WHY AND HOW SHOULD DATA BE VALUED

Dr. S. K. Gupta

Managing Director
ICMAI Registered Valuers Organization

&

Raja AK

Student – Graduate Insolvency Program, IICA

The Perspective

Over thousands of years, the importance of data then evolved – particularly in retail and trade, where a value was attached to certain assets, whether it be products or services. At that point it was only simple data – the supply and demand of a product or service – that would help to place a value on it. Since then, far more abstract assets have been and continue to be valued; from stocks, to brands, patents and trademarks. And yet, there has always been a difficulty in putting a value on the data itself, particularly as there are far more complex types of data, vast amounts more of the data, and because its use has become far more sophisticated.

In the knowledge economy, data has become a strategic asset that allows companies to acquire or maintain a competitive edge. For the shareholder, data embodies a financial potential. For the company itself, data can be used to optimize the way it does business: acquisition, retention, targeting, pricing, etc. In 1975, tangible assets comprised up to 83% of a company's valuation; but today up to 90% lies in intangible assets – data, intellectual property, brand, reputation and trust. Today's shared wealth lies in data – it is cited as the most valuable resource in the digital age.3 It promises commercial gain for business, improved public services for governments, better convenience and well-being for individuals, and positive outcomes for the planet and society.

Data is a very non-conventional asset and the valuation of data is a relatively modern practice that is still in its infancy. However, data valuation affects companies of all sizes—from newly emerged local start-ups to big multinational corporations. Companies must understand how to value their data to be able to monetize it accurately. The issue is, there are currently no common standard models for data valuation, and generally accepted accounting principles do not yet recognize data as an asset. Data valuation is complex, as the value of data can depend on several factors, and even the same data can have a different value for different users. This is a serious challenge for both potential investors and the company itself.

Forward-thinking executives have begun measuring the value of their organization's data assets to forge a data-driven culture that generates increased business benefits from data. These business leaders aren't allowing antiquated accounting standards to stand in their way. It doesn't matter that current accounting regulations such as IFRS and US GAAP generally prohibit reporting the value of data on balance sheets. They care about putting data to work.

Data as assets

Some people talk about data as the new oil, but this is too simplistic. Oil is a commodity–to be bought and sold. Data is an asset, an asset that grows in value through use. Intuitively, we all understand that data has value. It's why companies are investing heavily into data and why so many people have chosen data as their professional career path.

Think of data as an asset; organizations deploy assets to create value for different stakeholders. They also invest in assets to make them fit for purpose and, at any point in time, they have to consider which assets are worth investing in. You can think of this as the data value/data valuation cycle. You have to assess and understand what data you have (data assessment). You have to put a value on this data (data valuation) so your people recognize the value of data, treat it with respect inside your organization and work out how to make it more valuable. From this, you then have to invest (data investment) to make sure your data is fit for purpose. You have to ensure you have good governance in place, an appropriate data strategy, standards, systems and procedures to ensure you achieve good data quality.

And data can even be used as collateral, just like United Airlines and American Airlines did during the Covid-19 pandemic. Data from their customer loyalty programs were valued at around $20 billion for each airline.

Characteristics of Data

Data are raw alphanumeric values obtained and owned by data producers. When a data producer uses data for operations, they can be treated as **physical assets**. There are mature methods for assessing the value of physical assets. When data are used for other purposes (decision-making, regulatory, and research) they often are treated as **intangible assets**. These assets are more difficult to value and the methods are less mature and less precise.

Unlike a 'typical' asset that provides value to the organization that owns it, data also have immense value for secondary users creating value for multiple organizations for multiple purposes at the same time. Here, data behave as a **derived asset** (or non-rival good) whose value is tied to an end use. Methods for valuing derived assets are in their infancy.

Data assets are intangible, and generally

- **Are identifiable and definable** – Data assets may be made up of specific files, specific tables, or records within a database.
- **Promise probable future economic benefits** – To have value, data assets need to have a useful application. Identifying productive uses for data is often necessary to assign value to the asset.
- **Are under the organisation's control** – The organisation must also have rights to use the data in a way that is consistent with its rights under applicable law and any contractual licensing arrangements, while protecting the data and restricting access to it by others.

