Government
Office for Science

– Citizen data
– Governments
– Regional data systems
– Regulations
– Global trends
– Technologies
– COVID-19
– 2030 scenarios

Evidence and scenarios for global data systems

September 2020

The Future of Citizen Data Systems

Foresight

Ministerial foreword

The world is in the middle of a data revolution. With every passing second, the volume of digital information that flows between businesses, governments and people is expanding rapidly. Data now increasingly underpins our everyday lives: we use it to shop online, to monitor our exercise regimes and personal bank accounts. It has been a crucial weapon in the battle against coronavirus, and has become an increasingly important aspect of international trade and global relations.

This digital transformation represents a huge opportunity for the UK. By harnessing the power of data, we can boost growth and productivity, drive innovation, improve public services and inform a new wave of scientific research. Our upcoming National Data Strategy is an ambitious bid to make the most of this moment, so that data's value can be felt across the entire UK.

But we also need to understand the growing risks associated with the data boom, including cybercrime and disinformation. This report helps us identify those threats, and the government has been careful to reflect its findings in our own National Data Strategy, so that we can protect members of the public while they experience data's many benefits.

At the same time, we can't shape our own data regime in a vacuum. Emerging technologies, socioeconomic shocks, geopolitical changes and global crises – such as the pandemic we are tackling today – all have the ability to significantly alter the facts on the ground. That is why it is crucial that the UK uses its international influence and leadership, both to drive the global attitude to data while ensuring our own strategy is adaptable enough to respond to the inevitable shocks of the future.

We also want members of the public to be active agents in the thriving digital economy, and have confidence and trust in how data – including citizen data – is used. This will be especially important as we transform government's use of data to drive efficiency and improve public services – with a clear understanding that it is our duty to use data to deliver better outcomes.

A successful data system will need to be flexible and react quickly to changes. In future updates, we will lay out the steps that we will take to implement the strategy, and the way that engagement and evidence will have shaped our approach.

The UK is already one of the most competitive digital nations in the world. Through its recommendations and the evidence base it affords us, this report will help us build on that position of strength, fuelling a new era of growth and unlocking benefits for society as a whole.

The Rt Hon John Whittingdale OBE MP

Minister for Media and Data
Department for Digital, Culture, Media and Sport

This is not a statement of government policy

Preface

Information is a powerful and growing force in society. The collection and use of information about individuals and groups of people – citizen data – is accelerating particularly sharply. More than ever, we leave a digital footprint whenever we browse the internet, travel and shop; in virtually every aspect of our daily lives.

Citizen data presents enormous potential value to consumers, businesses and public authorities. In the digital age, a range of information about citizens can now be used far more easily for a wider set of purposes, and for purposes which were not initially anticipated. These can also involve malicious intent and, without careful management, harm to individuals, society and national security.

This report highlights how regional citizen data systems, and the business models that reflect them, have emerged across the world. Governments have taken, at times, strikingly different approaches to balancing their economic, social and national security goals when governing data. With strengthening links to global trade, these regional systems interact to create a 'system of systems', bringing complexity, uncertainty and risks of fragmentation. Demographic and technological trends can be powerful drivers of change. Conflict and political or economic shocks may also shape the future of these systems. The COVID-19 pandemic already looks set to have profound impacts on the use and sharing of citizen data with public authorities and others, and public attitudes towards this.

This report sets out four future scenarios for 2030, illustrating how global norms, business models and government approaches to citizen data could develop. In some scenarios there is convergence in governance of data, and it moves more freely around the world. In others, countries or regions take more divergent and often restrictive approaches, and data becomes increasingly localised.

We need to forge a path that best serves the interests of our society, whilst recognising that international collaboration and leadership will be crucial to success here and globally. The UK is well-placed to show such leadership, with a developed research system and strengths in emerging data-driven technologies which mitigate risks, preserve privacy and protect against misuse of data.

This report highlights the importance of having a clear vision for what we want to achieve with citizen data, and building understanding and confidence among citizens in how we will achieve it. It highlights opportunities and risks in the context of globally different approaches to citizen data. Given the inherent uncertainty described by the scenarios, it is clear that we will need to adapt our approach methodically when the situation changes and as new evidence emerges.

This report has been informed by a wide range of academic, government and industry experts. I am very grateful for their time and insights.

Sir Patrick Vallance

Government Chief Scientific Adviser

This is not a statement of government policy

Executive summary

Data about people, here referred to as citizen data, is increasing rapidly in volume and variety. The effective use and sharing of such data has the potential to bring huge benefits to the economy and society as a whole: boosting productivity and trade, enabling innovative products, improving public service delivery and informing scientific research.

However, the shifting data landscape is challenging for policymakers to navigate. Data collected for one purpose can be used many times over for a range of other ones, and government policy in one area can have unintended impacts elsewhere. Restrictions of data sharing can sometimes mitigate risks of privacy breaches and misuse. However, the linking together of new types of citizen data – across the boundaries of industry, government, and our personal lives – offers some of the greatest opportunities.

Data is internationally mobile, and increasingly tied to the provision of goods and services. This means data is an important consideration in international diplomacy, collaboration, and trade. The effectiveness of regulation and data security measures depend on enforcement activity which may involve agencies in other countries. Meanwhile malicious actors including hostile states are developing more advanced ways to use citizen data for their own interests.

This report therefore aims to inform public debate and government decisions with an international and whole-system view of citizen data. It considers interactions between data governance frameworks, public values and behaviours, technologies, and data-focused business models. We consider these components to form a 'citizen data system'. We contrast the evidence of how three major regional data systems reflect and determine developments in economies, security and societies.

Finally, the report explores factors that might drive future changes. We use this to build four plausible scenarios for the landscape of data systems across the world in 2030. These are intended to help decision-makers form ambitious strategies that are resilient to the uncertainties that prevail.

Data systems today

Governments can use several levers and policy stances to shape domestic data systems. These include regulation on privacy and data protection; competition policy; use of data for national security and law enforcement; and the use and sharing of public sector data. Regional approaches have tended to develop organically, building on existing norms but aligning with forward-looking geopolitical aims:

- **China has prioritised national economic and social security,** with strong government coordination and control of citizen data combined with restrictions on international transfers. These values are demonstrated by the Chinese Social Credit System, which is intended to aggregate financial, law enforcement, commercial, social media and other data in order to monitor citizens' compliance with various obligations, determine sanctions and encourage certain behaviours.
- **In contrast, the EU has prioritised citizen rights and competition within the internal market,** with variation between member states in national security approaches. Individuals can legally challenge the data practices of large international companies. Privacy campaigner Max Schrems' legal challenge to Facebook's policy of transferring user data to the US resulted in the invalidation of the entire EU-US Safe Harbour commercial data sharing agreement, and the adequacy decision for its replacement, the EU-US Privacy Shield, 5 years later. The EU has often been a first

This is not a statement of government policy

mover in developing rules and regulations on data, such as the comprehensive General Data Protection Regulation (GDPR). It appears that the new rights and responsibilities enforced by the regulation are being widely used; some observers worry that they may benefit incumbent dominant businesses but robust evidence of its impact on data markets is still scarce.

- **The US has, consistent with its wider economic stance, taken a generally less interventionist approach to data,** whilst actively seeking to use it in support of national security. For example, there is currently no comprehensive federal data protection law, although there is legislation relating to specific sectors and a marked variation in approach between different states. This has supported the growth of industries such as data brokerage, where data relating to individuals is amassed and passed on for profit. The US has levied large fines on companies for data-related reasons (such as the Federal Trade Commission's $5 billion fine on Facebook in 2019 regarding data privacy issues), but these tend not to be associated with competition concerns, in contrast with the EU.

Varying priorities have therefore contributed to the formation of somewhat divergent data systems in these three regions. As set out above, some of the differences have emerged as a direct result of government action. Factors not directly within government control, or which require agreement between different governments, will have also had an impact.

Citizen values may have also played a role in shaping data systems. For example, evidence suggests people in China typically indicate higher levels of trust and lower levels of concern over data use and privacy than those in the EU and US. However, it is unclear how far this is influenced by their governments' approaches in the first place. Values may change rapidly over time, and within most countries there is significant variation in attitudes towards privacy concerns and levels of trust.

Citizen behaviour does not always match reported values. Evidence also points to a disparity between people's expressed values around data privacy and the way they interact with services that require data sharing, for example only requiring very small incentives to share personal details, independently of reported privacy concerns. This 'privacy paradox', widespread mistrust of some business practices, and concerns about cyber criminals might imply that many citizens would welcome further government intervention to protect their data. However, surveys suggest, for example, that a large proportion of people worldwide see their government contributing to their distrust of the internet. This suggests that building greater consensus over policies which combine individual choice with trusted interventions to protect citizens should be a priority for building confidence in data systems.

The incentives to use citizen data to create successful or even dominant business models also shape data systems. The value companies are able to extract from large datasets, in targeting their services and training new programmes, incentivise business models based on aggregation of citizen data. These business models have dominated digital services and increasingly enable innovations in every-day products and devices, including via the Internet of Things. Large internationally operating companies with data-focused business models have grown in the US and China, often collecting and linking different types of citizen data from across multiple platforms and services in order to gain insights into citizen behaviour that inform product development, or are key components of products themselves.

This is less the case in the EU, where citizens typically rely on US companies for many online services. There, however, data-intensive models have still been a major driver of industrial development for large companies and SMEs alike, fuelling growth in sectors

This is not a statement of government policy

important to the UK such as e-commerce, financial services and telecommunications. As of July 2020, in most of the 25 ‘unicorns’ (private companies with a valuation over $1 billion) located in the UK, data is a major aspect of operations, incorporating fintech, AI, internet, e-commerce, data management, cybersecurity or hardware.

Innovative approaches to data will have an influence over the wider system. Data portability initiatives can enable greater consumer choice – the Open Banking initiative in the UK already has over 130 third-party providers of services including financial managers and account aggregators.

Trade negotiations are having an increasing impact on data systems. Regional citizen data systems interact most obviously through the international flow and use of data as part of business models and global value chains. In recent years, however, most new regulations have acted to restrict cross-border data flows. Many countries are introducing rules such as data localisation, with a preference for citizen data being stored or processed within national borders. This trend has been somewhat countered by key trade agreements and international frameworks promoting free flow of data, such as USMCA (United States-Mexico-Canada Agreement) and CPTPP (Comprehensive and Progressive Agreement for Trans-Pacific Partnership). Differences between major regional systems have led to international tensions, for example over data transfer between the EU and US.

Data diplomacy to shape or align with other systems is growing in importance. The geopolitical power of countries and regions can be used to spread their preferred models of data governance and potentially constrain the emergence of alternatives. Being ‘outside’ of a particular model can restrict data flows with countries within it, and evidence suggests that international restrictions on such flows can create considerable economic costs, particularly in relation to trade in services. Regulatory approaches similar to and influenced by GDPR have spread internationally along with market access arrangements. Meanwhile, China has expanded its influence on data systems with large investments in data infrastructure around the globe.

Future trends

Global megatrends will shape how data systems develop. The world’s economic centre of gravity is predicted to shift eastwards, with ageing populations in the West and rapid economic growth anticipated elsewhere. The distribution of internet traffic and users will also shift, with the potential for African nations in particular to hugely expand use. This may increase the global influence of data models different to those prevalent in the West, changing the dynamics of the geopolitical forces discussed above.

Wider geopolitical trends may play out in data systems, placing greater value on shaping international norms. If regional or other international data systems continue to grow in importance, individual nations may have reduced agency in designing their own policy frameworks if they wish to continue to benefit from data-enabled trade. Accordingly, the success of efforts to build global trade and expand it further in services may depend upon either building clearer global consensuses on data systems or improving interoperability between those that are different.

National prosperity is likely to be increasingly tied to the effectiveness of data systems. Almost all predictions suggest rapid increases in the volume and variety of citizen data, generated through increasingly varied devices and services, and held across the public and private sectors. Data systems that embrace this stand to benefit from higher productivity; improved public services; and a role in the advancement of global science.

This is not a statement of government policy

Data is an increasingly critical tool in addressing grand challenges, but the growing volume of energy-intensive processing raises its own sustainability issues. Without mitigation this proliferation in data will also contribute to increasing the energy demands of data processing. Whilst technological breakthroughs may help, enabling international data flows could come at the expense of domestic incentives to reduce data-related environmental footprints. The future of global data flows may therefore depend on making globally coordinated progress on climate change and sustainability policies.

New types of citizen data and novel uses bring new threats to manage. Evolving risks include micro-targeting of cyberattacks and disinformation, the exploitation of vulnerabilities in machine-learning systems, inadvertent introduction of biases, and harms associated with online targeting. A particular issue is the degree to which data can be genuinely made anonymous before being released or shared. More nuanced conceptualisations of openness and risk will be needed, and technological change will mean that judgements about what is 'safe' will be subject to continual revision. New privacy enhancing technologies will likely help facilitate more data-sharing for a given level of accepted risk, but this may come with performance trade-offs.

Potential disruptors

In addition to long term trends, recent experience suggests that future data systems are likely to be shaped by unpredictable shocks, and successful data systems will be those that can effectively and swiftly adapt.

New technologies might change what a successful data system looks like. New analysis approaches could reduce the importance of access to the largest datasets; instead the variety and type of data may become more important, potentially influencing business incentives to gather data. New computing technologies could address existing issues, for example with more compact computing power allowing more secure local processing of sensitive data; but they could also create new challenges, for example with quantum computers breaking some of today's cryptographic methods.

Economic shocks and the associated rise and fall of particular business models could change the incentives for business data use, reduce barriers to entry or entrench incumbent positions. Before the financial disruption caused by the COVID-19 pandemic, the technology company market showed some features similar to previous bubbles, such as the dot-com bubble of 2000. The long-term impacts of the pandemic on dominant business models are still to be seen, but initial evidence suggests incumbent, dominant technology market players have been well positioned to adapt.

Political shocks and conflict could change citizen beliefs and values, changing what national approaches can be sustained with public licence. For example, there was increasing support for security uses of citizen data in the US after 9/11, leading to the introduction of the USA PATRIOT Act.

New models of data governance may emerge. This could include different ways of addressing harms or shifting power balances, from new sets of data rights through to models of data ownership and data trusts. These changes could be led by governments, individuals, or businesses, and designed with a range of different aims in mind. Large-scale uptake of these models could improve some people's confidence in how their data is used and broaden access to its benefits. However, the proliferation of incompatible models could undermine innovations that require comprehensive, large-scale datasets.

The COVID-19 pandemic is changing the use of citizen data in ways which could have profound and long-lasting impacts. It represents a disruptor event that may significantly

This is not a statement of government policy

influence the future of citizen data systems. Its full impacts are still unknown; however, we can already see differences in national and international approaches to the use of citizen data in response to the pandemic, the role of technology companies in determining norms, and the debates being raised around prioritisations of individual privacy, security, and social aims.

Scenarios for global citizen data systems in 2030

This uncertainty and complexity implies a wide range of possibilities for how the future may look in 2030. We describe four scenarios, illustrating the scope of possible national and global outcomes, with potential implications for the UK economy, security and society in 2030. They all raise challenging questions for policymakers to consider.

- **Divergent data nationalism.** A world with low citizen engagement on data privacy and trust issues, a rise in nationalistic data policies, little technological innovation, and disruption of existing business models by regional government interventions.
- **Multipolarity.** A world with moderate but mainly government-led citizen engagement on data privacy and trust issues, hardening of the three main regional data systems, varied uptake of emerging data-driven technologies, and consolidation of market power for incumbent dominant players.
- **Deregulation.** A world with moderate but mainly individualistic and business-led citizen engagement on data privacy and trust issues, a relaxed global regulatory environment, high technological innovation, and consumer-led emergence of new business models.
- **Multilateralism.** A world with high collective citizen engagement on data privacy and trust issues, increased international collaboration on data policies, resistance to some emerging data-driven technologies, and disruption to existing business models by international interventions and a change in the value of large datasets.

These scenarios were produced in collaboration with a wide range of government, academic and industry-based experts. They were designed to help policymakers reflect on the uncertainty inherent in the global data system. They are intended to all be plausible, whilst sufficiently different and challenging to be useful to policymaking. For states and their citizens, outcomes will be determined by whether their policies are resilient to a range of potential futures, as well as whether the global arrangements that emerge meet their objectives and values in the first place.

Policy recommendations

What does this all mean for the UK's approach?

Navigating an uncertain future with appropriate agility is only possible with clarity about our aims. The UK government should seek to clearly articulate what it wants to achieve with its data system: what economic, social and security-related ambitions it has for better use of citizen data and what objectives for security, inclusion and individual rights it will prioritise.

It will be important to take a holistic approach to data systems. In developing its strategy, to avoid unintended consequences the government should take a 'whole system' view. It will be important to acknowledge the complex interactions between businesses, government, the wider public sector, the third sector and the public. Commitments made in one area, for instance on the use of data for the protection of national security, can have important implications for the assurances that can be given elsewhere, such as for privacy.

This is not a statement of government policy

Given this, the trade-offs between competing policy objectives for a data system need to be made consciously. Policymakers should be transparent and realistic about such trade-offs. In particular, governments should also recognise that seeking to maximise the benefits its citizens gain from global trade may mean not being fully free to set their own citizen data arrangements unilaterally. Coherence with regional data systems, for example the EU and regulations including GDPR in the UK's case, can be important for businesses seeking to export and consumer access to services. However, there are also important variations in domestic implementation of different policies and regulations, and the multilateral frameworks that have emerged do not necessarily preclude other forms of international coordination.

The UK should take opportunities to steer the formation of new global norms, as well as respond to them. Combined with domestic strength in data-intensive industries, showing leadership in developing forward-looking data regulation approaches, and ensuring wider economic policies are fit for the digital age, would put the UK in a strong position to do so. There may be opportunities to shape and support emerging data governance frameworks in countries with less developed systems, and aligning these with the UK's could help to underpin future economic partnerships.

Members of the public need to be an active and engaged part of the UK's data system. Given the lack of consensus within and between countries on the issues discussed in this report, and variable levels of trust, governments need to actively engage with the public about data. A reliance on supposed disinterest is unlikely to be sustainable long term. Governments should listen and respond to concerns, but also be willing to lead, educate and persuade where there is strong evidence in support of interventions. If a larger proportion of the public feel confident in our data system, more may engage with it in an informed way and access its benefits. The risks highlighted in this report need careful managing, and not all citizens will value economic gains equally, but an inability to harness data in a comprehensive way can, for instance, mean missed opportunities to help vulnerable families or improve public health.

A successful data system will need to be flexible and react quickly to changes. Given the uncertainties highlighted in this report, resilience and agility should be built into data policy development. All data policy should be developed with a range of futures in mind, and the scenarios developed in this report are intended to provide a starting point for this. More generally, it will not be possible for a strategy to foresee every eventuality. Error-correction mechanisms need to be built in. Some policies or regulations will need to adjust as new evidence emerges and as the global data system develops. This should not necessarily be taken as a failure of the original vision.

Finally, we will need to continually improve our understanding of the system. The most effective and valuable methods of integration of citizen data into businesses, public services and global interactions remains an emerging area of research. This should be prioritised by government and academia, building on the UK's existing strength in this field. This could support innovations, for example in energy-efficient computing and privacy-enhancing technologies, that would make the trade-offs described above easier to manage in future. This report highlights gaps, and some inconsistencies, in the available evidence. There is also a need for research into the impacts of our data system and alternative governance models on social, economic and security outcomes; the economic effects of diverging from trading partners' policy frameworks; and how to share the benefits of data-related innovations more widely.

This is not a statement of government policy

Contents

Ministerial foreword *1*
Preface *2*
Executive summary *3*
Contents *9*
1 Introduction *11*
1.1 Report aims, method and scope 11
1.2 What is citizen data? 12
1.3 What makes a data system? 13
1.4 Why is this important for government? 14
2 What governments can do to shape data systems *17*
2.1 Privacy and data protection 18
2.2 Competition law and policy 19
2.3 National security and law enforcement 21
2.4 Public sector data 22
2.5 International and trade policy 23
2.6 Strengths of intervention and enforcement 24
3 World data systems *25*
3.1 China: national, social and economic security 25
3.2 EU: fundamental individual rights and a stable internal market 28
3.3 USA: individual and economic freedom, and national security 32
3.4 International agreements and cross-border regulations 36
4 Wider data systems: people and businesses *40*
4.1 Citizen values and behaviour 40
4.2 Business models and use of citizen data 43
5 Impacts of data regulations *48*
5.1 Enforcement and effectiveness 48
5.2 Economic and wider social impacts 49
5.3 Impacts on data systems and interactions 49
6 Global trends are likely to be mirrored in data systems *51*
6.1 Shifts in economic gravity 51
6.2 Internet demographics 52
6.3 Geopolitics and data governance 55
6.4 Energy and environmental impacts 56
7 Data growth is likely to increase opportunities and risks *58*
7.1 Growth in data volume and variety 58
7.2 Increasing benefits from use of citizen data 60
7.3 Increasing risks from use of citizen data 62

This is not a statement of government policy

8 New technologies 66
8.1 New approaches to analysis, and the value of large volumes of data 66
8.2 Hardware and computing 68
8.3 Privacy engineering methods and technologies 69
9 Political, social and economic shocks 72
9.1 Business models and socioeconomic shocks 72
9.2 New models of data governance 74
9.3 Conflict and rapid political change 77
10 Case study: COVID-19 and citizen data 78
10.1 Divergence in digital contact tracing approaches from China, the US, and the EU 78
10.2 Tension between data nationalism, globalisation and technological dominance 79
10.3 Concerns around efficacy and equality of digital contact tracing 79
10.4 International approaches to COVID-19 and privacy 80
10.5 Potential future implications 80
11 Citizen Data in 2030 – four scenarios 82
11.1 Scenarios creation process 82
11.2 How to use these scenarios 82
11.3 Scenarios summary 83
11.4 Scenario 1 – Divergent Data Nationalism 84
11.5 Scenario 2 – Multipolarity 87
11.6 Scenario 3 – Deregulation 90
11.7 Scenario 4 – Multilateralism 93
12 Conclusions 95
12.1 Summary of main findings 95
12.2 Implications of scenarios 96
12.3 Policy recommendations 97
Glossary 99
Acknowledgments 103
References 104

This is not a statement of government policy

1 Introduction

1.1 Report aims, method and scope

This report is intended to stimulate thought and inform debate about system-wide and international issues related to citizen data, focusing on the future, and setting out plausible paths to 2030. It is not a comprehensive account of the evidence on technical or policy issues, and there are areas that we intentionally do not cover. We focus on the collection, processing and use of citizen data, rather than factors that affect or result from these things, such as physical communications infrastructure or specific applications of artificial intelligence (AI).

This report is based on interviews and engagement with experts on aspects of citizen data systems from technology to legislation. Expert input has been supplemented with desk research and evidence review by internal teams, but a comprehensive systematic review is beyond the scope of the report. The report has been also peer reviewed by several experts as listed in the Acknowledgements. The 2030 scenarios were developed through an anonymous survey of international experts, a development process with UK policy experts, and further expert review.

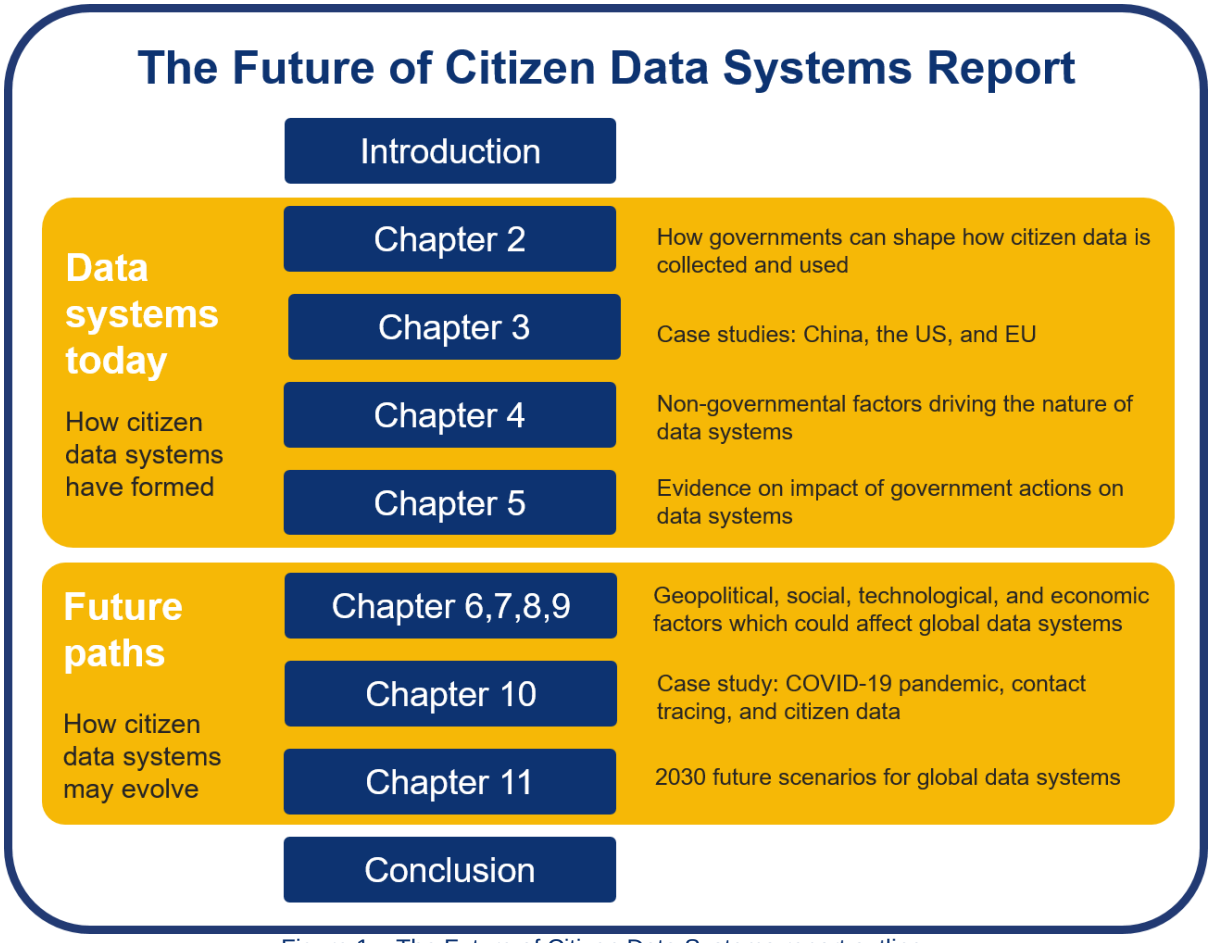

Figure 1 – The Future of Citizen Data Systems report outline.

Part 1: Data systems today **describes how citizen data systems have formed across the world.** Read Sections 2 and 3 for an overview of the ways that governments can shape how citizen data is collected and used, and how three different world regions have approached this. Read Section 4 for an explanation of how wider non-government factors such as commercial business models also determine the nature of data systems. Read

Section 5 for a discussion of the evidence around the impacts of government actions on data systems in different regions.

***Part 2: Future paths* describes potential future paths for citizen data systems across the world.** Read Sections 6, 7, 8 and 9 for a review of the geopolitical, social, technological and economic factors that could affect global data systems to 2030. Read Section 10 for a short case study on international approaches to responding to the COVID-19 pandemic, and what this might mean for the future of citizen data systems given the current limited and emerging evidence. Read Section 11 for four broad and varied scenarios of future global data systems in 2030, and policy questions for the UK in each.

1.2 What is citizen data?

Data about people is generated constantly through interactions with a huge range of public and private services, with and without knowledge and consent. Here, we use the term *citizen data* to refer to any data that could originate from or relate to any individuals or groups of people from any country, now or in future. Such data may be held by any number of organisations, including the private sector. We use this term to distinguish what we mean from specific legal or other definitions of *personal data*, which change over time and across jurisdictions and may imply specific rights associated with it such as ownership.

Examples of what we mean by citizen data include:

- Unique identifiers (e.g. NHS or passport number)
- Shared identifiers (e.g. name, date of birth, address)
- Biometric data (e.g. DNA, fingerprints)
- Medical, educational or other records
- Data generated or observed through interaction with services or devices, such as
    - Internet browsing history, tracking cookies, IP addresses
    - Video data (e.g. CCTV images)
    - Utility usage data (e.g. from smart meters)
    - Data generated though interaction with Internet of Things (IoT) devices (e.g. voice recordings)
    - Consumer data (e.g. online shopping behaviour)
    - Social media data
    - Location data (e.g. through fitness tracker apps)
- Inferences, predictions and assumptions derived from data about people (e.g. digital profiles used for targeted advertising)
- Metadata relating to data about people (e.g. when and how data was generated)
- De-identified data (e.g. medical records with identifying fields removed or changed, for use in scientific research and planning)
- Aggregate data such as census information, even if it is reportedly anonymised

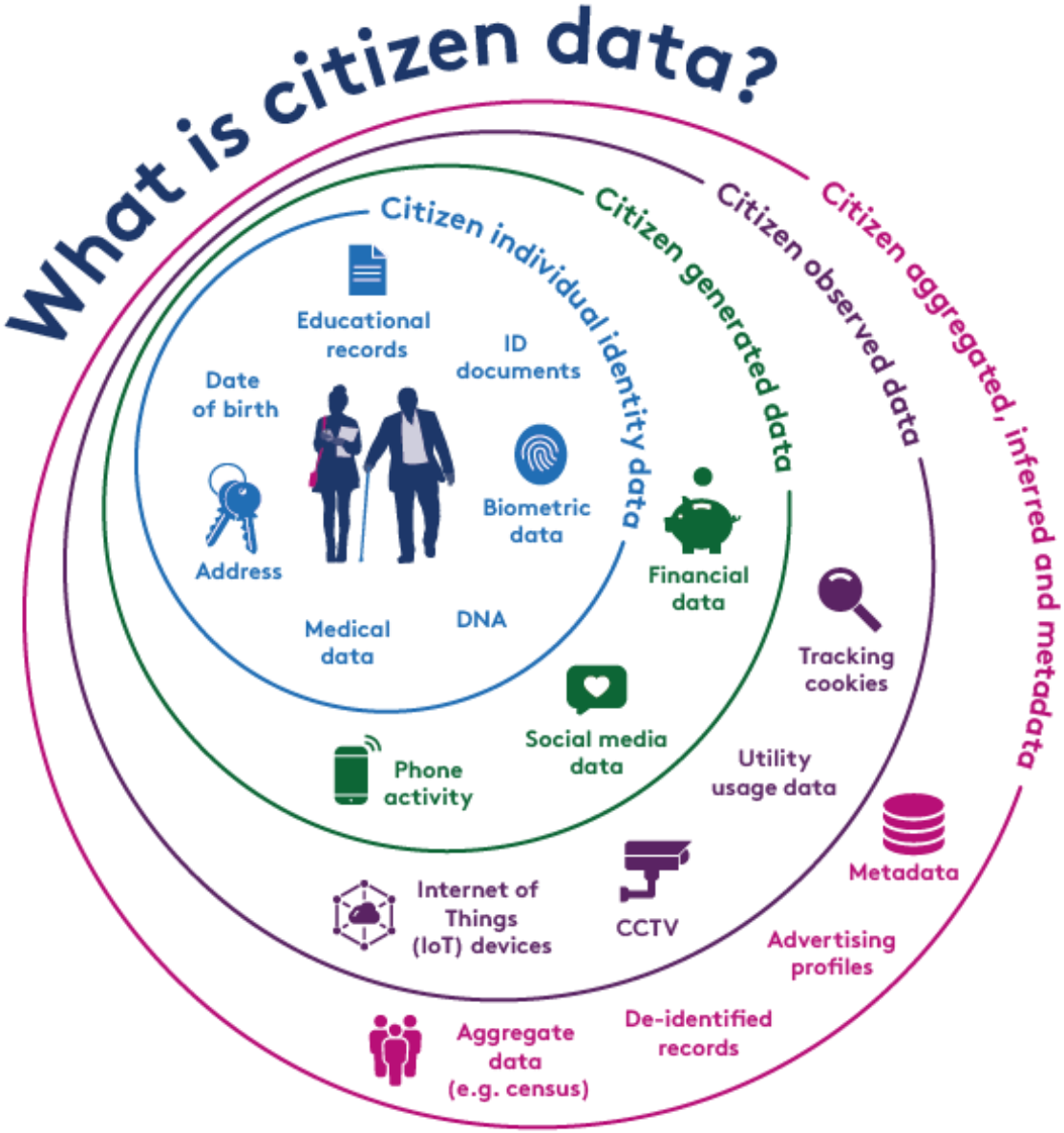

Figure 2 – A diagram showing some examples of types of citizen data, organised in layers of proximity from citizens, moving outwards from citizen individual identity data to citizen aggregated, inferred and metadata.

1.3 What makes a data system?

The term data system will have different associations for those with different backgrounds. Technologists and scientists may think of computing systems; business and IT experts may think of enterprise architecture; legal and policy analysts may think of the laws, regulation and policy which shapes what can and cannot be done with data. Others may have entirely different conceptions of the term. Such variation in viewpoints and language is a key challenge in understanding system-wide issues.

We take a broad view of data systems, which includes all the following, and more:

The people, processes and technologies involved in collecting, discovering, storing, analysing, linking, and sharing citizen data. This includes the individuals that collect data (e.g. in doctor-patient interactions), sensors and other data-generating technologies,

technical standards, systems that store and process data, and the degree to which datasets are consolidated and/or shared between different parties.

The legal, ethical, and procedural frameworks that shape how the above can take place, who is involved, and the degree of oversight or availability of meaningful challenge. This includes everything from laws and regulations, to codes of practice, ethical guidelines, and the standard operating procedures of individual organisations.

The wider incentives, values, behaviour and other dynamics of actors within these systems. This includes both the expectations and values of citizens and their behaviour related to data collection and use, the business incentives and business models that exist in particular regions, and the incentives and actions of other actors, from governments to lobby groups or malicious actors.

These components do not operate independently of each other, but instead interact in complex ways to form data systems. These interactions can be difficult to fully understand or predict, and have implications when designing interventions which attempt to address issues in specific areas. Therefore, throughout this report we take a whole-system view and consider the different components that make up data systems, and how they interact.

We also consider the interactions between systems, which may differ at local, regional, national and international level. The approaches of individual cities, countries and regions don't exist in isolation – globally there is a system of systems, with each potentially exerting pressure on others and leading to compromise or reaction elsewhere.

1.4 Why is this important for government?

1.4.1 Data is an increasingly important facet of our economy, security and society

The increasing volume and variety of data about people, generated by a growing number of digital devices, is a huge opportunity. It has the potential to boost productivity and trade[1], and support new businesses and jobs[2], through increasing the efficiency and scope of scientific research[3], and better targeting of policy and public services, bringing benefits to society and individuals.

Market forces alone are unlikely to ensure that data is used to the full benefit of society. This is because data is non-rival, meaning it can be used and re-used by many people at the same time without it being exhausted. In addition, there are externalities associated with using citizen data, both positive (for example, if datasets are aggregated and provide new insights), and negative (for example, unwanted intrusions into privacy). A market may therefore produce too much of some kinds of data and too little of others, even where there is potentially large public value[4].

Linkage and re-use of datasets mean that data issues cannot be viewed in isolation, or within individual domains or sectors. In current systems, access to the largest, highest quality and most varied datasets relevant for a particular use tends to increase actors' ability to achieve analytical, commercial, social or malicious aims. Big datasets collected for one purpose are becoming useful for others: consumer data can be used for policy, international development or public good research. Government interventions to increase or decrease data availability or sharing for one purpose are likely to affect all the others.

This also creates risks that must be considered holistically. Social media data could be used maliciously to direct cyberattacks, for example. The greater availability of data, the more likely it's possible to re-identify from supposedly 'anonymised' aggregate datasets. As the role of data in everyday life grows, familiar risks such as data breaches and privacy

scandals are increasing. New issues are emerging such as targeted data poisoning attacks manipulating automated systems, and the impact of data use on energy demands and the environment. Network effects in digital services mean that a concentration of users and data with a limited number of organisations could create a concentration of power and undermine competition.

1.4.2 Some interventions could be in conflict or have unintended consequences

Actions discussed or proposed now are often focused on tackling specific goals, issues or harms. These can include increasing data mobility and openness as a competition remedy[5], to tackling online harms and the potential for manipulation[6], to increasing individual privacy and security. But interventions are likely to have impact beyond their specific goal due to interactions between the components that make up a data system, as discussed above. Indeed, in any system, intervening by shifting information or data flows can have consequences felt elsewhere across the system[7]. There are costs to restricting access and use of data, and risks in opening it up. The global impacts of recent interventions are still unfolding and difficult to comprehensively evaluate. To take GDPR as an example, there is some evidence that the legislation has heightened security standards and may improve customer value to advertisers[8]. On the other hand, some rights under GDPR could be implemented in ways that increase risks to people and/or organisations[9,10], and some suggest GDPR may benefit large companies at the expense of smaller competitors and aspects of trade. Potential benefits and risks need to be understood and considered with nuance, as part of a complex system.

