WPP OPEN

DATA 2030
In these strange times it's hard to think about 2030; we don't know what the world will look like next week let alone in ten years.

Yet from a data perspective there are clear trajectories around its volume, variety, value, the role it plays in business growth, the skills required to optimise and the expectation from individuals on how ‘their’ data is collected and processed.

This paper sets out to explore some of these trends. It offers a review of where we believe we are heading but by no means promises that we know the future any better than anyone else. Our quest is to have as many people think about the power of data and be as excited about it as we are.

Let us know what you think.

Di Mayze
Global Head of Data & AI, WPP

“Our quest is to have as many people think about the power of data and be as excited about it as we are.

Di Mayze
INTRODUCTION

We are moving from the digital revolution to a data revolution; creating a data universe in which data infuses our everyday lives, the decisions we make and the way we behave. Until now, data has been a recorder and interpreter of the world, but by 2030 data will define the relationships between things, places and people, and decide much of our experience. Data will not yet be in charge, but it will be a constant influence.

This report explores the world of data in 2030 through four themes

01 The Dataverse
Data volumes will explode and data will track and influence everyday interactions, but the cost of storing and processing that data will be a major factor in how we use it.

02 Data the Decision Maker
Data will be used as a decision maker, although for the most part it will be restricted to augmenting human decisions rather than replacing them. Its new role will change how we process data and how we apply it.

03 Rebalancing Control
Regulation and social attitudes will transform the market in personal data and how we manage access to it.

04 The Professionalisation of Data
Regulation, energy costs and environmental taxes, as well as reduced access to personal data, will combine to create a more professional use of data.
Data volumes, already enormous, are set to explode. The IDC predicts 175 zettabytes of new data will be created annually by 2025, up from 33 zettabytes in 2018. 5G and the Internet of Things it will facilitate is imminent. With devices generating and consuming data directly from other devices, the human factor will no longer moderate how much data is produced. Machine learning will be commonplace, spawning its own data from the data it ingests. Africa and Asia will be fully online and those industries which are only tentatively digitising in 2020, such as education, agriculture, healthcare, and regulation, will have been transformed. By 2025 the IDC estimates humans will interact with data every 18 seconds and the infrastructure supporting this vast dataverse will need to evolve too. In the last 40 years we’ve already seen fibre optic loads increase from 45 megabits-per-second to the more than 1 terabit-per-second performance demonstrated by Nokia Bell Labs in March 2020. The speed of light, however, is finite so we will see an increase in the number of data pipes as well as continuing improvements in how much data can be carried.

---


---

Universe – Latin noun universum “all things, the whole world”, adjective universus “relating to all, turned into one” from unus “one” and versus “to turn, to be turned”

Dataverse – all things turned into data.
SUSTAINABILITY

As data volumes increase, so will the energy consumption required to store and process the data. The question is by how much. One best case scenario is that ICT will consume 8% of the world’s electricity demand by 2030, compared to 2% in 2020\(^3\). By then we expect standard global measures will exist to accurately assess such energy consumption.

Efforts to restrict the environmental cost of hosting and processing data will have spread beyond the few leaders such as Switch, Google and Facebook and be common across the globe\(^4\), driven as much by a need to contain the financial costs of the data explosion as by climate change regulation. Efficient hardware, 100% renewable energy use and almost universal use of centralised cloud computing will have transformed data hosting into a notably efficient industry.

By 2030 climate change concerns will have led to widespread use of externality taxes. Those companies which are not in the data storage business will be reluctant to host their own data and incur the new environmental taxes associated with energy consumption. Individuals will resist changes to their data consumption habits and governments will not be banning HD colour cameras on smartphones\(^5\), but they but will factor the efficiency of devices into their purchase decisions. By 2030 we will be well on our way to achieving the 6SPWh savings in energy use associated with digital smart meters and lighting\(^6\).

