3.0 Overarching Themes

Throughout the programme, six major cross-cutting themes emerged. They were discussed in different ways and there is a degree of overlap between them, but no matter what particular data-related issue was discussed, one or all of them is likely to be a key feature of the debate.

These themes are:

1. The issues relating to the collection and use of personal data – **data about me;**

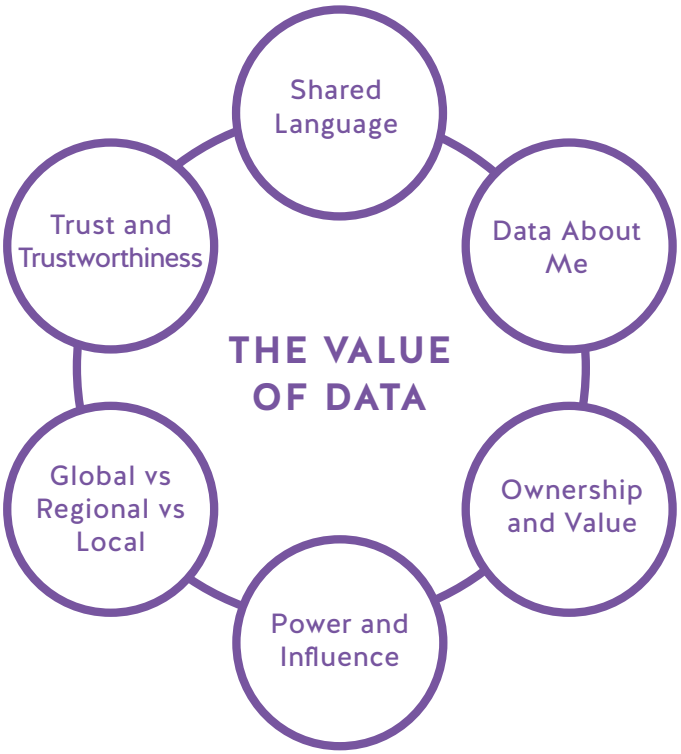
2. Topics linking to **ownership and value;**

3. Issues concerning the exercise of **power and influence;**

4. Matters relating to the level at which we are operating, such as **global versus regional versus local;**

5. Differing perspectives on **trust and trustworthiness;** and

6. The need for a **shared language** that avoids misunderstanding and confusion, and helps to clarify and advance the debate.



3.1 Data About Me

Rising concerns about personal data collection and use cover many issues. Pressure for solutions that inform and 'empower' individuals is growing.

In our workshops, much of the debate about data focused on personal data. This is not surprising. By definition, personal data relates most closely and directly to individuals' lives in many ways. Data about an individual may reveal intimate details about their lives. It could be - and is - used to bring them many benefits in terms of innovative, personalised services. But it could also render them vulnerable, especially if it gets into the wrong hands (for example via identity theft), or used 'against' rather than 'for' them (discriminating against individuals or groups of people based on what data reveals about them).

Personal data is also where debates about power and fairness is most acute. Huge amounts of money are being made by some profit-seeking companies via their collection and monetisation of the data of billions of individuals. Many individuals feel powerless in the face of these corporations and their intense concentrations of data power.

Such issues exercised the minds of many workshop participants, who wanted to analyse exactly what is going on in relation to the collection and use of personal data - and to find positive ways forward. It wasn't easy - partly because issues relating to personal data can be far more complex than they appear at first sight - starting with definitions.

Many people, when they talk about personal data, refer to very obvious bits of data such as name, address, contact details, payment card details, medical data, or personal purchase history. But the European General Data Protection Regulations (GDPR) go much further, defining personal data as: "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular 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 the physical, physiological, genetic, mental, economic, cultural, or social identity of that natural person."

By potentially including data points such as cookies ('online identifiers') and location data, this European definition of personal data casts the net much wider than many anticipate. As we will see in our discussion of the 'Internet of Things' and 'machine to machine' data, if it generates data that relates to an identifiable individual (for example, their usage of a device) in some jurisdictions, it will be seen as personal data. The border lines between 'personal' and 'non-personal' data are therefore not as clear as they may seem, especially when issues such as anonymisation and pseudonymisation are added to the mix.

This is important when we come to discuss the potential value of personal data. While much data 'about me' may include data that could be personally identifiable, there is also much data about people and their behaviours which is statistical in nature (i.e. not identifiable), but which is the source of important insights and of great potential in helping improve peoples' lives.

Many complex issues are therefore raised by how personal data is currently being collected and used. These include whether individuals know about or understand what data is being collected and what it is being used for, whether they would be comfortable about this collection and use if they did know, whether such collection and use of data infringes individuals rights to 'privacy', and whether they are receiving a fair share of the financial and other benefits that their data helps generate.

Multiple solutions are being proposed. These include:

- Ensuring greater transparency
- Questions about 'who we trust'
- User education
- Calls for regulation to empower individuals in their dealings with organisations
- Calls for regulation to restrict organisations' ability to collect or use data or exercise 'data power'
- Proposals to redistribute power and control by, for example, providing individuals with personal data stores which enable them to collect and control their own data independently of the organisations they deal with

"There is a need to find a balance between protection of personally sensitive data, and the value of sharing."

Bangalore workshop

It's not surprising, then, that many workshops focused their attention on issues relating to personal data. We will return to them in detail in specific chapters, but these quotes provide a flavour.

What We Heard

There was broad agreement that issues around the control of personal data are increasingly part of the public debate. In Dakar, it was observed, *"whatever happens, people still need to be at the centre of the system, not the machines. This will be difficult, because artificial intelligence is becoming more and more dominant."*

As understanding grows, many in our workshops felt that we are witnessing a swing away from corporate power, back to the individual. In Singapore, there was recognition that there is a conflict between what consumers understand to be ownership, and what companies understand to be access, but that *"people are taking data back – there may be a shift in power to control by the individual."* This sentiment was supported in Johannesburg, but with the proviso *"… it will depend on where ownership comes to rest."* In Tokyo, the view was that *"data will increasingly be owned by individuals and not by the government or corporates."* On the other hand, some felt that the whole issue is a bit of a red herring. In a student workshop in Pretoria, they proposed that *"no one should own data."*

Discussions around personal data highlighted a number of cultural differences. For example, in Europe, where privacy is held in high esteem, the view from London was that *"privacy is real – individually and nationally. We need a lack of compromise on this."* However, in Tokyo, the view was that *"most people don't really care about privacy – despite what the experts think."*

When it comes to the consideration of the value of data, the view in Bangalore was that *"there will be growing awareness of the value of personal data, and this will empower individuals…. But the appetite for monetisation will lead to more collaboration. There is a need to find a balance between protection of personally sensitive data, and the value of sharing."* In Copenhagen, they felt *"we have a willingness to sell data too cheap – it is a trade-off."*

3.2 Ownership and Value

Many link ownership with the right to extract value from data. But traditional notions of ownership don't apply, so new models are sought and tested.