1.4.3 The paths taken by other countries will affect the UK

Data is increasingly international. By one estimate, global data flows increased by a factor of 148 between 2005 and 2017[11]. International data flow enables business to operate more efficiently across borders, can allow SMEs (small and medium-sized enterprises) to access global markets, and is vital in the global exchange of knowledge and information[12]. International data flows can be integral to the success of businesses in many sectors, such as telecommunications, professional services, and advertising. The ability to analyse international datasets will be important for many objectives, such as addressing global development goals[13], improving image recognition[14], and in genomic medicine[15]. The dynamics can be complex. Data generated in one region may be cleaned and labelled by people in another[16,17], to be used by a business based in yet another region, to develop a product or service sold in multiple regions, use of which generates further data. Data and associated regulation form an integral part of global trading relationships, including those related to goods not conspicuously associated with the data economy, and will become even more important if trade in services grows. The UK's domestic and international policies, their compatibility with those in other regions, and the ability for regulators and others to cooperate across jurisdictions, will affect the ease of international data flow and analysis, costs and benefits to the economy, and privacy and security risks.

The global centre of gravity is changing. We will likely see an economic shift eastward and a shift in the internet and data economy, with the relative size of the US and Europe (in internet users, and as providers and users of data services) declining relative to Asia, Africa and South America. Many emerging data economies have the opportunity to leapfrog technological or governance stages – for example with the uptake of mobile internet in much of Africa without first going through a substantial landline stage, or countries where Chinese companies are exporting infrastructure. These economies may choose to adopt a particular governance system, such as the EU's relative focus on privacy rights, which will then have impact on that system's power and the global balance.

Different regions and systems may have differing priorities for particular risks and benefits, such as privacy, national security, commercial data availability and consumer trust. This will have implications for global society, security, prosperity, and analytical or commercial advantage.

1.4.4 We need to build resilience to an extremely uncertain future

All these factors point to an uncertain future. The choices made now in individual components of the data system will have long-term and hard to predict effects on our society, economy and security, and the evidence of the impact of different approaches is often weak. Because of this high uncertainty and limited evidence, we have used futures thinking techniques to develop plausible alternative scenarios out to 2030, to help in the development of resilient and robust policies that achieve positive outcomes and minimise risks associated with citizen data in the UK.

This is not a statement of government policy

Part 1: Data systems today

2 What governments can do to shape data systems

Many factors affect how data is collected, analysed and used, from technology to business models. There are a number of levers and stances that governments can use to shape data systems as a whole, such as regulation, procurement, and convening power. Here, we focus on some common ones: privacy and data protection regulation, competition policy, access to data for national security and law enforcement, public sector data, and international and trade policy. They have varying impacts, levels of agency and dependency on other actors, and similar interventions can be used for very different overall aims.

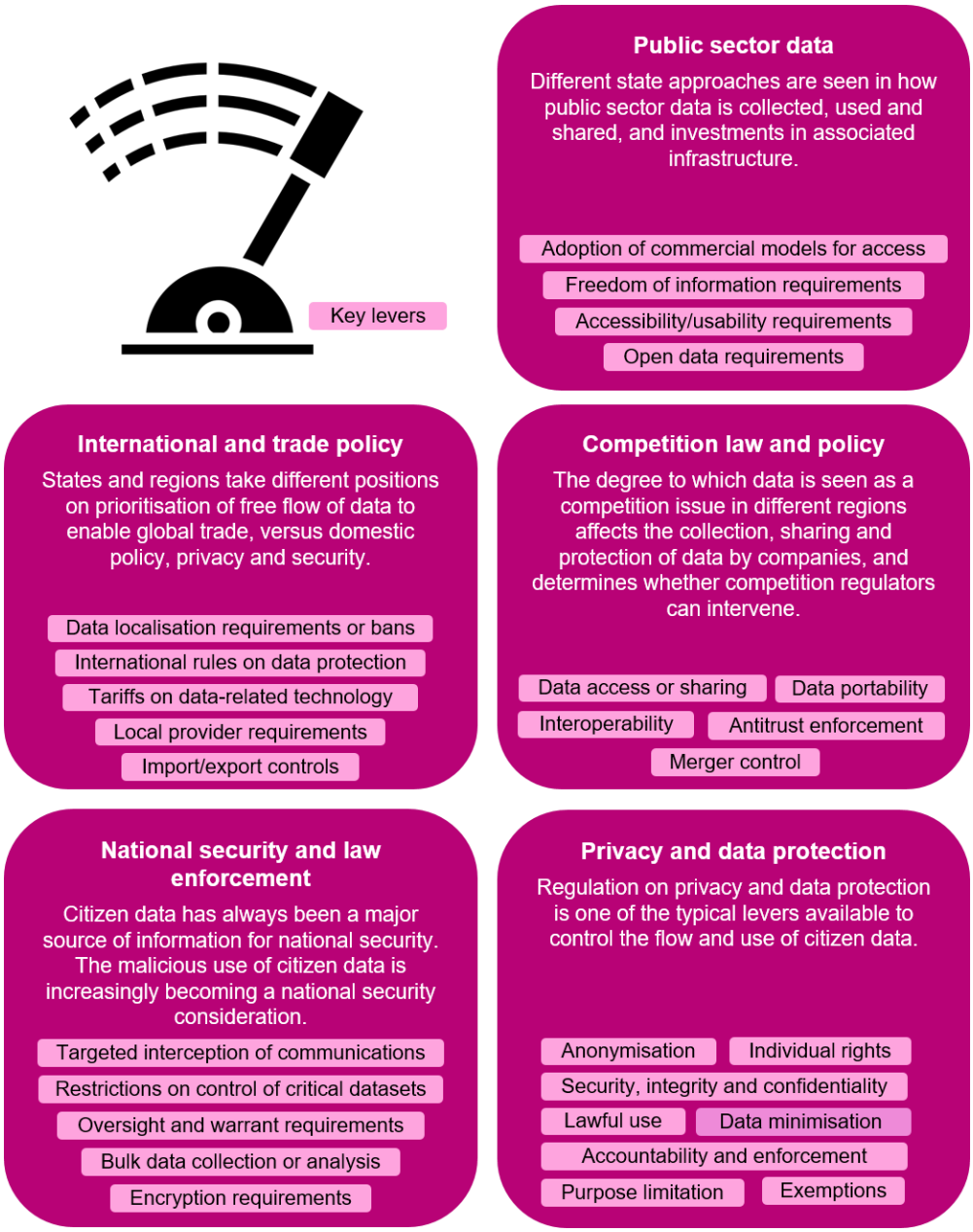

Figure 3 – Key levers governments can use to shape data systems.

This is not a statement of government policy

2.1 Privacy and data protection

Regulation on privacy and data protection is one of the typical levers available to control the flow and use of citizen data. Such regulations could protect individuals and groups against security threats, prevent intrusion and interference with private lives, empower people with rights over the content and use of data about them, and mitigate specific harms such as unwanted targeting or discrimination based on citizen data. Different requirements that could form part of privacy and data protection legislation are shown in Table 1.

Table 1 – Examples of government privacy and data protection requirements which can shape data systems.

Government requirement	Description
Lawful use	Defines lawful basis for collection and processing of citizen data (e.g. requiring consent or for specific purposes only)
Purpose limitation	Defines conditions for new use of data (e.g. if further consent is required or only specific uses are allowed)
Security, integrity and confidentiality	Defines minimum legal standards and obligations to maintain security, integrity and confidentiality
Data minimisation	Limits data collection to the minimum needed for a particular use; or storage to the minimum time needed
Anonymisation	Defines what types of data are considered anonymous, and therefore perhaps not covered by some other requirements
Individual rights	Defines what control individuals have over data about them held by others (e.g. rights to access, correct or delete data, to prevent or object to certain uses, to be informed of use)
Accountability and enforcement	Defines who is responsible for ensuring compliance, and the consequences of non-compliance (e.g. fines)
Exemptions	Exempts specific organisations and uses from requirements above (e.g. for national security or public good research)

The nature of these rights and controls has wider impact – and sometimes cost – to the economy and security. Without such controls, businesses and governments can freely collect, exchange, and use citizen data. This can have commercial and social value, for example by providing insight into behaviour. However, it also comes with risks of misuse, from unwanted use of data in another setting (e.g. health data used in insurance), to directly malicious attempts to steal identity or target citizens with disinformation. Where rights exist, they will to some extent restrict the collection and use of data by certain parties, involve a cost of compliance for data users and regulators, and could help or hinder the success of certain public services, research uses or business models using citizen data (see Section 4.2).

Current privacy and data protection rights vary widely between regions, perhaps due to differing government and societal views on their meaning and importance. There may be relatively comprehensive rights and requirements that apply to all actors in the data system, or frameworks may apply only to certain sectors, types of data or actors. These differences reflect values, history, economics, politics and public awareness – but in many cases, development of rights and regulation has been in reaction to negative events or risks.

‘Privacy by design’ is a framework which aims to embed the right to privacy into the design of information management systems and business practices. The concept was first developed by former Ontario Information and Privacy Commissioner Ann Cavoukian and establishes seven foundational principles including transparency, security over the full processing lifecycle, and privacy as the default setting[18]. The concept has since been adopted by various public and private sector organisations as part of laws or frameworks, including in Article 25 of the EU’s GDPR “Data protection by design and by default” (See Section 3.2.1 for a more detailed discussion of the EU approach). In this way, data protection regulation can protect citizens’ right to privacy, whilst also enabling data owners to trade and share data.

It is hard to evaluate how specific differences in approaches affect wider outcomes such as economic growth and national security. There could be advantages for regions where data access is easier, or for those where robust rights minimise harm to citizens. This will be entangled with wider social, political and economic factors. However, the compatibility of different systems is a key factor in enabling exchange of data and trade in data services.

2.2 Competition law and policy

Competition law and policy seeks to promote market efficiency by prohibiting businesses from engaging in anti-competitive conduct. Anti-competitive conduct may include anti-competitive agreements between firms (such as price fixing or collusion within cartels); abuse of dominant position (within the EU) or monopolisation (within the US) (such as charging excessively high prices, or excluding competitors from a market); and concluding mergers and acquisitions (M&A) that are likely to lead to a substantial lessening of competition (unless the potential negative effects on competition are offset by efficiency gains that ultimately benefit consumers and competition).

Control over the most valuable datasets is likely a substantial contributor to market power in some circumstances. The associated product developments can drive innovation, productivity and create better services, but could also lead to consolidation if competitors are unable to access the equivalent data inputs needed to compete effectively. For example, it has been speculated that several acquisitions of smaller companies have been motivated by a desire to acquire their data assets or complementary sources of data[19]. If companies can restrict competitor access to their user data, their ability to better target new offerings to their users may help them lock in customers and spread to new markets, alongside existing network effects. In principle, these possibilities could detract from some of the benefits for innovation cited above, if businesses focus on using user data to obtain market power, rather than improving services or reducing costs which may further reduce incentives to prioritise customers. This could also act as a barrier to entry and expansion for competitors, who may be able to offer more innovative products or services if able to compete on a level playing field. On the other hand, some policymakers, particularly those in the US, take the view that monopolistic tendencies in the tech industry are tolerable, following the argument of economist Joseph Schumpeter that they can lead to innovation that serves consumer interests, with the prize of winning dominance in a market a strong incentive to develop disruptive technologies in the first place[20].

There is evidence of market concentration (particularly for online platforms) which may be driven by data, and the evidence of impact on citizens is emerging. In relation to digital markets, evidence of potential benefits and harms for consumers was reviewed in the Furman Review of competition in the digital economy[5]. A recent market study of online platforms and digital advertising by the UK Competition and Markets Authority (CMA) found a range of concerns in these markets[21]. While large online platforms provide services to consumers for free, and higher volumes of user data may enable them to deliver better

personalised services, some have argued that an increasing volume of personalised advertising, misuse of data and privacy issues[22] are all evidence of low quality that may not exist in a more competitive market. As discussed in Section 8, access to the largest datasets may become less important in future as algorithms and computing power evolve.

The degree to which data is seen as a competition issue in different regions affects the collection, sharing and protection of data by companies, and determines whether competition regulators can intervene. Companies in the data economy are subject to traditional competition measures in most jurisdictions. EU and US cases against Microsoft in the 2000s provide an example: legal and technical restrictions Microsoft put in place that prevented PC manufacturers and users from uninstalling Internet Explorer and using other competing browsers, a technique known as 'tying', were found to represent an abuse of its dominant position in the PC operating system market[23]. But it is not always as easy to evaluate the use of data itself to consolidate market power in a wider range of industries. Particularly where 'free' services are offered to consumers in return for data (and paid for through advertising), it could be harder to use traditional variables such as price to establish anti-competitive behaviour[5].

If data is deemed an important input such that it can give rise to market power and competition concerns, regulators can already consider measures to widen or restrict access to it. Levers may be traditional (e.g. fines, blocking mergers or breaking up companies), or include data-specific measures (e.g. allowing consumers to move their own data between platforms, or mandating some level of data access for competitors). Examples of such levers are shown in Table 2. These measures either increase the number of parties who can access datasets or reduce the size/scope of the data available to perceived data monopolies. Several countries are considering updating competition frameworks to further address specific issues around modern digital economies and use of data, as discussed in later sections.

Table 2 – Government competition law and policy measures which can shape data systems.

Government measure	Description
Merger control	Prevent mergers where acquisition or linkage of large citizen datasets would be likely to cause a substantial lessening of competition, or to require merging parties to agree remedies to address competition concerns, e.g. the licensing or divestment of datasets, as a condition for merger approval
Antitrust enforcement	Fine businesses or impose criminal sanctions upon individuals that are found to have used data to restrict competition. Impose structural or behavioural remedies upon such businesses to divest assets, business divisions, and/or refrain from participating in certain conduct
Data portability	Mandate that consumers are able to access data about them and take it to another provider
Interoperability	Mandate standards that allow data and services from different companies to work together
Data access or sharing	Require companies to open or share their data with others

Use or not of these levers in different regions could have mixed impacts on citizens, the economy and security, which are hard to predict. For example, data portability or

more extreme measures to widen access could increase security or breach of privacy risks, for example if datasets are shared widely without properly mitigating the risks of re-identification. Fragmentation of platforms could negatively affect consumers and services. As with other forms of competition policy, states may face a challenge in balancing consumer welfare with the desire to promote domestic companies and protect national champions, particularly where citizens in one region rely on services from elsewhere.

2.3 National security and law enforcement

Citizen data has always been a major source of information for national security. This can range from targeted interception of communications (e.g. the calls or emails of a suspect), to analysis of broader datasets such as travel data or communications metadata (e.g. the times of calls). Such citizen data can enable both individual interventions (e.g. using data to monitor and identify suspects) and higher-level strategic decision-making (e.g. using data to target enforcement activity across institutions or areas, and analyse factors driving offending rates to inform crime prevention policies). Access through bulk powers or specific intercept powers may require authorisation (e.g. a warrant) and may be subject to independent oversight mechanisms to ensure proportionality and necessity. Further examples of potential measures are given in Table 3. The benefits to security and justice are usually recognised by society, but there is often seen to be a trade-off with privacy[24]. Views of this balance depend on the values of particular groups and regions, and the impact of recent events (from terrorism to perceived misuse of government power), which may not be stable or consistent within countries. See Section 4.1 for further discussion of this.

Table 3 – Government measures for national security and law enforcement use of citizen data which can shape data systems.

Government measure	Description
Targeted interception of communications data	Monitor a suspect's communications (e.g. phone data)
Bulk data collection or analysis	Collect communication metadata (where, when, etc) to inform decision-making and identify targeting identifiers
Oversight mechanisms and warrant requirements	Define whether a warrant is required before collecting or using data for security measures, and whether its use must be overseen or approved (e.g. by independent review)
Restrictions on control of critical datasets or infrastructure	Require that a key dataset (e.g. health or law enforcement records) is controlled by public or national companies
Encryption requirements or restrictions	Ban end-to-end encryption for certain uses to aid law enforcement investigations; or require the use of encryption for sensitive communications

The malicious use of citizen data, and control over data by foreign actors, are increasingly becoming national security considerations. This includes state-sponsored disinformation campaigns, cybercriminals using citizen data to target specific kinds of attack, and broader risks related to foreign companies gaining ownership or control over large citizen datasets and infrastructure. The openness of data systems may influence these risks. For example, features of the US data system often regarded as strengths (free speech, openness, limited regulatory intervention) could create asymmetries with other regions with contrasting regimes that increase cyber risk[25].

The choices made by regions in use of data for security purposes affect the wider data economy, and vice versa. Interventions in trade or competition that make data more available could increase security risks, without the right protection. The actions of external states and agents also determine security risk, and act as drivers for security measures elsewhere. Privacy-enhancing technologies and behavioural or societal drivers could reduce these risks or build resilience to malicious data use (see Sections 7 and 8).

2.4 Public sector data

Different state approaches are also seen in how public sector data is collected, used and shared, and in investments in associated infrastructure. Ethical guidelines, regulation or legislation may exist to direct when and how data should be used, and how issues such as bias in datasets should be addressed[26]. These have varied implementation and enforcement. Public data initiatives may be led by national governments, or by governance organisations at a more local level, such as cities or hospital trusts – or both may exist and interact. Examples of measures that can be taken by governments around public sector data are shown in Table 4.

Table 4 – Government measures around public sector data which may shape data systems.

Government measure	Description
Freedom of information requirements	Require that the government provide physical documentation or other information on citizen request
Open data requirements	Require that public or private sector organisations publish certain datasets online
Accessibility/usability requirements	Require that datasets be provided in common, machine-readable formats, e.g. with metadata to support use
Adoption of commercial models for data access	Enable public organisations to realise financial or other value from private sector use of data, through fees, licensing models, royalties or similar mechanisms

National statistical systems can play an important role in enabling effective use of citizen data to inform policy and academic research, and monitor UN Sustainable Development Goals. These systems are typically implemented by a national statistical office, usually an autonomous body with responsibilities established by legislation, in order to ensure independence, integrity and trust in the data that is provided. A national statistical office may collect relevant citizen data via censuses, through other government departments and agencies, and from charities and private sector organisations. Expert official statisticians then conduct and present analysis of this data that is trusted and useful for research, policymaking and monitoring. In this way they can play an important role in the wider data system of a region. Investment and capability in such national statistical systems varies widely between nations, with capability gaps in some developing countries. However, a 2017 OECD report noted that all countries have room to increase statistical capacity, transparency and use[27].

Sharing of data between regions and sectors is a key variable. There is great potential for public and economic good from opening public sector data. There are also risks to privacy and security, and the perceived risk of not effectively exploiting the commercial value of locally held datasets as intangible assets, which may bring direct benefits to the state and citizens.

The drivers and mechanisms for opening up government data differ widely, with considerable impact on how easy it is to access and use for different purposes. The motivations for opening up data can differ between regions and over time. The goal may be increasing the transparency and accountability of government, supporting devolved decision-making, improving service delivery (such as digital services and services provided by the public and charity sectors), or stimulating innovation in the wider economy. Some governments may simply make data available (for example publishing it on a website), while others ensure that data is findable, accessible, interoperable and reusable (FAIR), as well as easily machine-readable, such as through publication under an open licence[28]. While the former may satisfy basic requirements for transparency, the latter is likely to have far greater impact on re-use of data for research and innovation.

2.5 International and trade policy

International data flow enables trade in physical and digital goods and services, including data itself. It also underpins public-good uses of citizen data, for example in global health research and monitoring international development goals. In conjunction with other frameworks (e.g. privacy and data protection), trade and international data policy determines the ability of citizen data to flow across borders. Restricting or facilitating these flows could have considerable economic effects. One estimate is that international data flows contributed $2.8 trillion to the global data economy in 2014[29], while the EU estimated the value of its data economy at almost €300 billion in 2016[30]. In a UK context, it was estimated that EU personal data enabled services exports to the UK were worth approximately £42 billion in 2018, and exports from the UK to the EU were worth £85 billion[31].

Broader trade policy is often directed at data-related technologies. Tariffs and import/export controls can be used to restrict or control trade in technologies such as encryption that are dual use (i.e. may be used for civilian or military application) or seen as strategically important.

Free flow of data is generally seen as beneficial for global trade, but states and regions take different positions on its prioritisation versus domestic policy, privacy and security. These differences in approach can act as trade barriers, or as incentives for adoption of a particular data system. Some states require data localisation and other restrictions to keep control and access to certain data within their borders. Others (such as the EU) require evidence of a particular data protection or other standard before agreements that enable flow of citizen data. Examples of such conditions are detailed in Table 5. These may form part of wider international trade deals and negotiations. Specific international agreements and their key features are given in Table 7 in Section 3.4.

Table 5 – Government measures in international and trade policy which may shape data systems.

Government measure	Description
Data localisation requirements or bans	Require or prohibit that data (or a copy) related to citizens must be stored within national borders
International rules on privacy or data protection	Voluntary or binding privacy or data protection rules agreed or enforced between multiple countries
Tariffs on data-related technology	Tax the import or export of technologies related to data or communication infrastructure to/from certain countries
Import/export controls	Restrict the export or import of sensitive technologies such as encryption or surveillance devices to certain countries
Local provider requirements	Require that certain aspects of data storage or processing are conducted by local firms

2.6 Strengths of intervention and enforcement

Across all of these areas, governments can choose the style, strength and enforcement of each intervention. These can range from softer convening and stewardship measures, through funding or providing services, to harder laws, regulations and powers in cases of non-compliance. A mapping of UK data portability interventions found a wide range of approaches in use, from government acting as a convener in the financial services sector to primary and secondary legislation[32]. There is considerable debate as to what level of intervention is appropriate and effective for particular uses or sectors. For example, in AI (where issues often overlap with those of citizen data), many ethical principles have been agreed across the public and private sectors. However, questions remain over implementation and enforcement, perhaps requiring new regulatory or legislative frameworks[33,34].

Even where there are clear laws or regulations, enforcement capability and approaches can vary. The staffing and funding of data protection authorities varies widely across the world[35]. Similarly, different approaches to competition enforcement exist across different jurisdictions. For example, US competition law is enforced via a prosecutorial system managed by two federal authorities, while EU competition law is enforced by the European Commission (with National Competition Authorities within EU Member States responsible for enforcement both of EU and national law) (see Section 3.3.2). Differences in national approaches to enforcement can exist even in regions with harmonised rules such as GDPR[36], as data protection authorities within EU member states are responsible for enforcement and imposing sanctions for non-compliance. As data flow and use is often international, these differences can present a further challenge for governance. Emerging evidence on the impacts of such regulations and regional differences in enforcement are discussed in detail in Section 5.

Different regulators and organisations may perform similar or overlapping functions. For example, in February 2020 the UK announced an online harms regulator[37], whose role may overlap with other regulators. In the UK, the Information Commissioner's Office provides both guidance and enforces data protection regulations, while the National Data Guardian[38] issues official guidance specifically pertaining to UK health data.

This is not a statement of government policy

3 World data systems

This section reviews how China, the EU and the US have taken stances that affect the collection, use and sharing of citizen data within and between these regions. It looks into some of the factors that may drive differences between these regional approaches, and how the three systems interact with one another through international agreements and cross-border regulations.

We recognise that these are not the only regions with different or notable approaches to the control and use of data; that there can be substantial differences within these regions (e.g. at the level of EU Member States or cities); and that our areas of focus are not the only possible factors that affect data use. However, by considering these three contrasting and influential regional systems in depth, this section should serve to highlight the range of stances taken in mature systems globally, as well as the importance of different policy areas and the interplay between them.

	Data protection & privacy	Competition law	International policy	National security	Public sector data
	Some rules for businesses, but not government	Unclear if data considered a competition issue; may support domestic and state-owned companies	Extensive barriers to international data flows	Wide government access & control	Implementation varies, generally less open than other countries
	Fundamental individual rights	Data can be considered a competition issue	Free data flow within EU and adequate states	Each nation responsible; EU can overrule in certain circumstances	Directive initiatives for openness and transparency
	No comprehensive federal law; not historically prioritised	Data not typically seen as a competition issue	Promotes free data flow	Data for national security is a clear priority	Requirement for non-sensitive data to be open by default

Figure 4 – A summary of the different approaches taken by China, the EU and the USA regarding government policies that control the use of citizen data. These are discussed in more detail throughout this section.

3.1 China: national, social and economic security

3.1.1 National security aims drive data governance and use

Broad national security aims are often the primary driver for data governance frameworks. As early as the 1990s, China was developing approaches to data governance and use that focused on strong state control and security, in line with broader Chinese government aims. Examples include the CL97 law defining cybercrime, and the implementation of broad internet filtering – the so-called 'great firewall'[39]. The 2015 Counterterrorism Law requires telecommunications companies and internet services providers to assist the government, including sharing citizen data and "decryption and other technical assistance"[40]. The 2017 National Intelligence Law similarly obliges organisations and individuals to "support, assist and cooperate with state intelligence work"[41]. Chinese law, including the 2017 National Intelligence Law as well as the 2016 Cybersecurity Law discussed below, allows the state considerable control over Chinese companies, and potentially wider Chinese controlled technology groups, who are required to support the Chinese intelligence agencies in the interests of national security.

It has been suggested that the Chinese approach is driven by a desire to accelerate economic development through the opening of the economy, while maintaining social and national security[42]. This is further demonstrated by China's recently released draft version of an intended Data Security Law, which, if enacted, would likely become China's

highest-level set of rules for data governance, alongside the proposed Personal Information Protection Law referenced below. The draft Data Security Law includes developments such as a national-level data classification system for categories such as 'important'; data transaction markets which consider data as a 'factor of production'; and further principles for data access/flows between the state and private sector in China, and with foreign organisations and governments[43].

Government projects in China also focus on data for security. The Skynet programme uses data from nationwide surveillance cameras and facial recognition for law enforcement and counterterrorism[44]. This is being extended with the Sharp Eyes project, mainly aiming to expand coverage into rural areas, including use of phones and televisions[45]. Where these systems have been deployed, decreases in crime have been reported in some cases[46]. Reports suggest China is the largest market for security and surveillance equipment, with large-scale government procurement[47].

3.1.2 Privacy and data protection rights exist, but mostly without direct application to government

In China, privacy and data protection laws are developing, but currently exist only as part of wider frameworks for cyber- and national security. The main legislation in this area is the 2016 Cybersecurity Law, which mandates some data localisation and restricts cross-border transfers of data, as well as setting out measures for business to secure data against cyberattack and misuse[48]. It comes with a standard for data protection that provides guidance on consent requirements, deidentification, rights to deletion and data minimisation, among other things. A 2018 e-commerce law extends this to provide consumers with rights to correct and erase user information[49]. Authorities have also announced that they will soon introduce a Personal Information Protection Law, following the enactment of the country's first Civil Code in 2020, which includes preliminary measures on privacy and data protection[50].

Legislation generally provides the government with broad discretion over collection, access and use of citizen data[51]. For example, the data protection standard associated with the 2016 Cybersecurity Law provides exemptions for purposes related to state security, public security and major public benefits, among others. Another example of government use of citizen data is the Social Credit System (SCS). The SCS is a policy initiative consisting of multiple systems and pilot projects at differing stages of development, operating at various regional and national levels. It is reportedly intended to aggregate financial, law enforcement, commercial, social media and other data, in order to monitor individuals' and businesses' compliance with legal, moral and professional obligations. It assigns trustworthiness scores to individuals and businesses, and these can be used to determine sanctions such as restricted access to transport options or bank credit. The SCS is meant to foster trust between government, businesses and individuals, and strengthen social governance[52]. There is evidence that government SCS data is shared with business, with government blacklists used by Alibaba to prevent defaulters buying luxury goods, for example[53]. While the SCS may be less comprehensive and potentially less effective than often claimed in foreign media[52,54], there is clear evidence of the collation of large amounts of data related to individuals, for instance DNA databases and compulsory biometrics for residents of the Xinjiang region[55].

3.1.3 Competition policy is clearly aligned to the social and economic interests of the state

Chinese competition policy shares elements of EU and US approaches, but with wider alignment to state aims. China's Anti-Monopoly Law (AML) was passed in 2007, as the country's first comprehensive competition framework. It contains broadly worded measures to address monopolistic conduct, covering monopoly agreements between undertakings,

abuse of a dominant market position, and mergers and acquisitions (M&A) that restrict (or may restrict) competition[56]. In contrast to the EU and US, the AML allows for economic development and national interest to be considered within the competition authority's substantive assessment of whether certain conduct or a transaction may give rise to competition concerns[57]. Since 2018, one agency – the State Administration for Market Regulation – has been responsible for regulation and enforcement of the AML, which was previously split across three authorities focused on mergers, pricing, and non-price related issues[58].

The AML is perceived by some to benefit domestic and state-owned companies, but evidence supports this more for M&A than for other antitrust issues. The US Department of State has reported concern by foreign companies that antitrust enforcement in China is used to target them as an "extension of other industrial policies that favour state-owned enterprises and Chinese companies deemed potential 'national champions'"[59]. This may reflect a high level of reporting and concern among international business – in 2014, 84% of firms in one survey expressed this[60]. One study found that between 2008 and 2013, only 15% of merger decisions dealt with by the Ministry of Commerce concerned purely domestic deals, compared with 47% in the EU[57] and it has been reported that all deals approved with conditions or rejected to 2018 involved foreign companies[61]. However, antitrust enforcement appears mostly focused on domestic companies, although some reports claim foreign companies have faced higher fines[62].

How far data is considered as a competition issue in China is unclear and may be changing. A few, mostly domestic, internet companies have developed huge market share in China (e.g. Baidu in online search, Alibaba in online shopping). However, more recently an investigation was launched into potential anti-competitive practices of Tencent Music, a large domestic incumbent, although this was suspended in February 2020. The State Administration for Market Regulation asserted that "the Internet is not beyond the reach of antimonopoly regulations"[63], implying that such companies may be subject to competition interventions. Draft amendments to the AML published in January 2020 include additional factors for determining dominance in the Internet sector specifically, such as an entity's capabilities with regard to data manipulation and processing[64].

3.1.4 There are national rules for open government data, but implementation varies

National rules on open information have existed since 2008. The Open Government Information (OGI) regulations mandate access to government information from some agencies[65], and some local initiatives existed even prior to this, for example in Guangzhou[66]. The OGI regulations were revised in 2019, expanding their coverage, making clear that disclosure is the default approach, and clarifying exemptions (e.g. for information that may 'endanger national, public, economic, or social stability/security')[67]. 2018 regulations also require data from government-funded scientific projects to be available to the public[68]. It has been reported that the government proactively published more than 35 million records to 2018 and received millions of disclosure requests[69].

Open data has been stated as a priority, but implementation and standardisation can vary. China has typically ranked much lower than the US and UK on indices of open data availability, implementation and impact[70]. Despite this, multiple local and regional-level portals and services do exist, which may be rapidly growing[71]. Data may not always be machine-readable – for example Chinese open data on air pollution has been re-released in machine-readable formats by NGOs[72].

Such initiatives are reportedly designed to support law-based governance, transparency of government, and 'economic and social activities'[69,73]. More recently the importance of high-quality open data (versus simply provision of information) has been

This is not a statement of government policy

recognised. For example, open data is a key project in the Big Data Development Action Plan of 2015, and important for China's AI ambitions.

3.1.5 China has restrictive requirements for digital trade and international data flow

China is ranked highest on several global indices of digital trade restrictiveness[74],[75] and has many policies seen as barriers to international digital trade. For example, the 2016 Cybersecurity Law requires operators of 'critical information infrastructure' (a very broad definition encompassing many sectors) to store personal data within China unless there is a security assessed exemption[76]. This effectively requires many foreign firms to invest in local data centres, or contract/partner with domestic companies[77]. There are also limits on the ability to transfer certain data internationally. More recently announced draft measures on the cross-border transfer of personal information would make these restrictions even more stringent[78]. Imported and exported encryption products require state certification[79]. International businesses are also affected by the data-driven Social Credit System (see above), with for example several US airlines having received letters saying their trustworthiness score would decrease if they did not label Macau, Hong Kong or Taiwan as part of China[80].

Simultaneously, China is exporting its model of data governance to facilitate digital trade with other regions. As part of the wider Belt and Road Initiative[81] involving more than 100 countries, there is a major focus on Chinese firms providing digital infrastructure such as fibre-optic cables as part of a 'digital silk road'. Chinese companies have installed internet and mobile network equipment in more than 38 countries[82]; and more than 200 Chinese companies are involved in a Moroccan smart city project[83]. This may provide useful data for the Chinese companies – for example Guangzhou-based facial recognition company Cloudwalk have a reported agreement to use image data from Zimbabwe to train better facial recognition algorithms for darker skin tones[84] (see Section 8).

This approach could be driven by a focus on security and sovereignty across technical, industrial and social state policy. For example, there are requirements for the most sensitive critical IT infrastructure to use only domestic products[85], and China has led agreements to increase state control over internet regulations[86].

Overall, the Chinese data system is aligned to wider stated political and economic aims. Legislation discussed above is part of this, but also industrial initiatives, such as the 2017 AI development plan, Made in China 2025 and Internet Plus, aim to make China a global leader in advanced data-driven technologies such as AI and ICT[87]. The 13th 5-year plan mentions opening-up, sharing and coordination as key policy drivers[88]. The country's approach to data also aligns with reported aims to maintain social order and state security using a centralised systems-based approach, for example using technology and monitoring to encourage good behaviour[89].

3.2 EU: fundamental individual rights and a stable internal market

3.2.1 Privacy and data protection are fundamental rights

In the EU, privacy and data protection are fundamental rights, equal to others such as freedom of expression, and are protected as such in law. They are not *absolute* rights, and are balanced against other rights, EU laws and priorities such as national security – but they must be provided for by national governments[90] and are expressed in EU-wide regulations including the GDPR. These rights can be seen as integral to the formation of an integrated digital single market across the EU, enabling free flow of personal data, and a harmonised approach that provides stability for business operators, as well as empowering citizens and building trust in data use[91]. However, these rights can present a barrier to free

flow of data between the EU and third countries that adopt a different approach to data protection. They impose a cost of compliance and restrict the scope of data collection and use for affected organisations, including many public sector and government actors[92] (see Section 5). This position has been interpreted either as a signal that these rights are prioritised above other (e.g. economic) outcomes, or as evidence of a belief that their prioritisation can be 'win-win' and a competitive advantage for the region. It is also seen by some as a mechanism to export 'EU values' elsewhere (see Section 3.2.5)[93,94,95,96].

The 2002 Privacy and Electronic Communications Directive (commonly known as the ePrivacy directive) complements the EU data protection regime and was last updated in 2009. The directive sets out specific privacy rights on electronic communications such as marketing communications and use of online cookies, which are then implemented nationally[97]. The European Commission propose to replace this directive with a new ePrivacy Regulation, which will align with GDPR and will aim to further reinforce trust and security in the digital single market[98].

While values and opinions vary across Europe (see Section 4.1), a possible driver of the EU's robust approach to data protection regulation is the sensitisation of European citizens to the collection and use of personal data following WWII[99,100]. Concerns over intrusive census questions in Germany in the 1980s led to the German Federal Constitutional Court declaring self-determination over personal data as a fundamental right in 1983[101], which was mirrored in EU law.

3.2.2 Competition policy can consider data acquisition and use

EU competition policy has been developed alongside the Single Market. It is sometimes seen as wider in scope than the systems in other regions, with a greater focus on promoting competitive market structures[102], including fair competition across EU Member States such as maintaining an EU-wide state aid regime[103]. While National Competition Authorities have jurisdiction to apply both EU and national competition law in cases that produce, or may produce, anti-competitive effects within individual Member States, the European Commission has jurisdiction to investigate and impose penalties in relation to suspected anti-competitive conduct capable of affecting trade between EU Member States. The European Commission also has exclusive jurisdiction to enforce the State aid rules (albeit that national courts play a role with respect to illegal State aid)[104].

Data acquisition and use has been considered as a potential cause of harm by EU competition authorities. For example, Apple's acquisition of Shazam raised concern about acquisition of commercially sensitive data on users of rival services, although this was still approved[105]. The German competition authority found Facebook had abused a position of market power in the manner and scope of the collection and use of user data from third parties[106]. This decision was initially suspended on appeal[107], but the decision was upheld in June 2020 by the German Federal Court, which decided that the ban on Facebook processing user data without the further consent of private users imposed by the German competition authority could be enforced[108]. In the UK, the aforementioned market study of online platforms and digital advertising by the CMA found a range of concerns in these markets, and asserted that the inability of smaller platforms and publishers to access user data creates a significant barrier to entry[21]. They recommend that new legislation is needed to establish a pro-competition regulatory regime. A 'Digital Markets Taskforce' has since been established, led by the CMA, which will develop advice around what interventions, if any, are necessary to protect and promote competition and innovation in digital markets more broadly[109]. In addition, in June 2020, the European Commission launched a consultation for a Digital Services Act package of regulations, and a new competition tool intended to address concerns around competition in markets such as lack of access to data and data accumulation[110].

EU competition aims are also reflected in other areas of data policy. For example, measures such as data sharing requirements in Payment Services Directive 2 (PSD2)[a] and the right to data portability in GDPR are often cited as ways to improve competition by facilitating switching between different service providers, or expanding/creating markets for third-party services[111]. However, others have raised the risk that, more generally, GDPR may have raised regulatory barriers to entry in data-driven markets and entrenched incumbents' positions (see Section 5).

The EU's data strategy announced in February 2020 aims to create a single market for data[112]. Its stated aim is to improve the EU's global competitiveness while respecting EU rules and values of personal data protection and competition. The strategy includes plans for EU-wide common and interoperable data spaces for strategic sectors and domains (e.g. healthcare, mobility), and to make 'high value datasets' from the public sector available across the EU for free, particularly to improve data access for SMEs. The Commission also intends to publish extra guidance on existing EU competition law to enable compliant data sharing and pooling arrangements.