---

By 2030 data will be considered a decision maker, an influencer and an input to our actions, not just a record of what has happened. This is a long emerging trend and today’s challenge of ‘too much data, too little insight’ will be a thing of the past. Data will envelop our everyday experience.

**DATA PROCESSING**

With more restricted access to personal data (see our predictions in Rebalancing Control) and climate change fears driving up the costs of data storage, most companies will follow a data minimisation policy in which they use smaller sets of data with much greater care and to much greater effect. In the marketing industry, networks such as WPP, which have followed a policy of using data well rather than focusing on collection, will be admired for their foresight and ability to set rather than follow a trend.

The federated and virtualised data solutions already in market (such as Infosum and Blockgraph) will be in common use as organisations see more value in collaboration, both internally and externally. Such solutions will allow data to remain in its place of origin without being moved or replicated. Equally, it will be the norm to purchase machine learning models as a service rather than attempt to do-it-yourself. Such services will be based around business problems rather than model characteristics – for example, a marketing department will subscribe to a model that delivers acquisition growth of X%. There will be less need to maintain large volumes of data as organisations embrace collaboration and models as a service.

Automation and increasing data collaboration will create a pressing need for industry wide standards in interoperability and classification. Organisations will actively contribute to and comply with these standards within and across industries. The first signs of this sort of activity can be seen in the US healthcare interoperability initiative. This was begun in 2018, with participants such as Microsoft, Amazon, Google, IBM, Oracle and Salesforce.

Data availability and the speed of data decision making will become critical. We see this already in car safety features where data must be processed in real time. This need for on the fly decision making means that by 2030 data streaming will be a necessary part of the DataVerse. But edge and endpoint computing will also be critical as our data volumes increase and tolerance for latency decreases. Edge computing means that we will be able to process data locally on the same device that is collecting the data. So, for example, autonomous cars won’t be streaming all their data back to a central processing site, but rather will analyse and apply it on site.
The ethical dilemma of bestowing moral responsibilities on robots calls for rigorous safety and preventative measures that are fail-safe, or the threats are too significant to risk.

Hawking, Musk, Gates 2015

In 2030, general AI will still be a distant goal, but machine learning models will no longer be remarkable. Unexplainable machine learning models, however, will be a thing of the past as society demands to know how and why decisions have been reached. The warnings of its early technology leaders will have been taken seriously:

In spite of EU attempts to enforce it in 2020, human-in-the-loop oversight of these models will have been abandoned, replaced by more cost-effective automated verification and explanation. There will still, nevertheless, be reluctance to remove all human interpretation from data analysis and decision making. Machine learning will be most commonly used for mass administrative or low criticality data processing but human analysis (albeit machine assisted) will have become the gold standard.

With increased regulation and less access to personal data, software companies will not be able to tap into to the vast volumes seen in the years leading up to 2020. As a result, they will be increasingly dependent on synthetic data as an input to their models and the manufacturers of synthetic data will have become a critical component of the data industry. In addition, models will have evolved to reach decisions more efficiently with less data.

---

8 For example, NASA’s current preference is for human analysis. It only uses software for basic processing.
Our experience of digital will be immersive rather than interactive and data visualisations will no longer be confined to 2D dashboards and PowerPoint. Data visualisers will be exploring how to manifest data using other senses such as hearing and touch. These experiments will first be seen in entertainment and storytelling, but eventually will seep into more critical decision-making applications.

These immersive manifestations of data will sometimes be personalised, but more often they will be experienced at a social level. They will need to meet the needs of humans as a social species as well as the practical limitations of our daily lives: the digital display on our kitchen or office wall will need to inform everyone that is viewing it, not just one person.

Data and marketing creativity

In the marketing world we will no longer consider creativity and data as separate functions but will instead have ‘informed creatives’ who use data as a guide and an inspiration. They will use the different data signals to build up a picture of who has the affinity, interest and motivation to buy from their brand, and this knowledge will influence the creative work just as market research does today.

