DATA SHARING - PROS, CONS AND HOW TO LEVERAGE

Jim Browning and Sheldon Gorell
Texas Tech University

ABSTRACT

Success in the oil and gas industry comes with effectively juggling four key elements: money (made or lost), risk, technical capability, and competition. Information is key to managing this process. Data sharing is the controlled process of providing information to and obtaining information from your competitors in such a manner to ensure your success (and theirs, as well). When well executed, data sharing can help each participant optimally find and develop highly profitable properties reducing risk of failure. Unfortunately, poorly executed, the data sharing process can tilt the pursuit in the other direction, as well. This paper was prepared to provide the reader with an understanding of the data sharing process and how to effectively leverage information to succeed in such a competitive and technically challenging environment.

There are many data sources available, with varying degrees of cost and value. A great deal of data is available for free from public sources, in a variety of formats. There is also an entire industry made up of companies that, for a fee, provide consistent methods to retrieve public data. They also provide value-added services to validate, scrub, and, sometimes, interpret the data. There are also services to find relevant information or, if necessary, to generate data. Each of these methods incurs some cost, whether it be directly financial, in terms of effort, or risk (due to reliability concerns). A great advantage of these methods is that there is no need to release valuable data to one's competitors. The disadvantage is that a great deal of valuable information is not available via these avenues. This is where data sharing comes in, from consortia to directly sharing with potential competitors.

Data sharing can be extremely valuable, not only in obtaining data but also in developing relationships that build information conduits and can lead to profitable operations that can only be pursued with a partner. While there is considerable value to this approach, there are challenges and hazards that need to be navigated. This paper describes the various methods of retrieving, purchasing, and sharing data and how to utilize data sharing as a mechanism to effectively compete in a challenging environment.

INTRODUCTION

This period in history is known as the Information Age. It is also known as the Digital Age. Together these two monikers describe how information, in the form of data, is a differentiating resource that is vast and quickly accessible. Over the past 40 years, the competitive landscape has evolved from "mere" Data Processing to Business Intelligence to the current frenzy over Big Data, Data Analytics, and Machine Learning. In this rapidly evolving technical landscape, it is difficult to keep up while using emerging technologies profitably. In a technical frenzy, one can lose track of a long-term priority of a business – PROFIT. In the Information Age, it is apparent that long-term profits are achieved by a company that leverages its unique expertise through the appropriate use of (relevant) data and associated technologies. This paper focuses on acquiring, sharing and leveraging data, with minimal discussion of associated technologies. Description, selection and use of these technologies warrant their own separate discussion.

The most useful data in the petroleum industry is often proprietary, requiring contractual agreements to obtain. While this paper describes the drivers, strategies and mechanisms for sharing data, it should not be construed as legal advice. This is for informational use only. The author makes no claims of expertise in Intellectual Property law. Consult your attorney for legal advice and to prepare contractual agreements for sharing proprietary data. Also, please note that this paper is focused on U.S. sources of data. Other countries may have similar or very different methods as well as regulations for obtaining and sharing data.

DATA SHARING – WHAT IS IT AND WHY DO WE CARE?

While the terms data and information are often used interchangeably, it is important for this discussion to describe the differences. Data can be defined as discrete facts (numeric or textual). Context describes the conditions from which the data was obtained. Effective application and interpretation of context is necessary to transform the data into information useful for decision making (Figure 1). Entities will have varying levels of expertise in understanding data context and may have differing applications for the data. Consequently, different entities will have different perceived values for the same data. This is where data sharing comes in – data that is useless for one company may be very valuable for another.

Figure 2 lists the primary business processes for each of the sectors in the oil industry. Regardless of the sector, there are significant information requirements across the lifecycle of any asset (well, field, gathering system, etc.). Also, the definition of an asset's profitability depends on the owner's operations and cost profile. Considering this, one can see how relevant data can be important in the definition, development, operation, and custody transfer phases of an asset's life.