In the discussions, there was a strong desire for clear rules and frameworks to establish who is the rightful owner of what data; the common assumption being that once ownership becomes clear, so do the related rights, benefits, responsibilities, and so on.

In some cases, 'ownership' of data is obvious: for example, data generated by an organisation in its internal processes is 'owned' by that organisation. However, generally speaking, data doesn't 'work' in the same way as traditional tangible forms of private property. Very often it is co-created by two or more parties via transactions, interactions, and communications, thereby creating two or more potential 'owners'. Because data can be used without being 'used up', the same data can potentially be re-used by many parties for many different purposes. Data can also be replicated many times over for close to zero cost, which makes it economically limiting, or simply very difficult to enforce traditional proprietary restrictions on the uses of data.

"Data is not created by an individual, it's a joint effort; but it's not realistic to think that ownership is the proper debate to be having. There are multiple owners of data: think of bank transactions...Ownership is an inaccurate term; it's too loose to frame the question."

Bangalore workshop

Rights and Responsibilities

These complexities are driving the search for alternative ways of framing the debate by, for example, focusing on questions of rights of access and use, and on custodianship rather than 'ownership' per se. The workshops identified and distinguished the role of multiple actors in the supply of data: originators, custodians, processors, and users. A great deal of the discussion focused on defining the rights, responsibilities, obligations, and opportunities for each of these roles. The issues and dilemmas are particularly acute when discussing personal data, where, aside from complexities arising from data co-creation, issues of human rights often overlap and/or clash with narrow, legal notions of private property. This debate is also becoming increasingly important with the Internet of Things, where multiple parties, such as device manufacturers, device users, and devices themselves, all play a part in generating data.

Distributing Value

Many of the liveliest debates in several workshops concerned the distribution of value among these actors. Separating the 'ownership' and use of data by other parties was a recurring theme.

As a result, the emerging concept of data custodians was discussed at some length. It was suggested that 'data custodians' could have twin roles for which they would be rewarded: keeping data stores and sources secure (similar to a safe deposit box in a bank vault); and access and pricing control (similar to a literary agent). Some argued that the originator and custodian should essentially be the same actor, where all the data is both controlled and owned by the originator; others felt that the role is better suited to that of an intermediary or independent platform.

Managing Value

Although data manager business models are still emerging, the idea that some of us will gradually be willing to pay for our personal data to be looked after, shared against agreed preferences, and where appropriate, monetised, was often discussed. Whether there is a standard approach or whether there are different platforms with varied models for different sectors, cultures, and types of data, are as yet open questions. Many believed that if our personal data is worth something, then we should be able to see this, benefit from this, control it more effectively - and so also choose who else can access and gain from it.

"The value of data is very regional, and is largely focused on who benefits from it as much as who owns it."

San Francisco workshop

Several best practices for operating approaches and processes for data owners and custodians were also introduced into the discussions. These focused on areas such as payment for access to the data, and how ownership rights are transferred among the various stakeholders. Each of these models is different to those of today, where most of this activity is done by the processor.

Problems and Dilemmas:

- Is 'ownership' a useful/practical concept when it comes to certain types of data such as personal data?
- If not, what alternative concepts can we use to replace it?
- What other ways can we use to allocate rights, benefits, and responsibilities relating to data across stakeholders, including governments, technology companies, multinational corporations and individuals?
- In what circumstances does 'ownership' remain a valid notion?

What We Heard

In Frankfurt, the view was that in order to understand the value of our personal data, there must be a *"shift from a world where we have unclear views on data, lots of confusion, panic, and uncertainty, and no real alternative options for what to do with our data than what is provided by a few tech giants, towards a world with universal clarity of data value, ownership, and rights."*

Distributing Value

Type "who owns your data" into Google and you'll get dozens of interesting papers and articles – all with different opinions. But does it really matter? Many in our workshops thought not, and agreed with this perspective from Bangalore; *"data is not created by an individual, it's a joint effort; but it's not realistic to think that ownership is the proper debate to be having. There are multiple owners of data: think of bank transactions. Individuals interact with banks, creating at least a two-way process. Ownership is an inaccurate term; it's too loose to frame the question."* One way that this could be addressed is that individuals retain full ownership of their personal data in machine-readable format, but outsource its management and distribution to professional custodians, curators, or data brokers.

Managing Value

One way to manage the value of data is through personal data stores. These could allow individuals greater transparency on just how their data is being used. Essentially, this is a *"central repository for personal data, where individuals can access and control the access of others to their data."*[5] The creation of a new profession, privacy agents or data brokers, was also explored. In London, they were compared to the role played by asset managers, where *"in the main, we trust others to do it on our behalf – and can choose how (e.g. active, passive, ethical). The same may emerge in this space by trusted third parties (TTP), making it easy for the customer."*

Participants in our Kenya workshop built on the idea. In Nairobi, it was suggested that if there were a central repository for data, *"…allowing business and government to access personal information, but individuals to maintain control of their data and benefit from it,"* then *"… there will be wider access to information, without jeopardising personal privacy."*

Ownership to Custodian

There was general agreement that we will have to move on from 'ownership' to 'custodianship' within a decade. In Bogota, the suggestion was, although *"those who own data will continue to exploit its value…more data will be used for public benefit."* In Washington DC, they suggested that it would lead to *"better use of data from larger and more aggregated data sets"* that can have greater impact. Finally, in Sydney, it was suggested that we may well see more collaborative use with *"data being used to optimise social good – "data commons for social good," for example, focused on fewer car accidents, less teenage suicide, the ability to crowdsource health solutions, enhanced social belonging, more inclusive/less isolation and marginalisation – so data can make life better."*

This means there is a need for greater transparency, more information, better action, and a more widely shared informed view on data ownership and its implications. In a culture where everyone starts with trust as a default, the Danish view was that *"we can move on to community ownership of data – via cooperatives within society – that then provide the trusted platforms that can scale into broader ecosystems."* In San Francisco, a reflection was that *"the value of data is very regional, and is largely focused on who benefits from it, as much as who owns it."*

3.3 Power and Influence

Data is a means of exercising power, as well as a focus for multiple struggles for power. Regulation focuses on rebalancing influence between companies, government, and society.

Workshop participants around the world were acutely aware that with data comes power; that the more data an organisation can collect, use, or control, the more power it has at its disposal. This power can come in many forms. It could be the power to make decisions that affect peoples' lives by, for example, giving or withholding their access to services. Some organisations' use of data gives them the power to act as 'choice architects', deciding what information is to be presented to people and how. Concentrations of data can create concentrations of economic power, which in turn could affect the distribution of available benefits.