3.2.3 Data use for national security is balanced with fundamental rights and oversight

EU Member States must balance individual national security policy with EU laws and rights, but there is no harmonised approach to national security policy across the EU. In the EU, 'national security remains the sole responsibility of each Member State'[113], so each Member State is responsible for policy on law enforcement and surveillance, for example. However, laws and actions taken by EU Member States can be referred to the European Court of Human Rights (ECHR) or the Court of Justice of the EU (ECJ), particularly where there is interaction with fundamental rights.

The approaches of individual countries, and the findings of the ECJ and ECHR, have varied[114]. In 2014, the ECJ found the Data Retention Directive[b] to be invalid because of 'wide ranging and serious interference with fundamental rights'[115]. Similarly, both agreements for the transfer of personal data to the US (Safe Harbour and its successor, Privacy Shield) have been found to be invalid (see Section 3.3.3 for details). The ECJ has also found that EU Member States cannot impose a general and indiscriminate data retention requirement, as this may be incompatible with EU law and fundamental rights[116]. The ECHR has found that bulk interception of communications is a "valuable means"[117] to achieve the legitimate aims of preventing terrorism or serious crime, as long as there are adequate and sufficient safeguards against abuse[118]. In 2018 it found that Sweden's surveillance law provided adequate safeguards and was not in conflict with the right to respect for private life. In contrast, in 2016 it found that Hungary's legislation did create a conflict, as sufficient safeguards (such as assessment of the necessity of interception, and presence of remedial/judicial measures) were not provided. Some of these cases are being re-examined[119]. Recently the importance of addressing emerging threats such as online disinformation has been highlighted by the EU[120]. As the nature of threats constantly evolves, the measures pursued and acceptable to the courts and the public are likely to change.

The impact of recent events, public support, and a need to balance this with wider values and developments, have driven this balanced approach. Post 9/11, numerous European countries (and the US) passed legislation on national security, including data access. This is the first of the legitimate aims identified in the ECHR that could justify

[a] Implemented within the UK via the Open Banking initiative.
[b] Which aimed to harmonise EU Member States' approach to the retention of citizen communication data.

balancing with fundamental rights[121]. Public attitudes to privacy and security are likely to be nuanced and may be sensitive to recent events (see Sections 4.1 and 9.3)[122], or differ across demographic groups. There is evidence for this in both the EU and US[123], as discussed in Section 4.1.

3.2.4 The EU encourages open government data, but implementation varies

The 2019 Open Data Directive sets the EU-wide legal framework for government-held data. It replaces previous directives and sets rules to encourage EU Member States to "make as much information available for re-use as possible", for example through application programming interfaces and by increasing the transparency requirements for public-private agreements involving public sector information[124]. The European Commission has also clearly recognised the beneficial potential of data held in the private sector in both business-to-business and business-to-government contexts. However, action has remained at the level of principles rather than legislation[125]. Open data from EU institutions is available through a portal[126].

EU Member States, regions and cities vary in their approach to open data. For example, a recent OECD index ranks France and Ireland far higher than Lithuania and Sweden for the availability, accessibility and support for re-use of government data, with UK approximately in the middle, just below the OECD average[127]. Similarly, EU Member States can have different approaches to the intersection of open data and national security or be more concerned about public bodies losing out on benefits from commercial partnerships with data, for example with the UK's NHS[128].

Drivers of the EU's open approach include transparency and fair competition across the single market[129]. There is a long history of advocacy and support for open data in Europe from the perspective of government transparency and more effective governance[130]. Newer initiatives focus on the potential value of data for innovation and the economy, with the European Commission suggesting that high-value datasets should be provided free of charge to support industries such as AI.

3.2.5 International and trade policy aims for free data flow without compromising EU standards

The EU prioritises free trade within the internal market, and harnessing globalisation while preserving and promoting EU data standards globally. EU trade policy has evolved through the centralisation of trade authority at EU level and creation of the Single Market, which prohibits tariffs and limits and reduces non-tariff barriers between EU Member States. The EU Single market policy seeks to harmonise regulations for goods and services throughout the EU[131]. More recently, this has developed with plans for the digital single market, including measures to remove barriers to flow of personal (GDPR) and non-personal data within the EU[132].

In most cases, broad alignment with EU standards such as GDPR is required to enable international free flow of personal data with the EU, although there is variation on what this means in practice across EU member states. Some countries have adapted their privacy and data protection law to enable this (e.g. Argentina[133]) or have addressed adequacy alongside broader trade negotiations (e.g. Japan[134]). Some argue that the EU prioritises fundamental citizen rights (e.g. to privacy and data protection) over trade/economic outcomes[135], and that GDPR is an inadvertent trade barrier[136]. In any case, a desire to protect citizen rights and the integrity of the single market can be seen as key drivers of the EU approach to increasing international data flows and digital trade. This can be seen in a recent ruling of the ECJ, which upheld the validity of standard contractual clauses to allow continued international data transfers, but placed an obligation on

This is not a statement of government policy

businesses and regulators to suspend or prohibit transfers when there is a conflict with the law of the destination country[137].

EU Member States impose different national security restrictions on international data. While some measures such as GDPR apply across the EU, implementation and enforcement can vary across EU Member States; and there can be other national requirements such as restrictions on outsourcing storage or processing of sensitive public data, or retention/access requirements related to investigatory powers. In the UK, the GDPR has been implemented through the Data Protection Act 2018. As of August 2020, the level of protection provided by the UK Data Protection Act is currently being assessed by the EU, in a process known as a data adequacy assessment, in order to facilitate free flows of personal data from the EU to the UK[138]. The UK will be carrying out similar data adequacy assessments of the EU Member States and Institutions to facilitate flows from the UK to the EU.

The European Centre for International Political Economy (ECIPE) have developed a 'Digital Trade Restrictiveness Index' based on an assessment of digital trade policies in four broad areas they define as: fiscal restrictions and market access; establishment restrictions; restrictions on data, and trading restrictions[75]. They found that within the EU, France and Germany were overall the most digitally restrictive Member States and Ireland was the most digitally open, and suggested that overall, the most digitally open countries were small economies with larger services sectors.

3.3 USA: individual and economic freedom, and national security

3.3.1 Privacy and data protection rights exist in some states and sectors, with calls for comprehensive federal regulation

Privacy and data protection are not considered to be clear fundamental or constitutional[c] rights in the US, nor are they protected by comprehensive federal law. Some use- and sector-specific federal laws provide enhanced rights or protections for the use of certain types of data (see Table 6), but otherwise, regulations provide relatively little restriction on business and government in the collection and use of citizen data. The possibility of private litigation or reactive government action provides some deterrent against misuse[139], but case analysis indicates that US courts have tended to prioritise freedom of speech and national security above privacy and data protection[140]. The US was an early leader in privacy regulation such as with the Privacy Act of 1974[141], however this was focused on federal agencies' data use. At the same time, it is worth noting the substantial $5 billion fine levied by the Federal Trade Commission (FTC) against Facebook in 2019, for consumer privacy violations[142].

[c] Although an implicit right to privacy has been found by the supreme court in some cases, e.g. related to the 4th amendment right against "unreasonable searches and seizures"

Table 6 – Examples of federal privacy and data protection related laws within the United States, from the US Congressional Research Service[143] and references.

Federal law	Covered information *(Covered sectors)*	Nature of regulations
Children's Online Privacy Protection Act	Identifiable information collected online from a child under the age of 13	Consent for data collection and sharing; disclosure and security requirements
Communications Act	Personally identifiable information, customer proprietary network information *(Common carriers, cable and satellite operators/carriers)*	Consent for data sharing; disclosure and security requirements
Fair Credit Reporting Act (FCRA)	Consumer reports *(Credit reporting agencies, suppliers and users of consumer report information)*	Accuracy, use and disclosure requirements
Family Educational Rights and Privacy Act (FERPA)	Educational records *(Federally funded educational institutions)*	Consent for data sharing; disclosure requirements
Gramm-Leach-Bliley Act (GLBA)	Non-public personal information *(Financial institutions)*	Opt-out for data sharing; disclosure and security requirements
Health Insurance Portability and Accountability Act (HIPAA)[144]	Protected health information *(Healthcare providers and care plans)*	Consent for data sharing; disclosure and security requirements
Video Privacy Protection Act (VPPA)	Personally identifiable information *(Video tape service providers)*	Consent for data sharing
Other relevant laws	Authorisation requirements to intercept communications[145] or access data on a "protected computer"[146]; requirements for privacy/security policies not to be unfair or deceptive[147]	

Some US states have recently passed or are developing privacy laws which may be stronger or more comprehensive than federal laws. As of August 2020, at least 12 states were in this position, with signed laws in California, Maine and Nevada[148]. Of these, the California Consumer Privacy Act (CCPA) is the most comprehensive signed law, providing many similar rights and obligations to the EU GDPR although only applying to organisations 'doing business' in California and not applying to government. For example, it provides rights to individuals to access personal data and to opt-out of the sale of personal data[149]. There are also existing state laws focussing on specific types of citizen data, such as the Illinois Biometric Information Privacy Act, which regulates the collection and storing of biometric data specifically[150]. A law related to COVID-19 and use of citizen data is also being developed; see Section 10.

Possible drivers of the less interventionist approach to wider regulation in the US include a prioritisation of free speech and liberty, a historically narrower role for government in addressing social issues, and a desire to avoid stifling innovation and economic growth. Although there have been increasing calls for further regulation, in response to concerns over privacy and misuse scandals[151], this may be balanced by a

This is not a statement of government policy

desire to protect the competitiveness of the US tech industry[152]. Recently, the federal nature of the US political system, and differences in party control over various branches of government, may have also hindered attempts to pass more comprehensive legislation.

3.3.2 Competition enforcement has not typically considered data acquisition and use

Competition policy is based on a prosecutorial system that can lead to financial and custodial penalties against individuals. Federal enforcement is led by the Department of Justice (DoJ) and FTC, and actions may also be brought by US states or private parties[104]. Since 1979, US antitrust policy has focused on a 'consumer welfare standard' which finds an act anticompetitive if it "harms… both allocative efficiency and raises the prices of goods above competitive levels or diminishes their quality"[153]. There has been debate about whether this approach can be applied successfully to companies which occupy a dominant position in an industry while continuing to provide services to consumers cheaply or even for free (e.g. by selling advertising to create revenue instead)[154]. Here, the strategic advantages of monopoly power in a market may not take the familiar form of an ability to charge higher prices to end-users and restrict output of narrowly defined quality of service. It has been argued that it is indeed possible to assess the behaviour of and impact on consumers within digital markets using this framework, but that differences in decisions between the US and EU – which also uses the consumer welfare standard approach – are partly the result of differing judgements made about the importance of a wide range of potential consumer outcomes[155]. Such outcomes could include the risk of limiting future choice in one market by shutting out potential rivals in a different, linked market today.

The intensity of antitrust activity in the US has varied over time and has been declining in recent years. The number of monopolisation cases brought by US agencies was 15.7 per year between 1970 and 1999, falling to 2.8 per year between 2000 and 2014[156]. This overall trend of reduced antitrust enforcement has continued into the Trump administration[157]. Comparable cases against online platforms appear less likely to be brought in the US than in the EU. For example, in 2013 the FTC entered into a settlement with Google and closed an investigation into suspected anti-competitive conduct relating to Google's self-preferencing of its own content within the Google search results page[158]. In contrast, the European Commission fined Google €2.42 billion in 2017 for promoting Google's shopping comparison services by placing Google Shopping at the top of search results[159]. The European Commission found that this was an abuse of Google's dominant position in online search. Google is currently appealing the decision.

No cases have been brought by US antitrust agencies based solely on data acquisition[160]. Data has however been an important factor in some merger enforcement, where, for example, licensing of data to third parties has been required for deals to be approved[161].

There are now increasing calls for expanded competition rules or enforcement in the US, particularly for big tech companies and where there are perceived data monopolies. The FTC has conducted a series of hearings on competition in the 21st century[162], including the appropriateness of the consumer welfare standard[163]. Many commentators and some prominent political figures have called for updated rules or greater enforcement[164]. In July 2020, the CEOs of Amazon, Apple, Google and Facebook all testified at a Congressional antitrust hearing; this demonstrated increasing momentum for regulatory action around antitrust and the big tech companies, but specific details on what such actions might be are still unclear, and may be further complicated by party political differences in opinion on appropriate approaches[165].

This is not a statement of government policy

3.3.3 Data for national security is a clear priority

US authorities have tended to prioritise national security – a stance that has commercial and international implications. Following 9/11, legislation in the US expanded surveillance powers, particularly through the 2001 USA PATRIOT Act. The 2015 USA FREEDOM Act extended some of these powers, although with additional limits on bulk collection of phone metadata, following Edward Snowden's leaks about this in 2013. There is evidence that programmes under these frameworks have been effective, but the magnitude of effectiveness is debated. For example, a 2014 report by the Privacy and Civil Liberties Oversight Board (PCLOB) found that at least 12 of 54 counterterrorism success stories citied by the US intelligence community used the National Security Agency's (NSA) domestic phone metadata programme (under Section 215 of the PATRIOT Act), but suggested similar results might have been achieved with other approaches[166]. It has been claimed that the programme helped prevent at least 10 terrorist attacks[167]. However, the NSA have reportedly recommended dropping the programme as the logistical and governance burdens do not justify its impact[168], and proposed bipartisan legislation would reform oversight of Section 215 and end the legal authority for the programme[169].

Law enforcement data access requests have sometimes faced tensions with the privacy policies of companies. For example, Apple refused to allow FBI access to the iPhone of a suspect in the 2015 San Bernardino shooting, although access was eventually secured through other means[170]. More recently, an increasing emphasis on privacy by online platform companies has led to concerns that measures such as end-to-end encryption could jeopardise legitimate law enforcement access, and in turn public safety[171].

Perceived differences in approach to national security data use have caused some tensions between the EU and US data systems (see Section 3.4). For example, the European Court of Justice found in 2015 that the EU-US Safe Harbour agreement for transfer of personal data was invalid, because "national security, public interest and law enforcement requirements of the United States prevail [over the scheme]"[172] in a way that enabled "interference… with fundamental rights". The replacement for the Safe Harbour scheme was the Privacy Shield framework, but this was recently invalidated as a lawful transfer mechanism in the *Schrems II* case by the ECJ[137], on account of US national security laws and practices, such as a lack of effective redress mechanisms. The CLOUD Act clarifies US law enforcement access to data held by US companies overseas and enables bilateral agreements to facilitate cross-border data sharing in the investigation of serious crime. The first of these agreements was signed between the US and UK in 2019[173], and further discussions have begun with the EU, and formal negotiations with Australia[174]. The European Data Protection Board suggested in 2019 that a further international agreement may be necessary to clarify the legality of US access to data stored in the EU under the CLOUD act[175].

3.3.4 Open government data is a recent federal priority

The OPEN Government Data Act mandates federal agencies to publish information openly in standard, machine-readable formats[176]. This includes an expectation that data is open by default unless there are clear reasons for it not to be (e.g. confidentiality or national security). Initiated in 2017, the bill had bipartisan support and was signed into law by President Trump in 2019[177]. Before this, federal government data had been provided through data.gov since 2009, growing to over 200,000 datasets[178]. In the most recent available rankings, which pre-date the 2019 law, the USA was ranked 11th (Global Open Data Index[179]) and 9th (Open Data Barometer[180]). The latter, in 2017, ranked the US notably higher for economic than government or social impact[181]. In line with this, other metrics find that the US has a particularly high number of for-profit companies utilising open data[182]. Challenges in realising the ambition of recent legislation have been noted, including lack of

buy-in, shortage of capability and resource, challenges in working across silos, and difficulties in implementation of best practice[183].

Some local governments in the US also prioritise use and sharing of data. At least 34 cities in the US maintain open data portals, but there is significant variation in data collected, intended uses and outcomes, and the mechanisms by which data is made available[184]. For example, cities vary in the degree to which they monitor how and how often data is being used[185]. Some cities, including Seattle[186] and Louisville[187] are focused on public accountability through data and make data generated by the city openly available to the public. Some cities also aim to use data to support ambitions such as smarter traffic management. Standardised metrics may be required to further assess the impact of city data initiatives on social and economic outcomes.

Driving factors for open data have expanded from addressing calls for transparency to a broader focus on government effectiveness and the potential of data to support the economy. Freedom of information requirements have existed in the US since the 1960s[188], with a long history of amendments expanding or restricting coverage, for reasons of privacy or national security for example. More recently, the importance of accessible and usable electronic data for this purpose has been emphasised[189], and federal data strategy now clearly recognises secondary use by "researchers, entrepreneurs, and the public" as a benefit[190]. Potential benefits of use of open data are discussed in Section 7.2.

3.3.5 International openness and free data flows are prioritised in trade deals, but digital barriers are used for other strategic goals

The US has prioritised market-led approaches and removal of digital trade barriers domestically and globally, in line with domestic objectives. The US's international trade strategy seeks to include measures to promote an open, interoperable internet[191] within potential and existing trade agreements[d] of which the US is a part, with a view to ensuring cross-border data flows and preventing policies restricting this such as data localisation[192]. (See the next section for further detail.)

Alongside this, the US has restricted the import, export and use of certain data-related technologies, particularly those that carry a perceived security risk. For example, Chinese telecom company ZTE has been banned for 7 years from buying US tech/software[193] in response to alleged sanctions violations, and Huawei and its suppliers have also been barred from using US tech/software[194]. Historically strong restrictions on export of potentially sensitive technologies such as encryption tools have been somewhat relaxed over time but still exist[195], while further controls on data-related technology such as AI have been considered[196].

Historically, the US approach to digital trade has been characterised by a desire to protect the free market of the internet and digital services[197], but that is seen as consistent with privacy protection. For example, US-led trade agreements often require some level of privacy protection and recognise that governments may take different approaches to data protection (as discussed in the next section). This may reflect a balance between supporting the US tech industry and the values and rights of citizens[198].

3.4 International agreements and cross-border regulations

These discussions of the Chinese, EU and US data systems have highlighted the range of stances that can be taken by mature systems globally. They show that very

[d] Even where the US is no longer a party (e.g. the Trans-Pacific Partnership, which was replaced by the Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP)).

contrasting approaches are possible, through varying uses of government levers in order to achieve differing social, economic and security aims. These are however not the only approaches that are possible, and we have shown that there are also some substantial differences in national or local approaches within each regional system.

Data flows internationally, and individual regional systems don't exist in isolation. The approaches of individual regions exert pressure elsewhere, groups of countries may compromise or reach agreements to facilitate beneficial data use, and individual regions may take actions in an attempt to reduce perceived international privacy or security risks.

The number of international data regulations is increasing, driven by a variety of economic, security and privacy concerns, and affecting all domains. The cumulative amount of data regulation has been increasing more rapidly recently, including the modification of older regulations to account for new technologies, uses and risks[12].

Recent regulatory changes affecting cross-border digital trade have mostly been restrictive. According to the OECD, among measures affecting digital trade between 2014 and 2018, 79% were restrictive and 21% liberalising[199], and this overall trend continued up to the end of 2019[200]. Measures affecting infrastructure and connectivity, including data flows, are seen as the most important. Real and perceived differences in values and approaches, including the levels of protection and assurance offered to citizens, could drive some of this restriction on international data flow, as we have seen above between the EU and the US.

There are existing and proposed international frameworks to promote interoperability between data systems, but none are comprehensive (see Table 7 for details). World Trade Organization rules (WTO) could also apply to digital trade frameworks, although assessing the suitability of data regulation under the General Agreement on Trade in Services (GATS) or General Agreement on Tariffs and Trade (GATT) is complex and would be sector-dependent[12]. Proposed or considered changes at the WTO include reform to GATS to cover data flows, and a plurilateral[e] agreement on e-commerce between over 75 countries[201]. The latter could include measures on data similar to recent Regional Trade Agreements (see below) but could be opposed by emerging digital economies such as India looking to first develop their own independent national e-commerce policies[202] (see Sections 2.5 and 6.3). Similarly, while for now a WTO moratorium on tariffs for electronic transfers of data is maintained (and has been made permanent in some of the regional agreements discussed below), several WTO members have recently expressed concerns that this might not be in their economic interests[203]. There are also varying opinions globally on the use of digital services taxes, levied according to where service users are located. Multilateral efforts to develop a tax framework are ongoing at OECD level but progress has been slow, while some countries have recently proposed or implemented unilateral measures[204]. Such taxes and the extent to which they vary, might affect cross-border digital business models, with implications for international data flows.

Recent Regional Trade Agreements (RTAs) often include harmonising and/or liberalising data measures, at least between the parties involved. Such measures are present in CPTPP and USMCA[205], and other existing/proposed RTAs such as the Peru-Australia Free Trade Agreement, the Australia-Singapore Digital Economy Agreement, the US-Japan Digital Trade Agreement and the New Zealand-Singapore-Chile Digital Economy Partnership Agreement. International agreements tend to have some similarities but are not always consistent in their strength and direction (see Table 7 for some examples and comparisons with different international frameworks). EU-led agreements typically require full compatibility with international standards on personal data and privacy (and allowing parties to freely adopt these – e.g. GDPR), whereas the US is more likely to include stronger

[e] With a narrower group of signatories than all WTO members.

wording on prohibiting localisation, and referencing privacy frameworks such as APEC CBPR and OECD[206], which are often considered to be lenient in comparison to GDPR. These agreements often include exceptions for particular policy objectives such as crime or security.

The nature of international agreements and their harmonisation or fragmentation across the world will have profound implications. If international frameworks and trade agreements can lead to interoperable data regulation between regions with currently differing approaches, this could enable much greater flow and use of data – but it could also increase the risks and requirements around trust, and the need for verification of common standards. In the absence of large multilateral approaches, a proliferation of regional agreements and standards could strengthen regional blocs and their influence on emerging data economies. A continued trend towards restrictive regulations without multi- or plurilateral developments could support growing fragmentation and data nationalism. Developing countries and less developed countries may have specific governance issues to address. For example, it has been argued that developing countries should be more able to adopt restrictive data measures in order to create domestic business opportunities[207], although others claim that any benefit would be outweighed by wider economic losses[208].

This is not a statement of government policy

Table 7 – Key international agreements and frameworks with conditions on citizen data.

International Agreement	Signatories (Aug 2020)	Key Features
Asia-Pacific Economic Cooperation Cross-Border Privacy Rules (APEC CBPR)[209]	9	Voluntary, non-binding; requires certification with minimum enforceable standards. 9 principles: accountability; notice; choice; collection limitation; integrity of personal information; uses of personal information; security safeguards; access and correction; and preventing harm
African Union Convention on Cyber Security and Personal Data Protection[210]	14	Guidelines on consent, fair processing, purpose, accuracy, transparency, confidentiality and security
Council of Europe Treaty 108 (C108)[211]	55	Binding principles, requirements and rights for data processing. Parties "shall not, for the sole purpose of the protection of privacy, prohibit… transborder flows of personal data going to the territory of another party"
Council of Europe Treaty 223 (C108+)[212]	41	Binding principles, requirements and rights for data processing; limitation in international transfer to other parties only if serious risk of circumventing provisions, or party bound by harmonised rules of international organisation (e.g. EU and GDPR)
Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP)[213]	11	RTA; parties "shall allow the cross-border transfer of [personal] information"; prohibits data localisation requirements; requires legal framework for data protection
ECOWAS (Economic Community of West African States) Supplementary Act on Data Protection[214]	11	Regional act specifying required content of privacy and data protection laws of member states
General Data Protection Regulation (GDPR)[f,215]	43[g]	Comprehensive data protection framework, 7 principles: lawfulness, fairness and transparency; purpose limitation; data minimisation; accuracy; storage limitation; integrity and confidentiality; accountability. Requires adequacy with this to enable transfers to countries outside the EU/EEA
OECD Privacy Guidelines[206]	37	Non-binding. 8 principles: collection limitation, data quality, purpose specification, use limitation, security safeguards, openness, individual participation, accountability. Countries should "refrain from restricting transborder flows of personal data between itself and another country where (a) the other country substantially observes these Guidelines or (b) sufficient safeguards exist… to ensure a continuing level of protection consistent with these Guidelines."
United States-Mexico-Canada Agreement (USMCA)[216]	3	RTA; requires parties to have a legal framework for data protection, not prohibit or restrict cross-border transfer of information; recognises APEC privacy principles

[f] GDPR is an EU regulation rather than international treaty with signatories, but is included here for comparison.
[g] Includes countries in the EU/EEA and countries that have recognised adequacy with this data protection regime.

This is not a statement of government policy

4 Wider data systems: people and businesses

The evolving nature of citizen data use has played an integral role in economic development across the globe, including beyond the formal data economy. Citizen data is generated, held and used by many actors – not just governments – for a huge range of beneficial and malicious purposes. In this section we explore the role of citizens and businesses in shaping data systems, as providers and customers of citizen data. We investigate how variable citizen values are reflected in the data systems that citizens operate in, and the uncertain evidence around the relationship between reported values and observed behaviour. We also explore how business models of data-driven online companies have developed alongside differences in data governance, with US and Chinese companies particularly dominant.

4.1 Citizen values and behaviour

Differences in attitudes to data governance and use across the world – and underlying values such as privacy and trust – could be a key driver of differences in data systems. A large amount of survey evidence is available on attitudes to data use and privacy, particularly in the US, UK and EU. Responses have typically varied over time and between demographic groups. For example, one survey showed that concern over data privacy fell in the UK between 2012 and 2015, and a study from 2017 found that concern over data privacy correlated strongly with increasing age[217].

Fewer surveys have directly compared attitudes beyond Europe and the US. Where they have been done[218,219,220], China has often been among the countries where citizens report higher levels of trust and lower levels of concern in data use and privacy. Citizens in economically advanced countries sometimes report having less knowledge of how data about them is held and used by companies and governments, and some of these surveys report that citizens in developing economies have higher and rising levels of overall concern. In one recent global survey[219], only a minority of people tended to say they trusted a range of organisations with how they handle their personal data, with only 20% suggesting so for foreign governments and 39% for national governments. Only around a quarter said they had a good idea of what authorities held about them in the first place. Concerns and attitudes about privacy and data issues vary within and between countries, as shown in Figure 5.

Attitudes are also affected by the extent to which people believe they will benefit from sharing their data. Some of the surveys referenced above found that consumers in developing countries tended to value perceived benefits such as saving time and money, discovering relevant products, and better products and services, more than European consumers did. They also show significant variation between countries. One survey found that 40% of British respondents agreed that allowing companies to use data they collect about them "helps them provide you with products, services and information that better meet your needs", whilst only 27% did so in France and 58% did in India.

This is not a statement of government policy

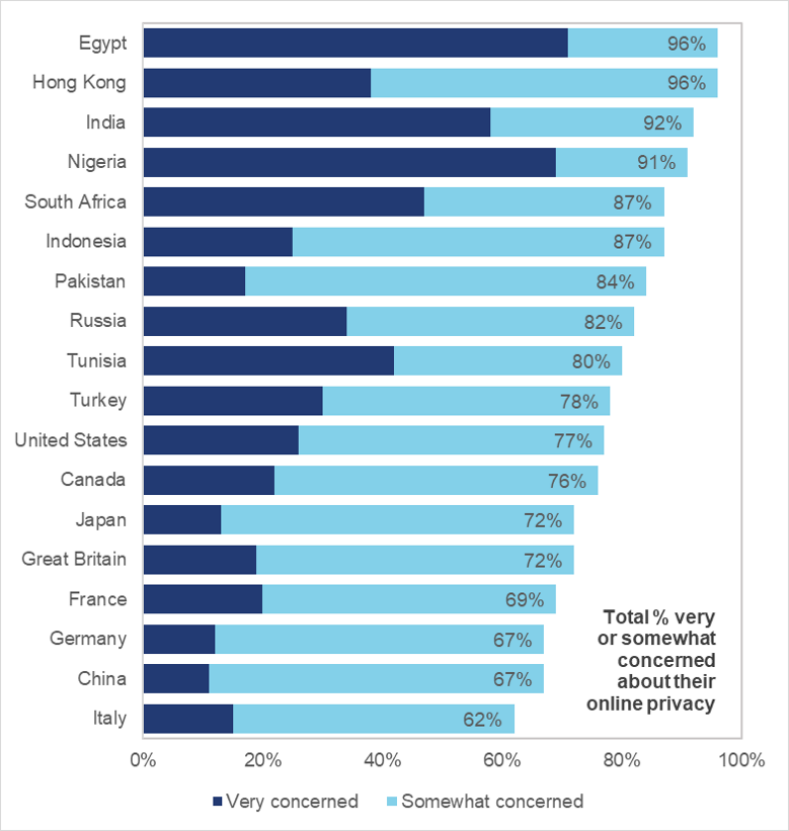

Figure 5 – Citizens' expressed concern about their online privacy for selected countries. Data from 2019 CIGI-IPSOS Global Survey[219].

Survey evidence has limitations. Questions are framed in different ways, there may be differences in the availability of contextual information to participants, and recent events or the existence of social desirability effects[h] could affect the accuracy of responses. A 2018 evidence review focused on the UK[217] found inconsistent terminology and definitions, a reliance on self-reporting for factors, such as level of understanding of data use, and a lack of qualitative data and demographic breakdowns. Where studies compare multiple countries, there are additional challenges: in some countries certain questions are not asked[i], and surveys in regions with low internet uptake can over-represent more urban and educated populations.

These limitations can be seen in conflicting views on Chinese citizen values regarding data use. There is survey evidence to suggest that Chinese citizens are more willing than others to share personal data in return for services[220], and that they have relatively high reported levels of interpersonal trust. However, it is also sometimes claimed that a deficit of social trust means that the population supports action by the government to address gaps and promote honest behaviour (as with the Social Credit System, Section 3.1.2)[221]. Qualitative evidence suggests that citizens may be less concerned about use of such systems as they assume data is already available to government[222]. Others have claimed that citizens may not always understand the risks related to privacy and data protection, that survey data is skewed by the setting (see above), or that events such as data breaches are underreported.

Asserting that 'Chinese citizens don't care about privacy' may be simplistic and misleading. Recent reporting from within China indicates at least some level of concern

[h] Where participants may be more likely to report a socially "correct" answer for their culture or region.
[i] E.g. questions about trust in government use of data in China and Saudi Arabia.

over blacklisting through the Social Credit System[223], and there have certainly been concerns around use of consumer data (for example, legal action against Baidu[224]).

Even where citizens indicate high levels of concern about privacy and data governance, it can be unclear whether this concern translates into to action. Some behavioural and empirical studies, mostly conducted in the US, identify a 'privacy paradox', with discrepancies between reported concern about privacy and observed actions. For example, one study found that that small incentives, such as free pizza[225], can significantly affect the likelihood of participants providing correct personal details, independent of their reported levels of privacy concerns. The cause is unclear, but explanations offered include a sense of resignation and common corporate practices around data use[226]. The aforementioned recent study of online platforms and digital advertising by the UK Competition and Markets Authority found many examples where online platforms' choice architecture and use of defaults inhibited users' ability to make informed choices about the way their data was used[21]. Where studies have attempted to measure the value of privacy to consumers, results are typically low. Estimates include a $0.75 premium to purchase from websites with greater privacy, to $44 to protect against improper access and use of personal data[227,228,229,230,231,232].

The way choices about data use are framed may affect behaviour. One study[233] found that participants were 5 times more likely to choose to protect their privacy for $2 when this was the default – that is, participants were more likely to be willing to *accept* the cost than to actively pay it. Also, several studies have indicated that citizens are unlikely to read terms and conditions before engaging with online services that may use their data[234,235,236]. This may be associated with a lack of understanding, and a lack of trust – which may contribute to the sense of resignation referenced above. Other contextual factors may also be important in determining citizen behaviours; a recent study from the American Marketing Association found that consumers were more willing to self-disclose personal information on smartphones than personal computers[237]. The authors suggested this might be due to feelings of comfort associated with smartphones, and a tendency to narrowly focus attention on the task at hand when using a smartphone due to the relative difficulty of working on a smaller device.

Privacy and security concerns could affect consumer behaviour, with potential economic impacts. An OECD analysis[12] suggests that privacy and security concerns play a key role in determining whether consumers shop online, although with significant geographical variation, for example with respondents in Switzerland and Slovakia far more likely to not order online due to privacy concerns (over 60%) than those in Ireland, Poland and Czechia (under 10%).

Citizen perspectives on the balance between privacy and security may be influenced by global events and change over time as historical memories fade. In surveys by the Pew Research Center between 2004 and 2016, most US respondents were more concerned that US anti-terrorism policies (see Section 3.3.3) had "not gone far enough to protect the country" than "had gone too far in restricting civil liberties" – with the notable exception of 2013, shortly after the Edward Snowden leaks[238]. However, the evidence is uncertain in this area. One UK survey in 2014 found that 90% of respondents thought that surveillance technologies improved national security, and 80% thought governments should use them[239], while another in 2016 found that a majority were not comfortable with surveillance of their communications without consent, in the context of a terrorist threat[240]. There may also be parallels with public perceptions around citizen data use for biosecurity and public health, and how this is balanced with privacy considerations, in the context of the COVID-19 pandemic. This is explored further in Section 10. More broadly, citizen values could also shift in response to increased awareness and experience of data use for public benefit, which could be felt differently by citizens across and within different regions.

Citizen attitudes, and behaviours around use of data about them, are likely to vary with the type of data being used, and the wider context of how it is used. Some different types of citizen data that might be collected, analysed and used by businesses or other organisations were discussed in Section 1.2 and illustrated in Figure 2. These can vary in (for example) sensitivity, value to different actors, and in how much information they can reveal about individuals or groups. A recent UK project based on focus groups and workshops found that participants expressed more nuanced values around the use of their data, once they explored the topic considering different types of data[241]. This variation is perhaps most clearly demonstrated by citizen attitudes to use of health and medical data, which can be complex, different to other types of data, and particularly dependent on who the data is being shared with and for what purpose[242]. Similarly, a recent study found that citizens assigned a higher monetary value to protecting financial and medical data than electricity usage and physical activity records[243]. As discussed elsewhere in this report, different types of data can be linked and aggregated to produce outcomes that may not be anticipated by citizens when consenting to sharing their data, further complicating attitudes and the ability to make an informed choice.

It is unclear whether reported attitudes and values are a driver of, or reaction to, existing differences in data systems. If genuine differences in citizen values related to data use across the world exist and persist, this could create challenges for convergence and harmonisation of data systems, for example through international agreements as discussed in Section 3.4. At the same time, countries can collaborate together and recognise that different cultures have different ways of delivering high standards of data protection and privacy.

A greater understanding of the likely behavioural response to measures intending to address data issues could be key to their success or failure. For example, policy interventions increasing transparency or individual control over use of data could have varied impacts depending on factors such as timing, framing, and existing norms. The ability of companies, governments or others to set the defaults around data use may play a major role in determining attitudes and setting data norms – particularly in emerging data economies and where existing models are exported through trade or by international governments (see Section 6.3).

4.2 Business models and use of citizen data

Service offerings and business models partly determine what types of citizen data are collected, how datasets are linked, and how data is used. Businesses may generate revenue directly from data, for example by selling targeted advertising space or aggregated data and associated insights. Businesses may also use citizen data to drive efficiency or maximise profit elsewhere, for example by informing product or service development, predicting demand, or identifying cross-selling opportunities. A company with broad service offerings across social media, e-commerce and payments will collect data in all of those areas. Data brokers may prioritise collection and linkage of broad citizen datasets that are of value to many sectors or focus on specific datasets of interest to a certain domain such as health and life sciences.

Citizen data is important to business models and is an input across many sectors, from automotive to life sciences. Businesses across sectors can also directly and indirectly create value by sharing the data they have collected with other businesses. This value could be from encouraging open innovation, building trust, improving market reach, addressing sector challenges, optimising supply chains, or gaining other insights through the shared data[244]. In the UK, regulated sectors, such as financial services, telecommunications, and energy provision, are examples of sectors where effective data sharing (when

customers want to switch providers) has made the customer experience better. Through Smart Data initiatives regulators and government are working to enable consumers and SMEs to make even better use of the data that firms hold about them, by allowing them to get help from third parties to make use of their own data, facilitating innovative new services, greater consumer choice, competition, and innovation[245].

Here we focus on technology companies and those operating online platforms, where the explicit use and analysis of citizen data forms a core part of their business models. Some emerging alternative business models, and the potential impacts of the COVID-19 pandemic on them, are discussed in Section 9.1.

Where revenue is generated from data itself, the need for customers can create privacy risks by widening access, or otherwise incentivise behaviour that could harm consumer welfare. It can incentivise actions such as selling access to datasets that could be de-anonymised (with or without the vendor's knowledge), or third-party data access for research or political advertising[246]. Where revenue is not directly linked to data, there may be more of an incentive for businesses to restrict access. For example, Apple relies comparatively less on advertising revenue than other tech giants such as Facebook and Google, instead making the majority of its revenue from hardware sales. Some have suggested they use privacy as a selling point to increase customer retention[247]. For businesses that offer free services and generate revenue from digital advertising, customers pay indirectly by providing their attention and data[248].

Today's dominant business models are shaped by the systems they grew in. Globally, US and Chinese companies dominate. US companies have typically relied on fewer revenue streams[249] at least initially, but their rapid rise to global dominance in certain areas has enabled them to diversify and consolidate data sources (see below). Some Chinese firms have developed wider offerings; coming slightly later to the market with favourable domestic conditions and a lack of existing infrastructure, many companies developed multiple revenue streams and data sources that link and reinforce (e.g. messaging, mobile payments, cloud, marketplace)[250].

