2.0 Setting the Scene

It has never been easier to gather and store information. Data is now a key raw material of business, government, and society. For governments, business, and even private citizens, data is cheap, widely available, and relatively easy to access, and its use influences almost all aspects of how our society works. Organisations and governments use it to conduct their operations - whether that is the delivery of services such as financial advice or healthcare, and the sharing of information in the media, or to make decisions such as what products should be made available to who and so on. People use data across the board: to access and use services, stay in touch with friends or family, administer things, make decisions, and to undertake multiple tasks such as take exercise or even (increasingly) to find romance.

Beyond this, data is opening up new frontiers in science and the humanities, from extending our knowledge of how the universe is built, to creating new understanding around climate change, to discovering the impact of a specific teacher on a specific pupil's performance. All this suggests that data has potentially enormous value.

What is surprising, however, is just how little consensus there is around exactly what this value is, or where this value comes from. Many fundamental questions remain unanswered and are subjects of continued debate and controversy.

Some of these questions include:

- Why exactly is data valuable? What are the benefits we derive from it?
- Is this value of data mainly financial/monetary, or does it derive from direct utility - the things we can do with it? If so, what sorts of things? Does the value derive from the ability to gain new insights, or perhaps to streamline or automate processes?
- What are the different ways in which this value is generated?
- Are the benefits being shared fairly? If not, how can a fairer share of benefits be achieved?
- What are the main barriers and obstacles to realising the full potential of data, and what is the best way to address them?
- Are there also harms as well as benefits? If so, what can we do to alleviate them?

Without a collective understanding of the distinctive characteristics of data and the multiple different ways it can be put to use, opportunities to extract value may be missed.

To this, we need to add the complexities of different cultures, different types of technology, and hugely differing stages of technology development and adoption around the world. This changes peoples' and societies' experiences of using data, their perceptions, and their priorities. The same uneven development also makes it unlikely that the benefits of data will be uniformly shared. And without a collective understanding of the distinctive characteristics of data and the multiple different ways it can be put to use, opportunities to extract value may be missed. Different innovations, perspectives, priorities, and initiatives from multiple sectors and across geographies, means that we are likely to see a myriad of future pathways. However, if we explore the different approaches to valuing data, and try to understand the rationale behind them, we may be better prepared for change.

This report highlights those issues and questions which were most debated during our research, and provides an overview of the different points of view around the world. We do not intend to cover every aspect of the Value of Data, or every argument or counter argument made in relation to the points we raise. However, we hope that we have managed to help clarify the issues that were important to those we engaged with.

2.1 What is Data?

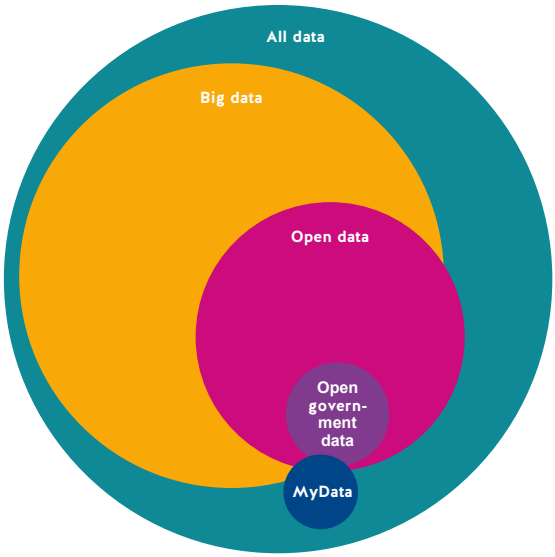
Before addressing these questions, we need to clarify terms. Some, if not all, hotly debated issues relating to data, stem from the fact that people are talking at crossed purposes. They haven't clarified what data actually is and they end up using the same word to describe different things. Many discussions about data may presume a shared understanding between stakeholders that may not actually exist. For the purposes of this report, the following definitions may help.