Several metaphors have emerged to help businesses, governments and individuals better grasp the unique nature of data – it has been compared to oxygen, soil and sunlight for its prevalence and exponentiality and negatively to carbon dioxide. There is debate about whether it should be treated as an asset at all, with proposals based on the fact that in economic models data acts more like labour than property. China has gone further in explicitly recognizing data as a factor of production in its own right. Ultimately, each of these metaphors and existing economic concepts offers something useful in communicating the foundational importance of data but has limitations in capturing data's unique characteristics. To date, there is little momentum on any one approach.

Data Characteristics Common to Other Tradeable Intangibles

- high initial creation cost but low replication cost for data in particular (although infinite reproduction may diminish commercial value in the long run)
- little inherent value: value potential depends on enabling value streams
- value creation potential depends on complementary business assets and context
- potential ability to create multiple simultaneous value streams in multiple context

Unique Characteristics of Data

- growing exponentially faster than any other category of intangibles
- for many organizations, value creation potential depends on emergence of relevant data value chains
- legal protections around data not as well defined as for IP. Some types of data are easy for competitors to emulate.
- risks related to privacy and protection of personal and corporate data • potential value time limited
- Internet-accessible data more exposed to theft and misuse than many other categories of intangible

What's Your Data Worth?

Many businesses don't yet know the answer to that question. But going forward, companies will need to develop greater expertise at valuing their data assets. The economy has been transformed by data in recent years. Data-driven firms made up seven of the global top 10 firms by stock market capitalisation in 2021. It is therefore obvious that data has value in an economically meaningful sense. However, despite the broad recognition of its value, and the need to develop appropriate policy frameworks, there is still no consensus method for empirically determining the value of data.

While there is broad consensus that data is foundational to new value creation in an increasingly digitized economy, there is little consensus about how to quantify the value of data. Data and data-driven value have a number of unique attributes that make them different from other types of goods or services in the economy. As such, current economic and accounting approaches are limited. Some new approaches are beginning to emerge.

Many companies, however, fail to understand both the value of their existing data assets and the underlying levers that can increase data value. This can mean, in turn, that they miss out on the competitive advantages and shareholder value that their data assets can generate. In order to capture and harvest the value of data over time, organizations must fi rst seek clarity on how to value data as an asset, then follow through with a comprehensive data strategy to drive value enhancement. Increasingly, data assets are the engine driving the total value and growth of modern organizations. As a result, building a framework to discover and realize the potential of your data is critical to increasing the value you provide to

shareholders, and to optimizing the future success of your organization.

Data from a Financial Reporting Perspective

In financial reporting, the cost of accumulating a data portfolio has traditionally been treated as an expense unless the data is acquired as part of a third-party transaction, in which case it can be recorded on a balance sheet at the fair value paid on the purchase date.4 Updated accounting standards now provide for recognition of certain self-generated intangibles (including data) as cost provided evidence of future economic benefits is strong and cost can be reliably determined. At present, and likely for the foreseeable future, balance sheets will continue to record data (and many other non-financial assets) at amortized cost less impairments. Traditionally, the purpose of transaction-centric financial reporting has not included accounting for the market or economic value of these [intangible] assets.

