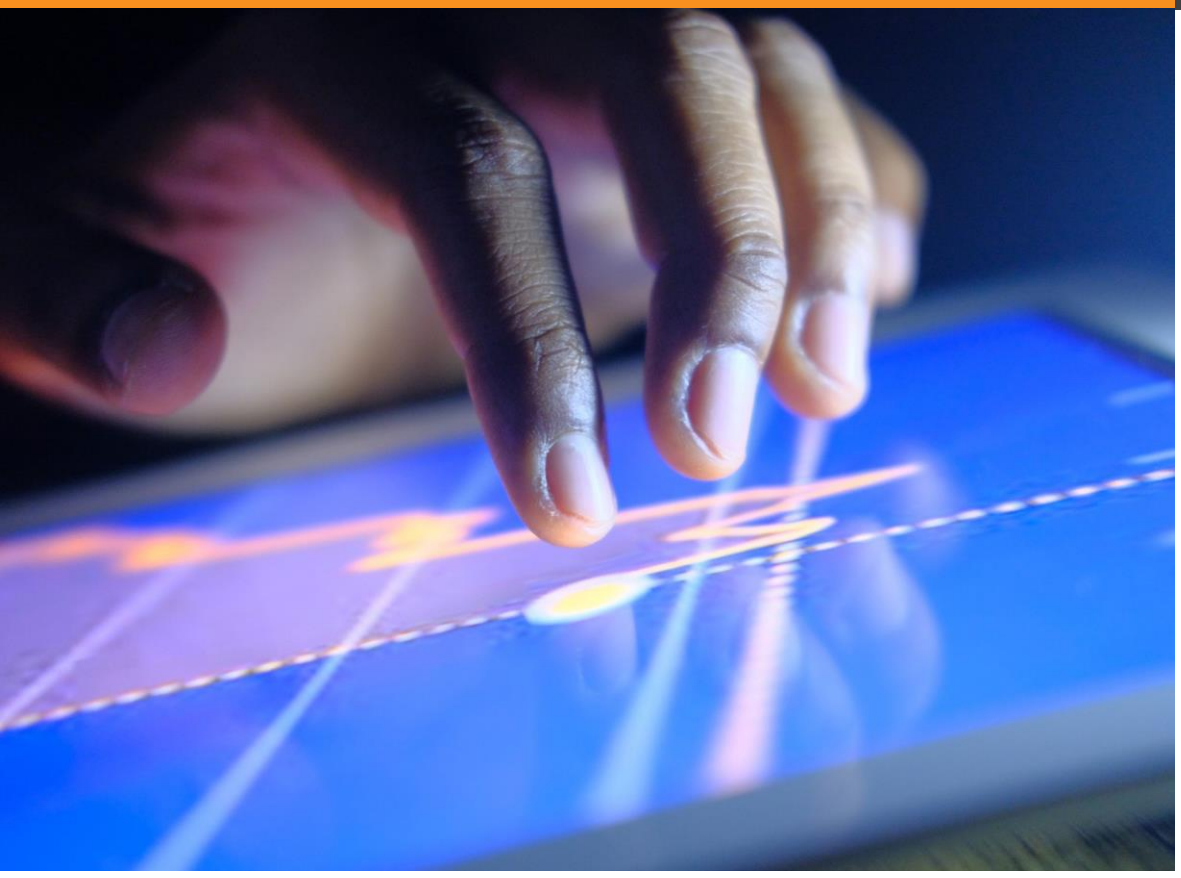
Putting a value on data



pwc

Introduction

The value of information assets has never been greater. According to the European Commission, by 2020 the value of personalised data – just one class of data – will be one trillion euros, almost 8% of the EU's GDP[1].

In a rapidly evolving data landscape, more and more organisations are looking for ways to create value using their or someone else's data (see Figure 1). Those that take the time to properly diligence and understand the value of their data will unlock significant value. Those that don't risk wasted investment in costly implementation programs with little or no upside.

In a recent paper[2], PwC's Data and Analytics leader, Neil Hampson, talked about strategies for creating value from data. In this paper we help organisations understand how to value data and some of the drivers that impact value.