All data begins with a **data point.** This is a discrete unit of information, such as the temperature (Celsius or Fahrenheit) at a particular location at a particular time. A **data set** is a collection of data points. For example, it could be the different temperatures in a particular location at different times, or in different locations at the same time. Combining different data sets, (temperatures in the same place but at different times; or in different places at the same time) helps us get a richer understanding - such as how temperatures vary between night and day and across seasons in one location, or the differences between this location and other locations. When combined, they could (for example) become vital elements of the study of climate change.

Meta data provides information about other data. Keeping with our example about temperature, this could be data about the device used to measure the temperature, or who did the measuring. Other examples of metadata include the time of receipt of an email or phone call, or the location where a picture was taken, but not the email, the phone call, or the picture itself.

According to the definition made by European data protection rules (the General Data Protection Regulation, or GDPR), **personal data** is any information relating to an identified or identifiable natural person (a **data subject**). This is a living individual who can be identified, directly or indirectly, by reference to an identifier, such as a name, an identification number, location data, an online identifier, or to one or more factors specific to their physical, physiological, genetic, mental, economic, cultural, or social identity. **Non-personal data** is any data that cannot be connected to an individual, such as a company email address, info@, or indeed a company registration number. It can also include **pseudonymous** or **anonymous data** that is personal data that has been, ideally but not definitely, irreversibly de-identified.

Structured data is made up of clearly defined data which is formatted so that it fits into a formal database and is easily searchable; while **unstructured data** is not organised in a pre-defined manner and is usually not as easily searchable; it could include formats like audio, video, and social media postings.



Relationship of Different Data Types

Big data usually includes data sets with sizes beyond the ability of commonly used software tools to capture, curate, manage, and process data within a tolerable elapsed time. It includes structured and unstructured data. It is normally defined by its volume, its variety, its velocity (the speed at which it is generated), and its veracity (where it originates from and how complete it is). While there is much debate about big data in day-to-day life, the vast majority of data-driven activities and processes are driven by '**small data**': discrete bundles of specific data points that are uniquely relevant to the task in hand. These currently make up most of the activities and transactions which use data.

Data comes from a vast range of sources. For example, in our daily lives we create a massive amount of digital information about ourselves. This ranges from data that is captured via technological means, barcodes, online systems, credit cards, and so on; it is 'volunteered' by individuals, sometimes formally through form filling or informally on social media; it is also generated as a by-product of operations such as banking, measuring pollution, or using security cameras. Some data is **personally identifiable**, such as bank details, while other data is **non-identifiable**, for instance, statistical data about pollution or traffic flow. Proxy data is data used to study a situation, phenomenon, or condition for which no direct information is available – for example, scientists use the measurement of tree rings as **proxy data** to estimate climate change variances.

Coincident with the revolution in the collection of data, we are experiencing dramatic changes in the technologies of **data capture, storage, analysis,** and **transmission**. Together, these technological advances are turning our society and economy from a data desert to data ocean. This is not only about the quantity of data available, but also around the ability to leverage it across organisations and industries.

'Data products' are created by the aggregation of data to establish a new, higher level data point that can be used for a particular purpose - for example, a profile, identity, or credit reference. Good examples of this are Google search and Amazon product recommendations, both of which improve as more users engage.

But 'data products' are just one of the many and varied uses of data. Data is essential for all record keeping and administration, and for organising and coordinating activities. For example, individuals use data when they say, 'I'll meet you at [this place] and [this time]', while retailers and producers use data to organise and manage their supply chains. Many industries' core operations are data driven. Without data, banks cannot operate, because they cannot track who is moving money, how much money, or where. Data is also used to measure and monitor our environment, be that temperature recordings, health records, or economic statistics. Interrogating data helps reveal patterns, trends, and variances that previously weren't visible. This in turn helps us develop knowledge and understanding - and the resulting insights help us make better, more informed decisions.