Given the many and varied ways in which data is collected and used by all the different parties, we found scope for multiple different power relationships, for example, between:

- Policy makers/regulators and large data-driven companies;
- Governments and their citizens;
- Companies and their customers;
- Different/overlapping political jurisdiction

"When companies mess with complexity too great to monitor or understand, disclosure becomes an empty gesture."

London workshop

There were also many different suggested ways of addressing unhealthy imbalances of power. The following generated particular interest:

Transparency: Many workshop participants were particularly concerned by what they saw as the unaccountable power of proprietary algorithms that are effectively immune from scrutiny, and give the organisations which develop them huge influence. The lack of transparency makes it almost impossible for anyone else to understand the economic, political, and cultural agendas behind their creation.

Accountability: There was also much concern about the ability of search engines and social networks to influence the information individuals are presented with. The power to include, exclude, and order the presentation of information, allows these companies to ensure that certain public impressions become permanent, while others disappear. Without knowing what a search engine actually does when it ranks sites, we cannot assess when it is acting in good faith to help users, and when it is biasing results to favour its own commercial, cultural, or political interests.

Ways of rebalancing power: Debate focused particularly on whether global technology companies have accrued too much power. Questions were asked as to whether they exercise this power responsibly, and what (if any) safeguards, regulations, and reforms are needed to create a healthier, fairer, safer, more innovative or resilient data ecosystem. Some workshop participants felt that the activities of those wielding disproportionate data power should be restricted by increased regulation. Others sought more radical responses by dispersing power more equally (via competition rules and anti-trust legislation, for example).

Problems and Dilemmas:

- Organisations collecting and using large quantities of data can generate significant value for individuals, society, the economy, and for themselves. At the same time, however, they may create excessive concentrations of power, and/or use the power they do have unfairly or inappropriately. How should these dangers best be addressed? By who?
- Moreover, by what criteria should we judge whether an organisation has accrued too much power, or is using this power unfairly or inappropriately? Who should be responsible for making such judgements?
- if a corporate entity is deemed to have too much power or to be exercising its power irresponsibly, what are the appropriate mechanisms for effective action?
- How should these decisions be implemented and enforced?
- How can/should disputes between different entities and jurisdictions (local, regional, global) relating to the collection and use of data be handled?

"Whatever happens, people still need to be at the centre of the system, not the machines. This will be difficult, because artificial intelligence is becoming more and more dominant."

Dakar workshop."

What We Heard

Questions relating to the exercise of power cropped up in most of our discussions. To provide a flavour of the discussions, we provide some examples here.

There is a growing sense that some companies are benefitting disproportionately from the collection, use, and frequently the sale of personal information. The Bangalore workshop pointed this out by saying, *"the consumers' rights are always fringe; they don't have the power of the likes of Google or Facebook."* This is driving a public desire to give individuals greater control over their data. It was recognised, however, that doing this could create a new dilemma; how to maintain control of our data without losing the benefits and conveniences that exchanging personal information for digital services undoubtedly provides.

Transparency: We heard many calls for more effective legislative frameworks to help shape the emerging data economy in a more equitable way, to increase transparency, and make technology companies more accountable. Many in Africa and Asia, inspired by the EU's stance on GDPR, were keen take up the challenge. In Mexico City, the view was that *"the biggest change will be in the way governments control data."*

"No one has yet worked out the extent to which patient data can compromise government security."

Singapore workshop

In Dakar, it was observed, *"as the power of data increases, it can be used to warp our sense of reality. Fake news is only an early sign of things to come…"* Across our workshops there were multiple calls for the need for greater digital literacy, so that individuals can choose what products and services they use, and have better control over their own personal data. Many argued for greater transparency and intelligibility around the use of data. They pointed out that if it is too difficult to understand what is being done with our data, it is impossible for individuals (or organisations) to have an equal relationship with the companies that exploit it. Some suggested that increased transparency would go a long way to addressing this, but it is not a solution on its own. One comment made in London was that *"when companies mess with complexity too great to monitor or understand, disclosure becomes an empty gesture."* For the power of data to be more equally spread, there needs to be greater public understanding about how data is being used. Some in London even suggested that transactions that *"are too complex to explain to outsiders, may well be too complex to be allowed to exist."*

Accountability: Across Africa and India, there was a strong sense of frustration about the dominance of primarily Silicon Valley American companies. Many saw this as a new form of colonialism, with personal data becoming the latest raw material exploited by the west. Participants in Singapore and Australia felt that managing the flow of national data was an issue of national security. In a workshop in Singapore, specifically focussed on patient data, we were told that the law restricts the sharing of health data beyond national boundaries because *"no one has yet worked out the extent to which patient data can compromise government security"*

In Bangalore, participants felt that the lack of transparency about how data is used and manipulated has led to a growing *"digital gap, both at country level and also for individuals."* This was also echoed in Madrid, where it was felt that this data divide will continue to grow, and will *"continue to be dominated by issues around transparency, ubiquity, and control."* Others reiterated the need for greater transparency about how data is managed and shared, in order to allow individuals to have greater control of their data.

Regulation: A number of mechanisms to ensure a more even distribution of power were discussed. This included greater interoperability and portability (spreading access), and the possibility of breaking up those organisations which have themselves become monopolies. In Bogota, it was suggested that public private partnerships could be the best way to create and implement better governance. Many advocated the establishment of a "Global Data Vision"[6], and a global body to develop and oversee the implementation of regulation. Sounds great - but when pressed, no one was really able to suggest how this should operate in practice, and where the ultimate responsibility should lie.

Finally, in Asia and the US in particular, we had conversations around geopolitics and how different ideologies might influence the use of data. In Hong Kong, the question was asked, *"what would be the implication of China winning the debate around data, and what would happen if it exports its values around the world?"* In Washington DC, the comment was, *"if you see this as competing modes, then it matters, because as China grows, more people/nations will try to emulate it."* Prosaically in Dakar, the view was, *"we don't mind if it's noodles in the morning or burgers in the afternoon; we need to create our own solutions."*

3.4 Global vs Regional vs Local

While many support further globalisation of data, others seek to assert stronger regional and national control to protect citizens and strengthen economies.

In many circles, there is a strong assumption that global 'Big Tech' firms can and will continue 'doing what they like'. But there is powerful sentiment, especially in fast-growing regions such as Africa and India, that governments should assert more control over data, to protect citizens' rights, develop the economy, and maintain a sense of cultural identity. This is creating potential conflict with those seeing global data flows as key to economic growth.