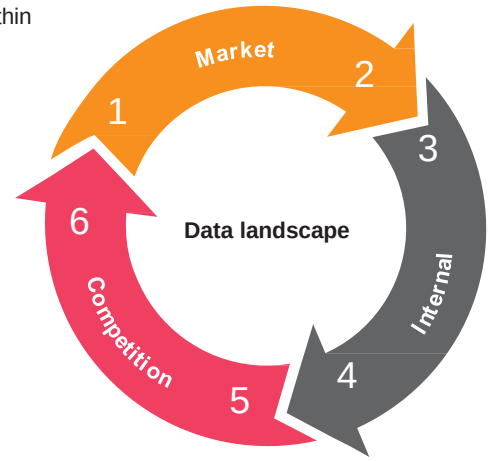
Figure 1 – Organisations are looking for value in a rapidly evolving data landscape

1 Organisations are starting to understand the value of being viewed as data-centric

- Valuation multiples of data-driven firms tend to be significantly higher than other industries
- Valuations of data-driven firms within the same industry also tend to be higher than their peers

2 Organisations are getting requests from others to use/buy their data

- Increasing use of data analytics to inform decision-making and expand capabilities
- Growing dependency on third party data sets
- Pace of change requiring new data sets and insights to meet evolving conditions (e.g. more sophisticated targeted marketing)



6 Organisations have rightly started to fear big data players

- High concentration of data aggregators; fragmentation of suppliers and consumers.
- Increasing potential for disruption from major data players (e.g. Facebook, Uber, Amazon, Google)

3 Organisations are continually on the look-out for new revenue opportunities

- Pressure to diversify revenue streams and grow margins
- Traditional customer acquisition/ retention methods are becoming less effective
- Monetising data from IoT sensors is becoming an essential element of many new business models

5 Organisations are witnessing the success of others with strong analytics capabilities

- Companies with good data analytics capabilities are twice as likely to be in the top quartile of performance within their industries

4 Organisations are finding technology is improving while costs are dropping

- Use of Open Source software for data storage and processing (e.g. Hadoop) is growing
- Cloud storage costs are only cents per GB per month

Technology and data trends

Mobility, connectivity, automation, big data, cloud computing and the proliferation of the internet of things enable new capabilities and data-driven opportunities

[1] https://www.weforum.org/agenda/2017/09/the-value-of-data/
[2] Creating value from data, March 2019

What is a data asset?

In simple terms data comprises – 'facts and statistics collected together for reference or analysis'[2].

There is a lot of it about. The world is producing about **2.5 quintillion bytes of data per day**, with 90% of all data having been produced in just the last two years[3]. With this explosion in the sheer volume of data being generated, an organisation wishing to understand the value of its data must start by identifying and classifying what it has – i.e. performing an inventory of its data. By doing so it can:

1. Assess the quality of its data
2. Identify gaps in the data (which it may wish to fill by buying data or partnering with others)
3. Highlight key issues which might hinder its data strategy (such as legal or regulatory restrictions over use of the data[4])
4. Think about 'use cases' for the data (i.e. what are the possible applications for its data)

The process of classifying the data will give an organisation a common understanding of the data it has across its different business units and functions (see Figure 2 as an example), but that doesn't mean the data is valuable.

Figure 2 – Illustrations of data source and data category dimensions – an inventory of data

Data source	Data category	Sub-categories	Illustration
Authored data • Typically created through some kind of creative, human process. • E.g. Architectural drawings, photographs	**Master data** • Describes people, places, and things that are critical to a firm's operations.	Customer data Supplier data Product data Employee data	Customer address Contact details Products, Features Employee name
User Provided data • Data purposefully provided by users into a system without any expectations. • E.g. Social media, ecommerce reviews	**Transactional data** • Describes an internal or external event of transaction.	Sales data Payment data Touchpoint data Geospatial data	Customer purchase history Payment date Call record Current Location
Captured data • Recorded from events occurring in the real world or in software. • E.g. Financial transactions, Web browsing logs	**Reference data** • Information that is used solely for the purposes of categorising data.	Jurisdictions Control data Currencies Industry standard data	Provinces Holiday calendar Currency codes Country codes
Derived data • Generated by combining, aggregating and otherwise processing other data. • E.g. Credit scores, aggregated transactions	**Metadata** • Characterises other data, making it easier to retrieve, interpret, or use the data.	Descriptive data Tables, columns Lineage data Audit trail data	Author, Abstract Type, Relationship Modifications Accesses, Changes
	Unstructured data • Data lacking a consistent format or syntax to describe objects and attributes.	Audio data Text data Video data Picture data	Recordings Reports Surveillance footage Social media postings