Europe generally does not have data-driven businesses of the same size, despite having a higher population than the US and having an economy comparable with the US and China on most measures. EU citizens are therefore often reliant on US companies for some services, such as online search. In the top 30 internet companies by market capitalisation, there is only one European showing, Spotify, with a valuation of $26 billion[251] in 2019[252]. Among the world's 2000 largest public companies in 2019, none of those focused on online retail or computer hardware were European, and large European companies in software and computer services were mainly focused on business customers (e.g. SAP and Accenture).

SMEs form a large part of the European data economy. One study indicates that the majority of data users and revenue for companies in the EU data economy was associated with companies of less than 250 employees. This is to do with the large numbers of SMEs in Europe (99% of all businesses)[253], as average investment and revenue is much higher for larger companies and their number is growing faster than SMEs[254]. As of July 2020, 59 out of a total of 479 'unicorns' (private companies with a valuation over $1 billion) were located in Europe (not just the EU), including 25 in the UK. In almost all of these, data was a major element of business or operations (34 of these companies are fintech, AI, internet, e-commerce, data management, cybersecurity or hardware focused)[255]. However, the US and China had far more than Europe, with 228 and 122 respectively, which were again mostly in data related sectors such as AI and fintech. The UK has strengths in AI in particular, with established players and start-ups such as Babylon Health in the healthcare sector and

Darktrace in cybersecurity. Increased access to data from both the public and private sector (as well as improved skills supply) could further support growth in this area[256].

There is debate as to how much the European data system is a driver of, or reaction to, its data economy. There is evidence that more restrictive data policies, including privacy and data protection, have a cost to trade and productivity (see Section 5). However, some have suggested that wider factors, such as access to funding and attitudes to risk, have played a more important role in the development of the European data economy[257]. One suggestion is that "a market economy cannot function without trust, and the data economy is no exception"[258]. The potential of trust and ethics as a competitive advantage for Europe is often emphasised, for example in the development and adoption of AI[259]. This is sometimes challenged by those who highlight the power of access to larger datasets (see Section 8.1), or the potential for regulation to create barriers to positive innovation[260]. To some extent, policy may also be a reaction to the state of its data economy: the EU may be able to afford to lead the way in data protection regulation because it lacks the dominant tech giants that may be vulnerable or resistant to such regulation (although see further discussion on this below).

In the US, many internet companies initially relied on narrow revenue streams to build global scale. In 2018, the majority of Google and Facebook's revenue was from advertising (>80%; >90%), Apple and Microsoft's from product sales (>85%; >70%), and Amazon's from retail (>75%)[261]. This is not always the same as major sources of profit. For example, in 2018 over 50% of Amazon's operating income was from cloud services[262]. US companies are dominant in UK online search, social media and mobile operating systems[5]. At a global level, these trends persist, with Google dominant in search (>90%) and Facebook in social media (>60%)[263]. The scale and user base developed through the core offerings of such companies has supported later expansion, shifts or pivots into e.g. cloud computing, subscription services[264], digital currency[265] and projects in sectors from health to mobility[266].

The US data economy has developed alongside permissive, free market policies in areas from privacy to competition. While US companies received early government support, often in the form of public R&D programmes[267], there has been relatively little intervention in the collection and use of citizen data, mergers and antitrust, and similar policy areas[1] (see Section 3.3). Many companies have grown by acquiring competitors or potentially valuable companies, often significantly increasing the volume of data held. From 2009-2019, the 5 largest firms (Google, Amazon, Facebook, Apple, Microsoft) made over 400 acquisitions, none of which were blocked by competition authorities, and very few of which were scrutinised[5]. Often large firms have acquired a highly capable firm operating in an adjacent or overlapping space, for example Google acquiring Youtube in 2006 and advertising technology business Doubleclick in 2007, and Facebook acquiring Instagram in 2012[268]. As discussed in Section 2.2, it has been speculated that some acquisitions of smaller companies have been motivated by a desire to acquire their data assets or complementary sources of data. Companies that offer a range of services have more opportunities to gather rich varied data on consumer behaviour, which can then be used to improve services and better target advertising and so may provide a competitive advantage.

More recently – alongside increased consumer and policy interest – the largest internet companies all support some increased regulation, including forms of federal privacy law in the US[269]. There are concerns that this support comes from the belief that such regulation could serve business rather than consumer interests, by acting as a barrier to entry for smaller enterprises and consolidating the market positions of big tech companies[270]. There are also concerns that that online platforms are incentivised to interpret existing data protection regulation in a way that entrenches their own competitive advantage, for example by denying third parties access to data that is necessary for targeting,

attribution, verification and fee or price assessments, while preserving the right of third parties to use this data within their own platforms[21].

In 2019, Chinese companies made up 9 of the top 30 global internet companies by market capitalisation, with Alibaba and Tencent both valued around $500 billion, up by about 100% since 2016[251]. Chinese business overall has seen huge growth: the share of total revenue among Chinese companies tracked in the Fortune Global 500 has increased from 6th ($1.1 trillion) in 2008 to 2nd ($7.9 trillion) in 2019[271]. Spending on all R&D, largely financed by business, has increased rapidly in China, reaching $463 billion in 2018 (second only to the US) but still only 2.1% of GDP[272].

Internet companies in China often developed wider infrastructure alongside their early core business, providing diverse data and revenue streams. For example, in the early 2000s many companies developed payment systems to overcome challenges with patchy online banking and credit/debit card coverage. Alipay overtook PayPal to become the largest mobile payments system in the world in 2013[273], and non-bank online payments exceeded debit card expenditure in China after 2015[274]. These factors may have played a part in China's recent plans for a central bank digital currency, although the role of Alipay and others in this is uncertain[275]. Similar businesses include Meituan which has expanded from a group-buying app to providing more than 30 services from food delivery to travel. This starting point has positioned many companies for further expansion, for example with Alipay's rebranding as ANT Financial in 2014[276] and subsequent moves into credit scoring[277] and facial recognition payments[278].

Chinese companies dominate the domestic market and are expanding and competing elsewhere. Domestic firms dominate virtually every citizen-data-related sector in China[279]. In many cases companies still rely on domestic activity for the bulk of their revenue (92% for Alibaba[280] in 2018) but recent shifts are towards acquisitions and expansion overseas. For example, Tencent and Alibaba are investing heavily in India and Southeast Asia[281]. Wider programmes such as the Belt and Road Initiative support this (see Section 3.1.5). More recently, companies such as ByteDance have had increasingly rapid success in Western markets (ByteDance's TikTok was the 4th most popular US iPhone app in 2019)[282]. This can come with political pressure and regulatory challenges – for example in the US, TikTok agreed to pay $5.7 million to settle FTC allegations of a violation of the Children's Online Privacy Protection Act[283].

Rapid growth and expansion of data businesses in China has occurred alongside protectionist international data and digital trade policy; and perceived risks may be a barrier to expansion in some markets. China has data and trade policies that can act as barriers, and it has been suggested that this, along with the large domestic user base, strong government support for "national champions", wider industrial policy, and favourable government procurement[284], have supported the growth and expansion of domestic companies. More recently, perceived economic security risks and political developments may be leading to trade policies in other countries that could slow expansion into their markets[285].

Data brokers operate in all these regions. Companies such as Acxiom, Oracle and Experian use citizen data from many online, offline and cross-device sources to develop insights and products from consumer segmentation[286] to credit scoring. In 2016 most large data brokers were based in the US and make the majority of their revenue there[287]. US data brokers are subject to scrutiny from the FTC[288] and others, and regulated in some specific cases (see Section 3.3) and states (such as Vermont[289]), but are not subject to comprehensive regulation regarding use of citizen data in all contexts. In the EU, where such activities are subject to regulation such as GDPR, some data brokers and those providing them citizen data have faced complaints[290], assessment notices[291] and fines[292].

These have been in response to perceived or actual violations such as sharing sensitive personal data with other organisations without obtaining informed consent. In China, companies are subject to provisions in the Cybersecurity law and others (see Section 3.1) regulating the collection and sale of citizen data[293], with penalties including fines and prison time for offences. Data brokers often claim that they keep individual consumers' identities anonymous. However, critics suggest this is misleading, as the sensitive data that brokers gather and link together, such as device location, can effectively be used to identify people and should be considered personal data under GDPR[294] (see Section 7.3 for a discussion of the evidence around this).

5 Impacts of data regulations

Regulations regarding the use of data can be effective in achieving desired outcomes or can have unintended negative consequences in other policy areas. Effectiveness may depend on the resources dedicated to enforcement and involve partners and agencies in other countries or regions. This section explores emerging evidence on the enforcement and effectiveness of recent data regulations, their economic and social impacts, and how they might affect business practices, international trade, and interactions between regional data systems.

5.1 Enforcement and effectiveness

Some evidence is emerging that major data regulations are being effectively enforced. In the first year of GDPR, most national authorities registered an increase in data regulation activity, with more than 280,000 cases, including more than 140,000 complaints and more than 89,000 data breach notifications, and over 440 cross border cases[295]. As of July 2020, almost €176 million of fines have been issued in at 24 countries, with the largest single fine being a €50 million decision against Google[296]. There has also been a reported increase in awareness about data protection rights among EU individuals compared to 2015[295]. A recent report from the European Commission evaluating GDPR after two years[297] suggested that overall it was considered to have successfully met its objectives, but expressed concerns also shared by others over country-to-country differences in enforcement and associated dedicated resources, and a lack of effective cooperation between national data protection authorities on cross-border cases[298].

There is evidence that the GDPR right of data portability has been implemented very differently by organisations. In response to requests in one study, some organisations provided machine-readable data, others provided screenshots or paper, and in one case the request caused a data breach as data relating to other people was provided[9]. Initiatives such as the Data Transfer Project (DTP) could help to standardise approaches to data portability. The DTP is a collaboration of organisations including Apple, Facebook, Twitter, Google and Microsoft, aiming to build an open-source, service-to-service platform to enable users to move their data easily and securely between service providers[299]. This could support new entrants to the market, as with the recent integration of Solid (Tim Berners-Lee's model of control for online social data) and Mastodon (an open-source social media platform similar to Twitter), but may not address all perceived concerns around effective competition and consumer choice[300]. Data portability initiatives and rights more broadly could also support uptake of some alternative models of data governance as discussed in Section 9.2, especially when collectivised.

Impacts of regulations on accepted standard business practices remain unclear. For example, the effect of GDPR on online advertising practices is still to be seen. Very early and limited data suggested that GDPR reduced the number of trackers per webpage. However, it was found to have a disproportionately negative effect on smaller advertising technology companies compared with the tech giants[301]. A more recent study found that there was a temporary reduction in the number of 'ID syncing connections' (e.g. cookie syncing) around the time of GDPR implementation, but that this later stabilised and that overall the amount of tracking was unaffected. The study authors suggest this may be due to companies not significantly changing their business practices but taking time to make their existing processes compliant with GDPR[302]. In June 2019, the UK Information Commissioner's Office (ICO) found that the sharing of personal information by advertising technology companies' 'real time bidding' practices remained "disproportionate, intrusive and unfair" and parts of the industry needed to improve in order to ensure compliance with GDPR and other data laws[303]. It is still to be seen what impact this will have in the industry,

and in May 2020 the ICO paused their investigation due to the COVID-19 pandemic[304]. A study on Internet of Things devices located in the US and UK found that most network traffic from both US and UK devices terminated in the US – although UK devices contacted slightly fewer third parties – while Chinese devices were more likely to send traffic to China (e.g. to cloud providers)[305].

5.2 Economic and wider social impacts

Evidence is also beginning to be explored around the system-wide economic impacts of different approaches to data regulation. Estimates generally suggest that restrictions on data flows and use impose costs on trade and productivity[306]. More recent evidence has attempted to look at these separately, with a study from ECIPE suggesting that trade in services is mostly negatively affected by restrictions in cross-border data flow[307]. Particularly restrictive countries such as China and Russia could be limiting services imports because of this[308]. Restrictions on both cross-border data flows and domestic data use may also have a negative effect on domestic productivity. Another ECIPE study found that more restrictive policy regimes had a significant negative effect, particularly for data-intensive sectors such as retail and information services[309]. Even in an internationally isolated system, there may be some productivity benefits from lower restrictions on data use. This could also apply for public good uses of data such as research.

The choice of regulatory model must also account for ethics, privacy and security, and how these could feed back to the economy and data use. Both privacy and security concerns could affect consumer behaviour (see Section 4.1), so there is an economic and business case for data protection. There are also costs and risks related to cybercrime and national security. Reports claim that cybercrime may have global annual costs of up to $600 billion[310], which may be partly mitigated by data policies that reduce vulnerabilities to such crimes[311], and that enable appropriate data sharing between enforcement organisations. A lack of controls to ensure data is used appropriately could increase other risks such as targeted disinformation (see Section 7.3). As such, some economists have suggested that deciding the optimal amount of privacy should not be led by economic, trade or innovation concerns alone[312].

We are still developing approaches to understand the impacts of different regulatory approaches, and particularly how to balance economic, social and security outcomes. Fundamental concepts, such as how to measure international data flows in trade statistics[313] and how to value data across sectors and uses[4], are not yet well defined and agreed. While there is emerging evidence on individual areas and outcomes as described above, there does not appear to be a clear way to compare outcomes across all these areas. Some have suggested that international policymaking could be based on a triangle between openness, efficiency and security. Well-designed data infrastructure (from datasets through to standards, and the organisations and communities involved with data) could help to manage potential trade-offs between citizen privacy, national security, economic productivity and international trade and competitiveness[314].

5.3 Impacts on data systems and interactions

Emerging regulatory models may be taken up in other regions, with impacts on how global data systems interact. If a major initiative such as GDPR is globally successful, then companies, business models and regions set up early to work within it could have first-mover advantages as the standard spreads. Alternatively, there could be a first-mover disadvantage if there are unforeseen negative effects such as trade barriers or higher costs of compliance for smaller businesses, restricting use of data and stifling innovation. This could lead other regions to take new approaches, potentially creating new regulatory blocs and requiring rollback or changes to existing models. At the moment, GDPR is already

having an impact on standards outside the EU; countries such as Argentina and Japan have reached adequacy agreements, and similar laws have been developed in California[315] and Brazil[316], as previously discussed. For developing countries in particular with more limited resources, there is a risk that they implement data regulations initially designed for higher capacity countries, but without the requisite capabilities and resources to put in place effective enforcement and accountability mechanisms; this could negatively affect the level of trust in and overall effectiveness of such regulations[317]. Potential future trends are discussed in more detail in Section 6.3.

This is not a statement of government policy

Part 2: Future paths

6 Global trends are likely to be mirrored in data systems

Shifts in demographics and politics, increased uptake of technologies and other megatrends will affect the dynamics of data systems. This section explores how such changes could interact and impact future paths, focusing on: how macroeconomic shifts could change the power of certain data blocs; how changes in demographics and internet uptake may affect the number of internet users and associated volumes of data; how wider geopolitical trends may play out in data systems; and how rising energy demands and environmental concerns could affect computing, communication and use of data.

Shifts in economic gravity

If current trends hold, the influence of emerging and developing economies, particularly China and India, will increase relative to the US and Europe.

The data systems that emerging economies are adopting now are likely to affect these dynamics, and the power of certain blocs.

Internet demographics

Growing internet uptake in developing economies will shift internet user demographics.

By 2030 Africa would only need to increase uptake to just over 40% to have more people online than Europe.

Energy and environmental impacts

Data-driven technology is seen by many as vital to enabling more efficient use of resources and supporting sustainability.

Communication, data storage and processing have energy demands and could lead to increased environmental impacts.

Geopolitics and data governance

An expanded model of data sovereignty may now be spreading, and elements of the regional data systems may be actively exported.

Businesses may have to comply with conflicting rules in different regions or miss out on market access in some countries.

Figure 6 – Summary of some key global trends that are likely to be mirrored in data systems discussed in this section.

6.1 Shifts in economic gravity

In recent years emerging markets and developing economies have become increasingly important in the global economy, and now account for over 75% of global growth in output and consumption[318]. China and India have had consistently higher GDP growth rates than the USA or Euro Area since 2001[319]. Projections from 2017 suggest the global economy will shift eastward to 2050, with the largest 7 emerging economies surpassing the G7 in their share of the global economy[320].

From 2005 to 2017, developed countries' importance as suppliers of both goods and services declined, although in services they still accounted for about two thirds of exports[321]. Similarly, the proportion of trade between developing countries has been higher than that with developed countries since 2011[322]. Between 2017 and 2018, the largest such trade growth measured by UNCTAD (United Nations Conference on Trade and Development) was seen in the East Asia region[j]. Much of the growth within this region was due to trade with China, and China itself outpaced all other regions tracked[321].

[j] Although the largest percentage growth rates relative to 2016 were seen in Sub-Saharan Africa, South Asia and West Asia/North Africa

If current trends hold, the influence of emerging and developing economies, particularly China and India, will increase relative to the US and Europe. The incentives for global businesses to operate within and trade with the data systems of these economies may increase in parallel. This could mean that other regions lose some of their soft power over data, or lead to protectionist moves in economic or data policy. The impact on data systems may depend on how far any regulatory measures affect digital services and data standards.

The models of data governance and use that emerging economies are adopting now are likely to affect these dynamics, and the power of certain blocs. For example, while Europe's overall share of the global economy might decline, the spread of GDPR elsewhere could give the EU a long lasting influence on rules and norms in data systems, potentially dependent on whether any updates and modifications to GDPR are also followed elsewhere. Alternatively, further uptake of a Chinese model of data governance would create a different scenario.

More recent evidence highlights uncertainty. A business-as-usual projection would assume China continuing to grow faster than other regions, and overall continuing international actions to reduce global trade barriers. However, both of these show signs of slowdown or change – GDP growth in China has slowed[k] from an average of 7.1% (2011-18) to 6.2% (2019), and a series of trade barriers (such as new tariffs) introduced since 2018[l] appear to be affecting trade flows and prices[323]. The ongoing impacts of the COVID-19 pandemic may also have long-term and as yet uncertain impacts on the global economy and growth in different regions.

6.2 Internet demographics

A large proportion of citizen data is generated through interaction with the internet. More than half the world's population is now online, according to 2018 International Telecommunication Union (ITU) estimates[324]. The impact of internet usage growth on the trajectory of global data systems is complex, because of variation at the level of regions, countries and users, and the relationship between internet uptake and data generation and use.

At the level of ITU regions (see Figure 7), internet use is near saturation in Europe, with 80% uptake in 2018, and joint lowest regional growth rate since 2005 along with the Americas[324]. In contrast, 2018 uptake in Africa stands at only 26%, but it is the region that has shown the strongest growth since 2005, when uptake was only 2.1%. In absolute terms, Asia-Pacific had far more internet users in 2018 than either Europe or the Americas, despite having the second lowest uptake (46%). However, there is considerable variation on smaller scales within these regions, between and within countries. For example, in 2017 Eritrea had 1.3% uptake while Morocco had 61.8%. In 2017, China and India were the countries with the largest number of internet users, and China alone had more internet users than Europe or the US[325].

In 2030, at median UN population projections[326], Africa would only need to increase uptake to just over 40% to have more people online than Europe (assuming no change in European uptake). Even with no change in uptake across the board, the difference between Europe's user base and the largest region, Asia-Pacific, would increase, while the gap between Africa and Europe would decrease. Figure 7 shows projections for number of internet users in different regions in 2030 for different levels of internet uptake. The Global

[k] Although other developing regions such as India and Brazil have not experienced similar slowdowns
[l] Particularly between the US and China

System for Mobile Communications Association estimate that Sub-Saharan Africa will have a mobile internet penetration rate of 39% by 2025[327].

As with the wider economy, the internet and global data economy are likely to experience a shift eastward and towards developing economies. This could mean businesses and governments must deal with differences in data governance between a wider range of countries or regions. This could increase the impact of these differences, particularly if they create barriers to exchange or processing of data, or privacy/security tensions as data flows across borders.

It is not clear how internet user numbers will translate into other factors, such as volume of data generated, breadth of activity, or market size. For example, in Q4 2018, Facebook reported far more total revenue and revenue per user in the US and Canada than any other region, despite having about 4 times as many monthly active users in Asia-Pacific[328]. This was because revenue per user was more than 11 times as high in the US and Canada. Similarly, bandwidth use may differ markedly between regions. More broadly, even if increasing volumes of data are produced by emerging economies, the commercial or other value extracted from such data may still be mainly realised by advanced economies with substantial existing data sources, infrastructure and dominant businesses[329].

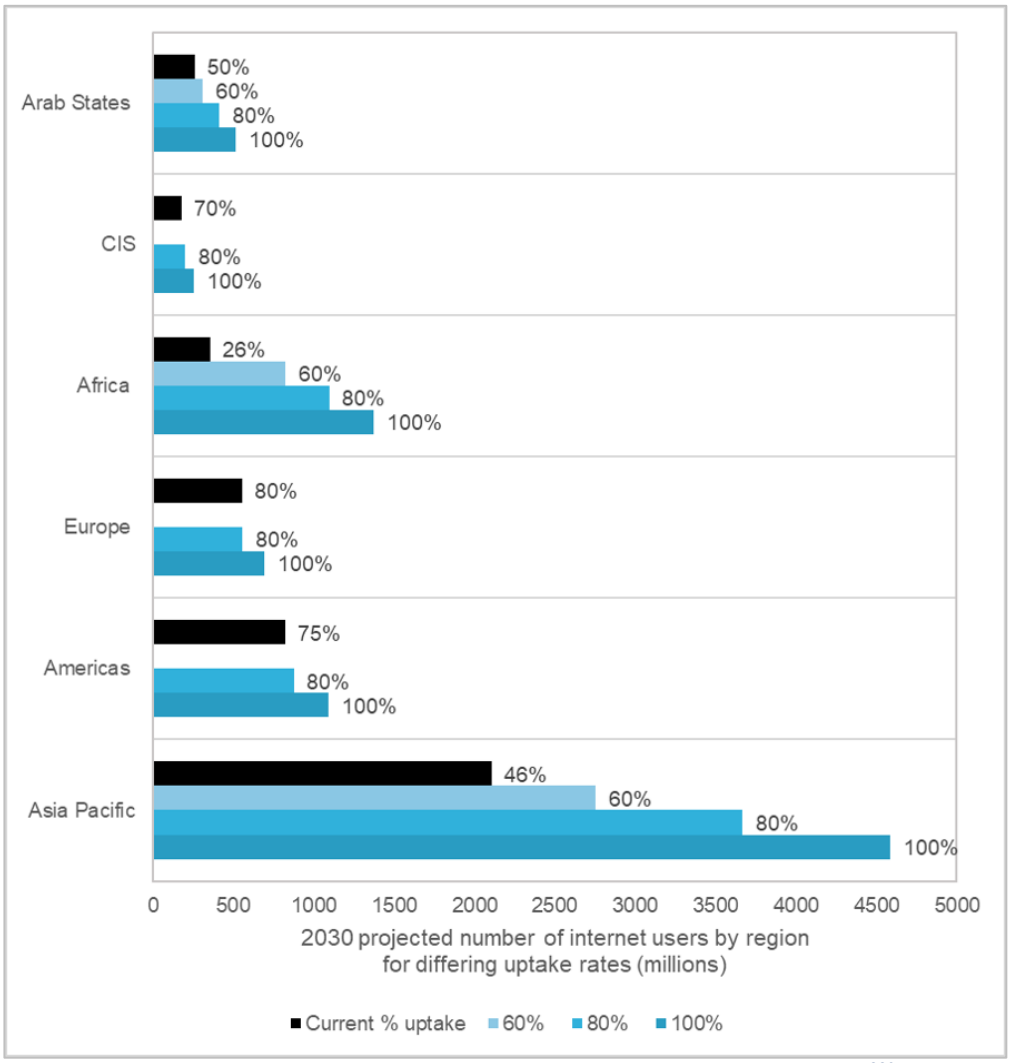

Figure 7 – 2030 projections for number of internet users in millions in each ITU region[330] at median UN population projections[326] and different levels of internet uptake. Note that ITU regions do not cover all countries included in UN population projections, and only those countries within each ITU region have been counted. CIS is the Commonwealth of Independent States.

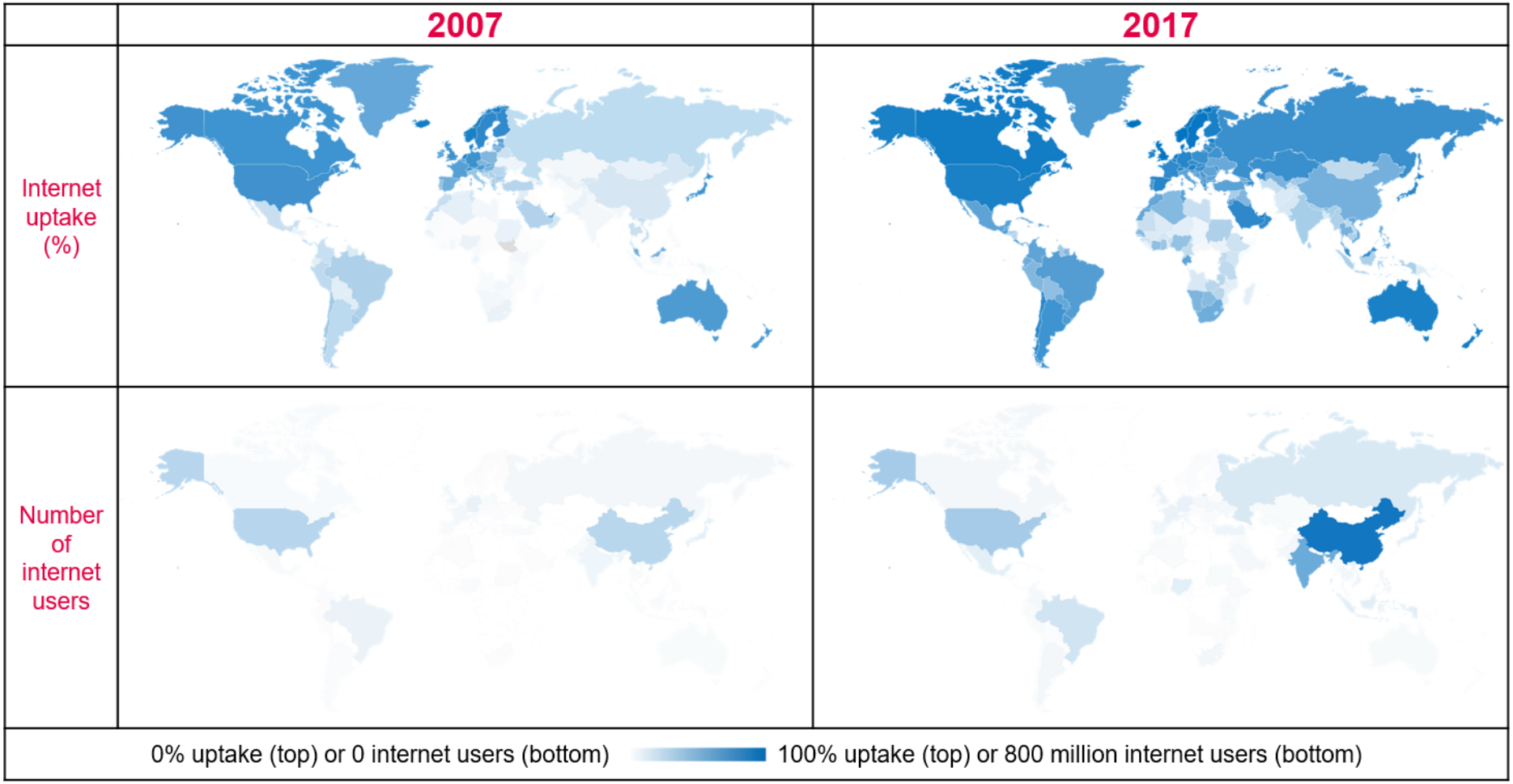

Figure 8 – Changes in internet uptake (top panel) and number of internet users (bottom panel) across the world in 2007 and 2017. Data from the World Bank[325]. Microsoft product screen shots reprinted with permission from Microsoft Corporation.

This is not a statement of government policy

6.3 Geopolitics and data governance

Data systems have developed partly as the expression of geopolitical aims, and partly in reaction to the actions of other regions. For example, at the time of the emergence of the global internet, which mostly grew from US initiatives, European and US governments and private actors advocated a multi-stakeholder governance model, while China and Russia advocated a data sovereignty model, giving more power to national governments[20,331]. The former approach prevailed when management of unique identifiers transferred from US oversight to the Internet Corporation for Assigned Names and Numbers (ICANN) in 2016[332]. EU and US values around data governance generally diverge more elsewhere (see Section 3).

An expanded model of data sovereignty may now be spreading, with measures that affect both internet freedom and wider data use. Such measures include data localisation and retention laws, restrictions on the use of encryption, requirements for stronger user identification, mandated access to citizen or company data, censorship and controls on international data flows. This is commonly described with reference to China, but elements can be seen in many countries, from a Russian sovereign internet bill[333] adding to existing localisation requirements[334], to proposals in India which originally required local storage of personal data[335], which some have seen as a way to counter US dominance in the data economy[336,337].

According to Freedom House, elements of the Chinese data system are being actively exported. Companies have exported network equipment to at least 38 countries, technology such as facial recognition to at least 18 countries, and government training on data governance to at least 36 countries[82].

Extraterritorial aspects of the GDPR[338], such as its application to data processing outside the EU when related to targeting of subjects inside, and adequacy requirements for international data transfer, could be viewed similarly as exporting EU data governance and values. Its influence on emerging regulation elsewhere, and as a factor for EU market access, can also be viewed through a geopolitical lens. Conditions on data have also formed part of major trade deals as described in Section 3.4.

The approaches of existing and emerging data economies may be influenced by pressure from external data systems and businesses. Of the countries tracked by UNCTAD[339], as of August 2020 66% have clear legislation on data protection, reduced to 55% in Africa and Asia. Far fewer have full or partial adequacy with GDPR[340]. Geopolitical or economic pressure, whether direct or indirect, could contribute to the spread of systems aligned to certain values; and the potential for reaction against this by other regions or systems.

The data collected in systems aligned to certain models may be more readily available to companies or governments within that system. For example, Chinese companies may gain access to more diverse data for algorithm training from external regions adopting the Chinese model (as mentioned previously, Guangzhou-based company Cloudwalk have a reported agreement to use image data from Zimbabwe to train better facial recognition algorithms for darker skin tones). Companies operating in regions with GDPR adequacy may be more easily able to consolidate and use European data for analysis.

Businesses may have to comply with conflicting rules in regions with different models or miss out on market access in some countries. Where market access has a financial burden, this may challenge newer entrants and strengthen the position of established players. In addition, in the case of stringent EU data regulations, internationally operating companies may adopt compliant policies across their global operations as the

price of participating in the large EU market, in order to avoid the costs associated with running separate compliance regimes[341]. For example, Facebook, Google and Microsoft have each adopted a single global privacy policy that mirrors the GDPR. This can lead to what law professor Anu Bradford has termed 'the Brussels Effect', where EU regulations become widely adopted even outside the regulatory region. This could be considered another form of exporting EU data governance and values internationally.

Shifts in values and politics across the world, and the success of new multilateral approaches, may affect all these geopolitical processes and interactions. If political leadership changes in one or more regions, that could drive harmonisation or divergence. For example, presidential or other election results in the US might affect the likelihood the US aligns with or diverges further from current or proposed approaches to competition of digital markets in Europe. If certain regulations are seen as having negative impacts on businesses or citizens, this could lead to rollback of existing regulations and lack of further uptake. This is explored in our 'Deregulation' future scenario in Section 11.6.

6.4 Energy and environmental impacts

Data-driven technology is seen by many as vital to enabling more efficient use of resources and supporting sustainability. For example, DeepMind have developed an AI-driven system to improve the efficiency of Google's data centres, with claimed energy savings of 30%[342]. This approach required large amounts of data, with more than 75 million training examples needed. Machine learning could benefit climate change mitigation, for example in forecasting energy supply and demand, supporting electric vehicles, and optimising systems within buildings to reduce the amount of energy used[343].

However, communication, data storage and processing also have energy demands and could lead to increased environmental impacts. If digitisation, communication and data use continue to grow, then specific technologies, infrastructure, policies and incentives will help determine these energy and environmental impacts, and how they feed back to shape other elements of data systems. By some recent estimates, global digital energy consumption is increasing about 9% a year, and its share of global greenhouse gas emissions could reach almost 8% by 2025, in a worst-case scenario[344]. Another study estimates it could reach 14% by 2040[345]. The energy intensity[a] of the digital industry may be increasing, one estimate suggests by almost 4% per year[344], while overall world energy intensity is decreasing. Related greenhouse gas emissions could vary considerably around the world, depending on energy mix and carbon intensities. For example, developing countries investment in clean energy fell in 2018[346], and the emissions of comparable digital infrastructure (data centres, laptops, etc) are estimated to be larger in China and the US than in Europe[344].

Data centres have been estimated to contribute around 19% of digital energy consumption[344]. Their energy demand could actually decrease in the near term despite increasing use, due to factors including a shift to larger scale data centres with higher capacity, bringing improvements in efficiency[347]. However, theoretical or practical limits may curb the potential for further progress[348]. Future improvements may require more radical developments in computing technology (see Section 8.2) or data centre infrastructure.

The interplay between data use and energy consumption could be both a driver and outcome of specific data systems. Increasing energy demands could drive innovation and efficiency, but these could lead to rebound effects further increasing demand[349]. If attitudes and approaches vary across the world, this could affect the dynamics between data systems. For example, domestic incentives (or stringent restrictions) to encourage

[a] Units of energy per unit of GDP

sustainable processes in specific regions could lead to offshoring of data processing and emissions, and/or cause an imbalance in approaches to data processing that are available in different regions, potentially leading to wider economic and commercial impacts. Alternatively, concerns over environmental impact and different approaches between regions could limit international data transfer. General global or domestic progress around wider energy and environmental impacts could reduce the need for data-specific considerations, for example if there is a large shift to renewable energy, or if broader energy policies incentivise ‘green’ computing.

This is not a statement of government policy

7 Data growth is likely to increase opportunities and risks

Recent measures and predictions suggest rapid growth in the number of data-generating devices, and the associated volume of citizen data they generate. This section explores the potential impacts of growth in data volume and variety, including economic, social and other benefits, and risks related to malicious and unintended misuses of data. We focus on managing privacy risk associated with larger datasets, the benefits of data use and re-use for the economy, research and public services, and risks related to data security, online targeting and system vulnerabilities.

7.1 Growth in data volume and variety

Most reports have recorded and continue to project significant increases in the volume of general data generated, and the number of devices involved – although estimates and projections are inconsistent and often revised. For example, in 2011 Cisco projected that there would be 50 billion connected devices by 2020[350], but in 2019 they revised this down to 28.5 billion by 2022[351]. The widely cited statement that "90% of data was created in the last two years" has been in circulation since at least 2013[352] and is hard to verify. Despite a general lack of high-quality evidence with clear sources and methodology, almost all reports agree that there will an increase in the number of devices and amount of data created. As an example, global internet traffic has increased significantly in recent years (see Figure 9). However, several of our interviewed experts highlighted an unexpected reversal of this trend as a potential significant disruptor to expected future trajectories.

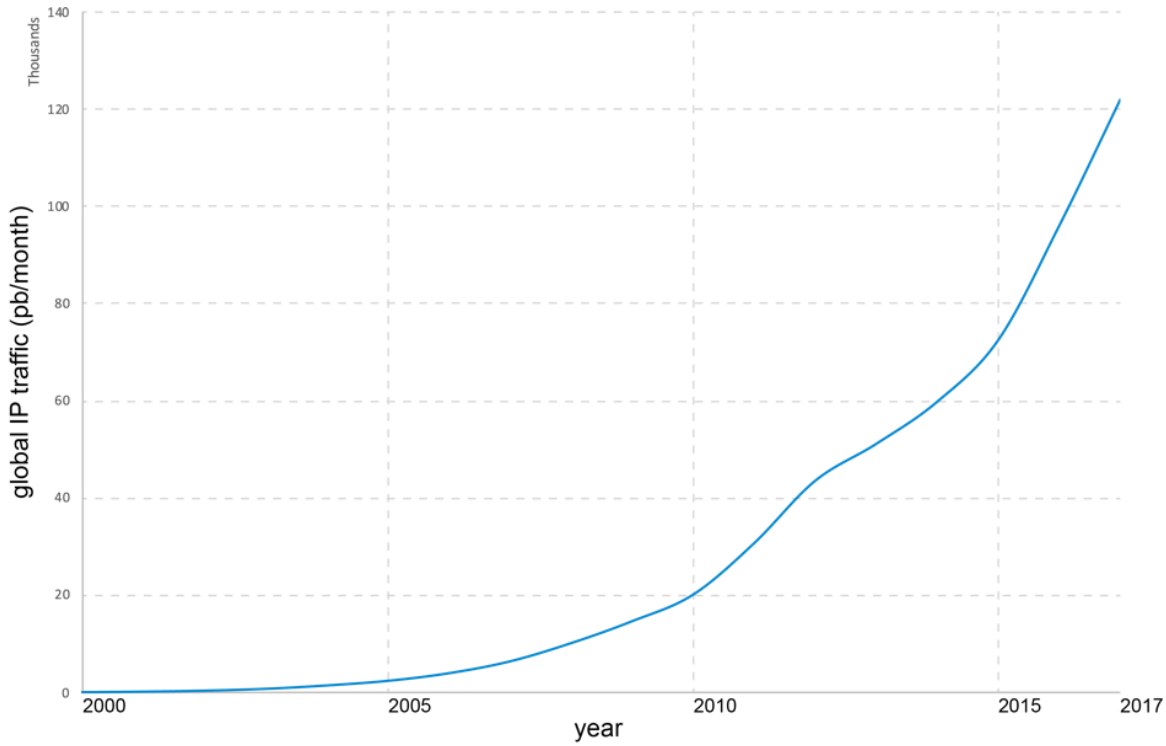

Figure 9 – Global internet traffic in pb (petabytes) per month, from the year 2000 to 2017. Data from Cisco[353].