If the world was ruled by a single authority making wise, legitimate decisions and capable of implementing them efficiently and effectively, life would be simple. But it isn't. Instead, our reality is extremely complex. We are governed by a myriad of different authorities with overlapping jurisdictions and widely varying histories and culture, definitions of who 'we' are, interests, incentive and priorities, and powers. The overlapping nature of these jurisdictions means there is often confusion or conflict about who should have, or who has the right to deal with specific issues, so that multiple parties all feel they should be the ones in charge. While on the other hand, some issues fall between multiple stools with no one taking responsibility.

The data revolution is unfolding in this context. It is creating an urgency for new understandings, rules of conduct, and so on, but confusion as to who is best to lead in their creation; triggering 'turf wars' as different parties seek power and influence, creating new arenas and flashpoints of conflict as well as new requirements and opportunities

"There needs to be a framework of common principles allowing public and private use of data across multiple jurisdictions. To achieve this, first there has to be global collaboration around a universally agreed set of standards."

Hong Kong workshop.

Problems and Dilemmas:

- When is it necessary/desirable for data to flow across national borders?
- What different rules should be applied to different types of data (e.g. personal, non-personal), different circumstances and use cases?
- Which bodies, at what level (local, regional, global), are best placed to take the lead on this?
- How to ensure a) their legitimacy in the eyes of key stakeholders, and b) their effectiveness?
- How to address key stakeholders' concerns (e.g. the dangers of a new 'data imperialism', the risks that constrained data flows could undermine innovation and economic prosperity)?
- How can countries ensure that they benefit from the data they produce?
- Do new innovations around AI and Machine Learning need a different form of governance and regulatory approach?

What We Heard

In workshops around the world, we heard the same basic refrain. Data has thrown up many new issues, and policy makers and regulators need to catch up. We heard calls for more regulatory action wherever we went. Likewise, the need for greater collaboration and coordination between government and industry. But there was no clear consensus as to who should, or is best placed to, address these challenges, and at what level: 'local' (i.e. national), regional (e.g. EU), or via some global body?

Various solutions were explored. They fell broadly into three different options:

- Global regulatory body
- Regional regulatory bodies: America, the European Union and a China-centric Asia
- National regulation

In a world of multiple overlapping jurisdictions, a common feeling was that: first, the management of data throws up issues that are so universal in their significance, for example around privacy, ownership, ethics, and 'fair shares' of value, that common solutions need to be found; and second, that no existing organisation is currently able to take this role. As a result, many suggested that we need a higher-level body which could set things straight, for example in terms of creating an ethical framework to establish principles and practices common to all.

The idea first came up in Bangalore, which suggested that *"the creation of a World Data Council may well facilitate international negotiations."* Such a Council could help develop consensus around issues such as *"data sovereignty, and to negotiate cultural differences around privacy, for example."* Some drew comparisons to the efforts made around establishing a collective approach to climate change. In Hong Kong, the suggestion was that there should be *"a framework of common principles allowing public and private use of data across multiple jurisdictions. To achieve this, first there has to be global collaboration around a universally agreed set of standards."* Workshops in Jakarta, Bangkok, Singapore, Mexico, and London all called for *"an independent global data regulation framework (maybe like the G20)."*[7] In Dakar, the call was for *"governments and nations (and perhaps even organisations) to start thinking seriously about the construction of a Data Vision… a strategic template for the use of data and data-driven technologies."* Whichever the favoured approach, it was clear that there is a common appetite for a higher, independent authority to set the standards, define the common ground, and ensure balance and independence.

But who, or which organisations, will be trusted, and able to take the lead on this? While across the discussions, there was a universal desire for 'someone else' to come and sort out how to regulate data, many in our workshops were aware that global alignment may be too hard to achieve, not just because of the scale of the challenge and the agreements required, but also because of mistrust between some governments and multinational corporations. This was particularly evident across Africa, India, and in some parts of Asia, but was also recognised in mainland Europe.

The World Economic Forum is just one of several major organisations trying to develop an international, collaborative, global approach, however, few in our workshops felt it would be effective.[8] In Madrid, for example, opinion was that *"dominant Western services, built by Western engineers, reflecting Western values, and built on Western data, will increasingly be seen as either imperialist, irrelevant, or inappropriate in different cultural regions."* Overcoming conflicting political imperatives and competing commercial interests will therefore remain extremely challenging.

Regional Regulation

A more practical option, perhaps, is a regional approach to data regulation. Regional bodies can deal with these complex issues more easily in a local cultural and political context. In Europe, the EU is already supporting new doctrines that are producing regional rules on privacy, data, and espionage. In Pretoria, it was suggested that a pan-African solution to data regulation could work; *"ideally this should emerge as a regional set of standards rather than just a local one, as this would both help to improve impact and prevent individual governments from increasingly using data regulation to drive top down state control of very powerful individual data sets."*[9]

Many we spoke to are keen to learn from others. For example, participants in both Asia and Africa are watching the progress of the EU's GDPR regulation with interest, and may well support similar measures. *"GDPR will change the data landscape in Nigeria, and bring in new standards"* It is not only Europe that is showing leadership here. China's economic clout and growing influence across Asia and Africa may mean that there is a swing towards their walled garden strategy. It will be interesting to see which will ultimately dominate.

`

Again and again across Africa, we heard that *"the liberal economy or capitalist / Western society currently has a stranglehold on the poorest countries,"*[10] and that *"African data should stay on African servers."*[11] The rationale behind this is so that local data can be more easily accessed and used to benefit the local economy, but also to prevent (largely US) multinationals from extracting the value of African data for themselves. Preserving cultural data was specifically prioritised in Kenya and Nigeria - *"cultural data is an asset store, and this should be licensed – it should be seen as intellectual property."* [12] In Dakar, there was a call for *"data to be used in the national interest, not simply for the benefit of international companies."* In a fast-growing continent, which has already had bitter experience of exploitation by the West, there is little appetite to allow data to become yet another resource which is extracted for another country's profit.

National Regulation

The pros and cons of national regulation were widely discussed and often seen through the lenses of data sovereignty and data localisation, both of which restrict the flow of data across borders. Data sovereignty makes data subject to the laws and governance structures within the nation it is collected, and data localisation restricts data flows across borders by either mandating companies to keep data within a certain jurisdiction, or by imposing additional requirements before it can be transferred abroad. The objectives behind these restrictions can be diverse, and include privacy, cybersecurity, national security, public order, law enforcement, taxation, and industrial development, amongst others. Both approaches appeal to a growing sense of national identity, and support for them is gaining traction in a number of markets we visited, particularly in Africa and Asia.