For the data to have value it must satisfy some basic premises:

1 It must be **identifiable and definable** – a data asset may be made up of specific files or specific tables or records within a database

2 It must promise probable **future economic benefits** – the data asset must have useful applications which are capable of driving incremental future cash flow either for itself or others

3 It must be **under the organisation's control** – the organisation must have rights to use and exploit the data in a way which is compliant with applicable laws and any contractual licensing arrangements, whilst also protecting the data and restricting access to it

[2] Oxford English Dictionary

[3] https://www.weforum.org/agenda/2017/09/the-value-of-data/

[4] Significant regulatory responses to widespread use of data – General Data Protection Regulation (GDPR) for example – have prompted many organisations to define what data they have and think more carefully about how it can be used.

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

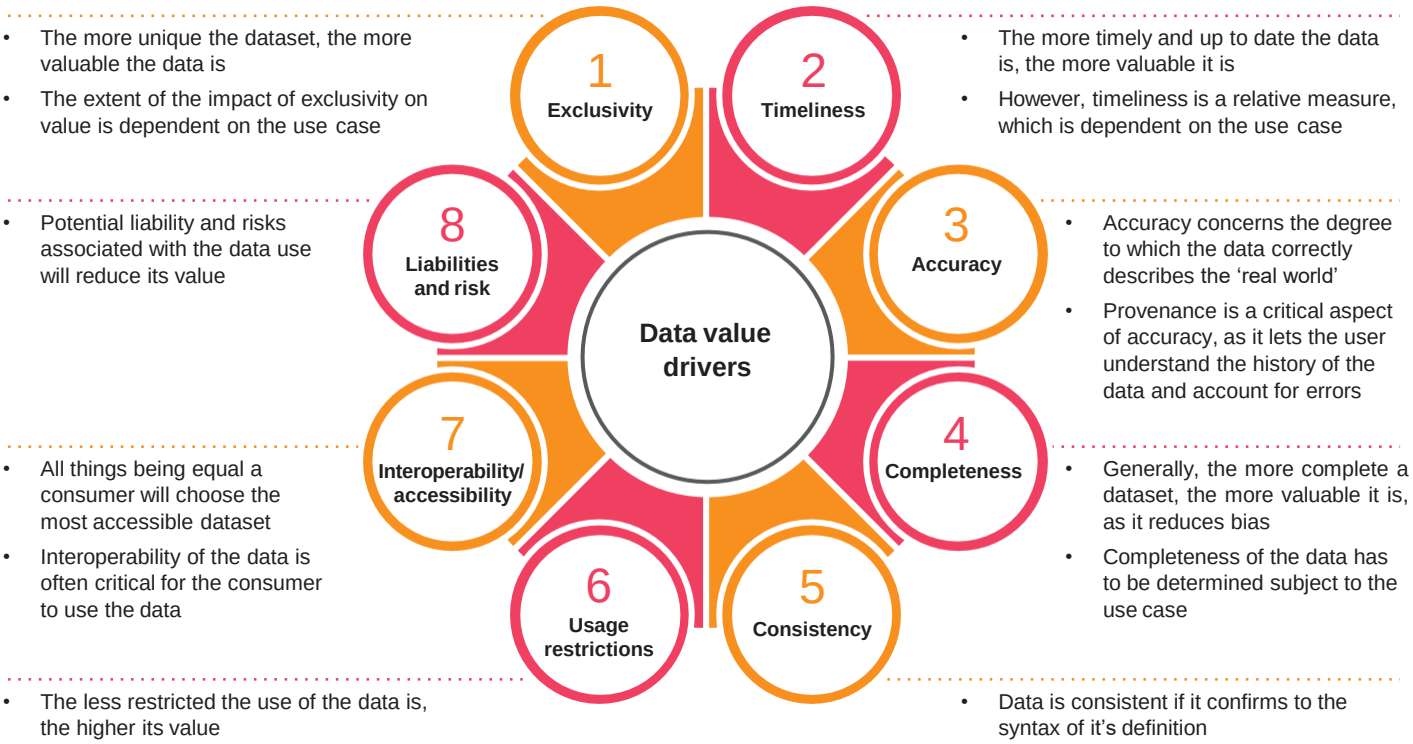
Strategy 1: Enhance current business		Strategy 2: Enter adjacent businesses	Strategy 3: Develop new businesses	
Leverage enhanced data for core business	**Generate new insights**	**White-label capabilities and infrastructure**	**Create new data**	**Create new offerings**
• Seek opportunities to enrich existing service through new data sources • Develop and leverage new platforms • Deliver enhanced services (e.g. in real time)	• Understand deep client insights • Enhance marketing campaign ROI and conversion	• Monetise existing analytics capabilities via white labelling to clients and other partners across the value chain • Commercialise infrastructure to sell platforms as a service	• Partner with adjacent players across the business value chain • Identify new sources of data (e.g. unstructured) to combine with existing data sets • Monetise new sets of data	• Develop new sets of analytics and data products (e.g. benchmarks, tools) • Develop new products that benefit from enhanced data and analytics (e.g. real-time net asset value, active non-disclosed exchange traded funds)