The variety of devices that collect, process or transmit citizen data is increasing, both in the consumer market and beyond. For example, it has been reported that over 180 million people in the US now use connected TV devices[354], and emerging smart city projects will use large numbers of data-gathering sensors monitoring aspects of human behaviour and activity[355]. This will contribute to what has sometimes been referred to as "datafication" of society, whereby an increasing proportion of peoples' everyday lives are recorded and

represented as data, for example through interactions with objects that previously were not digitally-enabled such as listening to music or reading books. This has potential implications not only for individual privacy, but also for the many industries and organisations which may benefit from access to this data, and wider society through how this data is used[356].

New technologies are likely to create entirely new sources and types of citizen data. For example, brain-computer interfaces could collect very sensitive citizen data. These have great potential for applications, from mind-controlled prostheses to enabling generation of speech from thought. However, they already raise ethical questions over the privacy of thought, individual agency and personal identity[357].

Existing data sources may also be used in new and unexpected ways in future. For example, image data from social media sites can be used in training facial recognition algorithms[358], and potential future uses of genomic data include predicting behavioural traits. New uses of existing data may come about as new approaches to analysis are developed, as discussed in Section 8.1.

This growth in data could bring enormous benefits (see Section 7.2), such as increased effectiveness and efficiency of public services, better outcomes from research and development programmes, and direct improvements in consumer welfare. The size and scale of these benefits may not be evident immediately, for example with the long-term research and health potential of collecting DNA data alongside routine care[359].

It may also bring increased risks (see Section 7.3), including security and privacy risks[360], and new malicious uses of data. If many more organisations are collecting and collating citizen data through more devices and sensors, that may create unforeseen challenges to interoperability, liability or control over data.

The balance of these benefits and risks will depend on the interaction between commercial incentives, consumer behaviour and regional regulatory and data governance environments. Advertising driven business models may enable more services to be delivered for zero price to consumers but may increase the likelihood of tracking or targeting using citizen data, including data sharing with third parties.

Growing volume and variety of data is likely to increase the potential of re-identification from datasets now considered anonymous or de-identified. Many studies have already demonstrated re-identification of individuals from supposedly anonymous datasets[361]. Well-known re-identifications include Netflix subscribers in a large “anonymised” ratings dataset[362], and the identification of the governor of Massachusetts by linking “anonymous” medical data with voter lists[363]. While many of these cases were due to poor anonymisation practices, even datasets published at a high level of aggregation (such as census data), using well-established statistical disclosure control methods[b], are increasingly at risk of re-identification unless effective measures to protect privacy are used. Technological developments to attempt to mitigate these issues are discussed in Section 8.3. Advances in computing power and algorithms mean that database reconstruction attacks, where underlying data can be computed from aggregate tables, are now feasible for some large releases[364]. Recent studies have also challenged the assumption that the release of partial or incomplete de-identified datasets provides plausible deniability and protects anonymisation, by demonstrating that it is possible to estimate the likelihood of correct re-identification even from heavily sampled datasets[365]. These developments may challenge current distinctions between personal and non-personal or anonymised data, which will have technical as well as policy implications. For example, GDPR has strict

[b] Such as suppressing summary data where there are small groups of respondents

definitions saying that if data can possibly be used to re-identify individuals, then it counts as personal data and is protected as such.

New techniques are being developed to enable data disclosure and analysis while protecting privacy[366,367] (see Section 8.3), but this may have to be continuously balanced with changing risk. In future it is unlikely to be sufficient to satisfy some pre-defined threshold of anonymity and then release a dataset, because the appropriate threshold may change with the availability of new data to link to, or new analysis techniques. The potentially variable definitions of personal and identifiable data across the world will affect the balance of benefits and risks felt in these regions.

7.2 Increasing benefits from use of citizen data

Increased volume, variety and linkage of data could support a wide range of social good and benefits to citizens – often far beyond the original purpose[c] of the collected data.

Data is an increasingly fundamental input to the economy, supporting jobs and productivity and benefiting consumers. While it can be hard to measure its specific contribution to such complex outcomes, information has always been used to improve performance, and data-related activity has become more central to many businesses (see Section 4.2). Recent estimates of the global digital economy range from 4.5-15% of GDP, accounting for 39 million jobs in 2015[1]. Studies have estimated the combined value of "free" services (often involving data-driven business models) provided to consumers in the US at thousands of dollars per person per year[368]. However, emerging literature is challenging these findings, particularly for social media services where it is complicated to assess overall value to users[369].

Citizen data is being used to develop new products and services across most sectors. For example, Transport for London's (TfL) open data supports development of apps such as Citymapper and brings estimated total savings and economic benefits to travellers, businesses and TfL of up to £130 million per year[2]. Data portability initiatives that enable standardised mobility of citizen data could enable greater consumer choice and drive innovation. For example, the Open Banking initiative in the UK already has over 130 third-party providers of services including financial managers and account aggregators. Further proposed uses for this initiative could have wider benefits, such as rapid determination of eligibility for legal aid, new services to help financially vulnerable people, and enabling charities to access data that support their activities[370]. Similar approaches are now being explored in other countries around the world, including the US, the EU, India, Japan, Mexico, Nigeria, and Singapore.

Administrative citizen data can drive research for the public good. In New Zealand, the Integrated Data Infrastructure links a wide range of data about citizens across education, health, justice and beyond, providing insights into society and the economy to aid policymaking[371]. In the UK, the Digital Economy Act gives the UK Statistics Authority and the Office for National Statistics (ONS) powers to provide accredited researchers with better access to administrative data to support research projects for the public good[372]. Administrative Data Research UK works with ONS to support use of de-identified data in research projects, providing evidence to address societal challenges, and where public service provision could be improved across areas such as education and crime[373]. Citizen data can be used to discover and understand potentially hidden inequalities experienced in

[c] Note the purpose limitation principle under GDPR requires that personal data is not processed for a new purpose unless it is compatible with the original purpose, there is new consent, or a there is a new legal obligation

societies. The UK government's Race Disparity Unit collects, analyses, and publishes government data about ethnic disparities, and supports departments to 'explain or change' any observed disparities.

Citizen data generated and held in the private sector can also be valuable beyond its original purpose. In the UK, the Consumer Data Research Centre manages research access to data from private and public organisations, to provide insights into consumer behaviour that can drive economic growth and improve society, for example using smart electricity meter data to identify and support vulnerable energy consumers[374].

Wider use and sharing of citizen data can have direct benefits to public service users and providers. Digitalisation of public services is already delivering benefits to governments and citizens, from Estonia's digital ID supporting 99% of services to be online[375], to use of AliPay reducing waiting times for public services such as booking hospital appointments in China[252]. In the UK, a report from the National Audit Office said that better use and sharing of data within government more broadly could bring a huge range of benefits for government, businesses and citizens, including: reduced fraud, better decision-making, improved efficiencies, innovation, and economic benefits[376].The report also discusses potential risks of harm associated with *not* using data effectively, giving the Windrush situation as an example, where data about individuals' status was not adequate to identify those without rights to live in the UK. Barriers to effective data sharing and use within government may be legal, technical, cultural, or be associated with lack of public trust, potentially due to lack of transparency, accountability, or citizen control in how their data is used[377]. Effective data sharing between government agencies and other public and charity sector organisations is also essential in protecting vulnerable children and families[378].

More sophisticated linking and analysis of citizen data may be particularly beneficial in healthcare provision and research. This could include linking of different types of citizen data from across the public and private sectors, for example linking clinical data with user data from commercial wearable or IoT devices, to provide new insights and enable innovation. Deep-learning approaches to medical diagnosis have the potential to match or exceed the performance of healthcare professionals[379]. Charities may facilitate the collection of data from individuals with specific diseases in order to inform research into that disease and improve patient outcomes; for example the UK Cystic Fibrosis Registry, to which over 99% of people in the UK with cystic fibrosis have consented to submitting their data[380]. New models of data governance may emerge (see Section 9.2) which further enable this kind of targeted data gathering for healthcare research. The global COVID-19 pandemic has demonstrated the importance of effective data use and sharing across borders for public health; see Section 10 for a case study describing the differing international approaches to use of citizen data in the pandemic, particularly around digital contract tracing methods.

Wider linkage and re-use of citizen data may be key to monitoring and evaluating development initiatives. For example, an analysis of over 170 household and census datasets in Africa has enabled estimation of female educational attainment at the district level, revealing within-country and regional differences that could help in planning better targeted interventions, supporting the Sustainable Development Goals[d],[381]. Mobile phone data has been used to support more accurate mapping of migration following natural disasters, allowing creation of proxy census maps and estimation of poverty levels[382]. More recently, location data has been used to develop higher resolution maps of economic inequality in the US[383].

While volume and variety of citizen data grows, it is likely that so will potential benefits from its re-use. What is uncertain is how this will interact with risks, externalities

[d] UN Sustainable Development Goal 4 – "ensure inclusive and equitable… education for all".

and other factors, and feedback to shape the nature of data systems. Higher citizen engagement and awareness of beneficial uses could drive support for wider data use or increase anxieties over risks. Even if data systems continue to differ between regions and nations, international agreements on data sharing for specific uses such as public health or natural disasters could develop, involving both public and private organisations[384].

7.3 Increasing risks from use of citizen data

Data breaches and incidents of malicious access to citizen data have been increasing overall over the past decade. According to one central monitor, data breaches have generally increased in frequency between 2005 and 2019 (see Figure 10). Recent increases are mostly driven by hacking, the largest cause of data breach each year since 2009. Even where the total number of recorded breaches declined in 2018, the number of exposed consumer records actually increased 126% from 2017, to over 446 million[385]. Large data breaches, of more than 30,000 records, have mostly increased year on year since 2009[386].

Use of citizen data to create and target disinformation appears to be increasing. There is increasing evidence of social media manipulation across the globe, with one estimate suggesting 70 countries had organised manipulation campaigns in 2019, up from 28 in 2017[387] – although this may be due to changing methods of analysis. Of these, at least 19 appeared to be using data-driven strategies such as the purchase of targeted advertising on social media, or illegal micro-targeting with online and offline data, to direct messages to specific groups. There are also geographical differences in the use of certain platforms – for example while Facebook is the most common platform overall, organised use of platforms such as WhatsApp and YouTube may be more common in low- and middle- income countries. Previous reports noted that disinformation originating in China mainly used domestic platforms such as Weibo, but more recently platforms including Facebook and Twitter have taken down accounts or attributed campaigns to China, including some with foreign influence aims[388]. As with other elements of data governance, there is evidence that countries currently pursuing data sovereignty are training organisations in developing data economies[389].

While disinformation is not a new phenomenon, the ability to use citizen data to attempt targeting at this scale and scope of impact was not previously possible. Using citizen data such as Facebook likes, it is possible to infer demographic and commonly used psychological profiles. The accuracy of results varies: in one study, some categories (e.g. openness) were predicted more accurately than others (e.g. satisfaction with life); but predictive accuracy generally increased with access to more data[390]. In some circumstances, the use of such inferred profiles to deliver targeted messages has been shown to be effective in changing behaviours, for example adverts targeted to introverted or extroverted audiences, as inferred from data, had significantly higher click-through rates[391]. It has been reported that targeted ads linked to the Russian Internet Research Agency had almost 10 times the click through rates of typical Facebook adverts[392].

However, the impact of data-driven targeting on complex behaviour such as voting is much less clear. Studies have reported that personality traits shape responses to different types of targeted online messaging[393], so for example some people may be more susceptible to specific targeting methods; but there is also evidence that the general effects of political campaigning might be very small[394]. In many cases, engagement with disinformation or fake news may be highest in subgroups already sharing some of these views[395,396], but studies have also reported that repeated exposure to such material can increase its perceived accuracy[397]. Attempts to evaluate the impact of targeted disinformation on often-discussed events such as the 2016 US election have reached conflicting conclusions[398], and there is a lack of high-quality causal evidence[399].

Malicious actors may be able to use citizen data to target other kinds of attack. For example, existing attacks are already combining multiple payloads that can be activated at different times. Access to citizen data such as emails could be used to automate and prioritise which type of attack may be more successful – for example prioritising silently monitoring (of e.g. email or financial) information on a machine used by senior executives vs. locking or installing crypto-mining tools in others[400].

Even in situations that are not illegal or directly malicious, online targeting using citizen data may create risks not just around individual privacy, but also autonomy, discrimination and beyond. A recent review of online targeting by the Centre for Data Ethics and Innovation discusses in detail the potential risks and benefits associated with online targeting, and how the proposed online harms regulator in the UK can attempt to mitigate some of these risks[401]. Some of these risks may be more significant for people in vulnerable groups such as children and those with poor mental health. People may be targeted (even unintentionally) based on such vulnerabilities, for example by repeatedly suggesting specific content with personal negative associations or supporting addictive behaviours such as gambling. The report also suggests that online targeting may lead to social fragmentation and polarisation by narrowing the range of content recommended to people, and further recommending content that reinforces existing beliefs. At the same time, there are benefits to online targeting that service users appreciate, predominantly in navigating otherwise overwhelming volumes of online content, and easily discovering relevant information.

New technologies, and increasing access to existing ones, are expanding the scope of risks and enabling entirely new kinds of attack. For example, approaches to AI such as Generative Adversarial Networks have potentially malicious uses, including to guess user passwords[402], and creating realistic false images or video (e.g. for deepfakes[403] or to manipulate medical imaging[404]). If these technologies continue to become more readily available and easy to use for an average person with limited resources, for example with editing based on simple text commands[405], this could expand the range of malicious actors and threats. At the same time, the volume of data they require could decrease, as for example in recent approaches that have enabled generation of video from a few photographs[406], further increasing the potential availability and use of such technologies.

Increased reliance on machine learning technologies, and widening scope of use, could also introduce new vulnerabilities[407]. Existing examples include "data poisoning" attacks to manipulate training data used by machine learning systems and using "adversarial examples" which are designed to be misclassified by deployed machine learning systems[408]. Both approaches enable malicious results. For example, a recent study showed the potential to manipulate an autonomous vehicle's classification system to mis-identify a stop sign as a speed limit sign if an attacker places a simple sticker on it[409]. These approaches could also be used where sensitive citizen data is analysed with machine learning, for example in facial recognition or genomic databases. Defences are developing: technical ones such as increasing robustness by using adversarial examples in training datasets[410], and governance ones such as pre-publication risk assessments[408]. The degree of openness of datasets and capabilities could have positive and negative impacts on risk: more open approaches in areas of high risk could allow a wider range of expert review of potential vulnerabilities but could also risk exposing data or information to malicious actors.

The future interaction between advances in technology, and wider social, behavioural and other changes is uncertain, and has important consequences for how such malicious uses of data might be addressed. For example, there are attempts to develop automated technological approaches to identify and mitigate disinformation[411], while programs to increase media literacy could help people identify and resist its effects[412]. Changes in the type and scale of attack could affect social behaviour. At present, people continue to engage with many online systems despite frequent data breaches, but a shift

This is not a statement of government policy

towards integrity attacks where data is manipulated – say large-scale changes to medical records – could potentially change that behaviour.

This is not a statement of government policy

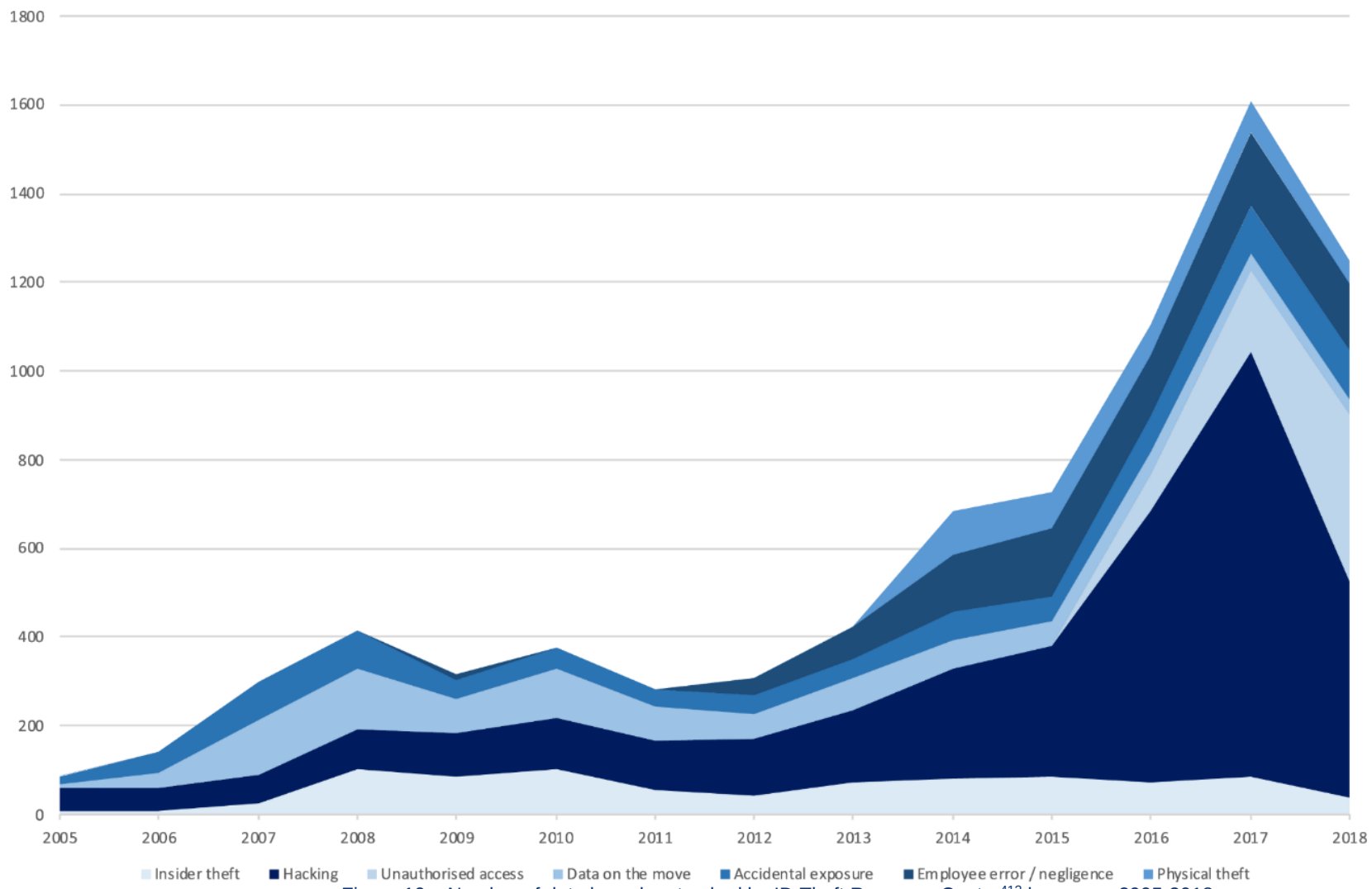

Figure 10 – Number of data breaches tracked by ID Theft Resource Center[413] by cause, 2005-2018.

8 New technologies

Developments in technology could change the challenges and opportunities related to citizen data. This section focuses on the relationship between analytical capability and access to large and varied data; how developments in computing could affect citizen data issues; and the potential of privacy-enhancing technologies and methods to manage risks around re-identification in anonymised datasets while enabling greater data use.

8.1 New approaches to analysis, and the value of large volumes of data

Access to larger volumes of data is generally thought to come with some analytical, commercial or other benefit to the data holder. Improvements in machine learning so far have been driven by increases in available dataset size, as well as computing power and the complexity of algorithms[414]. For example, increasing dataset size has been found to significantly improve performance of deep learning for vision tasks[415], although with somewhat diminishing returns. Organisations with access to the largest datasets are likely to therefore have some advantage. However, the blanket assumption that 'more data is better' does not always hold and is likely to depend on the context, timing and use of the data[4].

Now, dataset diversity is becoming increasingly important. Groups of people that are poorly represented in datasets used for research or training algorithms could see less accurate results when systems are deployed. Lack of diversity is a familiar problem in genomics: in 2009, 96% of participants in genome-wide association studies (GWAS) were of European descent, only decreasing to 81% in 2016 and 78% in 2019[416,417]. This reduces the quality of research based on such data, meaning important insights will be missed. Benefits arising from such studies may fall disproportionately to those who are overrepresented. In facial analysis, a 2018 study found that 2 datasets used for developing and benchmarking gender classification tools contained images from overwhelmingly lighter skinned subjects (80-86%)[418]. The study also showed that 3 commercial tools (from Microsoft, IBM and Face++ respectively) all had far higher error rates when classifying images of darker-skinned females (21-35%) than light-skinned males (up to 0.8%). A follow-up study published in 2019 found significant reductions in these error rates (2-17% for darker-skinned females), and that this was particularly stark when compared to other tools that were not targeted by the first study (from Amazon and Kairos). The authors suggested that by highlighting the issue and raising public awareness, the targeted companies had been motivated to address this issue through technical, governance and transparency changes. Another 2019 report evaluating over 100 available facial recognition systems[a] found that many had large differences in false positive rates between demographics, with higher (worse) rates for African and East Asian people. Many algorithms developed in China reversed some of these trends, with lower false positive rates for East Asian faces[419].

In many cases, acceptable error thresholds for such uses of citizen data have not been defined or accepted in policy or among citizens. Companies and governments are taking different approaches to address concerns, with potentially important implications for data systems. Example approaches include efforts to increase the diversity of commonly used datasets[420,421], investments in expansion of data access, data infrastructure and standards in different regions (such as the previously mentioned Chinese export of facial recognition software to Zimbabwe)[84], methods to measure bias and fairness in algorithmic systems[422,423], and restrictions by national or local government on particular uses of such tools[424]. The technical accuracy of such systems is only one factor to consider when deciding the appropriateness of their use. Even large, comprehensive datasets can be

[a] These differ from classification tools as they aim to identify whether an image is of a known person, rather than estimating a classification such as gender

biased by reflecting biased real-world practices that the data is based on, leading to the risk that such bias becomes embedded into algorithmic systems[26].

There is also much research on improving the generalisability of algorithms and their performance with smaller datasets. For example, in transfer learning, features of existing, pre-trained models are transferred to a new, related use, reducing subsequent data and training requirements. This is often used for new computer vision applications[425]. Research is also ongoing in few-shot or one-shot learning, where a new concept or rule is learned from just a few examples, or a single example. This is something that humans can easily do in many situations, but machine approaches have typically struggled with[426]. Between 2015 and 2019, a benchmark test showed progress on classifying handwritten characters using one-shot learning, but less progress on other tasks such as generating new characters[427].

Synthetic data is one approach to address lack of data for training or developing algorithms. Synthetic data is data which is generated using a model in order to have the overall properties of a real dataset, but composed of artificial individual data points rather than through real-world events or experiments. This has the potential to reduce resource intensive requirements to collect and/or label new data for a specific algorithm use. However, there are various concerns around the efficacy of synthetic data. A synthetic dataset is effectively only as good as the model underpinning it[366], and such models can be complex to create, requiring a detailed understanding of the system being modelled. If any important properties are not captured, analysis of the dataset can be misleading or inaccurate. A recent Royal Society report on machine learning methods suggested that "considerable caution is needed before relying too heavily on simulated data in many real-world settings"[428]. Synthetic data can still be used in ways that negatively impact people and potentially invade their privacy, even though their individual data is not in the dataset. For example, inferences, predictions or inaccurate decisions can be made about the groups of people being represented by the synthetic data.

In future, the approaches discussed above could reduce the need for large citizen datasets. However, this may not apply to all stages of deployment or development, as many of the methods above still require lots of data at least to develop initial models or approaches. In addition, these approaches are mainly focused on training algorithms rather than deployment, where real citizen data would presumably still be processed.

In some cases, more simple approaches that use less data could perform as well as newer data-hungry methods. One recent study of criminal behaviour argued that both a simple logistic regression model and human predictions, using as few as two pieces of information about a case, were as accurate in predicting reoffending as a commercial algorithm using 137 features[429]. While huge progress in data science and AI is likely to benefit society, some have taken results like these as an indicator that simpler approaches may be just as effective for some (often complex, social) problems[430].

As these methods and approaches develop, access to the most data may not be a key advantage for regions or organisations. Instead, the quality, diversity and type of information available could become more important, as could access to and effective use of the best algorithms or hardware. This would have significant implications for policy. For example there may be a need to improve the quality and diversity of public sector data (e.g. through improved collection mechanisms) rather than just enabling access, in order to achieve greatest public benefit; and competition interventions may need to target potential advantages other than sheer volume of data assets that benefit some online platforms. Timeliness of data may become particularly important; disruptive events such as the COVID-19 pandemic or climate change could have long-term impacts on people's behaviour, which would then limit the value of historically recorded data for AI systems and predictive technologies aiming to understand and model behavioural trends.

8.2 Hardware and computing

Over the last 50 years, the power of traditional computers grew exponentially, but now physical limits are slowing this trend, and new approaches to computing are being explored. Systems that use more than one kind of processor are becoming more popular, for example graphics processing units (GPUs), which are particularly useful for machine learning[431]. Many companies are developing and using application-specific integrated circuits, or new architectures to support tasks such as machine learning, for example Google's Tensor Processing Unit and Graphcore's Intelligent Processing Unit. In 2019, the US and China made up a just under 70% share of the performance of the top 500 supercomputers[432]. While such supercomputers are often used to process non-personal data, for example in physical sciences research, they also represent a huge capability in analysis of citizen data. One of the world's fastest supercomputers, Summit in the US, announced it had broken the 'exascale barrier' (speed of calculations greater than one billion billion per second) in 2018 analysing genomic data[433], and as of August 2020 was being used to support COVID-19 research[434].

Radically new types of computer could have a large impact on data systems in future. For example, quantum computers can in theory perform some types of calculation far faster than any classical computer. While small-scale demonstrations exist[435], a full-scale quantum computer is likely several years away[436]. Such a machine could have a profound impact on the use of citizen data, both negative and positive. It would be able to break many common encryption schemes based on factorisation, such that new methods to protect sensitive citizen data would be required. Such computers could also aid analytical tasks, as they will be able to search very large unstructured databases at high speed, which may become increasingly important as data volumes grow. Quantum computers may also have advantages in machine learning, optimisation problems and scenario planning. The benefits and risks that come about may depend on which organisations or governments first develop the technology, which due to the scale involved is currently only likely in global technology companies, major research institutions and national government programmes[437].

Another new type of computing is neuromorphic computing. This is inspired by biological nervous systems, for example with the use of electrical spikes similar to those in biological neurons. It has been suggested that neuromorphic approaches could reduce energy requirements and be particularly useful for some machine learning applications, supporting more powerful inference in smaller devices[438,439].

The history of computing has seen shifts in the use of centralised and decentralised models, from mainframes, to personal computers, and more recently to storage and processing of data in the cloud[440,441]. This could continue, with implications for where and how citizen data is stored and processed. For example, there is a resurgent focus on the potential of computing at the "edge" of the network: on or much closer to end devices and end users.

A shift towards more edge processing could address several problems. For example, increases in the number and variety of devices and the volume of data generated by these is a challenge, with potentially huge bandwidth requirements if this data is all transferred to central servers. Local processing could reduce the amount of data sent to the cloud. It could also enable processing that has to be done rapidly or in offline settings – for example in an autonomous vehicle processing vision data to avoid a crash, or the operation of drones in varied environments. Others propose that it could address privacy and security concerns, for example by reducing the amount of sensitive data held in large, centralised databases and enabling greater control over access. However, there are uncertainties around the burden

this could place on citizens to manage the privacy and security of their data (see Section 9.2).

With any shift away from the centralised cloud model, policymakers will need to consider several issues, such as ensuring the physical and cyber security of infrastructure that may be in less secure locations than cloud data centres; supporting development of standards for security and interoperability; and supporting methods to enable legitimate analysis use of data when it is distributed across devices and locations441. If increasing decentralised processing of data is beneficial for particular policy objectives such as energy or privacy, governments may be able to incentivise this via procurement requirements or other levers. Conversely, if the centralised cloud model continues to dominate, the location and scale of dominant cloud-based services may become more important from both a data governance and business competition perspective. For example, the Gaia-X project is a European cloud computing infrastructure initiative intended to enable European companies to compete globally and maintain EU data protection standards[442].

8.3 Privacy engineering methods and technologies

New privacy engineering methods and technologies could mitigate many risks and enable uses of sensitive citizen data that are currently impossible in some data systems. These focus on enabling wider processing, analysis and use of data without having to give access to all of it. Often referred to as privacy enhancing technologies (PETs), they are often very broadly defined – a recent Royal Society report notes a scope from "tape masking a webcam to advanced cryptographic techniques"[443] – and aim to preserve privacy in different stages of data collection, processing and release. The field of PETs is developing quickly, and while elements of some of these emerging technologies are very promising, further research will likely be needed before they can be utilised widely and effectively.

Differential privacy is a formal way of defining and managing privacy risk when releasing aggregate statistics and analysis relating to a dataset. It provides criteria for technologies and methods to satisfy[444], providing a mathematical guarantee against a wide range of privacy attacks, such as differencing[b] and reconstruction attacks (see risks previously discussed in Section 7.1). A particular criterion is that when an analysis of a dataset is released, it should not enable inference about an individual person that would not be possible if they were not included in the dataset. With repeated queries of a dataset, more information can be leaked, so there is often a prescribed limit of information disclosure after which a user is not allowed to make more queries. This limit is known as a privacy budget and is a quantitative measure of the level of accepted risk to an individual's privacy. Differential privacy is usually achieved by adding a carefully tuned amount of random noise to a dataset, either when collecting the data or when it is released – the latter requiring more trust in the data holder or intermediary. Differential privacy could be used to create differentially private synthetic data which retain properties of the real data, and can be used for repeated analyses or training of machine learning models.

Differential privacy could also be used to support release of national statistics such as census data. The US Census Bureau will use differential privacy to protect citizens' privacy for the first time for their 2020 census, although they have faced numerous challenges in doing so, including in sourcing qualified personnel, bespoke technical implementation, and in terminology[443]. Deciding an appropriate privacy budget has also been difficult, given the inherent trade-off between accuracy and privacy, both of which the Census Bureau are legally required to provide[445]. There are concerns that inaccuracies could particularly affect statistics relating to smaller populations living in remote areas, which could have equality impacts in areas such as funding allocations and health research[446].

[b] Using background knowledge about someone to infer their individual data from multiple statistics

Meanwhile in the UK, the National Statistician recently announced that the 2021 census might be the UK's last of its kind, as the Office for National Statistics are investigating whether existing sources of administrative data such as GP records can be used instead to get the same accuracy and richness of data at lower cost[447]. This approach will likely have its own methodological challenges in implementation, for example in quality assurance.

Some technologies are beginning to enable the useful processing of data without revealing sensitive information in other environments:

Homomorphic encryption, in contrast to standard encryption which protects data only during storage and transmission, enables some processing of data while the data remains encrypted. Fully homomorphic encryption would enable the widest range of computations to be performed on encrypted data but is currently at the research stage and is often inefficient. More restricted methods – referred to as somewhat, or partially homomorphic encryption – enable a limited number or type of computations. Partially homomorphic encryption is already used in a number of products, including by the NHS, to support de-identification of patient records across the health service[448]. Homomorphic encryption could enable more secure outsourcing of data processing to untrusted cloud providers or third parties, potentially between regions with differing data systems.

Secure multi-party computation, which may utilise homomorphic encryption, is a subset of cryptographic approaches that enable computations on data from multiple parties, without revealing the input data of one party to another. For example, data could be analysed from multiple banks or financial institutions to identify fraudulent activity, without the individual datasets being revealed to each other. This could enable a wider range of positive data use, such as proportionate security access to wider public sector databases, where private information is revealed only after agreed criteria, or only if there is a match between two datasets. It could also enable public good use of commercially or otherwise sensitive datasets, for instance joint analysis of public and privately held genomic data[449].

Trusted execution environments (TEEs) are secure areas inside main computer processors where code is isolated from the rest of the system. They have the benefit of having relatively low costs to utility and performance, as computation is performed on unencrypted data. This also means any computation can be performed easily. Products that use TEEs are widely available. However, there are security vulnerabilities, particularly 'side-channel' attacks which use other routes to gain information, such as memory and caches. Many current TEEs also have low memory, so only limited data can be processed at one time.

Federated learning is an emerging approach which can allow the training of machine learning models without centralising the data in a datacentre (e.g. while data remains on user mobile devices)[450].

These approaches have the potential to enable greater use of citizen data for good, but all come with trade-offs. They bring costs in utility, accuracy or performance compared with use of data without privacy protection. For example, with differential privacy, the need to add noise incurs a cost to accuracy (as some useful information is lost), and with homomorphic encryption more computing resources, time and power are required. They may also be less effective or appropriate for certain types of citizen data, such as unstructured data (e.g. unformatted video and text files).

Even with huge development and adoption, technology alone would not solve all problems regarding the use of citizen data to perform analyses relating to individuals. Where privacy is not the main or only concern, analyses may still be seen as unwanted or unethical even with use of PETs or synthetic data, as previously mentioned. For example,

several methods to deliver online behavioural targeting while preserving privacy have been suggested, from the use of metadata[451] to privacy-preserving targeting architectures[452]. While this may protect the privacy of individuals' data, this would not address often-cited concerns around manipulation or discrimination as discussed in Section 7.3. Development of global standards around socially acceptable uses of citizen data or the ability to audit may be required, particularly to build trust in international uses.

While differential privacy can provide guarantees for managing privacy risk, in many situations there is no agreed way to decide what level of risk is acceptable for a certain use. All the above means is that such technologies cannot be considered in isolation, and must be designed and operated in the context of the wider business or government processes they are being used for, considering the intended purpose and aims, types of data involved, and risk appetite. Further guidance may also be required to explain to organisations how these technologies can be used in compliance with existing and future privacy and data protection regulations.

There could be unforeseen or unwanted consequences from over-reliance on technological solutions. For example, a large sample size may be needed to implement differential privacy in practice. This could have implications where privacy and competition policy intersect, as incumbent dominant businesses with access to large datasets would be more easily able to implement this approach. Established privacy enhancing methods such as end-to-end encryption have already raised concerns over competition[453] (consolidating operations across a business in a way that makes interventions harder to enforce) and national security[454] (preventing access to data for law enforcement). New security threats or vulnerabilities could emerge alongside these technologies, and further emerging technologies or approaches could render the solutions proposed now ineffective in future. Without both the correct expertise and education around key concepts, such as the privacy budget, techniques could be misapplied in practice. Alternatively, the use of such technologies could face public resistance if, for example, a well-understood risk associated with a use of differential privacy occurs without effective communication of this risk beforehand.

Overall, the different emerging technologies presented in this section serve to demonstrate the range of potential disruptors to the nature of citizen data systems. These technologies variously have the potential to further enable citizens, governments and businesses to realise the benefits associated with citizen data, bring about new risks, and shift the balance of power between different actors in the system. It will therefore be important to pursue research, development and investment into key disruptive technologies, while being mindful of what they can and cannot achieve.

This is not a statement of government policy

9 Political, social and economic shocks

Before the COVID-19 pandemic, we asked experts to outline potential disruptors that were either likely or had the potential to significantly change the path of data systems (even if unlikely). Many were mentioned, from large-scale social or employee pushback against technology to conflict between major geopolitical powers. This section focuses on three areas that were frequently raised: economic shocks and changing business models; disruptive models of data governance which could shift power balances; and conflict or rapid political change.

9.1 Business models and socioeconomic shocks

Alternative models for data-driven businesses are constantly emerging. Existing business models have developed alongside rapid increases in the availability and value of citizen data, as previously discussed in Section 4.2. New models are being developed, some of which focus on addressing privacy and security issues for users. Table 8 gives examples of businesses that make money from citizen data in novel or shifting ways. Some of these use new models of data governance described in more detail in the next section. This business innovation is likely to continue – but the exact future shifts in models, and their degree of success competing with or replacing current approaches is to be seen.

Table 8 – Alternative emerging models for data-driven businesses, with differing revenue sources.

Company	Source of revenue
Hub of all things (HAT)[455]	Personal data store where third-party organisations pay a fee to access/use data
Digi.me[456]	Personal data store where third-party organisations pay a fee to access/use data; company shares in revenue from premium services
Brave[457]	Web browser that replaces targeted ads / trackers with privacy-preserving adverts; some revenue could be shared with users
Cocoon[458]	Social network limited to groups of max 12 (e.g. for family groups) with planned subscription business model
Jumbo[459]	Service that manages user privacy settings in other services and apps; “freemium” subscription model

Changes could be driven by the market, or by governments or citizens. Existing or future regulation in some regions could incentivise, or prohibit, certain business models. For example, as discussed in Section 5.1, real-time bidding in online advertising may be incompatible with GDPR and lead to changes in the industry. Commercial incentives could also cause shifts. Initial empirical evidence on the advertising revenue of a large publisher showed that the use of tracking cookies (which affects the ability to perform targeting) had a modest effect, increasing revenues by only about 4%[460]. As tracking measures are likely to have privacy compliance costs, the authors questioned whether these apparently minor potential benefits would always be enough to outweigh this. External pressure from consumers, lobby groups or wider citizen action could also have direct impacts – for example if citizens become more concerned over privacy and individual agency, and more resistant to targeted advertising. In January 2020, Google announced plans to phase out the use of third-party cookies on its Chrome browser in response to user calls for greater control over privacy[461].