In highly populated nations such as China and India, there was a view that confining access to national data will facilitate economic growth, build or protect political power, and increase local innovation. In Africa, this view was combined with a strong sense that there is a need to stop *"expatriate organisations grabbing the opportunity" and protect citizens from "data colonisation."*[13] Coincidently, in Europe, although there is a general desire for open data flows, there is also a sense that this has to be carefully balanced against the principle of privacy as a human right.

Proponents of cross-border data flows argue that local legislation undermines free trade by adding onerous and expensive obligations for businesses. These include building, operating, and maintaining data centres in multiple countries, as well as creating and updating separate data sets – even if they are a mirror of those held elsewhere. Add to that the inconvenience of having to go through a number of regulatory approvals to either operate in a market or comply with specific sector rules, and it's clear, they argue, that this restricts opportunity.[14] A 2016 report suggested that the effects of liberalising existing measures could add an estimated 8 billion euros per year to the European economy alone.[15] In emerging economies, some felt that the continued imposition of localisation measures will not only impact economic growth, but they will also have a negative impact on social development. In Dakar, it was observed that *"protectionism and boarded approaches to data could lead to a stifling of innovation, social uprising, mistrust in the potential for data to do good, suppression of whole segments of the world population, and large-scale state corruption."* Others pointed out that localisation potentially weakens national security – the more data centres there are, the more opportunities hackers have to target.

Keeping up with and capitalising on the growth and use of data will not be possible without the growing pains of adjusting regulation to account for this expansion. Looking ahead, it is clear that new techniques and legal constructs must be devised to ensure that we are able to extract value from data, while continuing to protect individuals' rights and acknowledging cultural differences. Quite how to achieve this in an effective and beneficial way is not quite so obvious.

3.5 Trust and Trustworthiness

Organisations seek to build trust in data use. This is increasingly about being more 'trustworthy', which is focused on being truthful and more transparent.

In the workshops around the world, there was a widespread sense that very few organisations, if any, can be trusted with data. Indeed, just as increasing levels of trust are needed, apart from some nations where trust in government remains high, the sense from most discussions was that levels of trust are in decline. The emerging challenge for organisations, policy makers, and regulators is, what does it take to demonstrate trustworthiness? On what basis can/should organisations be trusted with data?

Context

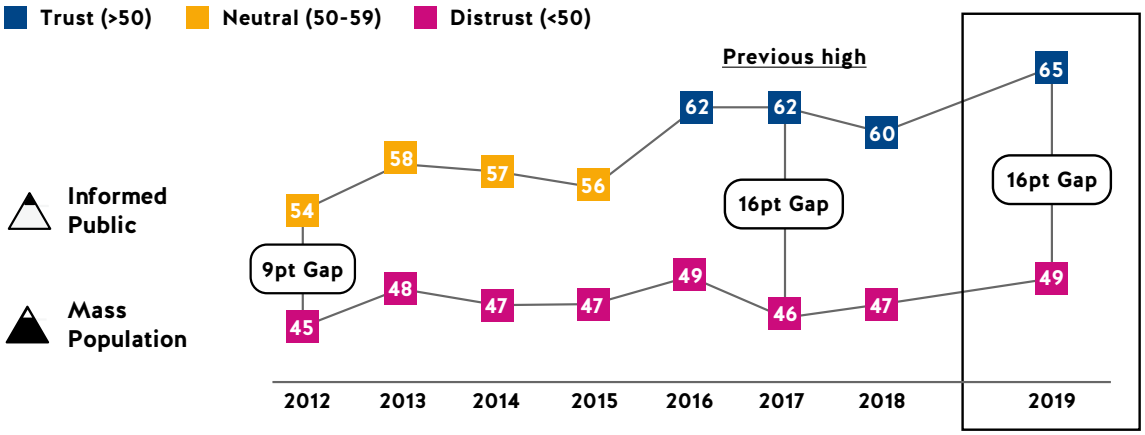
Trust is an economically potent force. When people trust each other, the costs of doing business fall (as less time and effort is spent negotiating, manoeuvring, strategising, monitoring, policing, and enforcing), while opportunities open up, because people are more willing to work and co-operate with one another, including sharing data. Likewise, low trust environments tend to create high operating costs (because of all that time effort invested in negotiating, manoeuvring, strategising, monitoring, policing, and enforcing), while opportunities close down as fewer people are prepared to risk working with others, or for example, to share data with them.

"As concern around security continues and the confidence of African developers increases, there is growing appetite for Ivorians to look after the data they produce and become less dependent on western (or other) nations."

Abidjan workshop

Globally, our workshops took place at a highly particular time: the Cambridge Analytica scandal was unfolding with clear impacts on the degree to which users trusted, not only Facebook with their personal data, but also organisations more widely, as questions were raised about the tech sector as whole. As one student put it, Facebook, Amazon, and Uber are all *"brands that we trust less than we used to."*[16]

The 2019 Edelman report found there is a wide gap between the more trusting informed public and the far-more-sceptical mass population, marking a return to record highs of trust inequality. The phenomenon fuelling this divide was a pronounced rise in trust among the informed public. Markets such as the U.S., UK, Canada, South Korea and Hong Kong saw trust gains of 12 points or more among the informed public. In 18 markets, there is now a double-digit trust gap between the informed public and the mass population



Edelman Trust Barometer (2019)

The Trust Divide: There is a 16-point gap between the more trusting informed public and the far more sceptical mass population, marking a return to record highs of trust inequality

This specific context added another layer of controversy to an issue which is already extremely complex. When it comes to trust, there are many dimensions to consider, such as:

- **Trust in who?** Are we talking about trusting big businesses, small businesses, national governments, supra-national organisations, or citizens? Each of these has different relationships with each other. Whether or not customers trust companies, or citizens trust governments may throw up very different issues and dynamics to regulators trusting/not trusting global companies, or global companies trusting/not trusting politicians.
- **Trust to do what?** We may have 100% trust in someone's capability and competence, but 0% in their motives, or vice versa. There may be multiple boundaries, where we trust a party within a certain range of constraints, but not beyond them.

Within this context, what it takes to earn and keep trust can differ greatly from situation to situation. Further complications arise from the dynamics of how trust works.

One of these complications is the relationship between trust and transparency. If one party isn't aware of another party's actions, their trust levels may be high, but misplaced. In such cases of 'ignorance is bliss', trust levels can fall precipitately as people are shocked to discover the truth. A climate of mistrust and suspicion can then set in, as the pendulum swings the other way, so that even good, trustworthy actors are not given the benefit of the doubt.

A common, but mistaken, assumption is that changing levels of trust translates directly into changing degrees of behaviour - for example, willingness to share information. However, multiple factors can intervene to break this connection. For example, one party might not trust another, but still feel they have to share information, because otherwise they would forfeit access to a service. In such circumstances, actors that are not trusted (and who may indeed be untrustworthy) are not directly 'punished' for not being trusted. The proliferation of new technologies such as the Internet of Things (IOT) and Artificial Intelligence (AI), may mean that for simple operational reasons, whether they like it or not, citizens will be obliged to 'trust' more.

