

# ValueD – 7 Considerations for Determining Data Value and Pricing

by **Mahesh Kshirsagar,**  
*CTO – Analytics & Insights, TCS*

## Introduction

A tremendous amount of data is continuously generated every day from machine and human business and non-business activities.

The terms Data Informed, Data Driven, Data Centric and Data Literacy are all linked to the varied degree of maturity in terms of how businesses use the data. Virtually all businesses are focusing on how to optimally leverage the data for competitive, compelling and differentiating decision making, across the operational, tactical and strategic decisions from rank-and-file employees to CxOs.

Businesses not only want to use the data available for themselves but also want to use the data in the business ecosystem. They may even want to use it beyond the business ecosystem, which typically they may need to purchase. This need has now led to data being termed a Future Currency. With the idea of data as a currency, there is an automatic link to price. Price refers to a marketplace, where there are buyers of the

data, and sellers of the data. However, it can be challenging to determine how sellers of the data price their data—the bid price—and similarly, how do buyers arrive at the offer price (the price at which they are willing to buy the data).

Data pricing has a number of dimensions from the seller and buyer perspectives. There is inherent subjectivity associated with it as there are a number of qualitative dimensions, which are contextual.

While these qualitative aspects cannot be eliminated, they can be significantly reduced to bring in more objectivity in pricing of the data. The seven considerations highlighted in this article around data pricing demonstrate how businesses can build a necessary framework to price data, both from the perspective of the seller (offer price) and the buyer (bid price).

## Typical considerations: the 7 Ps for data pricing

The typical considerations for data pricing can be categorized in seven sets: Primitive, Property, Purpose, Positioning – Seller, Positioning – Buyer, Pact and Parameters. Examining these seven Ps in detail can bring in the desired level of objectivity in data pricing.

### Primitive

The primitive consideration refers to category, source and velocity of the data.

They are three types of category: transactional data, managed through various business IT applications; master data supporting the business IT applications; and reference data, which is needed by the business operations.

The next dimension to category data is determining whether it is generated or authored by the business itself, or if the business has obtained it from an external source. Typically, when businesses obtain data from external sources, they often need to curate the data for their purpose before they can productively use it.

Velocity of the data is the other important dimension within this consideration. It refers to determining if the data is constant, meaning it does not change with time, or if it is dynamic, that is, it changes with time.

### Property

The property consideration addresses if the data is self-defining, i.e., can the data be understood without involving the creator of the data or subject matter expert associated with the creation of the data?

The self-defining aspect of the data differs when it is structured data vis-à-vis or when it is unstructured data.

For structured data, it is important that it has metadata associated with it both the technical and business metadata. It is also important to consider availability of measurement metrics like completeness of the data, correctness of the data, consistency of the data, and currency of the data (is it the latest data). Business data often goes through a series of

changes while it is processed by IT applications, and thus availability of data lineage is also key in understanding that data.

When it comes to unstructured data, semantics of the data (i.e. meaning of the data), ontology of the data, (i.e., about the properties of the data and relationship between various concepts present in the data), and representation in an easy-to-understand knowledge graph, if needed, become important considerations.

### Purpose

The next critical aspect driving value of the data is the purpose – that is, the business benefits this data can deliver. The benefits can be a one-time benefit, such as the information on a background check for candidate employment. The benefits can also be constant, such as knowing the current credit details of buyers while selling products/services on credit. The benefits can diminish with time, like the past credit history of the buyer as time passes. Or the benefits can increase with time; for example, as the business grows, underutilized data gets utilized and more effectively.

Another dimension to data value is linked to the focus of the benefits derived. The benefits can include:

- Workforce efficacy to achieve higher productivity and predictability of the workforce.
- Increased top-line and/or improved bottom line to contribute to financial performance.
- Customer experience to directly impact business outcomes.
- Adhering to compliance norms to help eliminate or significantly reduce financial loss and/or loss of reputation.

### Positioning - Seller

Another consideration is seller positioning. For the seller to arrive at the bid price of the data, it is important to understand the demand-supply situation.

The seller, who has monopoly, can demand a premium whereas the seller facing open competition will have to price the data at the most competitive price points. In the case of limited competition, the price is typically dependent on the negotiation ability of the seller.

The next criteria driving data value is its uniqueness in terms of its applicability. For example, whether the data is applicable only to a specific business process, or to a single industry, or whether the data is applicable across industries will drive its value. Whereas the value of the data, which is specific for a customer, will be driven by the need and availability equation.

The next aspect is the inventory cost to host the data. It is not really about the hardware storage cost, but the cost associated for being compliant and enabled to continually sell that data.

Merit is another important dimension. When the usefulness of the data is proven through customer testimonials, that data has more credibility than data carrying a perceived credibility basis on the hypothesis put forward by the seller of the data.

### Positioning - Buyer

When considering buyer positioning, it is once again useful to understand the demand-supply situation driving the buyer to arrive at the offer price for the data.

The buyer must necessarily consider how the data will impact its business. Will it enhance the internal operational efficacy, or will it lead to increased competitiveness? Or would it lead to differentiation in products/services offered?