A resurgent subscription model, alongside the continuation of free ad-driven services, could exacerbate concerns around “privacy as a luxury” and digital inequalities, as services

may offer extra privacy provisions at a price that is prohibitive to some users. Alternatively, measures to address concerns such as privacy might not actually change the main revenue models (e.g. encryption techniques enabling targeted advertising which preserves privacy[452]).

Unexpected macroeconomic shocks could lead to rapid shifts in market dominance and business models. Before the COVID-19 pandemic, the recent tech market shared some features with previous bubbles. Leading up to the dot-com bubble and crash in 2000, the proportion of US tech companies having an initial public offering (IPO) that was profitable declined rapidly, from around 70% in 1994 to under 15% in 2000 (see Figure 11). At the same time, the average price-to-sales ratio[c] of these companies increased from under 4 in 1994 to over 30 in 2000[462]. In 2018, the proportion of profitable IPOs was 15%, down from a recent peak of 71% in 2009. Notable within these were the IPOs of Uber (which lost $1.8 billion in 2018 and saw a share price fall of over 7% on the first day of trading[463]) and Lyft (which also had losses of almost $1 billion in 2018). However, the price-to-sales ratios of these were less than many of those seen prior to 2000[464], and the average was 7.6 in 2018, only modestly higher than relatively stable fluctuations in most of the years since 2002. Around the time of the dot-com bubble, there were also other likely contributing factors to the market crash, such as a rise in interest rates[465] and the 9/11 terror attacks in 2001[466].

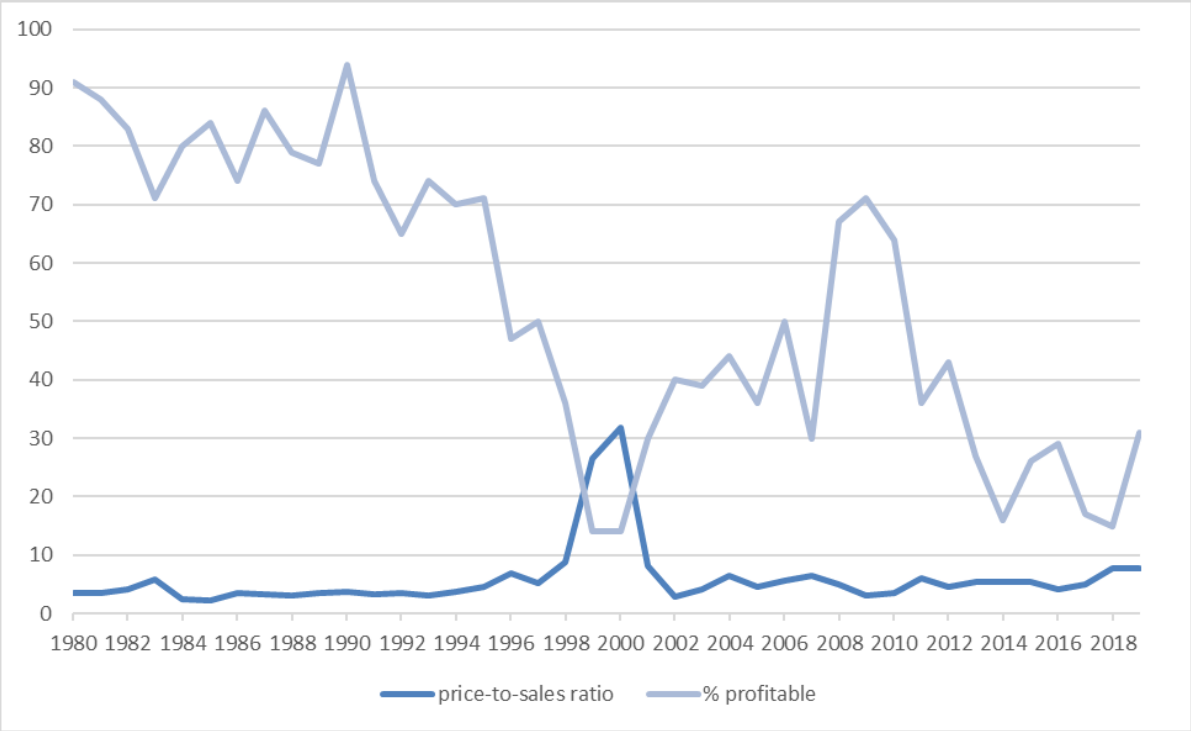

Figure 11 – Price-to-sales ratio and percentage profitable of US technology company IPOs, 1980-2019. Analysis of data from University of Florida[462].

Current large incumbents may survive an economic shock and be able to expand easily, or they may have a large enough user base to be able to pivot to new models – see for example Apple's shift towards services in recent years[467]. This could be particularly true for approaches that rely on a two-sided market, where a high number of consumers and businesses are needed for a platform to be successful. This may be the case for personal data stores and management platforms[468].

Remaining infrastructure following a shock could be repurposed by new entrants or global competitors. Following the dot-com bubble, it has been reported that the cost of

[c] Market capitalisation divided by last 12 months sales; based on offer price

bandwidth fell by more than 90%, with the majority of US broadband capacity unused[469]. A similar regional or global economic shock hitting the data economy could similarly reduce barriers to use of infrastructure for new companies.

The financial turmoil sparked by the COVID-19 pandemic has further caused uncertainty in this area. As of May 2020, the tech sector overall and large incumbents in particular were performing better than some market analysts anticipated following the initial economic downturn. Vastly increased numbers of citizens working and socialising from home increased demands for the services provided by companies such as Amazon, Facebook and Google[470]. This effect so far has mainly served to cement incumbents' market power, but the long-term impacts on dominant business models remains to be seen. The video-conferencing software company Zoom is an interesting example of a business operating a "freemium" model which experienced huge success since citizens' behaviour was changed by the pandemic, despite initial security and privacy concerns with the service[471].

9.2 New models of data governance

New models and mechanisms of governance of citizen data could disrupt current data systems. Recent shifts in governance and regulation around the world have expanded individual rights around data, and the responsibilities of data holders. However, in most cases the core features of who collects and holds data about citizens, and how it is used, have not changed radically. Companies and governments are still the major data holders, even if the scope and rights around this are constrained or expanded. In many circumstances citizens don't feel they have agency to exert control over how their data is used (see Section 4.1). However, this could change with proposed and developing ideas on data governance.

Data ownership is one commonly discussed concept. It usually refers to the idea that individuals should have control over data about them and share in the benefits that others derive from their data, and that a model analogous to property rights could be a way to ensure this. For example, a citizen may agree to share their data with a particular company only in return for some monetary or other benefit. Proposed models of data ownership have been particularly popular in the blockchain and distributed ledger community, as the decentralised nature of these mechanisms may be able to support fine-grained access controls without the need for a trusted third-party holding or managing the data[472], although this is debated[473]. This would also arguably align with the attitudes and behaviours of private companies which consider data to be a valuable strategic asset.

However, many experts see the concept of data ownership as problematic[474], and do not feel individual property rights to data are appropriate. Unlike most property, data is non-rival, meaning different people can use it many times for many different purposes, even simultaneously. Data may also be about many people – your date of birth represents information about your parents; your genome could be used to match relatives to another DNA sample – which would complicate such a concept. Advances in DNA technology and the increasing prevalence of commercial genomic databases make this an increasing privacy concern; a recent study suggested that in the near future nearly every US citizen of European descent would theoretically be able to be identified even if they have not participated in genetic testing, through comparison with distant relatives' results[475].

A model of data ownership could also place unreasonable burden on citizens to agree to or deny every single potential use of their data. This may be unwanted, and individuals may not always be in a position to give informed consent. The terms of this debate around data ownership might shift if other measures are taken to ensure that individuals who provide

This is not a statement of government policy

data experience the benefits of its use, for example through a better balance of benefits versus risks around online targeting as discussed in Section 7.3.

Relatedly, an individual can be affected by the data of other people. Inferences could be made about an individual who belongs to a specific group, by analysing data about that group, even if the individual is not himself or herself in the dataset being analysed. This could result in harm, for example through negative online targeting as discussed above. Because of this, an idea is emerging that we need to shift away from just focusing on individual rights and behaviours such as informed consent to data processing. Instead, some suggest we should also consider the potential collective rights of groups, for example 'group privacy' rights, and the legal and social implications associated with this shift[476]. As previously discussed, this can similarly still be an issue when using synthetic data, as even though individuals' data may not be directly stored or analysed, inferences can still be made about them that may impact their privacy or otherwise negatively affect them.

Many argue there is a need for new organisations or mechanisms to manage data use on behalf of citizens in a more trustworthy, accountable or democratic way, while encouraging and enabling data sharing and use for public and commercial good[477]. Such organisations may be variously referred to as data institutions or stewardship bodies, or other terms. We note that as this is a rapidly developing area of research, there is very varied use of language amongst the community, which can sow confusion[478]. Some commonly proposed mechanisms for this are data trusts, cooperatives and commons. Again, these terms are often used loosely and with varying degrees of nuance by different organisations to describe a range of mechanisms for data governance, management and access operating on behalf of data contributors, but some features typically distinguish different approaches:

Data trusts are inspired by legal trusts and are generally defined as involving a fiduciary duty (a legal responsibility of impartiality, prudence, transparency and undivided loyalty) for a trustee to manage data on behalf of those contributing data. The UK Biobank[479] could be considered to be a form of data trust. It is a charitable company with company directors who also act as charity trustees, and manages genetic and other health data contributed by 500,000 volunteer participants for health research purposes.

Data cooperatives are inspired by mutual organisations. They can be owned and democratically controlled by members. Data is then managed by those members in a delegated manner. For example, the MIDATA cooperative is a Swiss non-profit organisation where members have individual control over the use of their health and education data for research projects, and each has a single vote towards the organisation's decision-making[480].

Data commons are inspired by the management of common pool resources such as forests or fisheries. People or organisations collaborate to create and maintain shared data assets for mutual benefit. For example, the Data Commons for UK Tech is an open-access database holding information on start-up businesses, investors, accelerators and service-providers across the UK, which can assist both investors and entrepreneurs in business decision-making[481]. Users are also encouraged to contribute relevant data to the commons.

Data trusts, cooperatives and similar mechanisms have been proposed to deal with data misuse, overuse and underuse. This includes problems such as conflict of interest between data holders' duty towards citizens and their duty to others such as shareholders; and lack of beneficial and trustworthy data sharing and use, such as an input for AI algorithms[256]. They could also be used to better align trust and trustworthy behaviour, such that trustworthy agents are more likely to be trusted by users, through greater transparency and accountability[482]. Data trusts, institutions or stewardship bodies could be constructed in a top-down manner, for example mandated by law for certain kinds of data. They could also

be constructed bottom-up, created in response to the demands of citizens by individuals or businesses, and therefore can offer choice in line with varying user values, priorities and risk appetites[483]. It was recently reported that a group of UK Uber drivers are attempting to establish a data trust, which would be administered by a union, and into which drivers could port their personal data currently held by the platform for collective bargaining purposes[484.] As with many of these alternative data governance models, this would require effective portability of the data.

There are already many small to medium scale examples of these mechanisms[485,486], but little evidence on whether and how they could gain major uptake, for example becoming the main model of data governance in a region. Adoption is a challenge, despite anecdotal evidence of a huge appetite for such approaches[486]. A lot of data collection at present is passive (e.g. through online trackers) or a by-product of engagement with other services. It is not clear if there is currently enough interest for data trusts or stewardship bodies to gain use simply by existing as an alternative for the average citizen. However, this could change with potentially changing citizen values and behaviours, as discussed in Section 4.1. Even if uptake is seen, it might be skewed towards the most engaged and digitally literate groups, who may not be the people most at risk from misuse of citizen data. This may particularly be an issue with a bottom-up approach requiring citizens to join and potentially manage aspects of a trust or cooperative[482]. At the same time, a top-down approach could create new issues around the selected default data sharing and processing options, and ability to opt-out or switch.

Given this current uncertainty around the effectiveness of different models, a diversity of approaches in different contexts might be most appropriate. Further research may be needed to more clearly identify which models might produce social or economic benefits. More broadly, any effective large-scale changes to the existing complex data governance landscape are likely to require a high capacity for developing data infrastructures, skills and standards[487].

A wider set of data rights, or more radical changes in control of data, could also shift how data is shared and used. Some regions may take steps to restrict current or future uses of data seen as having an overall negative impact or high risk. For example, the German Data Ethics Commission proposes a total or partial ban on algorithmic systems that are likely to use citizen data, which have an “untenable potential for harm”[488]. Depending on how risk levels are defined, such approaches could disrupt current systems and business models.

Some have suggested a broader bill of data rights, with provisions such as “no person shall have his or her behaviour surreptitiously manipulated” and “no person shall be unfairly discriminated against on the basis of data”[489]. Such a bill could attempt to address the issues discussed above around the collective rights of groups, by focussing more on how citizen data is used rather than individual rights and protections. This could include increased transparency and accountability, empowered public participation, and strong sanctions for non-compliance[490]. Others focus on the potential for citizen data as a public good, perhaps owned and managed by a public organisation, to maximise social benefit or realise financial value[491]; but such a model could be in tension with existing individual rights, for example to opt out, as in many cases analysis and outcomes would be improved with all citizen data used. An extension of this is nationalisation of citizen data, where data is seen as a national resource and access could, for example, be sold to a company on behalf of the nation. This is suggested by some as a means to protect lower- and middle-income countries from perceived unfair collection and use of data by companies in other regions[492,493].

There is some evidence on the potential economic and behavioural impacts of new models of governance. These are mostly theoretical or modelling studies. For example,

This is not a statement of government policy

one economic modelling study investigated how different models of data control could affect how far data is shared when consumers prioritise both their privacy and gains from use of their data. A scenario where consumers control their use of data was closest to optimal, with maximised benefits from data sharing and privacy. Consumers kept some data private, but shared other data with many more organisations and companies, compared to a scenario where firms controlled data. In that scenario, individual companies used more data without respect for privacy, but did not share it widely with others[494]. Other studies have suggested that if this was a model of data ownership (see above), it may only lead to limited gains for consumers from selling their data, as the non-rival nature of data could limit the incentive for intermediaries to compete on price[495]. There have also been doubts over the effectiveness of competition interventions such as data portability[496]; and any intervention that increases access and re-use of data could increase security risks without the proper protections and oversight (see Section 7.3).

9.3 Conflict and rapid political change

The experts we consulted for this work frequently cited conflict and radical political shifts as among the events that could have the highest impact on data systems. This includes physical and cyber conflict, terrorism, reactions to perceived overuse of citizen data for security. In addition, political changes, such as increases or decreases in authoritarian governments across the world, could shift the balance and number of countries aligned to particular models of data governance and use.

Such events could radically change opinions and approaches to use of citizen data, or the level of influence of different groups. For example, a large-scale malicious attack on the integrity of important citizen data could shift support towards restricting access and data flows and lead to more restrictive regulations. In Section 4.1 we discussed evidence that in the US, survey respondents reported being more concerned about US anti-terrorism policies restricting civil liberties after the Edward Snowden leaks in 2013[238]. Increases in the frequency and severity of terrorism and conflict could increase support for surveillance and state use of citizen data at home and abroad – potentially with longer term consequences and reactions. In addition, a convergence in the types of political system (e.g. liberal democracy) and views of those in power across the world could increase the likelihood of international agreements and standards for data-sharing to be developed.

Data-driven businesses may also change their practices in response to political shifts. In summer 2020, after the growth of the Black Lives Matter movement, Microsoft announced it would not sell its facial recognition software to police departments until federal regulations were introduced, Amazon announced a one-year moratorium on police use of its facial recognition software, and IBM announced it would stop developing its facial recognition software altogether[497]. In Section 10 we discuss the potential impacts of the COVID-19 pandemic on opinions and approaches to use of citizen data by different groups.

This is not a statement of government policy

10 Case study: COVID-19 and citizen data

The COVID-19 pandemic has brought into sharp focus many of the key issues around citizen data systems that are discussed throughout this report. This includes individual privacy, national security and public health, technological advances, economic impacts and social values and behaviour. In this case study, we explore different international approaches to use of citizen data in the pandemic response, using the example of digital contact tracing to demonstrate the interplay of all the above factors. We also consider potential consequences of the pandemic for the future of citizen data systems, though at this relatively early stage any such assessment is subject to a high degree of uncertainty.

COVID-19 may be an event that causes a radical shift in global citizen data systems. Collection, processing and sharing of citizen data within and across domestic borders has been fundamental to the pandemic response, and divergent international approaches have highlighted existing differences between regional data systems. The emerging economic impacts have influenced the dominant business models that process citizen data, and citizens' values and attitudes may change in response to emerging events, and the perceived success or failure of different approaches. In answering whether public health requires a diminution of privacy, national responses show parallels to prioritisations between individual privacy and national security.

COVID-19 has raised numerous issues around the use of citizen data, particularly with the range of digital technologies that have been used in the response[498]. This case study focuses on the development of digital contact tracing as an exemplar of the interplay of economic, security, and privacy factors discussed in this report. Digital contact tracing – automatically recording when you are close to someone else, usually through mobile phones – has emerged as a key aspect of contact tracing strategies, which aim to maintain low transmission of the disease while lifting lockdown measures[499]. While also conducted manually, pervasive smart phone usage enables a digital approach, which shares the same aim but differs in the volume of data that can be collected, the scalability of resources, and information with which this data can be linked. Ways in which contact tracing approaches have diverged between states are rooted in differences in national data policies, technological capabilities and resources, and public acceptability of trade-offs between private data and public benefit.

10.1 Divergence in digital contact tracing approaches from China, the US, and the EU

The Chinese response is an example of strong state control of citizen data to monitor compliance and encourage desirable behaviour. Utilising a highly centralised design, movement in public places has been authorised by colour coded QR codes created by Alipay and WeChat with local and national governments. Algorithms assign restriction levels based on information such as location[500], travel history and body temperature that the users input[499,501]. Those without a code can be denied entry to public places, offices, malls, and transport facilities[502,503]. In some provinces, officials are reportedly looking to continue[504], adapt and expand the app after the pandemic to promote healthy behaviours[505]. Chinese authorities have also used facial-recognition software and location tracking to monitor quarantine violators[505].

The US data response has capitalised on low restrictions on business and government in the collection and use of citizen data. The US Center for Disease Control and Prevention has used data gathered by the mobile advertising industry on the location and movement of individuals. This data is pseudonymised, but not aggregated, and as such has, technically, potential for re-identification[506]. In response, despite a historical lack of protection of privacy and data by comprehensive federal law, a COVID-19 Consumer Data

Protection Act has been proposed to require data privacy and security measures from businesses that handle personal data specifically related to COVID-19[507,508]. Proposals appear to draw heavily on GDPR including opt-in consent, data minimisation, and right to deletion principles[507,509,510].

The EU response has reflected its value of privacy and data protection as fundamental rights. GDPR provides legal grounds for processing personal data in the context of COVID-19, for example for public interest in the area of public health[511]. In the UK notices were issued to further require certain healthcare organisations and local authorities to process information, also compliant with GDPR[512]. From the first weeks of the pandemic, many European countries were using anonymised or aggregated data from major telecommunications companies to understand population movement[506,513]. More than 20 countries and territories in Europe have launched or plan digital contact tracing apps[514]. Originally the optional connectivity of national apps was preferred by many nations, with the Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT) initiative aiming to allow different nations' tracing apps to communicate with each other[515]. Most EU nations that have launched apps have now opted for the Apple-Google collaboration discussed below, and the ambitions of PEPP-PT are now being taken forward by other institutions[516].

Apple and Google's collaborative COVID-19 exposure notification system is one of the most conservative designs, prioritising privacy. Its decentralised design transmits unique, rotating codes from phones via Bluetooth based on cryptographic keys that change daily[517]. Operational from May 2020[518,519], Apple and Google have made their interface available only to national, state, and regional authorities, which uses Bluetooth and disallows governments to collect location data[517]. The choice to maximise individual privacy reflects a prioritisation of consumers, which may contribute to a positive public image for these companies. However, some see this as a unilateral policy decision made by a private company which could have significant ramifications on public health globally, without effective routes for government challenge.

10.2 Tension between data nationalism, globalisation and technological dominance

National approaches to data sit in tension with the inherently global threat of the pandemic. The early impetus to develop national apps was driven by concerns about data ownership, data type requirements, national opinion, and timescales. For example, many national apps aimed to centralise data and include location information, to potentially identify 'hotspots' to support local health responses and regional decision-making. Connectivity between countries – a design strength of the Apple-Google model – was a secondary priority. Many countries, including the UK, have switched from developing their own apps to utilising the decentralised Apple/Google framework[520].

The primary driver behind the shift to the Apple and Google collaboration has been technical barriers. A pilot on the Isle of Wight revealed that the NHS app only recognised 4% of Apple phones and 75% of Google Android devices[521], which had been noted in Australia's deployed app[522]. By contrast, the Apple-Google model logged 99% of both Android mobiles and iPhones[523]. This was due to Apple operating system restricting Bluetooth communication by phones when an app was running in background mode, which has been rectified for the Apple-Google app only and not nationally developed tracing apps.

10.3 Concerns around efficacy and equality of digital contact tracing

The efficacy of these apps is yet to be proven, and dependent on high public uptake. Modelling studies that have estimated the potential impact of app-based tracing on transmission (e.g. reduction in the reproduction number of the disease) find any effect is substantially dependent on uptake level[524]. However, many countries have failed to achieve

high app uptake, with both Singapore and Australia achieving around one sixth of the population[505,525,526] and Germany achieving one fifth[527], though Ireland has achieved one third[528]; though overall it being challenging to ascertain exact numbers. The UK Government's Scientific Pandemic Influenza Group on Modelling indicated in April 2020 that under simple assumptions the proportion of contacts an app-based approach could detect is proportionate to the square of the population using it. They indicated that even if there was very high uptake it would be unlikely to detect more than 50% of contacts[529].

Public trust may affect uptake of any digital contact tracing apps and their corresponding efficacy. In the UK, the ethics advisory board reviewing the contract tracing app prioritise of trustworthy data use, including value and impact, security, accountability, transparency, and control[530]. There are also concerns surrounding equality, with reach of the software limited to people who own smartphones, which often excludes lower-income people, racial minorities and people over 65[531,532]. Some countries have circumvented this by lending phones or wearable devices the elderly or to those tested positive to COVID-19 during their quarantine period[533].

10.4 International approaches to COVID-19 and privacy

Internationally, government and local authority use of data during COVID-19 has potential to shape citizen values, particularly where re-identification, surveillance, and privacy breaches have occurred. Conversely, effective deployment may support future digital governance initiatives.

Public publishing of the movement of individuals to combat COVID-19 has resulted in re-identification. South Korea's response involved linking GPS phone tracking, credit card records, surveillance video and interviews with patients[531]. It published places that individuals who tested positive had been, resulting in a number of high-profile cases of re-identification in the media[534,505]. Singapore has also been publishing detailed data about every infected person including age, gender, workplace, and where they had visited[535].

State surveillance is being repurposed for COVID-19, being used to police lockdown, and to gather sensitive data on citizens. In Israel, the existence of mobile data which had been collected secretly to combat terrorism was made public and repurposed to trace contacts of COVID-19 cases[536]. Russia is using an app to police lockdown, which requests access to calls and location to ensure that individuals who have tested positive to COVID-19 obey quarantine[515]. The Colombian app reportedly asks people to answer questions about participation at protests and ethnicity[537].

Privacy breaches have occurred in a number of states. India's app tracks individuals using GPS, has been made mandatory for many workers[538], and the country has experienced several privacy breaches[539,540]. Pakistan, which has no data protection laws, has had similar breaches resulting in attacks on health workers[541]. Qatar's app, which linked to sensitive personal information and was mandatory to download, was found to have design flaws which made it vulnerable to hacking[542].

10.5 Potential future implications

It is likely that COVID-19 will cause a shift in data systems, as changes made during the pandemic have the potential to become embedded, and to change public perceptions. All the approaches outlined above reflect that there may be benefits for the economy, innovation or public services to be gained by sacrificing some level of privacy during a pandemic. This is explicitly acknowledged in the UK's COVID-19 approach:

> *"In the end, the choice you have to make is a balance between individual, group and national privacy, and the public health authorities having the minimum information*

necessary to manage the spread of the virus." – Ian Levy, National Cyber Security Centre, 4th May 2020[543]

In the UK, it is already legally specified that an infected person's privacy may be trumped by the risk they pose to other peoples' lives[d], with medical workers having a statutory duty to report incidences of certain diseases such as Cholera, Mumps and Rabies, as well as COVID-19[544,545].

It is possible that some changes, technologies, or systems will become embedded, and used in future development and delivery of public services. This could support the response to potential future health emergencies, and spill-over effects may increase the prominence and acceptability of behavioural and social measures in public health interventions. In the future, similar approaches could be adapted outside viral pandemics, which hypothetically could include managing antimicrobial resistance prior to novel drug development. Conversely, these could be repurposed for national security by nations where not already in use, with the technology used to monitor citizens with or without their knowledge for various domestic aims. Such systems will need to continue to be reviewed to ensure new data uses are ethical and proportionate, and the appropriate governance and legal frameworks are in place.

Some commentators raise concerns about the ability to reverse any privacy infringements embedded during the scaling up of systems during a pandemic[505,513]. More than 100 groups signed a joint statement setting out human rights conditions that must be met before the surveillance technology is used to fight the pandemic[546]. It is perhaps not insignificant that most of the technology quickly repurposed in early stages of the pandemic was originally used for national security and counterterrorism[505,547].

In many countries, it is still unclear as to how this will affect public perceptions of government data use. In most cases the full extent is yet to be recognised, but is likely to continue to evolve over the coming years, and may be affected by the observed efficacy and transparency of such approaches, and how information about them is communicated to citizens.

Overall, the COVID-19 pandemic has brought about significant changes to the way citizen data has been used across the world. These changes might have profound implications for state versus personal privacy, with a lasting legacy.

[d] Under the Public Health (Control of Disease) Act 1984 and the Health Protection (Notification) Regulations 2010.

11 Citizen Data in 2030 – four scenarios

Previous sections illustrate the high level of uncertainty in factors likely to affect the future of citizen data systems across the world. To consider these uncertainties in a manageable and structured way, we have developed four future scenarios. These illustrate possible alternative versions of the future in 2030. They are based on a range of possibilities for drivers of change and the interactions between them.

11.1 Scenarios creation process

To construct our scenarios, we anonymously surveyed around 40 international experts for their perceptions of the current and potential future states of global citizen data systems. These experts were drawn from across industry, academia and the voluntary sector. We took the results of this work, alongside our own research, to a workshop of around 35 policy experts from within government and the wider public sector. They were asked to identify the key future trends that are both uncertain and likely to affect the policy issues raised in this report. Key trends that are assumed to be true in all futures, such as increasing volume of data and data linkage, were also identified.

For each key uncertain trend, workshop participants identified a number of possible future states. Participants then considered which combinations of future states were theoretically possible and had the most significant consequences for global data systems and were therefore most worth exploring.

Based on these discussions, we distilled the range of uncertainties down to four key areas: social norms, values, and behaviours; domestic and international regulations and global data norms; technological advances; and business models and commercial incentives of data use. For each of these key areas we identified multiple potential future outcomes. For each scenario we then considered possible further drivers of change and implications covered in this report, including in malicious uses of citizen data, and how data use and energy demand might interact.

The four scenarios are based on combinations of the four key uncertainties that are internally consistent and create plausible futures relevant to UK policymakers. The outcomes within each key uncertainty used to construct each scenario are shown in Figure 12. The scenarios were also designed to reflect a sufficient breadth of future directions, and to be challenging enough to be worth policymakers exploring. As such the scenarios are not comprehensive, and we acknowledge many other futures are possible.

11.2 How to use these scenarios

These scenarios are not predictions and should not be used as such. Instead, they are tools to help policymakers rigorously consider the range of potential futures, understand potential long-term implications of different interventions, and how these might interact with each other. Doing this can lead to decisions and policies that are more resilient and coordinated. Policymakers can explore each scenario in turn, and consider questions in the context of their policy area such as:

- What strategic choices would the UK have to make in this scenario? What might our ‘red lines’ be?
- What UK systems and sectors would be vulnerable in this scenario?
- What impact would this scenario have on UK prosperity and security?
- How would UK citizens feel in this scenario?
- What could the UK do now to steer towards or avoid this type of scenario?

11.3 Scenarios summary

	Scenario	Social norms, values and behaviours	Regulations and global data norms	Technological advances	Business models and commercial incentives	Description
Scenario 1	**Divergent Data Nationalism**	Disengagement	Data nationalism	Consumer disinterest	Market interventions and fragmentation	▪ overall low citizen engagement on data privacy and trust issues ▪ rise in nationalistic data policies; more restrictions on international data flows ▪ lack of technological innovation ▪ disruption of existing business models by regional government interventions
Scenario 2	**Multipolarity**	Managed apathy	Regional blocs	Empowered engagement	Market consolidation	▪ moderate but mainly government-led citizen engagement on data privacy and trust issues ▪ hardening of the three main regional blocs of data systems; limited compatibility between them ▪ varied uptake of emerging data-driven technologies ▪ consolidation of market power for incumbent dominant players
Scenario 3	**Deregulation**	Managed apathy	Deregulation	Independent innovation	Reduced data barriers to entry	▪ moderate citizen engagement on data privacy and trust issues; high consumer choice (for those who can pay) ▪ relaxed global regulatory environment; high data flows with limited oversight ▪ high technological innovation ▪ emergence of new consumer-led business models
Scenario 4	**Multilateralism**	Active engagement	Globally interoperable systems	Resistance	Cottage industries	▪ overall high collective citizen engagement on data privacy and trust issues ▪ increased international collaboration on data policies; controlled international data flows ▪ resistance to some data-driven technologies ▪ disruption to existing business models alongside a change in value of large data sets

Figure 12 – A summary of the 2030 future scenarios, showing for each scenario the future outcome within each of the four key uncertain areas, and a short overall description of the scenario.

This is not a statement of government policy

11.4 Scenario 1 – Divergent Data Nationalism

11.4.1 Domestic and international regulation, and global data norms

It's 2030, and the global flow of citizen data is restricted. There has been a proliferation worldwide of data nationalism – domestic policies with the common feature of reducing international data flows. Example policies include data localisation, mandated government access to valuable citizen data from companies, and nationalised data stores in some countries. Less populous countries with significantly reduced access to data form strategic bilateral data-sharing alliances.

These broadly nationalistic policies are motivated by the rise of populism, commercial protectionism, security concerns, geopolitical tensions and a reaction to data use by large foreign companies. In more authoritarian countries, nationalistic policies have arisen due to a desire for greater social stability. In the US they have been prompted by protectionism, and a response to a perceived AI arms race. Some policies are in direct response to a large-scale malicious event influencing public opinion. For example, a data poisoning attack on NHS medical records by a foreign actor leads to an expansion of data sovereignty rules in the UK designed to protect critical infrastructure.

Across the world, the volume and types of citizen data being collected, processed and stored is growing, increasing energy demand and environmental impact. There is considerable variation between countries on whether and how this is mitigated. Some countries use increased access to local citizen data to reduce carbon emissions, for example by better forecasting consumer energy demand and supply, and optimising energy use in buildings. They also implement domestic policies limiting energy-intensive uses of data. In other countries, a lack of action leads to less sustainable data systems, which may be cheaper to operate. Concern over energy use is another driver limiting the transfer of data between countries, with different approaches and national priorities.

11.4.2 Social norms, values and behaviours

The experience of data use has become very different for people in different countries. In some, the state uses data-based surveillance and personalisation to nudge populations towards preferred behaviours, leading to increased state intrusion on personal lives. Similarly, some countries create mandatory nation-wide DNA databases of all citizens. Nominally for use in criminal justice, these have wider potential applications, for example, in determining citizen access to insurance, education and employment opportunities.

Some countries are almost completely disconnected from the global internet, with only limited government-controlled access points. This leads to an asymmetry with more open nations, that have more varied sources of information, and reduced global engagement between citizens. In a few countries, government attempts at more nationalistic data policies provoke anger and social unrest, leading to instability and a rejection of some government control of data.

There is more deceptive manipulation of data including deepfakes and online disinformation, in the UK and worldwide. Without globally effective mechanisms to counter them, such as agreed enforceable standards, individual countries to develop mitigations of their own. They do this with varying degrees of success depending on the available resources and level of priority.

Particularly for the few systems that remain relatively open, malicious and opportunistic actors exploit the divergence between systems to develop and test new data-intensive services and manipulative methods. Data-related crimes, including personalised cyberattacks, become more widespread and more effective in these open regions.

However, for the majority of citizens, individual behaviour and trust in data systems does not substantially change in response. This is partly due to apathy, a lack of understanding of potential consequences of effective alternative options, and a willingness to continue to use data services for personal benefit. This further reinforces nationalist behaviour by governments looking to restrict external access and protect citizen data.

In countries with mandated government access to citizen data, relevant data from private companies provide government with new insights into behaviours. These are used to improve and reform public services. For example, data on private taxi journeys are used to improve transport infrastructure management, and consumer food purchasing behaviour inform healthcare treatment models. However, the difficulties in sharing data internationally inhibit progress in more globally focused research and development.

11.4.3 Technological advances

The restrictive international regulatory environment and lack of global coordination stifles innovation in some data-driven technologies and associated personalised data services. These include advanced wearable or implanted devices for health monitoring. This leads to variable uptake of sometimes incompatible technologies in different regions. There is also limited consumer interest in such technologies overall, further reducing incentives for businesses to invest heavily in their development.

Governments focus more on developing and enforcing internal standards. Privacy enhancing technologies (PETs) are not widely used as a tool to mediate between different systems or manage trust across countries. System incompatibility and a lack of trust across borders and organisations undermine attempts to increase their use. However there is considerable variation in domestic approaches to privacy and security, from mandated use of cryptographic PETs for national databases, to bans on encryption of privately held datasets for security reasons.

Within individual regions, more isolated data-driven systems based on access to local citizen data flourish, such as smart cities in Japan and autonomous vehicles in the UK. Public services are increasingly digitised in many countries. Robust digital identification tools are used for citizens to access these services securely, improving the efficiency and access to these services for most citizens.

11.4.4 Business models and commercial incentives of data use

Companies with business models based on global expansion struggle to cope with the higher operational and compliance costs associated with divergent national systems. This particularly hits smaller international companies, although all are affected to some extent. Some companies split into more local operating companies acting semi-independently in different regions. There is increased collaboration between companies and governments to ensure compliance and gain access to data, for example in healthcare.

There are increased opportunities for home-grown companies to emerge and more effectively serve citizens of a single nation or region, particularly if they benefit from direct or indirect government support.

This is not a statement of government policy

A specific government intervention in one region, such as high taxes on digital services or the banning of behaviour-based targeting in the EU, further disrupts current business models. This leads to a shift away from free services towards subscription-based and 'freemium' models, and more aggressive monetisation strategies. In some cases, this results in a degradation of service quality.

Countries with more relaxed regulatory environments – so called data havens – become regional hubs for businesses based on development of tools and services built on citizen data. This exacerbates the differences between regions, perhaps increasing risks of misuse if these tools are deployed elsewhere on different populations. This is also associated with a migration of talented workers to these havens.

To consider in this scenario:

- What policies could achieve UK goals without international data flows?
- Which would be the UK's key partner countries?
- How could the UK boost prosperity when unilateralism and security issues dominate?
- Is there a way to support beneficial international data uses in this scenario?

This is not a statement of government policy

11.5 Scenario 2 – Multipolarity

11.5.1 Domestic and international regulation, and global data norms

It's 2030, and three main blocs of citizen data governance have emerged. Data flow is controlled in different ways in each bloc. Legal and technical interoperability between them is limited. The EU-led system continues to expand, with some African countries adopting rules modelled on EU regulations and individual privacy values. China's international influence also grows through the Belt and Road Initiative. Meanwhile in the US no comprehensive federal laws are passed around data protection regulation. This leads to greater tension with the EU system, and a more sharply defined choice for countries with emerging digital presences of which bloc, if any, they should align with. As populous countries such as India and Indonesia make this choice, there is substantial impact on the overall balance of power in wider geopolitics.

There is some progress towards international frameworks for legal and technological interoperability, but this is patchy and uncoordinated. International data-sharing agreements between blocs are generally limited to specific uses and types of data where consensus across systems is easier to achieve. For example, there is some progress in data for the management of natural disasters and that can be used to support the UN Sustainable Development Goals.

By contrast, within blocs there is a higher level of data-sharing, for example with increased business-to-government sharing in the EU via sector-specific European data spaces.

As the volume of citizen data grows, some blocs put increased effort into reducing the energy demands and environmental impacts of large-scale digital systems. This involves regulations and other interventions such as strategic investments, industry traceability requirements and data infrastructure standards. Data flow is further restricted between blocs with different approaches, for example with energy sustainability requirements being built into GDPR adequacy.

11.5.2 Social norms, values and behaviours

Between blocs there is some agreement on baseline citizen values around biases in algorithms, ethics and data manipulation. However, the legal and technological approaches used to enforce these standards vary widely between regions. This enables a moderate degree of trust in data systems by citizens in most regions.

In general, citizens still prioritise short-term gains and economic benefits over other concerns when interacting with data systems. Commentators continue to be surprised by citizens' willingness to share increasingly sensitive and diverse data with multiple companies in order to access novel products and services. This leaves governments with a difficult choice. Legislation, or other interventions, to protect against wider perceived harms risks lacking public legitimacy and being undermined by citizen behaviours. The appetite for this kind of intervention among governments also varies considerably between different blocs, with overall greater concern for individual privacy and ethics in the EU-led bloc.