"We need to recognise that data is not truth; it just presents information in different ways. We must learn to recognise bias or lose our freedom of choice."

Madrid workshop

Sometimes trust is bilateral: it's all to do with whether Party A trusts Party B to do something specific. But sometimes it's general: for example, a sense that 'no one out there can be trusted'. These different dynamics generate different behaviours. Levels of bilateral trust can influence whether and how two parties deal with each other. A general sense that 'no one can be trusted' is more likely to increase pressure for 'system-wide' political or regulatory interventions.

Issues and questions such as these came up time and time again in our workshops around the world. For example, there was widespread suspicion of the motives of some Big Tech companies and their desires to monetise data (Edelman's 2019 Trust Barometer shows that more than 60 percent of respondents, globally, believe "tech companies have too much power and won't prioritise our welfare over their profits").

There were also strong differences of opinion as to who is trustworthy: some cultures trust 'government', but not 'big business'; in other cultures, most noticeably in the US, it is the opposite. This is in stark contrast to attitudes in some parts of Asia, particularly Japan and Singapore, where there is confidence that the majority of government operates in the best interest of its citizens - but less confidence in business to behave in a similar way. The same is true in Canada and Scandinavia. Across Africa there was widespread acceptance that corruption is rife – both in government and in the private sector; trust there is effectively absent. (One issue this throws up, as we'll see later, is that in Africa, it's common for many individuals to lie when asked for data, creating a significant knock-on effect relating to the trustworthiness of data that is collected.)

What We Heard

Although there was widespread excitement about the way data is transforming society, and recognition of the multiple benefits this brings, some in our workshops expressed caution. There were fears that the mere fact that data is becoming so ubiquitous means that we will trust it too much and fail to question its accuracy or its provenance. *"We need to recognise that data is not truth; it just presents information in different ways. We must learn to recognise bias or lose our freedom of choice."* This was the view in Madrid, where participants argued that the issue is, in a way, "over-trust," as there is a growing disconnect between our dependence on data to manage our lives and our understanding of the ways it can be interpreted. They suggested that the public risks becoming increasingly vulnerable to exploitation both by political and commercial actors; *"increasingly data will be used to control emotions, particularly amongst the young and the susceptible. Brands and governments will be keen to exploit this, to exercise new ways of influencing consumers."*[17]

"Low levels of trust in government, institutions, and Big Tech, devalues data by making databases unreliable. Citizens are choosing not to share accurate information."

Washing DC workshop

This is all well and good if organisations behave responsibly, but if there is a *"trust Chernobyl"* [18] trust between is broken , the consequence may well mean that people are no longer prepared to share their personal data and are less likely to believe the information they receive from government or other organisations. *"It will be interesting to see to what extent we allow our intimacy to be breached (health, financial, personal information)."* [19]

In general, our conversations around trust were divided in two ways; trust in the management and control of data, and trust in the accuracy of data.

Who can we Trust?

In Madrid, it was observed that our increasing familiarity with technology and growing confidence in our ability to access data is re-shaping how we trust – rather than refer to an expert, for example, we use crowd-sourced data to make a broad range of decisions, from where to eat, to treatment recommendations. At the same time, the popularity of social networks has changed who we trust. *"We have seen the transition of power from nations to corporates, and now it is from corporates to the people."* Certainly, throughout our conversations there was a sense that trust has shifted to greater confidence in peer groups or communities, rather than in traditional institutions or in those of a supposedly superior status. Many who are searching for reliable alternatives to traditional trusted sources of news and information are going online to use social media and a network of "friends" or opinion-sharing communities to find what they believe to be true.

Cultural differences are also important when considering who to trust. In Abidjan, lack of trust in the intentions of Western organisations is galvanising support for the Communauté Economique des États de l'Afrique de l'Ouest (CEDEAO). This has coincided with increased confidence in the ability of African technology skills, *"As concern around security continues and confidence of African developers increases, there is growing appetite for Ivorians to look after the data they produce and become less dependent on Western (or other) nations."*

"Big data and AI provide a huge opportunity for intended and unintended discrimination."

Bangalore workshop

Inaccurate Data

Many workshops felt that trust in data that is publicly available and free to use is declining, because it is increasingly difficult to discern if the information that we are presented with is, in fact, accurate. This is true both for government data and also for information received on social media. There was acknowledgement that distinguishing truth on social media channels is particularly challenging, as it is often difficult to identify the original source for a post or news item. Given this, the recommendation from Hong Kong was that citizens need to become more adept at understanding what is factual and what is not; *"there is a need to recognise that data is not truth, it just presents information in different ways and we must learn to recognise the bias, or lose our freedom of choice."* Failure to ensure citizens have the sufficient skills to distinguish fact from fiction has the potential to lead to a breakdown in trust, and could potentially lead to disturbance and even civil unrest; *"there is a feedback loop – fake data leads to low trust leads to fake data. There are diminishing returns, and trust needs to be maintained in order to ensure a safe and successful society."*

The potential negative feedback between lack of trust in government and government's subsequent ability to provide trustworthy data was highlighted in a Washington DC discussion of people deliberately providing false information. The example given was about research into US Census data, which suggests that around 20% of the information given is false, because citizens do not trust government not to use the data against them. A comment was: *"low levels of trust in government, institutions, and Big Tech, devalues data by making databases unreliable. Citizens are choosing not to share accurate information."*

We heard the same in Lagos, where we were told that such is the level of distrust in both the national government and the private sector, that citizens are unwilling to share their personal data with anyone – this in turn renders government statistics so inaccurate that they are rendered almost useless for meaningful analysis. One suggested solution to this was to implement robust regulation around the collection and use of public data. *"Improved data policies will improve trust in government – currently there is limited trust, because there is limited accountability."* [20] However, certainly in Nigeria, there was little hope that this could be implemented any time soon.