Netflix provides a good example of Strategy 1 – using data to enhance its current business. Unlike many traditional broadcasters, it uses individual subscriber data to tailor its content offerings based on your particular preferences. You can see this whenever you let a family member use your Netflix account profile to watch films you'd never consider – you'll notice the streaming service's recommendations start to change a little. This use of granular data to customise the individual user experience makes subscribers sticky and reduces the costly risk of producing or commissioning unpopular content.

Delving one layer deeper, the extent to which data can drive future economic benefits for an organisation will depend on key characteristics of the data – **data value drivers** such as those shown in Figure 4.

Some of these value drivers define the quality of the data (Completeness, Consistency, Accuracy, Timeliness), others could either render the data valueless or create unique and valuable competitive advantages for the owner or rights holder (Exclusivity, Restriction, Liability, Interoperability).

Figure 4 – Data value drivers



As an example, think about these value drivers in relation to the value of geographic location data on individuals to a mobile phone company. This data might be used by a retailer in combination with data on an individual's historic purchasing habits to push real time product adverts to an individual's mobile phone:

1. **Exclusivity –** having exclusive rights to this data set would be far more valuable than merely being one of multiple licence holders.

2. **Timeliness –** data which describes an individual's location at a precise moment in time is significantly more valuable to a retailer than data which describes their location an hour ago.

3. **Accuracy –** data may pin-point a person's location to the nearest 10m or only to the nearest 200m. Less accurate location data may result in adverts being served when an individual has already left the shopping centre.

4. **Completeness –** data which encompasses location and direction of travel will be far more valuable than data covering only one aspect – location or direction of travel.

5. **Consistency –** the personal identification data which links purchasing habits to an individual needs to be consistent with the personal identification data which provides that same individual's location, otherwise they will be being served with the wrong data.

6. **Use restrictions –** local regulation may require an individual to 'opt-in' when it comes to pushing location based adverts to their mobile phone. Without this opt-in by the individual, the data will be worthless to a retailer.

7. **Interoperability –** the ability to combine real time location data with purchasing habit data creates significant value through the interoperability of the data sets.