This is not a statement of government policy

11.5.3 Technological advances

People feel empowered to engage with data-driven technologies in ways that benefit them. There is broad social acceptance of the role of such technologies in everyday life. This drives academic research, commercial investment and widespread uptake of some advanced technologies, such as personalised social robotic assistants based on advanced emotion recognition and prediction. In addition, the use of PETs becomes more widespread. This helps address issues around exchange and use of data internationally between systems with different privacy standards.

However, wider regulatory (and to some extent social) differences between blocs hinder the transfer of some innovative technologies between regions. For example, advanced biometric technologies and technologies aimed at children don't have global reach. This somewhat limits the extent to which advances gained in one region benefit citizens in others.

Governments in blocs with less commercial innovation are compelled to fund more research directly in order to remain technologically competitive, and for national security purposes. Data access controls limit the ability of scientists to exploit large multi-modal datasets from across blocs, which are required to understand and improve the resilience of digital societies and economies. Highly capable malicious actors are therefore better positioned to induce 'cascading failures' within digital societies and economies.

11.5.4 Business models and commercial incentives of data use

In the US and China, large tech and social media companies have grown and expanded laterally into new sectors such as financial services, with reduced operations in the EU. This further consolidates citizen data within the regional blocs in which these companies operate. There is only incremental progress in methods that increase machine learning performance with smaller datasets. As a result, businesses continue to be incentivised to combine large volumes of citizen data within blocs in order to be successful.

Large tech companies generally have strong relationships with the governments of the blocs in which they operate. They have a powerful influence over the data norms in those regions. This has led to an asymmetry in regulatory environments between blocs. Stricter competition interventions (such as mandated data sharing between companies) are brought forward in the EU-led regime to support EU companies' growth.

States find balancing citizen desires with the desire to promote domestic companies difficult in regions where citizens in one region rely heavily on services from another bloc. In addition, large tech companies and their associated research ecosystems act as significant attractors of talent. This limits the ability for growth in some regions without incumbent large companies.

In response to a perceived focus on privacy, particularly in the EU, internationally operating companies are increasingly protective of the citizen data they hold. Approaches include an increased use of encryption and employing differential privacy when collecting and sharing aggregate data. This reduces some privacy risks. However, in some cases it serves to consolidate market power – either by reducing interoperability or due to the requirements for large datasets for these approaches to be effective. This causes problems for legitimate security and law enforcement access, especially across jurisdictions where privacy regulations differ significantly.

This is not a statement of government policy

To consider in this scenario:

- Which bloc would the UK align with? What would the opportunity cost be of not aligning with each bloc?
- What could the UK do to bridge between different systems?
- How might policy interventions balance promoting competition, enabling domestic companies to be globally competitive, and services for consumers?
- How might the differences between each system affect security vulnerabilities or resilience?

This is not a statement of government policy

11.6 Scenario 3 – Deregulation

11.6.1 Domestic and international regulation, and global data norms

It's 2030, and globally citizen data flows freely between most regions. This has come about due to a pervasive free-market ideology and a perceived failure of previous regulatory efforts. These drove an uncoordinated rollback of regulations around data control and use in many digitally developed countries.

GDPR is perceived by many as not supporting the development of industry. Instead it's seen as having cemented the power of incumbent market players who have greater resource to dedicate to compliance. This reduces international support and adoption of similar data protection regulations. Some previously more restricted regimes are unable to maintain effective national online firewalls in the face of technological advances enabling new access routes. This further encourages sharing of data and information across regions.

A few countries choose to maintain or adopt more highly regulated, nationalistic data policies. These countries operate autonomously and become more economically and socially isolated from the main global digital system, potentially with negative consequences for prosperity and transparency in those regions.

11.6.2 Technological advances

This relaxed international regulatory environment encourages large internationally operating businesses to put increasing resources into innovation. They develop increasingly advanced data-driven technologies, such as brain-computer and brain-to-brain interfaces that enable a form of direct thought expression, with cross-sector applications. There is also increased globalisation, particularly into countries with emerging digital economies. This brings new service options and benefits for businesses and consumers.

However, increased globalisation of technologies also brings increased risks. For example, new types of very sensitive citizen data are being processed in regions with varying regulations and standards. Mechanisms to ensure due consideration of potential ethical and security consequences are scant. Some innovative, but potentially invasive and powerful, technologies are applied in sectors and geographical regions they were not originally intended for. Authoritarian governments, security organisations and other malicious actors seize on these new tools with limited oversight. This occurs without significant citizen engagement with or resistance to such technologies. Choice to opt out is limited in cases where there is a clear imbalance in the quality of technologies between regions.

As part of this increased but variable innovation, some businesses use advanced data-driven approaches to optimise their overall energy use. Machine learning methods to improve efficiency of their data centres are commonplace. This reduces the environmental impacts of data systems in some industries. However, without global coordination there is considerable variation between businesses, and transparency for consumers and governments is reduced due to complex global supply chains.

In countries with domestic sustainability incentives, data processing and the associated carbon emissions are sometimes exported to regions with lower standards. This reduces the overall effectiveness of data-driven solutions to climate change. There is also significant variation between regions in computer processing power and data analysis capabilities more

broadly. This has negative implications for local economies in regions with lower capabilities, where it is more challenging to realise value from data assets.

11.6.3 Business models and commercial incentives of data use

New digital tools such as high-quality synthetic data become more developed. This reduces the reliance on, and commercial power of, large citizen datasets in some sectors. Instead, there is increasing emphasis on access to diverse datasets and advanced analysis techniques. This somewhat reduces the incentive for companies to collect large volumes of citizen data to improve their algorithms. It also provides opportunities for smaller entrants to existing markets. A few of these have threated to grow sufficiently in economic power and skills to supplant established players, though their dominance may be unstable and vary significantly over time.

Associated with this, there is a moderate market-led shift in business models towards paid-for personalised services, based on curation of an individual's data. For example, new healthcare products and services have emerged. These are enabled through the easier sharing and linking of data from healthcare providers, wearable devices, fitness apps, public transport and even genetic information. More public sector data is open and freely accessible, allowing businesses to link datasets and gain insights that have applications beyond the original purpose of the collected data. The status of data held in the private sector is more varied. Some companies continue to protect datasets as competitive assets, others are now sharing as a means to realise wider benefits.

Companies have increased flexibility to design their data privacy and security policies, and so some choose to compete for customers based on this. This creates a luxury market for products and services where privacy is valued more highly. This principally benefits the relatively few highly engaged and wealthy citizens for whom this is a personal priority. This comes at a cost to overall quality of services. Access to such services is unequal, as more vulnerable citizens with less resources are disadvantaged, and overall citizen engagement is low.

To combat this, and other perceived market failures, some governments are compelled to act. They attempt soft domestic interventions to incentivise markets to behave in ways that benefit all citizens, setting up voluntary codes of practice for businesses, and funding citizen digital literacy programmes. Still, for the majority of citizens for whom privacy is not a priority, consumers consent to sharing increasingly sensitive data about themselves in return for access to advanced services and technologies. This carries the potential risks for where and how this data is used.

11.6.4 Social norms, values and behaviours

With sensitive data flowing freely between regions, some systems are more vulnerable to malicious attacks and deceptive uses of data. Identification of individuals in anonymised datasets and manipulation of data-driven autonomous vehicle systems both thrive. In the absence of strong national and international interventions, private companies are mainly considered responsible for responding to these threats and maintaining user security.

This drives industry and academic development and deployment of mitigating technologies to limit online harms. New approached include automated media forensics tools to flag deepfakes on social media, machine learning systems that are robust to attempted data poisoning, and technologies that can identify illegal behaviours on digital platforms. These technological, industry-led solutions are considered more successful than government interventions.

However, they do not address some deeper perceived issues around the incentives for data use in particular business models and may result in other systemic harms. Still, in part due to the relatively positive relationships between commercial service-providers and consumers, these industry-led mitigations give citizens overall a moderate degree of trust in the data systems they use.

In the absence of national and international regulations, city authorities and local governments play more of a role in determining the norms around data use, based on the values of the local citizens that elect them. This causes significant regional differences in technological and social environments even within individual countries. Some smart megacities prioritise interconnectedness and data-driven innovation, others individual privacy. The latter introduce local restrictions, for example on the use of surveillance technologies by local agencies and businesses. As people move to cities, this gives more citizens a choice about how and by whom their data is used in their local environment.

To consider in this scenario:

- What lighter touch interventions would be needed in the UK, for example in specific sectors or for critical datasets?
- Should the UK fund or set up organisations that support positive data use – for example managing governance and commercial models around use of UK citizens' data?
- In the absence of regulation, how could the UK incentivise the use of technological solutions to mitigate legitimate risks, such as PETs?
- What 'red lines' are there, indicating a case for the UK to maintain harder regulation, and how could these be enforced if such regulation is absent elsewhere?

11.7 Scenario 4 – Multilateralism

11.7.1 Domestic and international regulation, and global data norms

It's 2030, and globally there is a controlled flow of citizen data between most regions. This is enabled by a set of widely accepted data sharing standards and privacy frameworks, reached through considerable international cooperation including the US and EU nations. This cooperation has been driven by widespread global citizen engagement on data and privacy issues, brought about in part by two things. First, the generational shift to citizens who have mostly grown up immersed in a digital environment. And second, the emergence of strong global leadership with similar world views.

Even with this support, it is a challenge to maintain trust between such a disparate and wide group of nations and manage domestic risks. There is significant regional variation in capacity to put in place enforcement and accountability mechanisms. This makes it difficult to globally monitor compliance with agreed standards. There is also frequent compromise on appropriate and effective interventions in different policy areas. Progress towards agreements can be slow. Meanwhile some states remain outside this system entirely.

As part of this wider international collaboration, there are stronger international regulations around global carbon emissions and climate change policies. This reduces the requirements for specific regulations or interventions to mitigate the energy demands associated with data processing, transfer and storage, even in data intensive industries where the volume of citizen data being used is growing. This flexibility, and the broader regulations on carbon emissions, generates a market for more sustainable data processing systems, such as energy-efficient neuromorphic computing hardware.

11.7.2 Social norms, values and behaviours

To further enable the controlled sharing of data between regions, internationally operating data cooperatives and trusts have been set up. Led by the communities they serve, they give citizens more individual choice on how their data is managed by the organisations they interact with. This increases the alignment between trust in organisations and the reality of data use for individual citizens. However, it also increases the burden on citizens to make choices about their data. This has served to fragment access to datasets with clear public good uses, such as public health or climate research. Governments share their data using similar governance models, reducing the amount of truly open data available.

Internationally agreed technical standards and collaborative interventions make some systems more robust to some forms of malicious attack. For example, new critical infrastructure standards require that UK NHS computer systems be updated, reducing vulnerability to some types of data breaches. However, the high costs of compliance mean that strong local support is required in each country to justify this spending.

In addition, there are internationally agreed restrictions on businesses and governments using data-driven technologies to personalise or micro-target individuals. When combined with the increased choice citizens have around the services they use, this enables citizens to act with a high degree of trust in the data systems they engage with.

11.7.3 Business models and commercial incentives of data use

Competition policy related to data is part of multilateral agreements. There is compromise on measures to support smaller companies and companies in regions without large incumbents in order to maintain trade in services. These interventions include mandated shifts of dataset control to third parties who manage access, and in some cases the break-up of large companies, causing a loss of scale in some industries.

More advanced and highly effective algorithms are developed in areas such as few shot and transfer learning, which enable powerful analysis to be done on smaller datasets. This reduces barriers to entry for new businesses with less access to varied citizen data. Together with pro-competition interventions, these developments support a shift towards business models such as privacy-preserving advertising, subscription services and cooperatives. Overall, there is greater interoperability between platforms, but fewer consumers on each platform. This gives consumers more choice, perhaps at the cost of quality and cost of some services.

11.7.4 Technological advances

Linked to this, increased citizen engagement on privacy and data ethics has led to restrictive government interventions on some data-driven technologies. This includes widespread moratoriums on use of facial and emotional recognition in public spaces, and artificial intelligence for decision-making in justice and education.

This somewhat hinders technological progress, even in areas where appropriate use could be socially beneficial. However, risks of misuse have been minimised. This also makes the global system more vulnerable to those few countries and actors outside it, who are not limited by such restrictions. Such malicious actors make extensive use of new technologies to seek out and exploit vulnerabilities in a dynamic and adaptive manner.

In general, there is increased citizen resistance to technologies associated with intensive data-gathering, such as smart home technologies and human-computer interfaces. This limits research and commercial development of these technologies, which hampers the economy and technological progress for the benefit of citizens.

To consider in this scenario:

- How would the UK deal with malicious state and non-state actors in regions outside this dominant global system?
- What would be the UK's 'red lines', which, if crossed, would lead it to choose to diverge from this dominant global system? What would the costs of divergence be, for example on export opportunities and access to foreign services?
- How could the UK maximise influence in steering towards a preferred direction in such a large multilateral system, and monitor enforcement?
- How far could the UK, including local government and cities, take different approaches to reflect the needs of citizens while within a larger global system?
- How would the UK maximise positive data use?
- Is there a way to accelerate responsible technological innovation?

12 Conclusions

12.1 Summary of main findings

Citizen data is increasing rapidly in volume and variety. The effective use and sharing of it has the potential to bring huge benefits to the economy and society as a whole: boosting productivity and trade, enabling innovative products, improving public service delivery and informing scientific research. It has already formed a significant driver of economic development and innovation in public services in the UK and elsewhere.

The data landscape represents a challenge for policymakers, as data collected for one purpose can be used many times over for a range of purposes, and government policy in one area can have unintended impacts elsewhere. Many of the most promising uses of citizen data involve collaborations between the state and businesses. This is further complicated by the ease with which data flows across domestic borders, at least technically. Regional data systems need to interact with others successfully to achieve domestic goals, and trade negotiations are having an increasing impact on data systems.

Varying citizen data systems have evolved in different parts of the world, driven by geopolitical aims, social values, and the balance of power between individuals, governments, and businesses. Differing prioritisations across the economy, national security and individual privacy in the EU, US and China have contributed to the formation of divergent data systems in these regions. More widely, citizen perceptions and values may play a role in shaping systems, but it is unclear how far they in turn are influenced by governments or the product of individuals' experience of them in the first place. The views of citizens regarding trust in data use, privacy, and the role of governments vary significantly within and between countries.

These and other data systems are likely to evolve and emerge in future, due to a range of factors:

- **The world's economic centre of gravity is predicted to shift eastwards, and new growth in internet use will likely be concentrated outside of the West.** This could increase the global influence of data systems different to those prevalent there.
- **Interactions between national and regional data systems may change, placing greater value on shaping international norms**. If regional or other international data systems continue to grow in importance, individual nations may have reduced agency in designing their own policy frameworks if they wish to continue to benefit from data-enabled trade. The value of being able to influence international data institutions and norms is therefore set to increase.
- **National prosperity is likely to be increasingly tied to an effective data system**. Almost all predictions suggest rapid increases in the volume and variety of citizen data, generated through increasingly varied devices and services, and held across the public and private sectors. Data systems that embrace this stand to benefit from higher productivity; improved public services; and a role in the advancement of global science.
- **Data is an increasingly critical tool in addressing grand challenges, but the growing volume of energy-intensive data processing raises its own sustainability issues.** Whilst technological breakthroughs may help, the future of global data flows may therefore depend on making globally coordinated progress on climate change and sustainability policies.
- **New data and novel uses bring new threats to manage.** Evolving risks include micro-targeting of cyberattacks and disinformation, the exploitation of vulnerabilities

in machine-learning systems, or inadvertent introduction of biases, and harms associated with online targeting. More nuanced conceptualisations of openness and risk will be needed, and technological change will mean that judgements about what is 'safe' will be subject to continual revision.

In addition to long term trends, recent experience suggests that future data systems are likely to be determined by unpredictable shocks, and successful data systems will be those that can effectively and swiftly adapt:

- **Economic shocks and the associated rise and fall of particular business models could rapidly change the use of citizen data.** These events can change the incentives for business data use, reduce barriers to entry or entrench incumbent positions.
- **Political shocks and conflict could change citizen beliefs and values**, changing what national approaches can be sustained with public licence.
- **New models of data governance may take hold**, from new sets of data rights through to models of data ownership and data trusts. These changes could be led by governments, individuals, or businesses, and designed with a range of different aims in mind.
- **Most immediately, the COVID-19 pandemic is changing the use of citizen data in ways which could have profound and long-lasting impacts**. Its full impacts are still unknown; however, we can already see differences in national and international approaches to the use of citizen data in response to the pandemic, the role of technology companies in determining norms, and the debates being raised around prioritisations of individual privacy, security, and social aims.

12.2 Implications of scenarios

This uncertainty and complexity implies a wide range of possibilities for how the future may look in 2030. We described four scenarios, designed to be varied but plausible, which all raise challenging questions for policymakers to consider: *Divergent Data Nationalism*, *Multipolarity*, *Deregulation*, and *Multilateralism*.

These scenarios demonstrate the powerful influence of social values and norms in determining overall outcomes and the policy regimes, implying that it will be increasingly important to engage with the public and citizen groups over potential benefits of use of citizen data and acceptable levels of risk.

They also highlight the active role of emerging technologies such as new algorithmic tools and privacy enhancing technologies, suggesting that in future the research, development, investment and control of such technologies will be important in determining which benefits and risks emerge, and which actors experience them.

In most cases, there will be a need for data skills and knowledge across government, academia, industry, and the public, to enable effective domestic data use and innovation, and to potentially boost international influence. Cross-disciplinary skills that can be applied at the boundary between legal, ethical and technological fields will be particularly valuable in future.

These scenarios should help policymakers consider the potential interactions and longer-term implications of policy interventions, and so develop more resilient and coordinated policy approaches.

12.3 Policy recommendations

What does this all mean for the UK's approach?

Navigating an uncertain future with appropriate agility is only possible with clarity about aims. The UK government should seek to clearly articulate what it wants to achieve with its data system: what economic, social and security-related ambitions it has for better use of citizen data and what objectives for security, inclusion and individual rights it will prioritise.

It will be important to take a holistic approach to data systems. In developing its strategy, to avoid unintended consequences the government should take a 'whole system' approach. It will be important to acknowledge the complex interactions between businesses, government, the wider public sector, the third sector and the public. Commitments made in one area, for instance on the use of data for the protection of national security, can have important implications for the assurances that can be given elsewhere, such as for privacy.

Given this, the trade-offs between competing policy objectives for a data system need to be made consciously. Policymakers should be transparent and realistic about such trade-offs. In particular, governments should also recognise that seeking to maximise the benefits its citizens gain from global trade will mean not being fully free to set their own citizen data arrangements unilaterally. Coherence with regional data systems, for example the EU and regulations including GDPR in the UK's case, can be important for businesses seeking to export and consumer access to services. However, current international agreements do not encompass all aspects of domestic data systems. There are important variations in the way such agreements are implemented domestically, and the multilateral frameworks that have emerged do not necessarily preclude other forms of international coordination.

The UK should take opportunities to steer the formation of new global norms, as well as respond to them. Combined with domestic strength in data-intensive industries, showing leadership in developing forward-looking data regulation approaches, and ensuring wider economic policies are fit for the digital age, would put the UK in a strong position to do so. There may be opportunities to shape and support emerging data governance frameworks in countries with less developed systems and aligning these with the UK's could help to underpin future economic partnerships.

Members of the public need to be an active and engaged part of the UK's data system. Given the lack of consensus within and between countries on the issues discussed in this report, and variable levels of trust, governments need to actively engage with the public about data. A reliance on disinterest is unlikely to be sustainable long term. Governments should listen and respond to concerns, but also be willing to lead, educate and persuade where there is strong evidence in support of interventions. If a larger proportion of the public feel confident in our data system, more may engage with it in an informed way and access its benefits. The risks highlighted in this report need careful managing, and not all citizens will value economic gains equally, but an inability to harness data in a comprehensive way can, for instance, mean missed opportunities to help vulnerable families or improve public health.

A successful data system will need to be flexible and react quickly to changes. Given the uncertainties highlighted in this report, resilience and agility should be built into data governance frameworks. All data policy should be developed with a range of futures in mind, and the scenarios developed in this report are intended to provide a starting point for this. More generally, it will not be possible for a strategy to foresee every eventuality. Error-correction mechanisms need to be built in. Some policies or regulations will need to adjust

This is not a statement of government policy

as new facts emerge and as the global data system develops. This should not necessarily be taken as a failure of the original vision.

Finally, **we will need to continually improve our understanding of the system.** The integration of citizen data into businesses, public services and global interactions remains an emerging area of research. This should be prioritised by government and academia, building on the UK's existing strength in this field. This could support innovations, for example in energy-efficient computing and privacy-enhancing technologies, that would make the trade-offs described above easier to manage in future. This report highlights gaps, and some inconsistencies, in the available evidence. There is also a need for research into the impacts of our data system and alternative data governance models on social, economic and security outcomes; the economic effects of diverging from trading partners' policy frameworks; and how to share the benefits of data-related innovations more widely.

This is not a statement of government policy

Glossary

Adequacy decision (in context of GDPR)	*A status that allows personal data to pass freely between non-EEA countries and EEA countries without further safeguards due to the non-EEA country's level of personal data protection*
Aggregated data	*Groups of observations from data that have been gathered together and expressed in a summary form e.g. for statistical analysis*
Artificial Intelligence (AI)	*Systems that enable a machine to perform tasks that would ordinarily require human, or rational, thought processes or actions. This usually consists of computers running algorithms, drawing on data*
Algorithm	*A process or series of steps worked through to produce an output, often implemented by a computer*
Anonymised	*Information about individuals that has been modified in such a way that it cannot be related back to a given individual*
Brain-computer interface	*Systems that enable translation of brain signals to outputs in computers or physical devices*
Cloud computing	*A model of accessing computing infrastructure on-demand via the internet*
Confidentiality	*The agreed restriction of information about an individual or an organisation from being disclosed to certain parties*
Cryptography	*Secure information and communications techniques to keep information secret and protected from unintended observation using codes*
Data broker	*A business which collects certain valuable data (e.g. about consumer behaviour) sells it to other organisations*
Data cooperative	*An organisation through which members can store, manage and control access to the data that they contribute, in a democratic manner*
Data infrastructure	*Existing datasets, technologies, and standards, and the people or organisations associated with processing the datasets*
Data localisation	*A requirement to keep certain data, or copies of such data, within a particular jurisdiction or region*
Data minimisation	*Limiting the data collected and stored to what is adequate for a specified purpose*
Data nationalism	*A broad set of domestic policies and approaches with the common feature of reducing international data flows*
Data poisoning	*Introducing training data to a machine learning system that causes it to make mistakes*
Data sovereignty	*The concept that data is subject to the laws of the country or region in which it is processed*

This is not a statement of government policy

Data system	*The people, organisations, processes and technologies involved in collecting, storing, analysing, linking, and sharing citizen data; the legal, ethical, and procedural frameworks that shape how this takes place; and the wider incentives, values, behaviour and other dynamics of these actors*
Data trust	*A data governance mechanism usually involving a fiduciary duty (a legal responsibility of impartiality, prudence, transparency and undivided loyalty) for a trustee to manage data on behalf of those contributing data*
Datafication	*The process by which subjects, objects and practices of real-life behaviour are transformed into digital information*
Deep learning	*A subset of machine learning that typically uses large artificial neural networks with multiple layers*
Deepfake	*The manipulation of visual and audio content using advanced algorithmic methods. Can be used to, for example, replace an existing video or image of a person with another's likeness*
De-identified	*Data where information that links the data to an individual person has been removed or masked. The term 'depersonalised' is also often used in this way, particularly regarding health data*
Differential privacy	*A formal way of defining and managing privacy risk when releasing aggregate statistics and analysis relating to a dataset*
Digital infrastructure	*Information and communications technologies including hardware and software, which together underpin the functioning of the digital economy. This includes for example data centres, computers and fibre optic cables*
Disinformation	*Information that is false and deliberately created and spread to deceive or harm a person, social group, organisation or country*
Edge computing	*An approach to computing which involves data storage and computation close to the 'edge' of the network, nearer to end users and sources of data*
Encryption	*The translation of data from a readable format into an encoded format known as ciphertext, readable only to those with a special access method such as a key*
FAIR principles	*A set of guiding principles developed by various stakeholders for scientific data management and stewardship to make data Findable, Accessible, Interoperable and Reusable*
Freemium	*A pricing model that provides access to a service initially for free with the option to upgrade to a premium paid option*
Homomorphic encryption	*A form of encryption that enables computations to be performed on the encrypted data*
Internet of Things (IoT)	*The interconnection of large numbers of devices via the internet allowing them to exchange data*

Machine learning	*A form of artificial intelligence that allows systems to learn directly from examples, data, and experience. Instead of using pre-programmed rules, machine learning systems are trained on data, and use this to 'learn' how to perform complex tasks or detect patterns*
Metadata	*Information and documentation describing the properties of data and datasets*
Network effects	*When a product or service gains value as more people use it*
Neuromorphic computing	*Computing hardware that emulates the neural structure and operation of the human brain*
Personal data	*(In relation to GDPR) Any information relating to an identified or identifiable individual natural person*
Personal data store	*A system that provides an individual with access to data about them, and control over how it used and shared with others*
Personally identifiable information	*Information that can be used to identify a specific person*
Privacy	*A concept that can mean different things to different people, including the right to be let alone, and freedom from intrusion into matters that are considered personal*
Privacy budget	*(In relation to differential privacy) The quantitative measure of risk to an individual's privacy due to that individual's data being included in the inputs of an algorithm which is then released*
Privacy Enhancing Technologies (PETs)	*A collective term for a broad range of technologies and approaches that support in mitigating security and privacy risks associated with processing citizen data*
Privacy paradox	*The apparent discrepancy between an individual's stated intentions to protect their privacy and their actual personal information disclosure behaviours, especially online*
Pseudonymised	*Data where information that links the data to an individual person has been replaced or modified, such that the individual can no longer be identified without additional information (e.g. replacing names with reference numbers). The additional information is stored separately to prevent re-identification. See also 'De-identified'*
Quantum Computing	*Computational devices based on quantum phenomena, which would enable certain computations to be done orders of magnitude faster than existing digital computers*
Re-identification	*A process by which de-identified data can be linked back to an individual, sometimes through combining several sets of data*
Secure multi-party computation	*A subfield of cryptographic approaches that enables computation of combined data while concealing private input from different parties*
Supercomputer	*Powerful high-performance computer that can process large amounts of data very quickly and allow complex problem-solving*

This is not a statement of government policy

Synthetic data	*Data not generated by direct measurement, but artificially through algorithmic or other approaches. The data may be designed to replicate the statistical properties of a real-world dataset without including identifiable information*
Transfer learning	*A machine learning technique where features of existing, pre-trained machine learning models are transferred to a new, related use*

Acknowledgments

The Government Office for Science would like to thank the many government officials, academic and business experts and stakeholders who contributed to the work of this project, and generously provided their advice and guidance. This includes the Chief Scientific Advisers and specialists across government.

The project team in the Government Office for Science was led by Dr Arne Blackman and included Dr Arianna Sorba, Dalia Majongwe, Kiran Sidhu, Elizabeth Killen, Sepi Latifi, Peter Sellen, Dr Simon Whitfield, and Dr Tom Wells.

We are especially grateful for the time and input of Sam Cannicott, Centre for Data Ethics and Innovation; Professor Diane Coyle CBE, University of Cambridge; Professor Jon Crowcroft, University of Cambridge; Dr Hamed Haddadi, Imperial College London; Professor Neil Lawrence, University of Cambridge; Professor Sabina Leonelli, University of Exeter; Professor Carsten Maple, University of Warwick and the Alan Turing Institute; Dr Kieron O'Hara, University of Southampton; Valentina Pavel, Ada Lovelace Institute; Hetan Shah, Chief Executive, the British Academy; Jeni Tennison, Jack Hardinges and Renate Samson, Open Data Institute; Professor Steven Weber, University of California Berkeley; and Dr Philippa Westbury, Royal Academy of Engineering, for providing expert review of content in relation to this report.

References

[1] UNCTAD (2019). *Digital Economy Report 2019. Value creation and capture: implications for developing countries.* United Nations Conference on Trade and Development. https://unctad.org/en/PublicationsLibrary/der2019_en.pdf

[2] Deloitte (2017). *Assessing the value of TfL's open data and digital partnerships.* Transport for London. http://content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf

[3] Beagrie, N., Houghton, J. (2016). *The value and impact of the European Bioinformatics Institute*. EMBL-EBI. https://beagrie.com/static/resource/EBI-impact-report.pdf

[4] Coyle, D. et al. (2020, February). The Value of Data: Policy Implications. *The Bennett Institute for Public Policy, Cambridge in partnership with the Open Data Institute.* https://www.bennettinstitute.cam.ac.uk/publications/value-data-policy-implications/

[5] Digital Competition Expert Panel (2019*). Unlocking digital competition: report of the Digital Competition Expert Panel.* HM Treasury. https://www.gov.uk/government/publications/unlocking-digital-competition-report-of-the-digital-competition-expert-panel

[6] Information Commissioner's Office (2018). *Democracy disrupted? Personal information and political influence*. https://ico.org.uk/media/action-weve-taken/2259369/democracy-disrupted-110718.pdf

[7] Meadows, D. H. (2008). *Thinking in systems: A primer*. Chelsea Green Publishing.

[8] Aridor, G., Che, Y. & Salz T. (2020). The Economic Consequences of Data Privacy Regulation: Empirical Evidence from GDPR. *National Bureau of Economic Research No w26900.* http://dx.doi.org/10.2139/ssrn.3522845

[9] Wong, J., & Henderson, T. (2018). How portable is portable? Exercising the GDPR's right to data portability, *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers. ACM* (pp. 911-920). Fourth International Workshop on Legal and Technical Issues in Cloud and Pervasive Computing (IoT), Singapore, 8/10/18. https://doi.org/10.1145/3267305.3274152

[10] Singh, J. & Cobbe, J. (2019). The Security Implications of Data Subject Rights*. IEEE Security & Privacy, 17(6),* 21-30*,* https://doi.org/10.1109/MSEC.2019.2914614.

[11] McKinsey Global Institute (2019, January). Globalization in transition: The future of trade and value chains*. McKinsey and Company*. https://www.mckinsey.com/featured-insights/innovation-and-growth/globalization-in-transition-the-future-of-trade-and-value-chains

[12] Casalini, F. and López González, J. (2019). Trade and Cross-Border Data Flows, *OECD Trade Policy Papers, No. 220*, OECD Publishing. https://doi.org/10.1787/b2023a47-en.

[13] Annan, K. (2018). Data can help to end malnutrition across Africa, *Nature, 555,* 7, https://doi.org/10.1038/d41586-018-02386-3

[14] Doshi, T. (2018, September 6). Introducing the Inclusive Images Competition. *Google AI Blog,* https://ai.googleblog.com/2018/09/introducing-inclusive-images-competition.html

[15] Topol, E. & Lee, K. F. (2019). It takes a planet, *Nature Biotechnology, 37,* 858-861, https://doi.org/10.1038/s41587-019-0214-z

[16] Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., West, S.M., Richardson, R., Schultz, J. & Schwartz, O. (2018). *AI Now Report 2018*. AI Now https://ainowinstitute.org/AI_Now_2018_Report.pdf

[17] Kaye, K. (2019, August 7). These companies claim to provide "fair-trade" data work. Do they? *MIT Technology Review.* https://www.technologyreview.com/2019/08/07/133845/cloudfactory-ddd-samasource-imerit-impact-sourcing-companies-for-data-annotation/

[18] Cavoukian, Ann (2011, January). Privacy by Design: The 7 Foundational Principles. Implementation and Mapping of Fair Information Practices. *Information and Privacy Commissioner of Ontario.* https://www.ipc.on.ca/wp-content/uploads/Resources/pbd-implement-7found-principles.pdf

[19] Argentesi, E. et al. (2019, May 9). Ex-post Assessment of Merger Control Decisions in Digital Markets. *Lear.* https://www.learlab.com/publication/ex-post-assessment-of-merger-control-decisions-in-digital-markets/

[20] O'Hara, K. and Hall, W. (2018, December 7). Four Internets: The Geopolitics of Digital Governance. *Centre for International Governance Innovation. Paper No. 206.* https://www.cigionline.org/publications/four-internets-geopolitics-digital-governance

[21] Competition and Markets Authority (2020). Online platforms and digital advertising market study final report. https://www.gov.uk/cma-cases/online-platforms-and-digital-advertising-market-study

[22] House of Commons Digital, Culture, Media and Sport Committee (2019). *Disinformation and 'fake news': Final Report.* House of Commons. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf

[23] Jennings, J. P. (2006). Comparing the US and EU Microsoft antitrust prosecutions: how level is the playing field. *Erasmus L. & Econ. Rev.*, *2*, 71. http://www.eler.org/include/getdoc.php?id=70&article=16&mode=pdf

[24] Friedewald, M., van Lieshout, M., Rung, S., Ooms, M., & Ypma, J. (2014, September). Privacy and security perceptions of European citizens: A test of the trade-off model. *IFIP International Summer School on*

Privacy and Identity Management (pp. 39-53). Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-319-18621-4_4.

[25] Goldsmith, J. & Russell, S. (2018). Strengths Become Vulnerabilities. *Aegis Series No. 1806,* Hoover Institution. https://www.hoover.org/sites/default/files/research/docs/381100534-strengths-become-vulnerabilities.pdf

[26] Centre for Data Ethics and Innovation (2019, July 25). *Interim report: Review into bias in algorithmic decision-making,* gov.uk. https://www.gov.uk/government/publications/interim-reports-from-the-centre-for-data-ethics-and-innovation/interim-report-review-into-bias-in-algorithmic-decision-making

[27]OECD (2017). The role of national statistical systems in the data revolution, In *Development Co-operation Report 2017: Data for Development*. OECD Publishing. https://doi.org/10.1787/dcr-2017-8-en

[28] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Bouwman, J. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data, 3.* https://www.nature.com/articles/sdata201618

[29] McKinsey Global Institute (2016). *Digital Globalization: the new era of data flows*, McKinsey Global Institute. https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Digital%20globalization%20The%20new%20era%20of%20global%20flows/MGI-Digital-globalization-Full-report.ashx

[30] IDC & Open Evidence (2017). *European Data Market, SMART 2013/0063.* Directorate-General for Communications Networks, Content and Technology, European Commission. https://ec.europa.eu/digital-single-market/en/news/final-results-european-data-market-study-measuring-size-and-trends-eu-data-economy

[31] UK Department for Digital, Culture, Media and Sport (2020, 13 March). Explanatory framework for adequacy discussions – Section A: Covering Note. https://www.gov.uk/government/publications/explanatory-framework-for-adequacy-discussions

[32] ODI (2019). How we used the 'styles of government' tool to explore trade competitiveness. *Open Data Institute Blog.* https://theodi.org/article/styles-of-government-how-governments-use-interventions-to-boost-trade-competitiveness/.