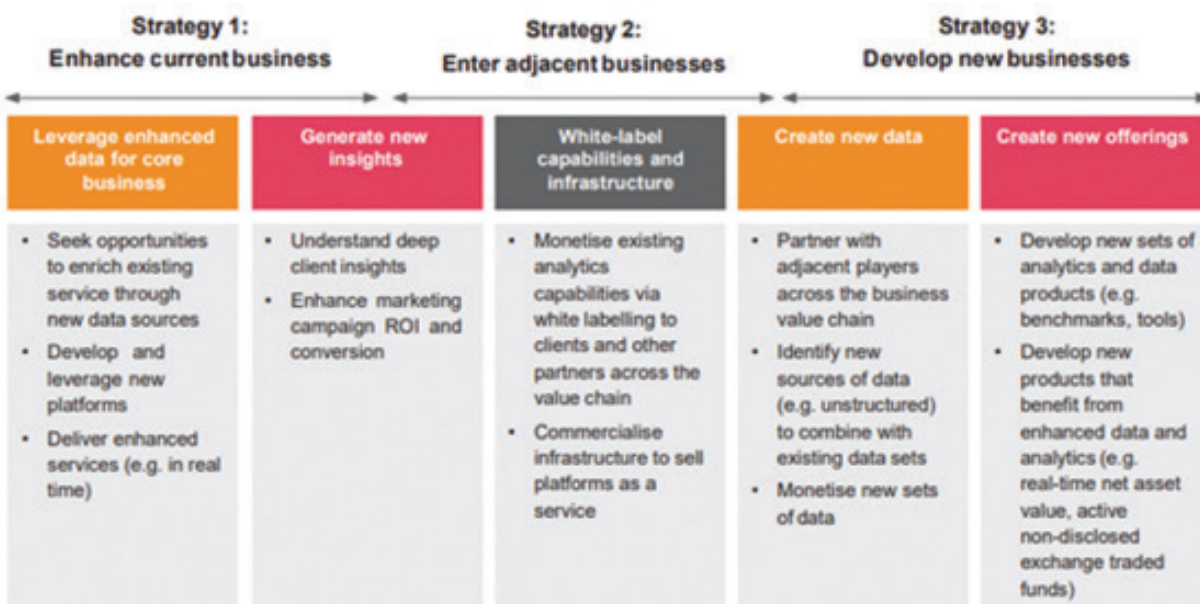
Data Value Drivers

Organisations are under increasing pressure to invest in data assets, either organically or through acquisitions. Assessing those investments robustly and making the right strategic decisions requires an understanding of data valuation methods and value drivers.

Data value drivers

Assuming an organisation invests the time to create an inventory of its data, the value of that data lies in its ability to allow an organisation to generate future economic benefit ('data monetisation'). Some examples of data monetisation strategies are shown in Figure 3 below:

Figure 3 – Typical data monetisation strategies



- **Completeness :** Completeness refers to how much of a known universe the dataset covers. In general, the more complete a dataset, the more valuable it is due to the increased accuracy of insights derived. Enhancing the completeness of data increases the value of the data. Consistency
- **Consistency :** Data is consistent if it conforms to the syntax of its definition. For example, structured data such as storekeeping inventory and business transactions are data that conform to a pre-defined syntax and format. These data have a high degree of organisation, making analysis and processing easier. On the other hand, unstructured data such as images and sounds may require some degree of processing to conform to defined rules and syntaxes so as to enhance their consistency. In such cases, the more processing is done, the more valuable the data is.
- **Accuracy :** Accuracy describes the degree to which data correctly describes an object or event. Reliability of the data significantly impacts the value of the data. Inaccurate data produces unreliable insights, which makes it ineffective for any organisation seeking to utilise the data. In addition,