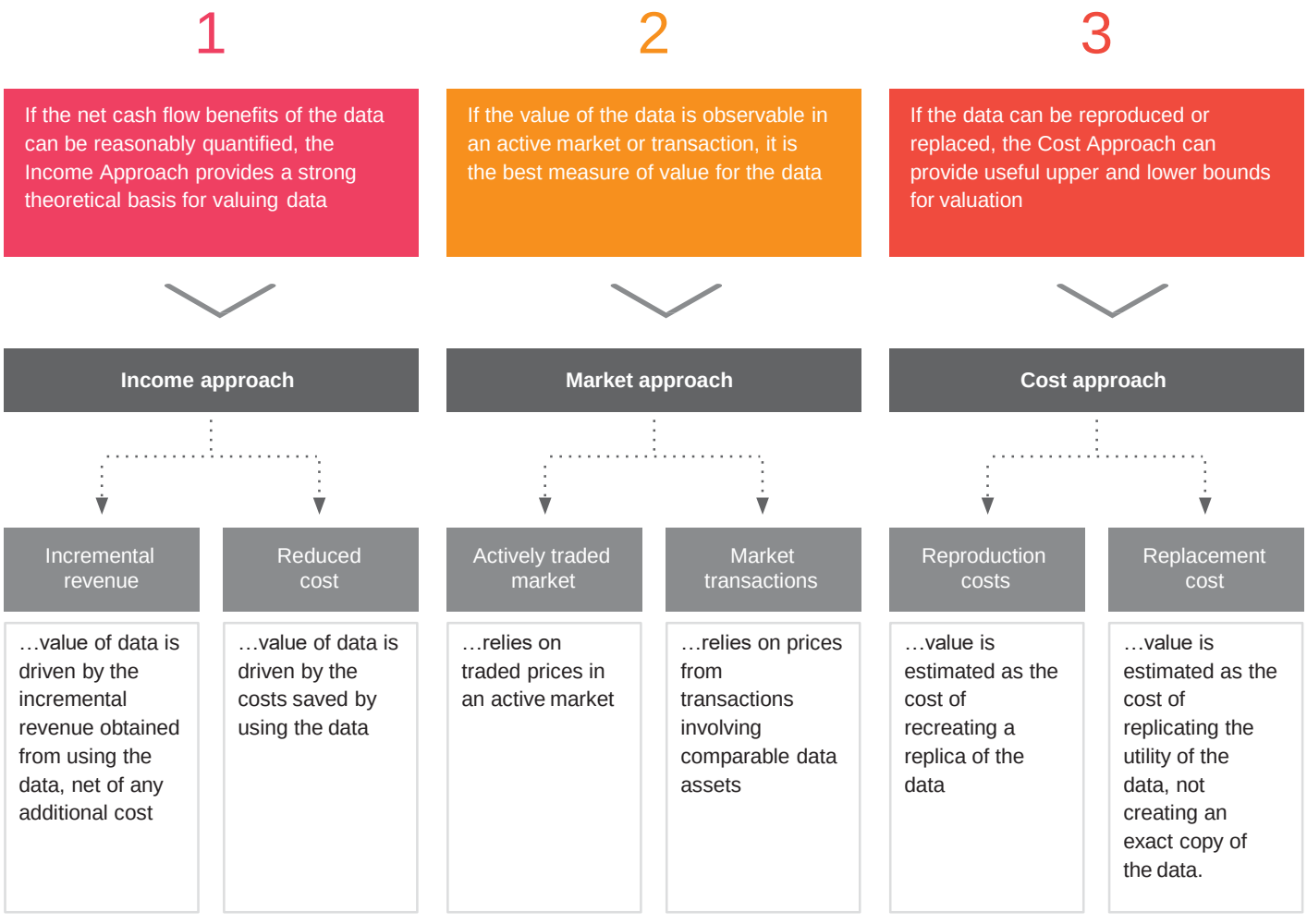
8. **Liabilities and risk –** the reputational consequences and financial penalties for breaching new data regulations, such as GDPR, can be severe. The greater the risk associated with the data use, the lower its value.

Assessing data against these key value drivers in the context of different applications/use cases will force an organisation to challenge itself on the data's potential to generate future economic benefits (net of any investments required).

The right method for valuing data

There are not many publicly agreed methods for valuing data, posing challenges for organisations, regulators and tax authorities alike. But the good news is that the same three approaches typically used to value any asset – namely Income, Market and Cost approaches – are still appropriate when it comes to data. Figure 5 below sets out the three approaches:

Figure 5 – Data valuation methodologies



1

The income approach

In a data valuation context, an Income Approach measures the incremental cash flows which the use cases are expected to generate. These incremental cash flows are derived from a comparison of the organisation's cash flows **'with' and 'without'** the data. With the data, the incremental cash flows could be in the form of incremental revenue, decreased costs, or both. Take these simple examples:

Incremental revenue for a beauty salon using data on consumer buying behaviours

For example, a spa and beauty salon chain purchases data on consumer buying behaviours in the beauty industry from a data intermediary. The study highlights a correlation between the customers' age and the type of marketing campaigns to which they are most receptive. Women below 30 years old are more receptive to social media influencers while women above 50 prefer famous actresses as marketing ambassadors. As a result, the spa and beauty salon chain is able to use the data to run a more targeted advertising campaign. Social media influencers are hired to advertise products and services more commonly used by younger women, such as the latest cosmetic products or make-up trends, while famous actresses are engaged to advertise services which are more popular among older women, such as face lifting and skin tightening treatments. The company is able to reach out more effectively to its customers and sell more products and services, increasing its annual revenue by 20% compared to years when only mass marketing campaigns were used.