2.2 What Makes Data Uniquely Special?

All of these varieties and differences are important when discussing data and how to unleash its value. But there is one special thing about data that unites them all. As a 'resource', data is different to traditional physical resources. Like knowledge and ideas, when data is 'used', it doesn't get '*used up*'. This means the same piece of data can be used for multiple different purposes by multiple different parties. And far from being a depleting resource, it is an accumulating one.

The implications and ramifications of these unique characteristics of data are vast, as this report will show. They up-end our notions we often see as 'fundamental', such as that of 'ownership'. They force us to challenge many of the assumptions that lie at the heart of economic analysis. They transform both the possibilities data creates and the dilemmas it generates. They change relationships between stakeholders, be they individuals, communities, networks, organisations, government, or wider society.

2.3 Debating Data

Given the multiple different types, forms, and varieties of data, and the equally multiple range of different uses, when debating issues relating to data, it is almost impossible to avoid the classic fable of the blind men and the elephant, each one touching a different part of the beast, drawing entirely different conclusions about its overall nature. To help them describe what they are experiencing, they naturally turn to analogies. These can be useful, as they capture certain aspects of the way that data behaves in relation to certain aspects of the economy and society, at certain times. But they can equally mislead, causing us to associate data with the wrong things, leading to wrong, even dangerous, conclusions. Here are some of the analogies that were discussed in our workshops.

Data as 'the new oil': Well, data is mined and refined, like oil. Vast hordes of it can make its owners (or 'controllers') very wealthy and powerful, like oil. We might even go to war over it, like oil. But there are also many ways in which data is not like oil. Unlike oil, data is not a finite, exhaustible resource. We have just seen that, unlike oil, when data is used, it doesn't get used up. Indeed, in many cases, data is replicable or reproducible, with the very process of using data creating *new* data - for example, the meta data about what data has been used, and for what purposes. Also, unlike oil, the material costs of extraction, collection, and movement of data are not high, and are falling rapidly. And data ownership is not easily defined, unlike oil.

These differences are important, since they point to a set of end-points for the data economy that are completely different to those of an oil economy, and so demand a different set of societal responses.

Data as currency: Data can certainly serve as a medium for exchange, as it does when a consumer, for example, shares their personal data in exchange for so-called 'free' services. It can also be used as a store of value, even in quite a literal (albeit unstable) sense when it comes to crypto-currencies. So yes, data can be used like a currency in some circumstances.

But describing data as currency really doesn't tell us much. It just tells us that data has exchangeable value in certain contexts. In that sense, many things operate like currency. The economic value of data might have risen in recent times, and more people might be aware of that value, but the same might also be said of quinoa. Describing data as currency can simply edit out many of its most important features.

Data as 'a periodic table': One suggestion, made in Singapore, that subsequently gained widespread support, was that data should be considered like the periodic table: *"Data is similar to the elements on the periodic table. They can act independently, but interact with each other to create new combinations."* Moreover, from a value perspective, different organisations have varied ideas of the value of the individual elements, depending on where they are looking from and what their field of activity is. But when a number of elements are combined into a compound, then it can be more useful with more value to more people, organisations, and society. As a metaphor, this seemed to work for many.

> "Data is similar to the elements on the periodic table. They can act independently, but interact with each other to create new combinations."
>
> Singapore workshop

2.4 Getting Value From Data

Generically speaking, participants talked about different types of value that data is used to generate:

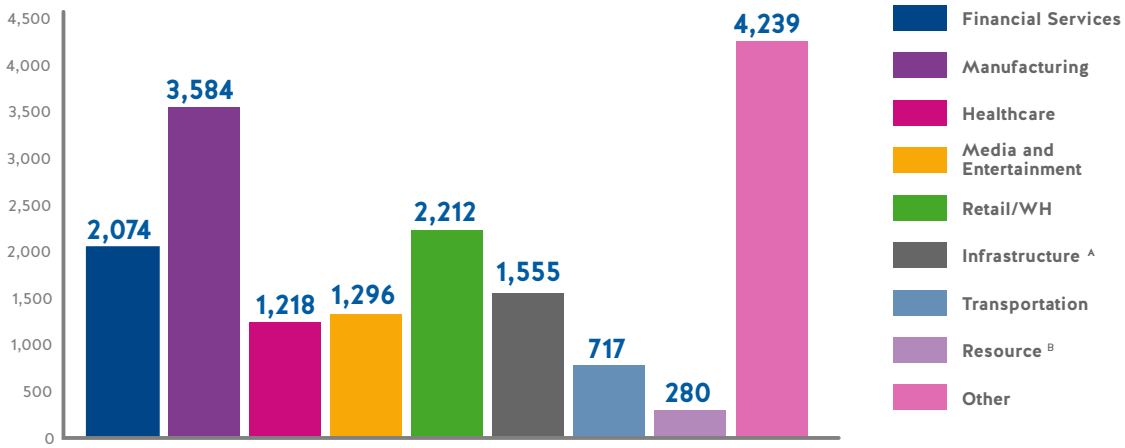
- use value: all the different uses to which data is put, as described above (administration, organisation and coordinating, analysis, decision-making)
- exchange value: the different ways that people can make money directly from data by selling, renting, trading, or other ways of charging for access to data

Stated boldly like this, the concept of value may seem simple. But, of course, it generates much complexity. In societal terms, much debate about 'value' is a proxy of what people think is 'good' versus what they think is 'bad': subjective judgement is never far away from debates about value. Other perspectives include those of different stakeholders, such as are we talking about value to a customer or to a firm? More sophisticated analyses highlight the difference between *potential* value and *realised* value. According to these arguments, value resides not in a product or service per se, but rather in a human being's experience. For example, a company's offering, whether it takes the form of a product, a service, or some mixture of the two, can be described as unrealised value (i.e. a store of potential value). This value is only realised when a customer uses it, and this is invariably an act of co-creation in a particular context.

In terms of practical application, value can be both positive and negative. Sometimes, data can provide unequivocally positive value, such as when large data sets are used to build smart energy or water grids, to improve travel safety, or to search for new cures for diseases. On other occasions, data may be used to generate unequivocally negative value, such as identity theft, cyber-attack, data blackmail, or the proliferation of false information ("fake news").

Other uses seem to allow for mixtures of both positive and negative value at the same time - or positive vs negative depending on how you look at it. Connecting people on a massive scale, for example, can enhance human relationships, allow ideas to flourish, and give voice to those who may not otherwise have one. But it can also enable bullying, criminality, or terrorism, give strength and credibility to bad ideas or ideologies, and encourage mob rule. Data-harvesting for surveillance purposes can help the development of new kinds of consumer products, be used for the purposes of delivering targeted advertising, or for feeding sophisticated algorithms that underlie the efficient delivery of services (policing, insurance, access to government services, etc.). But they can also be seen as incursions of, and threats to, privacy and civil liberties.

Much debate about 'value' is a proxy of what people think is 'good' versus what they think is 'bad'.



Source: IDC's Data Age 2025 study, sponsored by Seagate

[A] Infrastructure includes Utilities, Telecommunications

[B] Resource includes Oil and Gas (Mining), Transportation of oil & gas through pipelines or shipping, Resource industries, Petroleum and coal _ manufacturing/refining

Where the Data Lies: Global Enterprise Data by Industry (2018)

As will be discussed later, some suggest that a multi-capital view of data value should be used in line with what is proposed and adopted for the Integrated Reporting of an organisation's activities.[3] Others are proposing methods of valuing data against UN Sustainable Development Goals.[4] The IMF has recently hosted conferences exploring how to measure the value of an organisation's data. And in a world of 'free' and 'open' data, it is also looking at the implications of capturing digital impact within national accounts and GDP.

