventional gas, biofuels, and other areas. However, two major issues will mute and limit that success:

1. **Manpower.** As demand increases, the need for manpower increases and manpower availability decreases. Simultaneously, experienced people will retire and exit the industry in record numbers.
2. **Data and knowledge.** Our current knowledge of existing reservoirs and our practices are based on partial data. It is a truism in the business world that 80% of business-relevant information originates in unstructured forms. In the E&P industry, where "natural data" cannot be ordered and constricted to the confines of "manmade" databases, the percentage of unstructured data can be substantially higher.

In the domain of criminal justice, The Innocence Project has shown that partial data can lead to partial truths, in some cases leading to incarceration of innocent people. In many of those cases, DNA evidence not considered in earlier jury trials has exonerated innocent people who spent decades in prisons. Partial data can lead to partial truths.

The E&P industry has had success working with partial data and partial truths but it is imperative that it work with more and better data, or "Big Data," to get closer to the whole truth, and to ensure greater success so vital to meet projected future demand of a growing global population.

Both challenges mentioned above could be met by focusing attention on creating a unified data store that is a ready platform to apply computational algorithms and analytics to extract patterns from both structured and unstructured data. These patterns can then be used to create models with forecasting, anticipatory, or predictive capabilities that reduce the cones of uncertainty or increase the probabilities of success of actions in the E&P industry.

Why patterns? The human brain is a pattern-recognition organism. Our brains create meaning from patterns we see or at least think we see in nature. Humans,

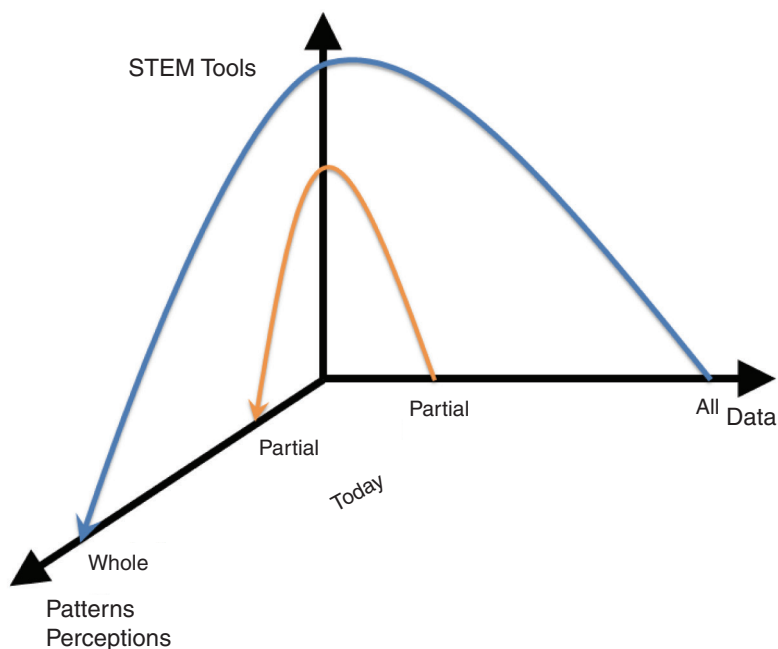


Fig. 1—By expanding the scope of data, existing and new STEM tools can be applied over larger datasets.

especially in the high-risk E&P industry, are uneasy with chance, let alone chaos, and have a tendency to see patterns everywhere. Patterns are important when making decisions and judgments, and in acquiring knowledge.

Often, patterns are real; sometimes they are chance manifestations. However, it can be better to see patterns where none exist than to miss them when they do exist. This is especially true in the area of safety. Throughout human history, our ability to recognize patterns has helped us to survive and grow into a modern civilization.

This basic human facet of pattern recognition has been used extensively in E&P to drive decisions. Graphs and logs are depictions that make visual pattern recognition easier. About 25 years ago, computing technologies enabled the creation of 3D seismic interpretation visualization systems, where palettes of colors were used to depict values of seismic frequencies. The human eye and brain could discern subtle changes in hues that escaped the analytical left brain, to reach more superior reservoir exploitation conclusions. That 3D visualization also brought the unseen geologic world

into the world of human perception by replicating the unseen in familiar forms and patterns that we see on the surface of our planet.