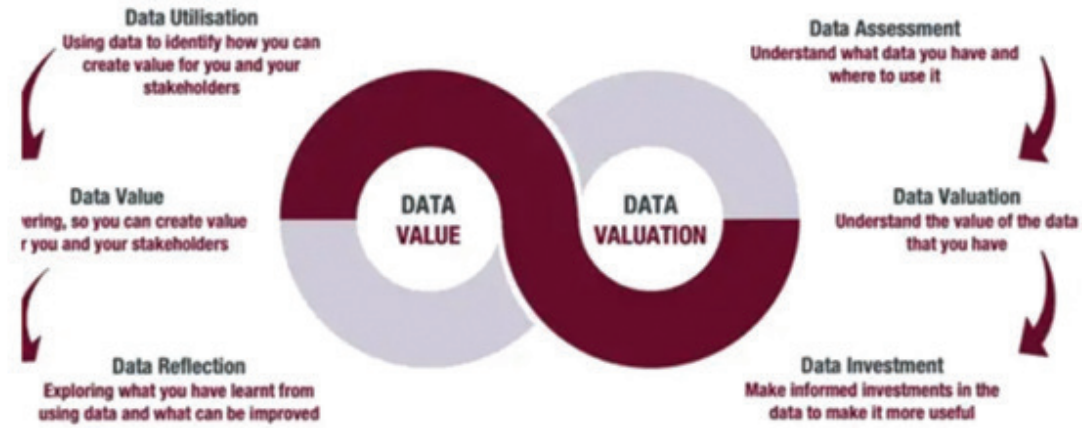
knowing the data's

- **Timeliness :** Timeliness refers to the degree to which the data is up-to-date at the required point in time of use. In general, the more up-to-date the data, the more valuable it is. However, timeliness is a relative measure, which is dependent on the intended use case for the data.
- **Exclusivity :** Exclusivity refers to the uniqueness of the data. In general, the fewer existing alternatives for the data, the more valuable the data is. The key driver of value for exclusivity lies in the competitive advantages and revenue opportunities afforded by the data. Exclusivity can be enhanced by: a. creating unique datasets by integrating and enriching existing data with data from other sources; b. controlling access to the data through technical and procedural means; and c. identifying new sources of data or creating new means of capturing data.
- **Interoperability/Accessibility :** Interoperability/Accessibility is critical to the value of the data. This is because in many cases, the value of data lies in its potential to be combined with an internal dataset. Without the ability to combine and enrich the data, the data is generally of little value to potential consumers.
- **Restriction, Liability and Risk Usage :** restrictions have to be compatible with the sharing use case for the data to have value. In general, the less restricted the use of data, the higher its value. In addition, potential liability and risks associated with the data could reduce the value of the data, and are very often the main deterrents to data sharing. Sharing arrangements between organisations would require organisations to design contractual obligations to satisfy internal risk management policies. provenance (or source) is a critical aspect of determining its accuracy, as it informs the data consumer of the history of the data and account for errors, if any.

Figure 4 – Data value drivers

Data Valuation Methodologies

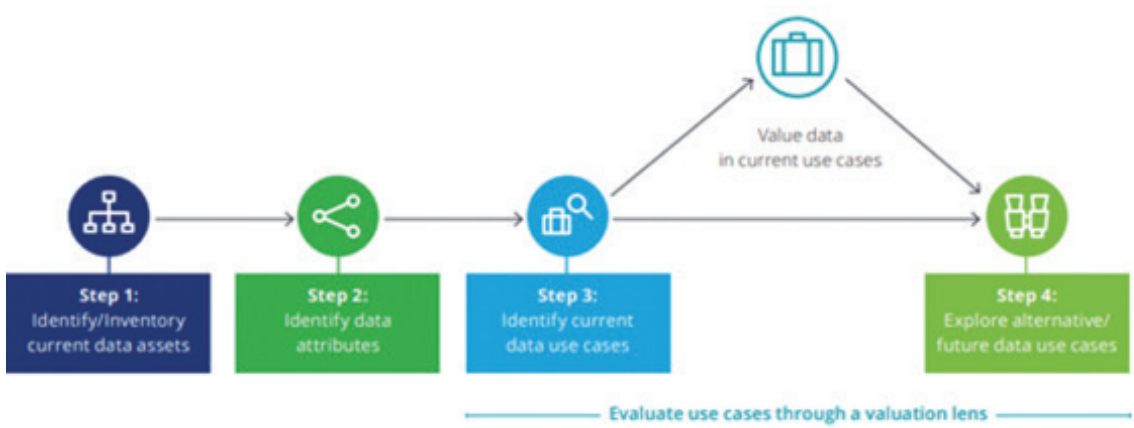
The distinct characteristics and dynamics of data – contextual, relational and cumulative – call for new approaches to articulating its value. This requires a mindset shift – businesses should value data based on cases that go beyond the transactional monetization of data and take into account the broader context, future opportunities to collaborate and innovate, and value created for its ecosystem stakeholders. Assessing data against key value and cost drivers, in the context of different use cases and with attention to shared value for stakeholders, will encourage companies to think about the future value data can help generate, beyond the existing data lakes they sit on, and open them up to collaboration opportunities.