Some of the data used in the lifecycle of an upstream asset include:

- Leasing
- Infrastructure
- Geophysical
- Petrophysical
- Drilling and Completion
- Production

Midstream assets have similar data requirements. There are a variety of sources for obtaining data:

- Internal
- Public Sources
  - Government/Regulatory Agencies
  - County Courts
  - Trade Associations
  - Research Institutions
- Subscription Services
  - Online Services
  - Private Libraries
- Industry Publications
  - Government/Regulatory Agencies
  - Trade Associations
  - Research Institutions
  - Professional Societies
  - Independent Publishers
- Other Companies/Institutions (sharing)

PUBLIC SOURCES

While generally trustworthy, internally generated data can be expensive and difficult, if not impossible, to obtain. Plus, preexisting internal data does not often meet the needs of the current effort. Upon finding gaps in internal data, free and convenient sources come to mind first.

There are quite a few sources of free data. There is an immense amount available online from regulatory filings in the United States. The Interstate Oil & Gas Compact Commission (iogcc.publishpath.com) serves the interests of individual states and has a list of state agencies with their associated internet addresses.

The U.S. Department of Energy (www.energy.gov) and the U.S. Energy Information Administration (www.eia.gov) provide industry-wide data, analyses and reports online. The Department of Transportation regulates and provides information regarding pipelines (https://www.transportation.gov). The Securities and Exchange Commission (www.sec.gov) regulates and provides filings for researching publicly traded companies. There are also quite a few trade associations (e.g. API, INGAA, IPAA, GPSA) that serve various sectors of the industry by lobbying and providing data, analysis, and technical standards.

There is a considerable amount of freely accessible information. The above entries provide just a sampling of the many sources of free data. However, when considering "free" data, it is important to keep an old acronym in mind: TANSTAAFL ("There ain't no such thing as a free lunch."). In many "free" does not necessarily mean FREE. One must consider the reliability, relevance and completeness of the data accessed. One caveat on regulatory data is that regulatory filings contained legislatively required data – and ONLY the required data. Consequently, the data may be incomplete or misleading for one's purposes. Data may be aggregated in a form that precludes quantitative analysis (e.g. reporting production on a lease basis rather than by well). Also, what is "sufficiently accurate" for the filing may not be acceptable for additional analysis. This must be considered during analysis. While freely available online, regulatory data is often cumbersome to download and transform into a directly useable format. To resolve the challenge associated with obtaining large data sets, agencies or associations may offer bulk downloads or publications – for a fee.

While the concept of "data" generally draws one to think of numbers, physical samples are very important in the oil industry, particularly in the upstream sector. There are facilities that offer free access to view, and in some cases analyze (non-destructively), core and cuttings samples. They may also offer access to existing sample analysis, log data and maps. Additional value-added services, such as slabbing and core photography, may be available for a fee.

SUBSCRIPTION SERVICES

Data subscription services seek to resolve the challenges associated with collecting regulatory data and charge for their services on a monthly or volume basis. They gather data consistently from state regulatory agencies and county courts. The data is stored in a database and is then accessible online via their interface (web page, mobile app or another of their software products). These subscription services generally add some other value-added services, such as data scrubbing, analytics and mapping. For this discussion, data scrubbing is of particular interest because the service employs a team of personnel who "scrub" the regulatory data through a variety of methods, some proprietary. The service may also supplement the publicly acquired data with data through other sources, such as through companies or consortia with other, custom, pricing schemes.

There are also subscription-based libraries that provide access to petrophysical data and samples to paying members. Log libraries provide access to well logs and other related data, generally including: scout tickets, completion tickets and maps. Log libraries generally acquire their collections from regulatory agencies, private collections, and negotiated trades. In addition, they may offer other additional value-added services such as retrieval, copying and digitizing. Sample libraries provide access to drilling cuttings and/or cores. They may also provide access to lab analysis results.

INDUSTRY PUBLICATIONS

Most of the public sources listed above publish data, analytics, and reports. Some of the publicly available information can be downloaded for free, while the rest is only available for purchase. Professional societies and independent publishers also publish journals as well as data collections and reports available for one-time purchase or via subscription. Industry journal articles generally provide the

methodology and results of a particular investigation, with derivative (results) data being published. Raw data is generally only available by contacting the author of an article, who may or may not provide the raw data. If raw data is provided, the author will generally insist on constraints for its use and sharing.

When using derivative or raw data, it is necessary to consider the source. Professional societies generally require a rigorous peer review of each article published. However, independent publishers cover the whole spectrum of review, from very vigorous review all the way down to verifying that the publishing fee clears. Caution is recommended, particularly with obscure publications.