Data usage also plays an important role. Is the data needed at a granular level, i.e., for specific values of the data, or it is enough to only know about the pattern, correlations, trends in the data?

The next important thing for the buyer is the opportunity cost. It can be considered as the cost needed to create the data, which the buyer does not have today, and the business impact it will have because the buyer does not have the required data.

For the buyer, exploiting the data to its fullest is another consideration. And hence, it is important for the buyer to understand how the data is formatted and if it allows interoperability. In addition, can the data usage be repeated once the buyer is in possession of the data, and once the buyer experiences the first benefit from that data?

## Pact - Contract

The pact, or contract terms of how the seller and buyer negotiate on their part to agree on the price points, is the next typical consideration for data pricing.

From the seller perspective, it is important to factor in the kind of rights the seller is giving to the buyer for data usage. How well is the buyer protected in terms of the data usage being purchased, and is the seller indemnifying the buyer from the risks and liabilities associated with the data. The contractual terms of how the data will be kept up-to-date and apt for usage also play a key role. Is the seller going to provide periodic refreshes for the dynamic data? If yes, will there be incremental charges, and what would those charges be?

From the perspective of the buyer, the freedom to use the bought data without any risk and liabilities, as the business demands, is of paramount importance to the buyer. Whether the buyer can use the same data for a specific business process, or for an entire business unit or for the entire enterprise are other key considerations for the buyer. The time period for which data is available for use is important for the buyer. And once that time period is over, on the expiry of the data usage, what are the obligations of the buyer? Is the buyer expected to purge the data or preserve the data, as demanded by the specific business scenarios?

## Parameters

Parameters is the seventh and final P of the typical considerations for data pricing. In this article no formula is given to calculate the value of the data. The previous six Ps, discuss considerations in arriving at the value of the data. Both the seller and the buyer can independently do their calculations and arrive at their best bid price and the offer price respectively.

The market conditions, namely the competition and the uniqueness, play a key role here. The potential business benefits associated with the data is another dimension to consider. The next dimension is about the data itself –whether the data is easily available, does it have proven merit, what would be the usage terms of the data (to be used as-is or masked?), and is the data self-defining, all of which carry importance.

From the buyer perspective, aspects linked to opportunity cost, usage freedom, usage period, associated risks and liabilities are critical. From the seller perspective, inventory cost, and indemnification from any risks and liabilities are key.

## Four-quadrant view of data and decision making

Earlier sections of the article highlighted the importance of data in decision making. In any business scenario, the business stakeholders have to deal with both "Identified" and "Unidentified" which translate to "Identified-Identified", "Unidentified-Identified", "Identified-Unidentified" and "Unidentified-Unidentified", as shown in Table 1. Again, the data takes center stage here.

<p><b>Radical innovation</b></p> <p>Exploring new avenues</p> <p>Managing "Identified – Unidentified"</p> <p><i>Example: New products for acquiring new customers</i></p>	<p><b>Disruptive innovation</b></p> <p>Demystifying unidentified</p> <p>Managing "Unidentified – Unidentified"</p> <p><i>Example: Products for game-changing customer experience</i></p>
<p><b>Complementary innovation</b></p> <p>Expand products/services portfolio</p> <p>Managing "Unidentified – Identified"</p> <p><i>Example: Customer analytics on ecosystem data</i></p>	<p><b>Incremental innovation</b></p> <p>Exploring new avenues</p> <p>Managing "Identified – Identified"</p> <p><i>Example: Compliance analytics on enterprise data</i></p>

In essence, it is about the business knowledge of the "scenario" and the "response" to that action.

- **Identified- Identified:** The business can analyze what/where/how/why to enjoy benefits of incremental innovation.
- **Unidentified - Identified:** The business can analyze the ecosystem data, which can lead to complementary innovation.
- **Identified - Unidentified:** The business can analyze data beyond the ecosystem and the result can be radical innovation.
- **Unidentified -Unidentified:** The business aims at demystifying uncertainties, which may lead to disruptive innovation.

As it leads to innovation and as innovation fuels the business growth and transformation journey, developing the “Identified x Unidentified” understanding of the business scenarios, in determining the value of the data, is very important for the business.

## Conclusion

With data becoming more pervasive, and solutions like data marketplaces in the system that channelize buying and selling data activities, understanding the value of the data, especially in the context of “Identified x Unidentified” business scenarios has become extremely important. Having a set of considerations to arrive at a bid price and an offer price of the data asset will become increasingly critical across the business ecosystem.

### **About Tata Consultancy Services Ltd (TCS)**

Tata Consultancy Services is an IT services, consulting and business solutions organization that delivers real results to global business, ensuring a level of certainty no other firm can match. TCS offers a consulting-led, integrated portfolio of IT and IT-enabled, infrastructure, engineering and assurance services. This is delivered through its unique Global Network Delivery Model™, recognized as the benchmark of excellence in software development. A part of the Tata Group, India’s largest industrial conglomerate, TCS has a global footprint and is listed on the National Stock Exchange and Bombay Stock Exchange in India.

For more information, visit us at [www.tcs.com](http://www.tcs.com)