There are many different methods for determining the value of data. Regardless of which method you choose to determine the value of your data, the first step will always be to understand what your data assets are. From there, the next step is understanding how data drives value.

What are the aims of Data Valuation?

- **quantify** the **value** and the **costs** of data assets,
- understand how to **improve** data management,
- **identify** innovation **opportunities**
- **promote** more **data-oriented** business culture.



The cost approach

This method is based on the cost to produce and store data, as well as the cost to replace lost data and what the impact on cash flow would be A method that uses the concept of replacement cost as an indicator of value. The premise is that an investor would pay no more for an asset than the amount for which the utility of the asset could be replaced, plus a required profit/return to incentivise a third party to replace the asset.

One of the main advantages to this overall method are that it is easier to execute than many of the other methods. Additionally, it provides part of the answer to quantifying the return on data, because it captures some of the costs. However, this method is extremely subjective. While it does allow an organisation to conceptualise the value of their data, in part, it falls short of providing a reliable economic picture. Put simply, cost-based methods of data valuation will always undervalue data because it is only concerned with one aspect of value, it ignores the question of how does data becomes business value. There's much more to the question of what is data valuation, than just cost.

With-and-without method

A method for estimating the value of data assets by quantifying the impact on cash flows if the data assets needed to be replaced (assuming all of the other assets required to

operate the business are in place and have the same productive capacity). The projected revenues, operating expenses and cash flows are calculated in scenarios "with" and "without" the data, and the difference between the cash flows in the two scenarios is used to estimate the data's value.

Relief from royalty method

A method built on the assumption that if the company doesn't own the data asset, it might be willing to license the data from a hypothetical third party who does. In this method, the company would forgo a certain amount of profitability to license the data from a third party over a certain lifecycle.

The market value approach

This approach is based on what others pay for comparable data on the open market, by observing those selling data (thus drawing on an example of data value) and calculating the data selling price. Today, companies are using advanced analytics to more fully understand their data, and to identify ways to license it to third parties. In addition, within various ecosystems, data exchanges are being developed so market participants can aggregate and trade data assets, and participating companies can exchange data to create even more value for their enterprises. As companies continue to mine their data and develop models to transact in this asset category, these transactions can be used to derive market indications of value.

While simple to calculate, it has some significant drawbacks. Some data is simply not traded – there may be no comparable examples of business data – either because others are not interested at this time, or because a company is keeping its data to preserve a competitive edge. Additionally, some data is one-of-a-kind, so there will be no comparable examples to study. Getting a true price of the data relies on there being an efficient market, which at present, there isn't. Alongside this, users of this method must understand that price is not the same as value.

Multi-period excess earnings method (MPEEM)

An income-approach methodology that measures economic benefits by calculating the cash flow attributable to an asset after deducting "contributory asset charges" (CACs), which are appropriate returns for contributory assets used by the business in generating the data asset's revenue and earnings.

The economic value approach

On the economic value approaches, there are two key methods.

The first is **income or utility valuation**, which tracks the impact of data on the business' bottom line, therefore it can identify value added to the business by data and can be used to identify value add for specific business functions or use cases. However, this is hard to measure, particularly distinguishing value added by data from value added more broadly. Much like the other approaches, a lot of this is subjective and it is incredibly hard to predict the future value of data.

The second approach is around **use case valuation** – and there are two separate techniques here.

The first is the business model maturity index (Internet of Water), which calculates the value of data by identifying a number of business use cases, estimating the value of each of these use cases, and calculating how much of this value is contributed by data. The benefit of this approach is that it values the data based on a thorough analysis of multiple use cases within the business, and ties it to real business outcomes. However, it is one of the most subjective as the contribution of data assigned to each use case is through surveying, based on hypothetical scenarios rather than real use cases. The margin for error is large.