Leveraging data is not only about incremental revenue generation. Organisations can use data to reduce costs through better planning and optimisation of operations, as well as reducing and managing risk. Examples of cost reduction strategies include using data to enable better management of fraud risks, or sharing inventory data with suppliers to optimise inventory management. Here's another example:

Decreased costs for an online retailer using data on consumer returns

A large percentage of product returns is one of the greatest challenges an e-tailer faces as returns cost 1.5 times the actual shipping fee. An online clothing marketplace purchases a market study from a data analytics company to understand more about product returns such as 'what products tend to be returned most', 'what are the most prevalent reasons for return (e.g. size mismatch)?', 'which suppliers' products are returned most frequently?', etc. With this understanding, the company takes actions to reduce returns, such as providing size charts in different standards, and proactively calling up customers if products are more prone to returns such as shoes. These efforts are predicted to translate into a 60% cost saving.

These incremental revenues or reduced costs may be earned or saved over a future period of time limited either by legal rights to the use of the data, or by the useful economic life of the data from a purely commercial perspective. Once the data owner has determined the time period over which incremental revenues or reduced costs can be earned or saved respectively, they can be present-valued, net of any investment required.

2

The market approach

A Market Approach can lay claim to providing the best evidence of value if the relevant information is available. **Active markets** for data are still relatively rare though. The ones that exist are either trading in illegal or illicit data, or are relatively nascent and not particularly liquid, so while such markets should in theory offer the best evidence of value, sourcing such information is often impossible. Examples of such markets would include:

- **BDEX** which is a real time marketplace that enables consumers to sell their data (including spending habits, browsing profile, purchase history, etc.) to retailers, brands and agencies.
- **Shutterstock** and **Flickr** are examples of marketplaces where digital images are bought and sold on a regular basis.

Transactions may also inform the value of certain types of data. For example, in July 2018 a California-based DNA-testing company, 23andMe, sold exclusive access to its database, comprising the genomes of over 5 million people, to the pharmaceutical giant GlaxoSmithKline for US$300 million. It was reported that 23andMe's database, one of the world's largest private genetic databases, could be used for 'research and development of innovative new medicines and potential cures.' From this transaction a 'value-per-data record' of US$60 could be used as an indicator of market value for personal genetic data (US$300m divided by 5,000,000 people). As with any comparable transaction approach, if the observed value is being used to estimate the value of another data asset, then the two assets must of course be comparable.

So, while examples of market values for certain types of data exist, it is currently very challenging for organisations to obtain such information, and so the ability to apply a Market Approach in valuing data is currently very limited.

3

The cost approach

The Cost Approach is a more straight forward approach, but the problem with it is that it typically fails to capture future economic returns that may be earned from the data by the owner or rights holder. For that reason, although it can still provide some useful benchmarks, it would not be recommended as a primary approach.

It may, however, set a base value in the sense that an organisation selling data should, at a minimum, seek to recover its cost of investment in the data asset, plus some fair return. From a buyer's perspective they may determine that the most they are prepared to pay for a data set is either the cost they would have to incur to recreate that data, or the cost of buying it from someone else. That assumes of course that these two options are available.

Conclusions

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.

The most robust way of valuing data is currently the Income Approach. It also has the advantage of forcing an organisation to think explicitly about the incremental cash flows that different use cases for the data can generate and the impact of data value drivers on those cash flows.

As transactions involving data assets become more prevalent, the Market Approach will grow in importance, but this will take time. Until then it is back to fundamentals – what incremental cash flows does the data investment generate, for how long and with what level of risk?

Contacts

Nick Rea
M: 07713 793612
E: nick.rea@pwc.com

Nick advises clients in the technology and financial services sectors on value-related issues for M&A, disputes, tax and commercial negotiation purposes. He leads PwC UK's valuation practice within PwC's Deals business and is a member of PwC's global valuation leadership team.

Adam Sutton
M: 07483 407724
E: adam.sutton@pwc.com

Adam provides specialist valuation advice to clients in the technology, media and telecoms sectors covering areas such as M&A, tax, disputes, financial reporting and pension covenant issues.

[5] An organisation's ability to exploit the value of data is also contingent upon having the right technical infrastructure and management processes, as well as the right talent in place. These are topics which we don't cover in this article, but which are a vital component of successful monetisation.

pwc.com

190306-135602-AS-OS