Smart marketers across the industry will be inspired by data as much as they are by blog posts and TED talks today. They will move on from talking about the mechanical benefits of targeting and optimisation. Instead they will use the stories data tells us about markets and customers to describe the opportunities our clients have and the challenges they face.

At the same time, consumers will take a much more central role in the creation of their customer experience. Google’s predictions for Hypertelling give a glimpse into a world where an individual’s access to data and devices as varied as music, images, videos, VR and location mean he will be able to curate his own interpretation of an advert. This will of course be merely an appearance of freedom: marketers will still want to limit choices to retain control of the brand and message.

The Use of Data

Data visualisation

Although most rudimentary data decisions will be automated and therefore unseen, the importance of data visualisation will be fully acknowledged. For most of us, data does not ‘exist’ unless it is visualised. Visualisations will be required to entertain and educate, tell stories, convey insight and aid decision making. In 2030 visualisations will no longer be tethered to personal devices and augmented reality will be commonplace. The data distractions of Apple’s windshield will be a reality – although one hopes only for autonomous cars.
**The impact on autonomy**

So far, our predictions have been largely positive, but the shift from data as an interpreter to data as a decision maker will have an insidious effect on human autonomy. We will live in a world of constant surveillance; one in which a disparate set of commercial organisations encourages individuals to adapt their behaviour to optimise their experience and obtain rewards and discounts.

It will be argued that this optimisation of behaviour is positive for society: for example, car owners will drive more safely to reduce their insurance premiums, knowing that they are being monitored at all times not just by the occasional speed camera. But autonomy is a central element of human identity and as the history of any authoritarian regime shows, its loss leads to resentment and sometimes revolution. Shoshana Zuboff describes this new world as surveillance capitalism “a new global architecture of data capture and analysis that produces rewards and punishments aimed at modifying and commoditizing behavior for profit”. It is unlikely, however, that in 2030 individuals will resist, or even notice, a system that rewards them for conformity:

“There is no putting the genie back in the bottle ... Everyone will expect to be tracked and monitored, since the advantages, in terms of convenience, safety, and services, will be so great...continuous monitoring will be the norm”

Pew Research, 2014

---

12 Zuboff, Big other: surveillance capitalism and the prospects of an information civilization, 2015
03 REBALANCING CONTROL

Between 2020 and 2030 the imbalances in data processing capability between companies and between individuals and companies, will come to a head. EU and US competition laws will have broken up data monopolies, helped by the multitude of data protection regulations dismantling the personal data market that has sustained them.

At the same time, individual awareness of how their personal data is collected and used and their tolerance for this use will have changed. This will be part prompted by the increase in biometric (retinal, fingerprint, facial) and emotional data (motivations, emotions). Until now biometric data has been largely restricted to discrete security functions such as unlocking phones and passport control, but its use is becoming more widespread. For instance, the UK police force is already trialing mass facial recognition and in Seoul and Tokyo fast food restaurants use facial recognition to customize food orders. Research indicates increasing discomfort with such data use:

"Consumers have drawn a line in the sand on what information they are and aren’t comfortable being collected and shared about them...and it’s not related to data identifying the individual, rather how personal and intimate the information is.”

WundermanThompson Data, 2019

By 2030 society will no longer tolerate a business model that relies on mass transactions of increasingly sensitive personal data: a quite different system will be in place.

PERSONAL DATA AND CONSENT

Personal data will no longer be the default payment currency for digital services. Organisations will still want to use personal data – profile demographics, behaviour, biometrics, emotions, interests and motivations – but they will now have to convince individuals to part with that data.

By 2030 cookies will be a historical curiosity; a temporary solution to the stateless Web that was badly misused and best forgotten. MAIDs and other tracking activities will be similarly obsolete. As organisations will no longer be able to immediately track online behaviour through cookies, a new customer experience will develop in which the case in favour of sharing one’s data will be built up over the course of a customer journey, much as the case for product purchase is built today. Once that consent is given, organisations will be able to accurately identify users and collect the data they have chosen to share.