The **decision-based valuation** method is similar but has an increased degree of sophistication as it models frequency of data collection, accuracy and how fit for purpose the data is. However, once again there is a degree of subjective estimation. It is also a complex model to apply for data assets as it requires the ability to conceive and project use cases.

There is also an issue with 'unknown unknowns' – in other words, using this method businesses can only model use cases and desired outcomes that can be thought of from inside of the business. This relates back to the importance of what question a business is asking – sometimes if it is too specific, and if the data set is also very specific, a business will get the answers it wants, but this discounts many of the other factors and unknowns.

The stakeholder value approach

Value is in the eye of the beholder. The stakeholder value approach goes right to the source of value, by measuring the economic value created for each stakeholder. Not just shareholders, but customers, employees, suppliers, communities and the environment.

This makes it a more modern approach, aligned with the shift from shareholder to stakeholder capitalism much discussed at the World Economic Forum 2020, and mirrored by the growth of environmental, societal and governance (ESG) factors in investing. It's not perfect, but it does overcome many of the problems of previous data valuation methodologies. While other data valuation methodologies race towards data monetisation, they ignore the broader context, to focus

on data in use, or not. The stakeholder method works from an understanding of the total economic value the organisation creates for its stakeholders. Valuation isn't an end in itself, it's a means to achieve better management and decisions. Decisions are never taken out of context, so data valuation shouldn't be either. The most difficult part of this methodology is attributing the right portion of the organisation's total value to specific activities, and from there, into the data that underpins them.

Intrinsic Value of Information

The IVI allows you to evaluate the innate quality of data assets. This method describes how complete and accurate your information is and how likely it is that other organizations have this data.

$$IVI = Validity \times Coverage \times Scarcity \times Useful\ Life$$

The formula for the Intrinsic Value of Information (IVI)

- **Validity:** The percent of records with correct values
- **Scarcity:** An estimate of the percent of other organizations who **don't** have this data.
- **Coverage:** The number of records in the dataset as a percentage of the total universe of potential records
- **Useful Life:** The number of periods (months, for example) that each record can reasonably be used or is relevant

Business Value of Information

Unlike the IVI, the BVI recognizes the relevance of the information to business activities, as well as the quality and timeliness of that information. It can be used as a quick-and-dirty way to measure the potential real-world benefit information assets provide.

$$BVI = \sum_{p=1}^{n} Relevance_p \times Validity \times Coverage \times Timeliness$$

The formula for the Business Value of Information (BVI)

- **Relevance(p) —** The potential usefulness (0 to 1) of the information to the business process **p**
- **Validity —** The percent of records with correct values
- **Coverage —** The number of records in the dataset as a percentage of the total universe of potential records
- **Timeliness —** The probability that at any time, the information is current (matches real-world facts). This is a more easily measured version of the time-lag between real-world events and the appearance of those events in a dataset.

Performance Value of Information

The PVI approach defines the value of information by its impact on improving some business performance driver, as measured by a Key Performance Indicator (KPI). Some examples of potential KPIs are "Conversion Rate" or "Order Fulfillment Time."

IVI and BVI are leading indicators of business value, whereas this measure is a lagging indicator because we have to run a controlled experiment to determine the impact of the information on KPIs.

$$PVI = [(\frac{KPI_i}{KPI_c}) - 1] \times \frac{T}{t}$$

The formula for the Performance Value of Information

- **KPI(*i*) —** The KPI for the business process *with* the information
- **KPI(*c*) —** The KPI for the business process *without* the information (control group)
- **T —** The usable life of any datum
- ***t* —** The time over which the KPI was measured

The KPI ratio provides a measure for the lift in KPI when using this information asset and the time ratio projects the lift over the useful life of the data.

A positive PVI indicates a net benefit for the process, whereas a negative PVI detracts from the KPI.