OTHER COMPANIES/INSTITUTIONS (SHARING)

Despite the immense amount available, free or fee, it may still be necessary to get data from other companies – who may be peers, competitors, or service providers. This is where data sharing becomes necessary. The following methods will be discussed:

1. Data Sharing Between Two Entities
2. Service Agreement
3. Research Agreement
4. Consortium

Table 1 below defines parties to an agreement for this discussion. These definitions are based on commonly accepted usage of the terms 1st Party, 2nd Party, 3rd Party, etc.

1. DATA SHARING BETWEEN TWO ENTITIES

In this agreement, the one seeking data (2nd Party, the customer) goes to the owner of the data (1st Party) to obtain access to the data. This is the type of agreement used by the Subscription Services listed above. In the case of an agreement with a Subscription Service, the Subscription Service (1st Party) retains ownership and the user (2nd Party) can use the data under specific terms. Note that the customer paying for the subscription never owns the subscribed data and is thus prohibited from giving (or sharing) the subscribed data to anyone not indicated in the agreement.

This type of agreement is also used between two companies. Even if they are competitors, the two companies can benefit from sharing data without the exchange of funds. For example, if two companies have properties (or markets) near each other they may define terms for sharing data that are acceptable to all. Since the two parties in the agreement may have different amounts and different types of data available, they may negotiate the specific data that will be shared to assure an acceptable balance of value in the exchange. Once again, each company retains ownership of their data.

2. SERVICE AGREEMENT

When dealing with sensitive data, service agreements generally contain language where the owner (1st Party) retains all rights to the materials and results, the 3rd Party contractor (consultant, lab, etc.) can only share the data with parties named by the owner. One or more 4th Parties (subcontractors) may participate in the creation or use of the contracted data. 4th parties represent the subcontracting 3rd Party and are subject to the constraints of the subcontracting 3rd Party, at a minimum. The 3rd Party contractor (and associated 4th Parties) also may only use the information for efforts defined in the agreement by the owner of the materials/data.

3. RESEARCH AGREEMENT

Research agreements are much like service agreements with the exception that the 3rd Party (Research Institution) may have the ability to utilize the derivative data (results of data analysis) for purposes approved by the owner. Examples include publishing research and using the derivative data in further research. Such utilization of derivative data requires approval by the data owner. 4th Party

subcontractors are subject to similar constraints as those described above in the Service Agreement section.

4. CONSORTIUM

In a consortium, all the above agreements may be expanded to include more than two companies. The same terms as above apply. In a consortium, data may be anonymized to sufficiently protect ownership of the data while still contributing to the mission of the consortium. If the consortium is large enough, one or more of the members may not have an equivalent amount of data to share and thus the terms need to be flexible enough for each member's input to balance the acceptable shared value. Also, membership fees may be necessary to cover the costs of managing and operating the consortium.

SUMMARY

The American petroleum industry is very dynamic with the ongoing development of new opportunities to discover, operate, sell, and acquire productive assets. There is a great deal of competition in this environment as there is substantial risk to accompany the potential for substantial reward. As a consequence, relevant data can be extremely valuable to the company with the appropriate expertise and ability to turn the data into information useful for decision making. Such data can be obtained from the following sources:

1. Internally Sourced or Generated
2. Public Sources
3. Subscription Services
4. Industry Publications
5. Other Companies/Institutions (sharing)

Whether it be due to data cost, quality, or even availability, data sharing often becomes the best option. Data sharing agreements with subscription services, service providers and research institutions have been commonplace for many years due to the role that they have in collecting, generating and analyzing data. Data sharing with peers and competitors has come to fore more recently due to the climbing costs of collecting, generating and analyzing enough quality data to support the technologies and analytical methods currently in place. Also, companies have found that it was worth partnering with other companies so that they could share data and knowledge in a way that benefitted both companies financially and technically. Eventually, data sharing grew to be a common practice among successful companies.

While sharing data between two companies can be useful, sometimes a larger group of companies need to contribute data for a common cause to understand a particular challenge, and they gather to form a consortium to gather and learn from their collective data. A research consortium may be formed by adding a research institution to the consortium so that additional analysis, research, and evaluation may be formed to benefit the participants in the consortium.