Concern was also expressed on growing reliance on AI, especially relating to the delivery of government services. Workshop participants were particularly concerned about programmers' ability to exclude bias in the selection of data used to train AI, or indeed identify it quickly should it occur. In Bangalore for example, it was felt that *"Big data and AI provide a huge opportunity for intended and unintended discrimination."* In Johannesburg the view was that if pubic concerns around data bias grows, there is a chance that they will no longer trust the products and services that are delivered, and certainly would not wish to participate in sharing their personal data. To address this, it was suggested that data should be labelled with, *"data dignity metrics," which could be used to measure and monitor the use of data for the common good, while maintaining the "dignity" (appropriate levels of privacy, for example) of individuals."*

Irresponsible Use of Data

The main actors in the data-driven economy, large tech firms and governments, were both widely criticised in our workshops. Time and again we heard discussions on the way that the many technology firms, particularly social media companies, exploit the data that we share, with little regard to personal safety or privacy. Few believed that lessons had been learned from the Cambridge Analytica scandal and that in the future we could be more confident in the organisations which have control over our personal data. In Hong Kong, it was observed that *"as understanding of the current Big Tech companies grows, expect more disagreement about their current business models."* In Bogota, they said, *"manipulation of the people will continue."*[21]

In addition, there was clear frustration with what was seen as a lack of leadership within the technology sector. In London, the perception was that it is this that has generated the real crisis in trust; *"it's not a crisis of trust – more a crisis of leadership. We can't impose trust downwards."*[22] The conversation went on to focus on the importance of trustworthy behaviour – and the need to make it accountable, *"…it's about confirmation, not trust."* Similar views were expressed in Singapore and Toronto.

Sometimes we heard debates about national security and the need to protect citizens from bad actors. This mistrust can seep into many, perhaps unexpected, areas. For example, participants in Singapore and South Africa both stated that one reason why DNA data is not shared with the US, is national security.

A number of different alternatives were identified, which could help rebuild trust in the use of data and data organisations. These are some of the solutions:

- **Greater transparency.** In Dakar, it was agreed that the public revelations around data lapses and the exploitation of personal data by some technology companies, have demonstrated a failure of self-regulation. The consequence of this is that *"tech companies will be obliged to be transparent about the data they collect, and the uses they make of it. This will be driven by increasing consumer pressure, and a competitive environment in which transparency and responsible data use become a point of differentiation."* Others agreed; from Madrid to Hong Kong, Singapore to Bogota, it was felt that social media companies in particular, should be more proactive in helping to distinguish between truth and inaccuracies on their platforms. There were numerous examples about how misinformation has influenced behaviours in both rich and poor countries, including overt bias in elections and online scams.

Full transparency may not, however, be a silver bullet; too much information can also be confusing. In London, it was observed that *"full transparency is only really needed if trust is absent. It certainly does not mean a requirement to share mountains of information as a means to ensure ethical behaviour."* In Frankfurt, the view was that as private citizens become aware of just how much personal data about them is being accumulated and traded, the demand for greater transparency will grow, and regulation will likely follow; *"if there is no transparency, it will block acceptance of online services."*

- **More Accountability:** There was universal consensus that greater accountability could increase trust, but there were differences in opinion about how this could be achieved. Ensuring that government data is accurate was of particular importance in the Washington DC, Tokyo, Singapore, Lagos, and Copenhagen workshops. In Lagos, the view was that the only way to achieve this is through open multi-party collaboration. *"Improved data policies agreed by multiple stakeholders will improve trust in government - currently there is little trust in government use of data, because there is limited accountability."* A suggestion from Denmark was that there would be greater public confidence in public institutions if there was *"Data NATO or a UN organisation, which could develop and oversee guidelines, codes of conduct, and shared standards."*

- **Technical solutions:** Some in the workshops suggested that new technologies such as blockchain may go part of the way to providing a reliable safeguard against abuse, and therefore help rebuild trust. In Tokyo, the view was that it *"spreads responsibility and increases trust in the system."* Creating a distributed, immutable record of information — which can never be deleted or modified — would at least provide a degree of transparency. Data could be recorded and distributed in a more transparent fashion, and could not be changed without amending all records across most users. Content creators could use distribution channels that guarantee that their content does not get altered, filtered, or blocked by a third party. Equally, a distribution channel leveraging blockchain could make it more difficult to censor and limit access to information.

"Improved data policies agreed by multiple stakeholders will improve trust in government - currently there is little trust in government use of data, because there is limited accountability."

Lagos workshop

- **Consumer influence:** In Bangalore, it was felt that regulation has been too slow to control the behaviour of some of the technology companies in their exploitation of data. Therefore, they suggested that public opinion is more likely to drive change in advance of any regulatory response to the decline in trust. *"The public response to unethical behaviour often happens before the law is enforced, or indeed appropriate regulation created,"* and from a personal data perspective, *"growing public understanding of potential harm to the individual will lead to increasing demands for better rights and greater accountability."* Those in Copenhagen built on this idea, suggesting that encouraging greater citizen involvement in monitoring the use and accuracy of data might help build trust. *"Is there, maybe, a role for something like Wikipedia in the mix here?"*[23]

- **Digital literacy:** Many felt that greater public education around the use of personal data would both help to build public trust in open data for public services, and give citizens sufficient skills to be able to identify when that trust could be misplaced. In Santiago, the hope was that recognition of this, alongside some hefty fines, would moderate corporate behaviour; *"when the public is more involved, accountability becomes "horizontal" rather than vertical."* As awareness grows, the ability to *"watch the watcher" and "critically understand"* will mean that large organisations of all kinds will temper their actions and take greater account of what is considered to be acceptable – both off and online.

- **Generational shift:** There was a recognition that trust in technology and the data that it delivers, is dependent on generational expectation. Some, for example, suggest that millennials are much more likely to be data-savvy around security and privacy and so on, than older generations, and may be less likely to be concerned about it. *"Gen X is the last pre-digital generation – the generation after will have better understanding of the importance of management and control."*[24] However, some fear that the next generation will be so dependent on technology, that "any data will be believed to be fact, and its veracity will not be questioned.[25] But we also heard the opposite view; *"in ten years' time, things will be more balanced. We are currently in a transitional phase and in a state of flux – people were scared of the car when it was first invented."*

Looking Forward

Throughout all our workshop discussions, it is clear that we are at a point of transition. Technology innovations, powered in the main by a select number of hugely powerful global organisations, several of which are not widely known, are triggering dramatic changes across all sectors of society, and influencing how millions of people live their lives. Often these changes are for the better, but not always. Such is the momentum, that many citizens feel that these changes are being 'done to them', whether they like it or not. This makes trust all the more important.

Building trust is not one, single challenge. It is multi-faceted. Data-based systems rely on the accuracy of the data that is fed into them. They only work effectively if enough people trust them to share accurate data, and believe in the accuracy of the information that they get in return. When incidents occur which reveal irregularities, corruption, or incompetence, trust is damaged, making individuals less confident about the benefits of participation. The risk is that growing numbers decide to scale back participation, or provide inaccurate data. If enough people do this, the system fails.