These visual pattern recognition methods have an underlayer of science, technology, engineering, and mathematics (STEM) that applies reason and logic to data. When this STEM layer is applied to partial data, a dire consequence can be limited patterns, which may then lead to myopic perceptions and distortions of the truth. By expanding the scope of data, existing and new STEM tools can be applied over larger datasets, creating and allowing the discovery of better patterns and perceptions that are closer to the whole truth (Fig. 1).

When libraries of patterns have been generated, predictive models can be created. After testing for false positives and false negatives, these models can be applied to optimize rewards and to anticipate risky situations (e.g., where safety is compromised). Over time, an arsenal of patterns and models, across all areas of E&P activity, can be built to create a much-needed foundation for the next generation of technologists to build innovative technologies that can meet future

demand for hydrocarbons. This assemblage of patterns and models would partially compensate for the retirement of deep experience from the industry. After all, good experience is an agglomeration and learning of human pattern recognition of the highest order.

Many farsighted companies in the E&P industry have invested in business intelligence, analytics, and informatics talent at their organizations. These new resources are charged with discovering new routes in uncharted seas and oceans. However, they are working with partial data, which can deliver only partial truths that may not help in discovering new continents of knowledge and insight. The E&P industry needs to embrace both structured and unstructured data, discover new perspectives of reservoirs, and invent new processes to drill, complete, and produce from them. We should be seeking the whole truth that, in all its complexity, lies buried in the data.

So far, pattern seekers and model builders have been challenged by the immense velocity, volume, and variety of data the industry produces and stores. They have been limited by the availability of a unified data store, a pattern recognition platform that not only combines structured and unstructured data but also addresses the varying complexity of data. Recent technologies have made this unifying pattern recognition platform a reality. It has been created in many industries with the following steps:

1. Choose technologies that have been used successfully in other industries, such as open source components for distributed file storage systems and analytics.
2. Though the repositories carry the source data, mnemonic normalization and unit harmonization of collected data are needed for analysis and cross-comparison. This will create an environment for analytics.
3. Extract metadata and apply analytics for pattern extraction to enrich metadata.

4. Create models with discovered patterns that guide and drive better decisions.

So, how do we get started to realize the immense value embedded in Big Data? Different firms have different risk thresholds for adopting new technologies. To pick their first Big Data projects, firms might consider the following two dimensions:

1. The probability of a bad decision in a domain. A decision maker has to deal with the three “V’s” of decisions—variety of decisions being made, volume of decisions, and velocity of decisions.
2. The magnitude and cost of bad or untimely decisions in a domain of interest.

Operators, asset owners, and oil-field services companies could apply

these two dimensions in seeking projects to use recent but proven unifying pattern recognition platform technologies to improve critical facets of their businesses—technical, operational, and business. The combination of technologies could be a game changer. Closer to home, it can significantly improve the competitive advantage of firms in the E&P industry. It can be applied to improve safety, production, success rates, performance, operations, innovation, customer intimacy, document and records management, and a host of other business focuses that suffer from the ill effects of decisions made with partial data.

For starters, Big Data solutions can be used to simply organize both structured and unstructured data to provide simple static views of data and infor-

mation that were hitherto unavailable because of their disparate, inaccessible distribution. The next step up would be to create a unified data store that is a pattern recognition platform, which provides deeper descriptive analytics for better, timely decisions. The third step would be to offer real-time, forward-looking insights, with statistical and scientific bounds.

There is no doubt that the ability to create pattern recognition platforms on unified data stores in order to create new models for better, faster decisions will help the next generation of E&P professionals by accelerating their acquisition of professional insights. This will go a long way in helping the industry overcome the challenges of satisfying the growing global demand for oil and gas. **JPT**