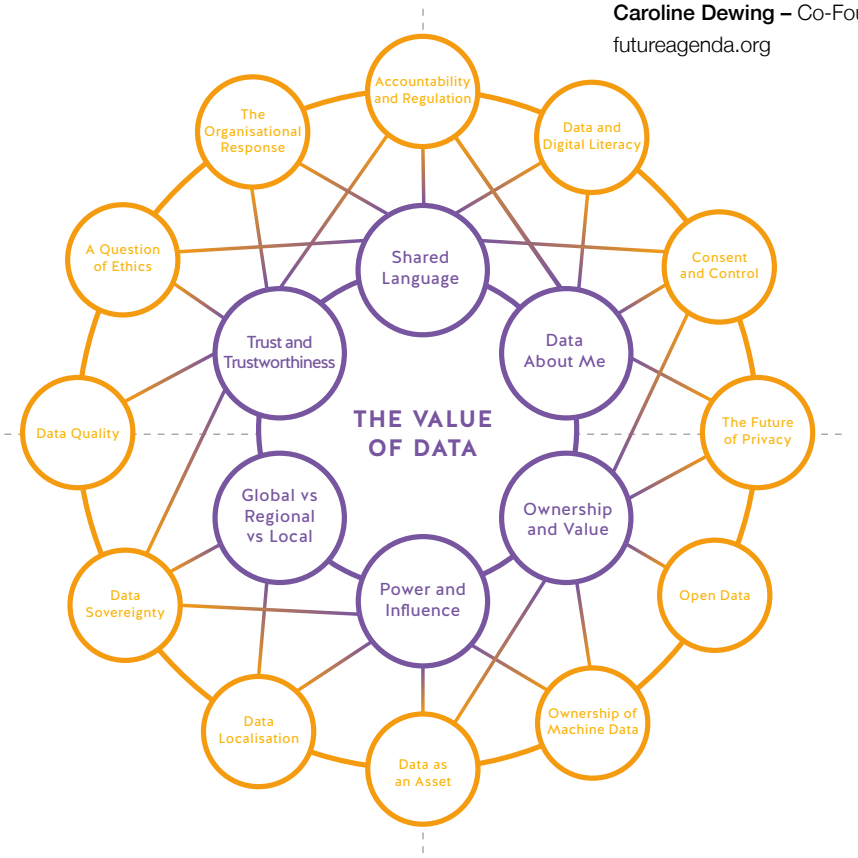
We will discuss these rich and varied views and nuances in this report, but the core of it remains relatively simple: the value of data (both positive and negative) lies in how it is used and/or how it is exchanged, and the effect that this can have on the economy, society, and individuals.

Context

Throughout 2018, Future Agenda canvassed the views of a wide range of 900 experts with different backgrounds and perspectives from around the world, to provide their insights on the future value of data. Supported by Facebook and many other organisations, we held 30 workshops across 24 countries in Africa, Asia, the Americas, and Europe. In them, we reviewed the data landscape across the globe, as it is now, and how experts think it will evolve over the next five to ten years.

The aim of the project was to gain a better understanding of how perspectives and priorities differ across the world, and to use the diverse voices and viewpoints to help governments, organisations, and individuals to better understand what they need to do to realise data's full potential.

From the multiple discussions 6 over-arching themes were identified alongside 12 additional, related future shifts as summarised in the diagram below.



Details of each of these, a full report and additional supporting information can all be found on the dedicated mini-site: www.deliveringvaluethroughdata.org

About Future Agenda

Future Agenda is an open source think tank and advisory firm. It runs a global open foresight programme, helping organisations to identify emerging opportunities, and make more informed decisions. Future Agenda also supports leading organisations, large and small, on strategy, growth and innovation.

Founded in 2010, Future Agenda has pioneered an open foresight approach bringing together senior leaders across business, academia, NFP and government to challenge assumptions about the next ten years, build an informed view and establish robust growth strategies focused on major emerging opportunities. We connect the informed and influential to help drive lasting impact.

For more information please see:
www.futureagenda.org

For more details of this project contact:
Dr Tim Jones – Programme Director,
tim.jones@futureagenda.org
Caroline Dewing – Co-Founder, caroline.dewing@futureagenda.org

First published November 2019 by:
Future Agenda Limited
84 Brook Street
London
W1K 5EH