This aspect of trust is primarily technical - of creating systems that are fit for purpose. A second, more complicated and controversial dimension of trust relates to the motives and intentions of different stakeholders. The challenge for those organisations and institutions leading the transition to a data-driven economy and society, is to demonstrate that they are *trustworthy.*

Being trustworthy is not the same as being trusted. It means that organisations accept they should be held to account; that they demonstrate that they have 'good intentions'; that ethics are not something to talk about for PR purposes, but actually shape what decisions are made and how they are implemented. Greater transparency helps, but is not the only answer, particularly when trust in corporates is at a low ebb. A robust regulatory framework, either developed globally or regionally, would do much to create standards, along with checks and balances to curtail the power of the large corporates, which many we spoke to felt are still largely unaccountable. Individuals also have a role to play by becoming more aware of their rights and responsibilities online. If successful, and we create a lasting, robust, and trustworthy system, then the next generation can only benefit from its potent force.

3.6 Shared Language

People are unclear on where the value in data comes from or what form it takes. A key step is a common language about data that provides clarity of terms

Mounting discussion in the media and politics about data, its ownership, use, and its value, highlights a lack of consensus around how to describe fundamental concepts. In government, business, and civil society, this undermines the ability to build alignment and develop robust ways forward. A simple, shared, accessible terminology is increasingly being called for, in order to establish a common understanding of what the key issues are, and what options are available to address them. This lack of a common language and understanding is a major impediment to attempts to build cooperative or regulatory endeavours. Without it, the possibility of reaching an agreement or deciding on an appropriate course of action is limited, if not impossible. Given this, there was widespread consensus in our workshops that time and energy must be spent to define and agree terms around the use and value of data.

Problems and Dilemmas:

- Is it possible to create a 'common language' where, across the world, key stakeholders all use the same terms and definitions to describe what is happening with data?
- Is it possible to create a shared understanding of what the issues and options are, even if there are disagreements as to how important these issues are, or what the most desirable courses of action are?
- If it is not possible to create such a common language and shared understanding, how to advance debate and understanding of the multiple issues being raised by the emergence of a data-driven economy?
- If it is possible to create this common language and shared understanding, what is the best means of doing so, and who should lead/take responsibility for this quest?

What We Heard

Beyond the varied metaphors for data (sunshine in Tokyo, the periodic table in Singapore, religion in Madrid), myriad views on the definition of key issues, such as informed consent or digital literacy, were expressed everywhere. In the vast majority of workshops, the lack of agreement around precise, common terms for the key elements of the digital world was highlighted as a major concern. These were not just at a holistic cross-society and cross-industry level, but also within individual sectors. For example, our preceding 12 discussions on the future of patient data in 2017/18 highlighted how little is understood by professionals within healthcare on the differences between aggregated and anonymised data, ownership and control; machine learning and artificial intelligence (AI), and artificial general intelligence (AGI); as well as between data bias and data quality. Other sector-based discussions on automotive data in the UK, US, and Germany showed similar different interpretations.

In our workshops, examples such as these were all repeated in varied locations. Different definitions were used for data sovereignty and data localisation, between a data tax and digital taxation, and between data literacy and digital literacy – even by regulators. There was widespread acknowledgement of this and resounding support for the need to develop a global, cross-sector agreement for the terminology of data in multiple locations around, including Jakarta, Bangkok, Dakar, Mexico City, Toronto, and even Washington DC. Those in Singapore voiced the view of many, when they suggested that the rationale for this is to deliver *"a more clearly articulated government data strategy to enable community-driven initiatives which have wide public benefit."*

Language is not only about policy, however. It is about understanding. Without an agreed language around data use, it is difficult to see how populations can become digitally literate. Concerns about this sparked a total of nineteen separate discussions on Digital Literacy during the programme. Irrespective of geography, age, employment, or method, the message is clear; *"the divide between the technology literate and the technology illiterate will be a huge challenge, and will have grave consequences if not addressed."*[26] The reasons for this are not hard to uncover. As access to connectivity increases apace, and governments increasingly rely on data to connect with their citizens, managing cyber risks, ensuring individuals have the skills necessary to engage with the state, and building a workforce fit for a digital economy, are all priority areas. Failure to address digital literacy will have consequences, not least widening the digital divide, creating skills shortages, and extracting value from data. But, how will governments be able to extend a digital literacy programme if the lack of clarity around the language of data remains unresolved?

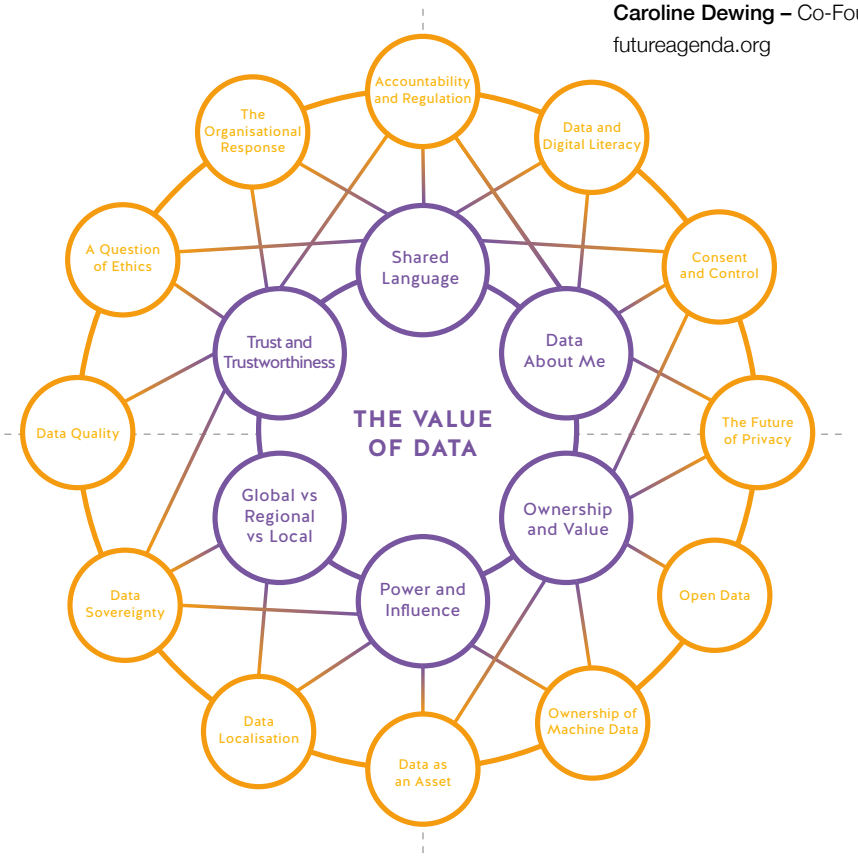
> "The divide between the technology literate and the technology illiterate will be a huge challenge, and will have grave consequences if not addressed."
>
> Tokyo workshop

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