The principle of notification and consent will remain, but the onus will shift from the individual having to deny access, to the individual granting access. Dark pattern consent mechanisms will be banned and there will be a more rigorous implementation of the GDPR’s transparency requirement which requires information about how data will be processed to be "concise, transparent, intelligible and easily accessible…using clear and plain language". Organisations will not be allowed to limit basic services such as the use of a website if access is not granted (an interpretation of the GDPR’s Recital 45 which is already used by the UK’s ICO). Instead, organisations will reward those who share their data with value-add services – much like a loyalty scheme. Those organisations will also need to find ways to avoid collecting only the data of low value prospects and customers.

State oversight will also increase as the impact of privacy loss on individuals and society becomes better understood. Many governments will limit how much personal data an individual can share, particularly data that is considered more intrinsic to the self, such as behavioural, emotional and biometric. Other governments may forbid payment for data, taking the view that personal data and the privacy it guards an inalienable right.
**A NEW BUSINESS MODEL FOR ACCESSING DIGITAL SERVICES**

The paid media market will not disappear, but it will be based on views and non-personal contextual information such as weather and location. Organisations will be able to use the personal information they collect from consenting individuals but only to optimise their own experience and engagement channels such as web, app, and messaging. With the loss of advertising revenue, paywalls for content will be the default business model for publishers – a growing trend in 2020 but one that will be dominant by 2030. Social platforms will have become the newest category of publishers as by 2030 governments will hold them responsible for the content they host. Along with digital services such as maps, storage and search, social platforms will also use paywall models. This trend is already emerging in 2020 as social influencers begin to charge for content. Subscription based bundles of digital services included in the provision of internet access will be common. The old libertarian dream of a free internet might be over, but so will the wild west of fake news and trolling so characteristic of the early 21st Century.

**SELF-SOVEREIGN ID AND PROFILE MANAGEMENT**

The data oligarchs whose businesses will have been so disrupted by regulation, will turn their efforts to ID management, but this time as a paid service for individuals. In this field they will compete with other trusted institutions such as banks and government systems. The battle for control of this market will still be underway in 2030 with the inevitable security breaches causing concern and column inches, but the principle of users managing their own ID and access to their profile will be accepted. The current system of every digital offering managing its own set of data profiles cobbled together from unreliable 3rd party, device graphs and tracking data will seem bizarrely inefficient in retrospect.

The switch to self-soverignty will be part motivated by the desire to rebalance control, but the need for the large data processors to find an alternative business model will be the real driver. The side effect will be more accurate personal data. Individuals will want to keep their data up to date so that they can benefit from the added value services organisations will provide them in exchange.

---


20 https://www.voguebusiness.com/ https://www.voguebusiness.com/companies/influencer-paywall-what-it-means-for-fashion-brands “Some charge a monthly fee to become a “Close Friend” on Instagram, while others are trialling WeChat’s new paywalls. In effect, content quality is becoming a focus.”

21 This principle, which is only just emerging in 2020 with initiatives like Sovrin (personal management of digital IDs) and Inrupt/Solid (store and manage access to your own data in your own pod).
Methods for managing access to profiles will still be maturing in 2030. Managing which attributes can be shared and how long the data can be used will necessitate sophisticated data management software and potentially a new data advisory profession similar to financial advisors. Associated industries providing ethical and data security indexes, and ranking the rewards of company data exchange schemes, will be developing as individuals seek advice about whom they share their data with.

One significant question will be whether companies who have been given access to collect personal data must share the data they obtain back to the individual.