Key elements in data sharing include:

1. Assuring that confidential data remains confidential.
2. A partnership attitude between the participants.
3. Assuring that all participants sufficiently benefit.
4. Legal review of the data sharing agreement to avoid future debate or legal issues.

In conclusion, the current state of the oil industry requires considerable data at every phase of an asset's lifecycle. Data sharing agreements help assure that the right data is available for making informed decisions and help develop industry relationships for mutual benefit, both in the short term and long term.

REFERENCES

1. ContractStandards, “Data Sharing Agreement”, https://www.contractstandards.com/public/contracts/data-sharing-agreement

2. NContracts, “First Second Third Fourth and Fifth Parties How to Measure the Tiers of Risk”, https://ncontracts.com/articles/first-second-third-fourth-and-fifth-parties-how-to-measure-the-tiers-of-risk/

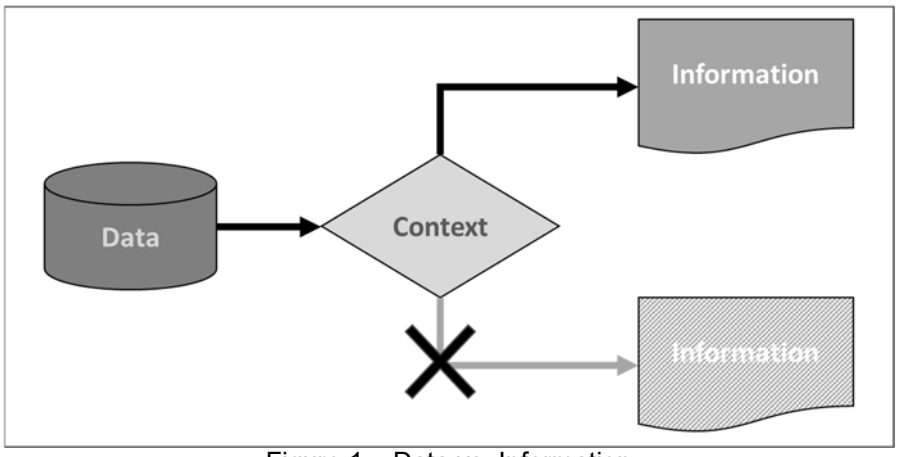

Figure 1 – Data vs. Information

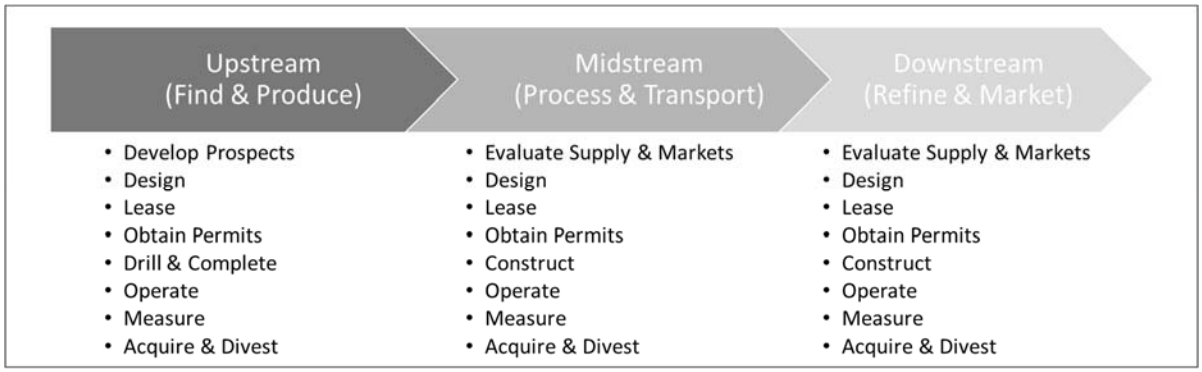

Figure 2 – Primary Processes by Sector

Parties to an Agreement	
1st Party	Owner, Seller, Lessor
2nd Party	Customer, Buyer, Lessee
3rd Party	Contractor (e.g. Research Institution, Lab, …)
4th Party	Subcontractor to 3rd Party

Table 1 – Parties to an Agreement