Choice of an appropriate method

Some key considerations for the most appropriate method which have been mentioned in this paper are:

- what is being valued?
- who is valuing the data?
- when is the valuation taking place?
- what is the purpose of the valuation?

Conclusions

Organizations are under increasing pressure to collect, process and assess data – by 2025 company values will reflect their information portfolios; going public, mergers and acquisitions require companies to answer questions about data valuation. In the US, lawmakers have introduced legislation requiring companies to disclose what data they collect from consumers and how they benefit from it; while others have proposed a data dividend to share the wealth created from consumer data.

It's hard to estimate the company's business value and future potential accurately. This is especially true if neither the investor nor the company properly understand the potential future value of the company's data assets.

However, what makes data valuation difficult is that data is an asset not yet recognized by generally accepted accounting practices.

The distinct characteristics and dynamics of data – contextual, relational and cumulative – call for new approaches to articulating its value. This requires a mindset shift – businesses should value data based on cases that go beyond the transactional monetization of data and take into account the broader context, future opportunities to collaborate and innovate, and value created for its ecosystem stakeholders. Assessing data against key value and cost drivers, in the context of different use cases and with attention to shared value for stakeholders, will encourage companies to think about the future value data can help generate, beyond the existing data lakes they sit on, and open them up to collaboration opportunities.

References

- *AEG. (2020). Recording and valuation of data in National Accounts. 14th Meeting of the Advisory Expert Group on National Accounts. 5-9 October. Virtual Meeting. Agenda item: 5.3.1. Available at: https://unstats.un.org/unsd/nationalaccount/aeg/2020/M14_5_3_1_Recording_of_Data_in_SNA.p df*
- *Anmut, (n.d.). Data Valuation| Why It Matters & How It's Done, viewed 28 April 2022. Available at: https://www.anmut.co.uk/an-introduction-to-data-valuation/*
- *Arrieta-Ibarra, I., Caruso, T., Hernández, D.J. and Weyl, E.G. (2020). An empirical study of the value of data. Mimeo.*
- *Heckman, J.R., Boehmer, E.L., Peters, E.H., Davaloo, M., and Kurup, N.G. (2015). A Pricing Model for Data Markets. In iConference 2015 Proceedings. Available at: https://www.ideals.illinois.edu/handle/2142/73449*
- *Laney, D.B. (2017). Infonomics: How to monetize, manage, and measure data as an asset for competitive advantage. Routledge.*
- *Neely, A., "Why You Should Be Treating Your Data As An Asset", Anmut, https://www.anmut.co.uk/why-you-should-betreating-your-data-as-an-asset*
- *The Economist, "The world's most valuable resource is no longer oil, but data", 6 May 2017, https://www. economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data*
- *The Economist, "Are data more like oil or sunlight?",20 February 2020, https://www.economist.com/specialreport/2020/02/20/are-data-more-like-oil-or-sunlight*
- *ODI. (2018). Data's value: How and why should we measure it? [Blog] Open Data Institute. 9 August. Available at: https://theodi.org/article/datas-value-how-and-why-should-we-measureit/*
- *Schmarzo, W., "3 Theorems on the Economic Value of Data", Irion, https://www.irion-edm.com/news/3-theorems-onthe-economic-value-of-data-by-bill-schmarzo/*
- *Slotin, J. (2018). What Do We Know About the Value of Data? Global Partnership for Sustainable Development Data, 2018, available at: http://www.data4sdgs.org/news/what-do-we-knowabout-value-data. CGMA*
- *Data valuation: Understanding the value of your data assets, Deloitte, 2020, https://www2.deloitte.com/content/dam/ Deloitte/global/Documents/Finance/Valuation-Data-Digital.pdf*
- *Coyle, D., Diepeveen, S., Wdowin, J., Kay, L., and Tennison, J., The Value of Data, Bennett Institute + ODI, 2020, https://.cam.ac.uk/media/uploads/files/Value_of_data_summary_report_26_Feb.pdf*