The state and the not for profit sector will be notable beneficiaries from this new model. By 2030 we expect individuals to feel duty bound or even compelled to share (anonymized and aggregated) personal data to help national healthcare providers, social and economic policy development. In fact, 2020’s unexpected COVID-19 crisis is accelerating this trend, with the telcomms sector and Google sharing location data in order to help track the spread of the virus. Individuals will also be encouraged to share personal data for more specific social good such as charities they want to support. The devices and apps created to collect this type of data will also share advice and nudge desirable behaviours – the UK’s NHS Couch to 5K app is just an early example. Of course, this is another example of the tension between good intentions and loss of autonomy.
The increase in data volumes and data-influenced decisions gives those using the data more power and more responsibility. Data will be the foundation for how we live, and we will see the data industry rapidly mature over the coming decade, moving on from its buccaneering big data days and settling into a more thoughtful, exploratory and scientific mode. The need to demonstrate data value, comply with regulation, and treat personal data with care, will result in a more professional attitude to the collection and processing of data, and the application of insight.

DATA PROFESSIONALS

The data job market will have transformed by 2030. The number and variety of data specialist jobs will increase and confusion over what makes a data scientist versus a data analyst will have disappeared. At the same time, familiarity with data will become the norm; everyone will have a base understanding of what data can and cannot do, and how to interpret it. Saying ‘I don’t understand data’ in 2030 will be like printing emails in 2010.

Who are the professionals

Data management professionals (the administrators and the architects) will continue to sit within IT organisations, working with data engineers to build and administer the infrastructure and databases.

In the analytical world a class of data scientists will have emerged, who will tell the stories data reveals and seek to answer business questions. They will work with the machine learning specialists who create the tactical solutions – coding the algorithms and running the machine learning models. Data analysts and scientists will no longer find the majority of their time taken up with administrative work. Tedious data cleansing and processing tasks will have been fully automated, leaving these professionals free to focus on extracting insight and value from their data.

A third category of professional, data visualisers, will have far greater prominence by 2030. They will use their expertise to ensure insights will be understood rapidly and easily, whether the user is driving their car, defining public policy or planning their business strategy. Data visualisers will combine creativity, data knowledge, industry expertise and behavioural science to bring data to life. They will also have a greater responsibility to accurately interpret and convey insight – data visualisers will need to be statisticians as well as creatives, and there will be more rigorous scrutiny of the work they produce.
The new CEOs

By 2030 data will be foundational to business decisions and data expertise will be required in every department. As a result we are likely to see the first data professionals rise to CEO positions. Much like finance professionals, they will be seen as employees who thoroughly understand their organisation, its industry and its markets, as well as being individuals whose decisions are guided by information rather than intuition. The Chief Data Officer role will evolve from being a champion of data use to being an overseer of data quality and standards – managing the data professionals embedded in every business department.

DATA INTEGRITY

By 2030 data integrity will become a core element of data professionals’ work and training. Data will be processed to academic standards and professionals will be expected to work with accurate data and statistical rigour. Because organisations will be working with much smaller and more reliable sets of data, accuracy and completeness will become obtainable goals.

Chains of trust will be established whereby oversight or partner or supplier data use will become standard elements of procurement and contract negotiation. In 2030 data auditing will be a common business activity, partially automated but usually overseen by 3rd party specialists. Unlike financial auditing, it will still be an internal affair, but it will be in company interests to publish their results to increase the likelihood of individuals sharing their data.

DATA AS A BUSINESS ASSET

In 2017 a McKinsey report estimated that using data to assess market conditions, improve marketing targeting and implement operational efficiencies, could increase operating profits by approximately 6%.22 In 2030 data will even more valuable, but as personal data becomes less accessible and more expensive to store and process, its contribution to that value will be more rigorously assessed. Companies, especially digital start-ups, will no longer focus their energy in building data moats but will balance their investment alongside more traditional business strategies such as product and service development. Organisations will seek to understand what value personal data actually gives them – in terms of product development, operational savings, marketing, network effects and customer retention – and will use this knowledge to determine their investment in customer data consent activities.

Marketing Data ROI

As it gets harder to obtain personal data, achieving a positive ROI from personalisation – that is the value generated from using personal data versus the cost of processing it – will become more difficult. As a result, the ROI will be more rigorously assessed. Factors such as the cost of data retention versus the value of historical data, and the cost of storage versus the value of each attribute will be scrutinized. Because, however, organisations will be using profile data that has been willingly shared and will only be using this data on their own platforms, attributing its value will no longer be a dark art, but a reliable calculation.

23 Companies seek to protect themselves from disruption and competition by building moats. Personal data can be an effective barrier and is typically referred to as a data moat. Network effects (seen in social platforms where the product becomes more valuable to users as more people use it) and personalisation (where customisation of a product makes it less likely a user will switch to a competitor product) are common examples.
By 2030 the battle between a free for all internet and a monitored environment will be over. Social platforms will block fake news and tweak their algorithms to avoid the recommendation spiral of increasingly extreme content\textsuperscript{24}. In the data world, a tension will emerge between the need to maintain data integrity and the need to avoid discriminatory decision making.

Data can hold a distressingly accurate mirror to society; whether through the bias inherent in the data set being analysed or the bias of those writing the algorithms. In such cases the data outcomes and decisions can be unwanted and extremely damaging to corporate reputations. Recent examples include Apple’s difficulties applying equal credit card limits to its female customers\textsuperscript{25} and US policing and parole solutions which are accused of perpetuating racial bias\textsuperscript{26}. Unable to fix society, organisations will increasingly use tools such as Google’s WhatIf\textsuperscript{27} to change their data and avoid the reputational damage of prejudiced data decisions. In 2030, there will still be debates as to whether this compromises the accuracy and integrity of the data, or whether it avoids the unwelcome prejudices of the past, but the practice will be common. Some will argue that because data decisions are always based on the past, they cannot ‘move on’ as a human would, making a little human nudge necessary. Others will argue that data whitewashing, tweaking algorithms and banning fake news is essentially censorship, and they will question the wisdom of allowing such censorship to become a commercial task.

\textsuperscript{24} New York Times (2018) \url{https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html}
\textsuperscript{26} Smithsonian (2018) \url{https://www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968332/}
\textsuperscript{27} \url{https://pair-code.github.io/what-if-tool/}
HUMAN DATA

Our last prediction is the most heart-warming. In the decade leading up to 2030 the slow data movement will go mainstream. This movement is similar in ethos to the slow food and slow TV movements of the early 21st century. It is a reaction to the increasingly noisy digital world and the speed of modern decision making made possible through digital data and devices. Slow data advocates a love of data for itself; taking time to play with it and discover the beauty of data storytelling. One of its leading proponents; Giorgia Lupi, talks of embracing data’s complexity, personalising data visualisations, being happy with small data and remembering that data is imperfect. She asks us to immerse ourselves in data and take the time to really understand what it’s telling us.

“

We have to reclaim a personal approach to how data is captured, analysed and displayed, proving that subjectivity and context play a big role in understanding even big events and social changes—especially when data is about people

Giorgia Lupi, 2017

CONCLUSION

Data has long allowed us to look back in time and it now allows us to anticipate and even influence the immediate future. As yet, however, no-one can predict the long-term future. And of course, all of the above presupposes a non-quantum world, where bits still rule and decision certainty is required. Quantum computing has been just around the corner for some years now, but perhaps by 2030 it really will have arrived. If so, then all bets are off. The speed of analytics, modelling and processing will have undergone a quantum (sorry) leap. Data decisions will be unseen and will seem anticipatory. Data decisions will replace rather than simply augment human decisions. The application of quantum decisioning in every aspect of our lives will make user interfaces for humans to point and click on seem as obsolete as flint arrowheads.

The ideas above are just some of the many possibilities for the future of data; they are described to entertain and to stimulate cooperative debate. Such critical reasoning will inspire our business so please join the discussion at WPP’s new Data and AI Community site.

“I always avoid prophesying beforehand because it is much better to prophesy after the event has already taken place.”

Winston Churchill