[33] Mittelstadt, B. (2019). Principles Alone Cannot Guarantee Ethical AI. *Nature Machine Intelligence,* https://dx.doi.org/10.2139/ssrn.3391293

[34] Hao, K. (2019, December 27). In 2020, let's stop AI ethics-washing and actually do something. *MIT Technology Review.* https://www.technologyreview.com/s/614992/ai-ethics-washing-time-to-act/

[35] Fazlioglu, M. (2018). *How DPA Budget and Staffing Levels Mirror National Differences in GDP and Population.* IAPP. https://iapp.org/media/pdf/resource_center/DPA-Budget-Staffing-Whitepaper-FINAL.pdf

[36] Hodge, N. (2019, July 19). GDPR enforcement varies widely by country. *Compliance Week.* https://www.complianceweek.com/gdpr/gdpr-enforcement-varies-widely-by-country/27436.article

[37] UK Department for Digital, Culture, Media & Sport & Home Office (2020). *Online Harms White Paper.* Gov.uk. https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper

[38] UK National Data Guardian (2020). National Data Guardian https://www.gov.uk/government/organisations/national-data-guardian

[39] Clayton, R., S. J. Murdoch & R. N. Watson (2006). Ignoring the great firewall of china. *International Workshop on Privacy Enhancing Technologies*, 20-35) https://www.cl.cam.ac.uk/~rnc1/ignoring.pdf

[40] Bissel, B. (2015, December 30). What China's Anti-Terrorism Legislation Actually Says. *Lawfare.* https://www.lawfareblog.com/what-chinas-anti-terrorism-legislation-actually-says

[41] Tanner, M. S. (2017, July 20). Beijing's New National Intelligence Law: From Defence to Offense. *Lawfare.* https://www.lawfareblog.com/beijings-new-national-intelligence-law-defense-offense

[42] Taubman, G. (1998). A not-so World Wide Web: The Internet, China, and the challenges to nondemocratic rule. *Political Communication*, *15*(2), 255-272.https:// doi/10.1080/10584609809342369

[43] Sacks, S. Chen, Q. And Webster, G. (2020, July 9). Five Important Takeaways from China's Draft Data Security Law. *New America.* https://www.newamerica.org/cybersecurity-initiative/digichina/blog/five-important-take-aways-chinas-draft-data-security-law/

[44] Yusha, Z. (2018). 'Sky Net' tech fast enough to scan Chinese population in one second: report. *Global Times.*: http://www.globaltimes.cn/content/1095176.shtml

[45] Yu, Z. (2018). China turns televisions and mobile phones in villages into surveillance terminals. *Global Times.* http://www.globaltimes.cn/content/1108589.shtml

[46] Xuanzun, L. (2018). Ubiquitous surveillance cameras in a Beijing district reduce crimes by nearly 40%. *Global Times.* http://www.globaltimes.cn/content/1113386.shtml

[47] Weber, V. (2018, July 17). The Rise of China's Security-Industrial Complex. *Council on Foreign Relations.* https://www.cfr.org/blog/rise-chinas-security-industrial-complex

[48] Sacks, S. & Li, M.K. (2018, August 2). How Chinese Cybersecurity Standards Impact Doing Business in China. *Centre for Strategic & International Studies.* https://www.csis.org/analysis/how-chinese-cybersecurity-standards-impact-doing-business-china

[49] Zhang, L. (2018). China: E-Commerce Law Passed. *Global Legal Monitor.* Law Library of Congress. https://www.loc.gov/law/foreign-news/article/china-e-commerce-law-passed/

[50] Fang, S., Bigg, C. and Zhang, J. (2020, June 9). New Chinese Civil Code Introduces Greater Protection of Privacy Rights and Personal Information. *DLA Piper.*

https://www.dlapiper.com/en/uk/insights/publications/2020/06/new-chinese-civil-code-introduces-greater-protection-of-privacy-rights-and-personal-information/

[51] Chen, Y. & A. S. Cheung. (2017). The transparent self under big data profiling: privacy and Chinese legislation on the social credit system. *Journal of comparative law,12*(2). https://dx.doi.org/10.2139/ssrn.2992537

[52] Lewis, D. (2020, May). Separating Myth from Reality: How China's Social Credit System uses public data for social governance. In *The AI Powered State: China's approach to public sector innovation.* Nesta. https://www.nesta.org.uk/report/separating-myth-reality/

[53] Udemans, C. (2018, October 23). Blacklists and redlists: How China's Social Credit System actually works. *Technode.* https://technode.com/2018/10/23/china-social-credit/

[54] Daum, J. (2017, December 24). China through a glass, darkly. *China Law Translate.* https://www.chinalawtranslate.com/en/seeing-chinese-social-credit-through-a-glass-darkly/

[55] Human Rights Watch (2017, December 13). China: Minority Region Collects DNA from Millions. *Human Rights Watch.* https://www.hrw.org/news/2017/12/13/china-minority-region-collects-dna-millions

[56] Fels, A. (2012). China's Antimonopology Law 2008: An Overview. Review of Industrial Organisation, *41*, pp 7–30, http://doi.org/10.1007/s11151-012-9343-y

[57] Marinello, M. (2013, November 9). The dragon awakes: Is Chinese competition policy a cause for concern? *VOX CEPR Policy Portal.* https://voxeu.org/article/chinese-competition-policy#fn

[58] Brumfield, N. A., Gidley, J.M., Zhang, Z.A., Ying, Y. (2018, March 29). China Merges Antitrust Enforcement Agencies into One, as its Anti-monopoly Law Approaches 10th Anniversary. *White & Case.* https://www.whitecase.com/publications/alert/china-merges-antitrust-enforcement-agencies-one-its-anti-monopoly-law-approaches

[59] US Department of State (2019). 2019 Investment Climate Statements: China. U.S. Embassy Beijing Economic Section. https://www.state.gov/reports/2019-investment-climate-statements/china/

[60]USCBC (2014). Competition Policy and Enforcement in China. *The US-China Business Council.* https://www.uschina.org/reports/competition-policy-and-enforcement-china

[61] Tabeta, S. (2018, August 2). China's decade-old antitrust law still vexes foreign companies. *Nikkei Asian Review.* https://asia.nikkei.com/Economy/China-s-decade-old-antitrust-law-still-vexes-foreign-companies

[62] Masson, J. (2019, July 19). US slams China's antitrust enforcement for bias. *GCR, Law Business Research.* https://globalcompetitionreview.com/article/1195408/us-slams-china%E2%80%99s-antitrust-enforcement-for-bias

[63] Li., Y. (2019). Tencent Music probe opens up whole new avenue for China antitrust enforcement in digital sector. *Mlex Market Insight.*https://mlexmarketinsight.com/insights-center/editors-picks/antitrust/asia/tencent-music-probe-opens-up-whole-new-avenue-for-china-antitrust-enforcement-in-digital-sector

[64] Baker McKenzie (2020, January 20). China Unveils Draft Amendments to Anti-Monopoly Law. https://www.bakermckenzie.com/en/insight/publications/2020/01/china-unveils-draft-amendments-antimonopoly-law

[65] Paul Tsai China Center (2020). Open Government Information in China. *Yale Law School* https://law.yale.edu/china-center/resources/open-government-information-china

[66] Horsley, J.P. (2003). Guangzhou's pioneering foray into open government. *The China Business Review.* https://law.yale.edu/sites/default/files/china-law-documents/2003_horlsey_guanghou_open_govt_cbr_article.pdf

[67] Zhang, L. (2019, September 10). China: Open Government Regulations Revised. *Law Library of Congress.* https://www.loc.gov/law/foreign-news/article/china-open-government-regulations-revised/

[68] China. The State Council. (2018, April 2). State Council releases regulation on scientific data management. *The State Council, People's Republic of China.* http://english.www.gov.cn/policies/latest_releases/2018/04/02/content_281476099479814.htm

[69] Horsley, J.P. (2019, July 1). Open government developments in China: Implications for US businesses. *China Business Review.* https://www.brookings.edu/opinions/open-government-developments-in-china-implications-for-us-businesses/

[70] World Wide Web Foundation (2018). Open Data Barometer – Leaders Edition. *World Wide Web Foundation.* https://opendatabarometer.org/leadersedition/report/

[71] Gao, F. (2016). Open Government Data in China: Lesson Learnt and New Approaches. *Proceedings of the 17th International Digital Government Research Conference on Digital Government Research,* 501-502. https://dl.acm.org/doi/10.1145/2912160.2912219

[72] Hsu, A., Yan, C., Cheng, Y. (2017) Addressing the gaps in China's environmental dta: the existing landscape. *Yale Data Driven Lab.* https://datadrivenlab.org/wp-content/uploads/2017/01/ThirdWave_Data_Gap_Analysis_Final.pdf

[73] Horsley, J.P. (2011). Regulations of the People's Republic of China on Open Government Information. https://law.yale.edu/sites/default/files/documents/pdf/china/ogi_regulations_eng_jph_rev_9-11.pdf

[74] OCED.Stat (2020). *Digital Services Trade Restrictiveness Index.* Organisation for Economic Co-operation and Development. https://stats.oecd.org/Index.aspx?DataSetCode=STRI_DIGITAL;

[75] Ferracane, M.F., Lee-Makiyama, H. & van der Marel, E. (2018*). Digital Trade Restrictiveness Index.* European Centre for International Political Economy. https://ecipe.org/dte/dte-report/

[76] Bird, R. and Yi Quah, P. (2019, September 11). Where are we now with data protection law in China? *Freshfields Bruckhaus Deringer.* https://digital.freshfields.com/post/102fqnd/where-are-we-now-with-data-protection-law-in-china-updated-september-2019

[77] Livingston, S. & Greenleaf, G. (2016, September 29). Data Localisation in China and Other APEC Jurisdictions, Research Paper No. 17-11. *Privacy Laws & Business International Report, 143,* 22-26; UNSW Law https://ssrn.com/abstract=2895610

[78] Kennedy, G. and Lee, K. H. F. (2019, August 22). More Changes on the Horizon: New Cross-Border Transfer Restrictions and Personal Information Requirements in the PRC. *Mayer Brown.* https://www.mayerbrown.com/en/perspectives-events/publications/2019/08/more-changes-on-the-horizon-new-cross-border-transfer-restrictions-and-personal-information-requirements-in-the-prc

[79] White, V.& Cheung, I. (2015, January 20). China's rules on encryption: what foreign companies need to know. *Freshfields Bruckhaus Deringer.* https://www.lexology.com/library/detail.aspx?g=c20a0a51-a667-417a-8e96-c473b1eecfaf

[80] Stevenson, A. & Mozur, P. (2019, September 22). China Scores Businesses, and Low Grades Could Be a Trade-War Weapon. *The New York Times.* https://www.nytimes.com/2019/09/22/business/china-social-credit-business.html

[81] Chatham House (2019). China's Belt and Road Initiative (BRI). *Chatham House.* https://www.chathamhouse.org/research/topics/china-belt-and-road-initiative-bri#

[82] Shahbaz, A. (2018). Freedom on the Net 2018: The rise of digital authoritarianism. *Freedom House.*https://freedomhouse.org/report/freedom-net/2018/rise-digital-authoritarianism

[83] Morocco World News (2019, April 28). Morocco Partners with Chinese Companies to Develop Tangier Tech City. *Morocco World News.* https://www.moroccoworldnews.com/2019/04/271663/morocco-china-tangier-tech-city/

[84] Gross, A., Murgia, M. & Yang, Y. (2019, December 1) Chinese tech groups shaping UN facial recognition standards. *Financial Times.* https://www.ft.com/content/c3555a3c-0d3e-11ea-b2d6-9bf4d1957a67

[85] United States Trade Representative (2014). 2014 Report on Technical Barriers to Trade. *Office of the United States Trade Representative.* https://ustr.gov/sites/default/files/2014%20TBT%20Report.pdf

[86] Levin, D. (2015, December 16). At U.N., China Tries to Influence Fight Over Internet Control. *The New York Times.* https://www.nytimes.com/2015/12/17/technology/china-wins-battle-with-un-over-word-in-internet-control-document.html

[87] Nouwens, M. & H. Legarda (2018). Emerging technology dominance: what China's pursuit of advanced dual-use technologies means for the future of Europe's economy and defence innovation. *International Institute for Strategic Studies: China Security Project.* https://www.merics.org/sites/default/files/2018-12/181218_Emerging_technology_dominance_MERICS_IISS.pdf

[88] China. Central Compilation and Translation Press (2016). The 13th five-year plan for economic and social development of The People's Republic of China (2016- 2020). *Central Compilation and Translation Press.* https://en.ndrc.gov.cn/newsrelease_8232/201612/P020191101481868235378.pdf

[89] Hoffman, S. (2017, December 12). Programming China: The Communist Party's autonomic approach to managing state security. *Mercator Institute for China Studies (MERICS) China Monitor.* https://www.merics.org/sites/default/files/2017-12/171212_China_Monitor_44_Programming_China_EN__0.pdf

[90] EU. European Data Protection Supervisor (2020). Data Protection. *European Data Protection Supervisor.* https://edps.europa.eu/data-protection/data-protection_en

[91] European Commission (2019, July 24) *Communication from the Commission to the European Parliament and the Council: Data protection rules as a trust-enabler in the EU and beyond – taking stock. COM (2019).374.* European Commission. https://ec.europa.eu/info/sites/info/files/aid_development_cooperation_fundamental_rights/aid_and_development_by_topic/documents/communication_2019374_final.pdf

[92] Chowdhury, S. & Moës, N. (2018, June 28). Trading invisibles: Exposure of countries to GDPR. *Bruegel.* https://bruegel.org/2018/06/trading-invisibles-exposure-of-countries-to-gdpr/

[93] Frumholz, 2000; Kobrin, 2004; Reidenbeg, 2000, cited in Movius, L.B. & Krup, N. (2009). U.S and EU Privacy Policy: Comparison of Regulatory Approaches. *International Journal of Communication, 3* pp178. https://ijoc.org/index.php/ijoc/article/view/405/305

[94] Ciriani, S. (2015). The Economic Impact of the European Reform of Data Protection, Regulatory Affairs, Orange, France, *Digiworld Economic Journal, 1*(97)*,* pp41. https://ideas.repec.org/a/idt/journl/cs9702.html

[95] Curtis, T. (2016). Privacy harmonisation and the developing world: The impact of the EU's General Data Protection Regulation on developing economies. *Washington Journal of Law, Technology & Arts, 12*(1). https://digitalcommons.law.uw.edu/wjlta/vol12/iss1/5/

[96] Scott, M. & Cerulus, L. (2018, January 31). Europe's new data protection rules export privacy standards worldwide. *Politico.* https://www.politico.eu/article/europe-data-protection-privacy-standards-gdpr-general-protection-data-regulation/

[97] Information Commissioner's Office (2020). Guide to Privacy and Electronic Communications Regulations. https://ico.org.uk/for-organisations/guide-to-pecr/

[98] European Commission (2020). Proposal for an ePrivacy Regulation. https://ec.europa.eu/digital-single-market/en/proposal-eprivacy-regulation

[99] Waxman, O.B. (2018, May 24). The GDPR is just the latest example of Europe's caution on privacy rights. That outlook has a disturbing history. *Time Magazine.*https://time.com/5290043/nazi-history-eu-data-privacy-gdpr/;

[100] Bach, D. (2001) The new economy: Transatlantic policy comparison, industry and self-regulation in the economy. *Berkeley Roundtable on the International Economy.*

[101] Germany. Bundesverfassungsgericht (1983, December 15). Abstract of the German Federal Constitutional Court's Judgment of 15 December 1983, 1 BvR 209, 269, 362, 420, 440, 484/83 [CODICES https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/1983/12/rs19831215_1bvr020983en.html

[102] Manne, G.A. (2018). Why US Antitrust Law Should Not Emulate European Competition Policy, *International Centre for Law & Economics.* https://www.judiciary.senate.gov/imo/media/doc/Manne%20Testimony.pdf

[103] European Commission (2019, April 29). Overview: making markets work better. *European Commission.* https://ec.europa.eu/competition/general/overview_en.html

[104] Erbach, G. (2014). EU and US competition policies: Similar objectives, different approaches. *European Parliamentary Research Service, Briefing.*http://www.europarl.europa.eu/RegData/bibliotheque/briefing/2014/140779/LDM_BRI(2014)140779_REV1_EN.pdf

[105] European Comission (2018, September 6). Mergers: Commission clears Apple's acquisition of Shazam. *European Commission.* https://europa.eu/rapid/press-release_IP-18-5662_en.htm

[106] Germany. Bundeskartellamt (2019, February 2). Bundeskartellamt prohibits Facebook from combining user data from different sources. Bundeskartellamt https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2019/07_02_2019_Facebook.html

[107] Germany. Bundeskartellamt (2019, August 26). The Decision of the Higher Regional Court of Düsseldorf (Oberlandesgericht Düsseldorf) in interim proceedings, 26 August 2019, Case VI-Kart 1/19 (V). *Bundeskartellamt*. https://www.d-kart.de/wp-content/uploads/2019/08/OLG-D%C3%BCsseldorf-Facebook-2019-English.pdf

[108] Chazan, G. (2020, June 23). Germany says Facebook must comply with antitrust ruling on data use. *Financial Times.* https://www.ft.com/content/a169921d-4744-4c16-8ae8-028d52bb655c

[109] Competition and Markets Authority (2020, July 1). Digital Markets Taskforce: launch of call for information and stakeholder engagement, *gov.uk.* https://www.gov.uk/cma-cases/digital-markets-taskforce#launch-of-call-for-information-and-stakeholder-engagement

[110] European Commission (2020, June 2). Antitrust: Commission consults stakeholders on a possible new competition tool. https://ec.europa.eu/commission/presscorner/detail/en/IP_20_977

[111] Hardinges, J. & Whitworth, G. (2018, February 15) Will GDPR and data portability support innovation? *Open Data Institute.* https://theodi.org/article/will-gdpr-and-data-portability-support-innovation/

[112] European Commission. (2020) *A European strategy for data.* https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy [accessed May 2020]

[113] EU. EUR-Lex (2016, June 7). Consolidated version of the Treaty on European Union. *European Union Law.* https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.C_.2016.202.01.0001.01.ENG&toc=OJ:C:2016:202:TOC#C_2016202EN.01001301

[114] Väljataga, A. (2018). ECtHR: When not backed by strong safeguards, mass surveillance violates privacy. The NATO Cooperative Cyber Defence Centre of Excellence (CCDCOE). https://ccdcoe.org/incyder-articles/ecthr-when-not-backed-by-strong-safeguards-mass-surveillance-violates-privacy/

[115] Court of Justice of the European Union (2015, April 8). *Press Release No. 54/14, The Court of Justice declares the Data Retention Directive to be invalid.* https://curia.europa.eu/jcms/upload/docs/application/pdf/2014-04/cp140054en.pdf

[116] Court of Justice of the European Union (2016, December 21). *Press Release No 145/16, The Members States may not impose a general obligation to retain data on providers of electronic communications services.* https://curia.europa.eu/jcms/upload/docs/application/pdf/2016-12/cp160145en.pdf

[117] European Court of Human Rights (2018, September 13). *Case of Big Brother Watch and others v. the United Kingdom.* https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-186048%22]}

[118] European Court of Human Rights (2019, September). *Mass surveillance.* https://www.echr.coe.int/Documents/FS_Mass_surveillance_ENG.pdf

[119] Azarmi, M. (2019, March 4). European Court of Human Rights to Reexamine Bulk Collection. *Centre for Democracy & Technology.* https://cdt.org/insights/european-court-of-human-rights-to-reexamine-bulk-collection/

[120] European Commission (2019, October 29). *Code of Practice on Disinformation one year on: online platforms submit self-assessment reports.* https://ec.europa.eu/commission/presscorner/detail/en/statement_19_6166

[121] European Court of Human Rights (2013). *National security and European case-law. Council* of Europe. https://www.echr.coe.int/Documents/Research_report_national_security_ENG.pdf

[122] Patil, S., Patruni, B., Lu, H., Dunkerley, F., Fox, J., Potoglou, D. & Robinson, N. (2015). *Public Perception of Security and Privacy: Results of the comprehensive analysis of PACT's pan-European Survey*. RAND Europe. https://www.rand.org/randeurope/research/projects/pact-security-privacy.html#findings

[123] Pew Research Centre (2016, September 21). *The state of privacy in post-Snowden America.* FACTANK, Pew Research Centre. https://www.pewresearch.org/fact-tank/2016/09/21/the-state-of-privacy-in-america/

[124] European Commission (2020, March 8). *European legislation on open data and the re-use of public sector information.* Data Policy and Innovation, European Commission. https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information

[125] European Commission (2020, March 12). *Guidance on private sector data sharing.* Data Policy and Innovation, European Commission. https://ec.europa.eu/digital-single-market/en/guidance-private-sector-data-sharing

[126] European Union (2020). *EU Open Data Portal.* https://data.europa.eu/euodp/en/home

[127] OECD (2019) *Government at a Glance 2019,* OECD Publishing. https://doi.org/10.1787/8ccf5c38-en

[128] Harwich, E. & Lasko-Skinner, R. (2018). *Making NHS data work for everyone.* Reform. https://reform.uk/sites/default/files/2018-12/Making%20NHS%20data%20work%20for%20everyone%20WEB_1.pdf

[129] European Commission (2018). *Commission Staff Working Document: Impact Assessment: Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information.* https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=SWD:2018:0127:FIN:EN:PDF

[130] European Data Portal (2016). *Open Data Goldbook for Data Manager and Data Holders.* European Commission. https://www.europeandataportal.eu/sites/default/files/goldbook.pdf

[131] European Commission (2020). *Internal Market, Industry, Entrepreneurship and SMEs: The European single market.* https://ec.europa.eu/growth/single-market_en

[132] European Commission (2020, February 24). *Free flow of non-personal data.* Cloud and Software, European Commission. https://ec.europa.eu/digital-single-market/en/free-flow-non-personal-data

[133] EU. EUR-Lex (2003, July 5). *2003/490/EC: Commission Decision of 30 June 2003 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequate protection of personal data in Argentina (Text with EEA relevance).* European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32003D0490

[134] European Commission (2019, January 23). *European Commission adopts adequacy decision on Japan, creating the world's largest area of safe data flows.* https://europa.eu/rapid/press-release_IP-19-421_en.htm

[135] Movius, L.B. & Krup, N. (2009). U.S. and EU Privacy Policy: Comparison of Regulatory Approaches**.** *International Journal of Communication, 3,* pp19, https://ijoc.org/index.php/ijoc/article/view/405

[136] Ciriani, S. (2015) The Economic Impact of the European Reform of Data Protection *Communications & Strategies, 97*, 41-58. https://ssrn.com/abstract=2674010
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2674010

[137] Court of Justice of the European Union (2020). Press release No 91/20 on Judgment in Case C-311/18; Data Protection Commissioner v Facebook Ireland and Maximillian Schrems. The Court of Justice invalidates Decision 2016/1250 on the adequacy of the protection provided by the EU-US Data Protection Shield. https://curia.europa.eu/jcms/upload/docs/application/pdf/2020-07/cp200091en.pdf

[138] UK Department for Digital, Culture, Media and Sport (2020, March 20). Policy paper: explanatory framework for adequacy discussions, *gov.uk*. https://www.gov.uk/government/publications/explanatory-framework-for-adequacy-discussions

[139] Watanabe, P.J. (2017). An Ocean Apart: The Transatlantic Data Privacy Divide and the Right to Erasure. *Southern California Law Review, 90*(5), https://southerncalifornialawreview.com/2017/07/01/an-ocean-apart-the-transatlantic-data-privacy-divide-and-the-right-to-erasure-note-by-paul-j-watanabe/

[140] Dimitrova, A. & Brkan, M. (2017). Balancing National Security and Data Protection: The Role of EU and US Policy-Makers and Courts before and after the NSA Affair. *Journal of Common Market Studies, 56*(4), 751-767, https://doi.org/10.1111/jcms.12634

[141] US Department of Justice. *Privacy Act of 1974.* https://www.justice.gov/opcl/privacy-act-1974

[142] Federal Trade Commission 2019, July 24). FTC Imposes $5 Billion Penalty and Sweeping New Privacy Restrictions on Facebook. https://www.ftc.gov/news-events/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions

[143] Mulligan, S.P., Freeman, W.C., Linebaugh, C.D. (2019, March 25). *Data Protection Law: An Overview.* Congressional Research Service. https://fas.org/sgp/crs/misc/R45631.pdf

[144] U.S. Office for Civil Rights (2013). *Summary of the HIPAA Privacy Rule.* U.S. Department of Health & Human Services. https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html

[145] U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Assistance. (2019). *Electronic Communications Privacy Act of 1986 (ECPA),* Justice Information Sharing, U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Assistance. https://it.ojp.gov/PrivacyLiberty/authorities/statutes/1285

[146] Thomson Reuters Practical Law (2020). Glossary: Computer Fraud and Abuse Act (CFAA) https://uk.practicallaw.thomsonreuters.com/2-508-3428?transitionType=Default&contextData=(sc.Default)&firstPage=true&bhcp=1 [Accessed June 2020]

[147] U.S. Federal Trade Commission (2006). *Federal Trade Commission Act.* https://www.ftc.gov/enforcement/statutes/federal-trade-commission-act

[148] Noordyke, M. (2020). *US State Comprehensive Privacy Law Comparison*. International Association of Privacy Professionals. https://iapp.org/resources/article/state-comparison-table/ [Accessed August 2020]

[149] State of California (2018, November 8). *The California Consumer Privacy Act of 2018, AB-375 Privacy: personal information: businesses.* California Legislative Information. https://leginfo.legislature.ca.gov/faces/billCompareClient.xhtml?bill_id=201720180AB375

[150] Illinois General Assembly. *Biometric Information Privacy Act.* https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=3004&ChapterID=57

[151] Auxier, B., Rainie, L., Anderson, M., Perrin, A., Kumar, M. & Turner, E. (2019, November 15). *Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information.* Pew Research Centre. https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/

[152] World Economic Forum (2019). *The Global Competitiveness Report.* http://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf

[153] Melamed, A. D. (2019, March 15). *Congress Hears Challenges to The Consumer Welfare Standard.* Stanford Law School. https://law.stanford.edu/press/congress-hears-challenges-to-the-consumer-welfare-standard/

[154] Larywon, M. & Hatch, J.H. (2019, March 15). *Congress Hears Challenges to the Consumer Welfare Standard.* Patterson Belknap. https://www.pbwt.com/antitrust-update-blog/congress-hears-challenges-to-the-consumer-welfare-standard

[155] Melamed, A. D., & N. Petit (2019). The Misguided Assault on the Consumer Welfare Standard in the Age of Platform Markets. *Review of Industrial Organization*, *54*(4), 741-774. https://doi.org/10.1007/s11151-019-09688-4

[156] McLaughlin, D. (2017, October 4). Tech's new monopolies. Livemint. https://www.livemint.com/Opinion/s6HVwWR1AOLe92nkhytqHI/Techs-new-monopolies.html

[157] Shubber, K. (2018, November 28). US antitrust enforcement falls to slowest rate since 1970s*. Financial Times.* https://www.ft.com/content/27a0a34e-f2a0-11e8-9623-d7f9881e729f

[158] U.S. Federal Trade Commission (2013). *Google Agrees to Change Its Business Practices to Resolve FTC Competition Concerns in the Markets for Devices Like Smart Phones, Games and Tablets, and in Online Search*. https://www.ftc.gov/news-events/press-releases/2013/01/google-agrees-change-its-business-practices-resolve-ftc

[159] European Commission (2017, June 27). *Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service*. https://ec.europa.eu/commission/presscorner/detail/en/IP_17_1784

[160] Breen, L.M., Schöning, F. (2018, September 27). *Exploring the contrasting views about antitrust and big data in the U.S. and EU.* Hogan Lovells. https://www.hoganlovells.com/en/publications/exploring-the-contrasting-views-about-antitrust-and-big-data-in-the-us-and-eu

[161] Abbot, A.F. (2019, July 6). *Big Data and Competition Policy: A US FTC Perspective. Presentation by General Counsel, U.S. Federal Trade Commission, at Penn Wharton China Center, Beijing.* U.S. Federal Trade Commission https://www.ftc.gov/system/files/documents/public_statements/1543858/big_data_and_competition_policy_china_presentation_2019.pdf

[162] U.S. Federal Trade Commission (2019, June 12). *FTC Hearing #14: Roundtable with State Attorneys General*. https://www.ftc.gov/news-events/events-calendar/ftc-hearing-14-roundtable-state-attorneys-general

[163] U.S. Federal Trade Commission (2018, November 1). *FTC Hearing #5: Vertical Merger Analysis and the Role of the Consumer Welfare Standard in U.S. Antitrust Law.* https://www.ftc.gov/news-events/events-calendar/ftc-hearing-5-competition-consumer-protection-21st-century

[164] Kolhatkar, S. (2019, August 20). *How Elizabeth Warren came up with a plan to break up big tech.* The New Yorker. https://www.newyorker.com/business/currency/how-elizabeth-warren-came-up-with-a-plan-to-break-up-big-tech

[165] Kang, C. and McCabe, D. (2020, July 29). Lawmakers, United in Their Ire, Lash Out at Big Tech's Leaders. *The New York Times.* https://www.nytimes.com/2020/07/29/technology/big-tech-hearing-apple-amazon-facebook-google.html

[166] U.S. Privacy and Civil Liberties Oversight Board (2014). *Report on the telephone records program conducted under Section 215 of the USA PATRIOT Act and on the operations of the foreign intelligence surveillance court*.https://www.pclob.gov/library/215-Report_on_the_Telephone_Records_Program.pdf

[167] Alexander, G.K (2013, June 18). *Hearing of the House Permanent Select Committee on Intelligence on How Disclosed NSA Programs Protect Americans, and Why Disclosure Aids Our Adversaries*. US Intelligence Community. https://www.intel.gov/index.php/ic-on-the-record-database/results/43-hearing-

of-the-house-permanent-select-committee-on-intelligence-on-how-disclosed-nsa-programs-protect-americans,-and-why-disclosure-aids-our-adversaries

[168] *Volz, D. & Strobel, W.P.* (2019, April 24). *NSA Recommends Dropping Phone-Surveillance Program.* The Wall Street Journal. https://www.wsj.com/articles/nsa-recommends-dropping-phone-surveillance-program-11556138247

[169] Wyden, R. (2020). *The Safeguarding Americans' Private Records Act.* https://www.wyden.senate.gov/imo/media/doc/The%20Safeguarding%20Americans%20Private%20Records%20Act%20of%202020%20One%20Pager.pdf

[170] Khaney, L. (2019). The FBI wanted a back door to the iPhone. Tim Cook said no. *Wired.* https://www.wired.com/story/the-time-tim-cook-stood-his-ground-against-fbi/#

[171] U.S. Department of Justice (2019, October 3). Attorney General Barr Signs Letter to Facebook from US, UK, and Australian Leaders Regarding Use of End-To-End Encryption. *Department of Justice Office of Public Affairs.* https://www.justice.gov/opa/pr/attorney-general-barr-signs-letter-facebook-us-uk-and-australian-leaders-regarding-use-end

[172] Court of Justice of the European Union (2015). Press Release No 117/15 on Judgement in Case C-362/14; Maximillian Schrems v Data Protection Commissioner. The Court of Justice declares that the Commission's US Safe Harbour Decision is invalid. https://curia.europa.eu/jcms/upload/docs/application/pdf/2015-10/cp150117en.pdf

[173] U.S. Department of Justice (2019, October 3). U.S. And UK Sign Landmark Cross-Border Data Access Agreement to Combat Criminals and Terrorists Online. *Department of Justice Office of Public Affairs.* https://www.justice.gov/opa/pr/us-and-uk-sign-landmark-cross-border-data-access-agreement-combat-criminals-and-terrorists

[174] Daskal, J. & Swire, P. (2019, October 8). The U.K.-U.S. CLOUD Act Agreement Is Finally Here, Containing New Safeguards. *Lawfare.* https://www.lawfareblog.com/uk-us-cloud-act-agreement-finally-here-containing-new-safeguards

[175] Jelinek, A., Buttarelli, G. (2019, July 10). LIBE Committee letters to the EDPS and to the EDPB regarding legal assessment of the impact of the US Cloud Act on the European legal framework for personal data protection. *European Data Protection Board.* https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_edps_joint_response_us_cloudact_coverletter.pdf

[176] U.S. Congress (2019, January 14). H.R.4174 - Foundations for Evidence-Based Policymaking Act of 2018. *Congress.gov* https://www.congress.gov/bill/115th-congress/house-bill/4174

[177] U.S. Whitehouse. (2019, January 14). Bill Announcement. https://www.whitehouse.gov/briefings-statements/bill-announcement-18/

[178] Kim, H. (2019, May 31). Data.gov at Ten and the OPEN Government Data Act. *Data.gov.* https://www.data.gov/meta/data-gov-at-ten-and-the-open-government-data-act/

[179] Global Open Data Index (2016). United States. *Open Knowledge Foundation.* https://index.okfn.org/place/us/

[180] World Wide Web Foundation (2017). The Open Data Barometer. https://opendatabarometer.org/?_year=2017&indicator=ODB

[181] World Wide Web Foundation (2017). The Open Data Barometer – Leaders Edition: ODB Methodology v1.0. http://opendatabarometer.org/doc/leadersEdition/ODB-leadersEdition-Methodology.pdf

[182] Centre for Open Data Enterprise (2016). North America. The Open Data Impact Map. https://opendataimpactmap.org/na

[183] Austin, T., Mader, D., Ravichandran, M., Rumsey, M. (2019). Future of Open Data: Maximizing the Impact of the OPEN Government Data Act. *Data Foundation.* https://www.datafoundation.org/future-of-open-data-maximizing-the-impact-of-the-open-government-data-act

[184] Zhu, X., Freeman, M.A. (2018). An evaluation of U.S. municipal open data portals: A user interaction framework. *Journal of the Association for Information Science and Technology* (*asis&t), 70*(1), 27-37, https://doi.org/10.1002/asi.24081

[185] Stone, A. (2018, March). Are Open Data Efforts Working? *Government Technology.* https://www.govtech.com/data/Are-Open-Data-Efforts-Working.html

[186] City of Seattle (2020). Open Data Portal. https://data.seattle.gov/

[187] Louisville Metro Government (2020). Louisville Metro Open Data. https://data.louisvilleky.gov/

[188] United States Government (1966, July 4). Public Law 89-487, To amend section 3 of the Administrative Procedure Act, chapter 324. https://www.govinfo.gov/content/pkg/STATUTE-80/pdf/STATUTE-80-Pg250.pdf

[189] Tauberer, J. (2020). The Annotated 8 Principles of Open Government Data. https://opengovdata.org/

[190] U.S. President's Management Agenda (2020). Federal Data Strategy 2020 Action Plan. *US Federal Government.* https://strategy.data.gov/assets/docs/2020-federal-data-strategy-action-plan.pdf

[191] President of the United States (2017). National Security Strategy of the United States of America. https://www.whitehouse.gov/wp-content/uploads/2017/12/NSS-Final-12-18-2017-0905-2.pdf

[192] Gao, H.S. (2018). Digital or trade? The contrasting approaches of China and US to digital trade. *Journal of International Economic Law.* http://doi.org/10.1093/jiel/jgy015

[193] Mayeda, A. & King, I. (2018, April 16). U.S. Cuts Off China's ZTE From American Tech for Seven Years. *Bloomberg.* https://www.bloomberg.com/news/articles/2018-04-16/commerce-blocks-china-s-zte-from-exporting-technology-from-u-s

[194] Swanson, A. (2020, May 15) U.S. Delivers Another Blow to Huawei With New Tech Restrictions. *New York Times.* https://www.nytimes.com/2020/05/15/business/economy/commerce-department-huawei.html (updated 14 July 2020)

[195] Goren, D.E. & Townsend, T. (2018, May 2). Are You an Exporter? You Might Be: The Often Overlooked Controls on Software with Encryption Capacity. *The National Law Review.* https://www.natlawreview.com/article/are-you-exporter-you-might-be-often-overlooked-controls-software-encryption-capacity

[196] US. Bureau of Industry and Security, Commerce (2018, November 11). Review of Controls for Certain Emerging Technologies. *Federal Register.* https://www.federalregister.gov/documents/2018/11/19/2018-25221/review-of-controls-for-certain-emerging-technologies

[197] e.g. Federal Communications Commission (2013). Telecommunications Act of 1996. https://www.fcc.gov/general/telecommunications-act-1996

[198] Gao, H.S. (2018). Digital or trade? The contrasting approaches of China and US to digital trade. *Journal of International Economic Law.* http://doi.org/10.1093/jiel/jgy015

[199] Ferencz, J. (2019) *The OECD Digital Services Trade Restrictiveness Index, OECD Trade Policy Paper No. 221.* OECD Publishing http://dx.doi.org/10.1787/16ed2d78-en

[200] OECD (2020, January). OECD Services Trade Restrictiveness Index: Policy trends up to 2020. *OEC Publishing.*

[201] European Commission (2019, January 25). 76 WTO partners launch talks on e-commerce. *European Commission.* http://trade.ec.europa.eu/doclib/press/index.cfm?id=1974

[202] Suneja, K. (2019). Global ecommerce talks strike at roots of WTO, says India. *The Economic Times.* https://economictimes.indiatimes.com/news/economy/policy/global-ecommerce-talks-strike-at-roots-of-wto-says-india/articleshow/68246646.cms?from=mdr [accessed April 2020]

[203] Lee-Makiyama, H. and Narayanan, B. (2019, August). The Economic Losses from Ending the WTO Moratorium on Electronic Transmissions. *European centre for International Political Economy.* https://ecipe.org/publications/moratorium/

[204] Asen, E. (2020, June 22). What European OECD Countries Are Doing about Digital Services Taxes. *Tax Foundation.* https://taxfoundation.org/digital-tax-europe-2020/

[205] Mitchell, A. & Mishra, N. (2019) Regulating Cross-Border Data Flows in a Data-Driven World: How WTO Law Can Contribute. *Journal of International Economic Law. 22*(3), 389-416. https://doi.org/10.1093/jiel/jgz016

[206] OECD (2013). The OECD Data Privacy Framework. *OECD Publishing.* *https://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf*

[207] UNCTAD (2018). Trade and development report 2018: power, platforms and free trade delusion. *United Nations Conference on Trade and Development.* https://unctad.org/en/PublicationsLibrary/tdr2018_en.pdf

[208] van der Marel, E., Lee-Makiyama, H. & Bauer, M. (2014). The Costs of Data Localisation: A Friendly Fire on Economic Recovery. *European Centre for International Political Economy.* http://ecipe.org/publications/dataloc/

[209] Cross Border Privacy Rules System. *About CBPRS.* http://cbprs.org/about-cbprs/ [accessed August 2020]

[210] African Union. *African Union Convention on Cyber Security and Personal Data Protection.* https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection [Accessed August 2020]

[211] Council of Europe. *Chart of signatures and ratifications of Treaty 108.* https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/108/signatures [Accessed August 2020]

[212] Council of Europe. *Chart of signatures and ratifications of Treaty 223.* https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/223/signatures [Accessed August 2020]

[213] World Trade Organisation. *Regional Trade Agreements Database.* http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx [accessed August 2020]

[214] Greenleaf, G. and Cottier, B. (2020, April 22). Comparing African Data Privacy Laws: International, African and Regional Commitments. *University of New South Wales Law Research Series.* http://dx.doi.org/10.2139/ssrn.3582478

[215] European Commission. *Adequacy Decisions.* https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en [Accessed August 2020]

[216] Office of the United States Trade Representative. *Agreement between the United States of America, the United Mexican States, and Canada 12/13/19 Text.* https://ustr.gov/trade-agreements/free-trade-agreements/united-states-mexico-canada-agreement/agreement-between

[217] Black., C., Setterfield, L. & Warren, R. (2018). Online Data Privacy from Attitudes to Action: an evidence review. *Ipsos MORI Scotland and Carnegie UK*

*Trust.*https://d1ssu070pg2v9i.cloudfront.net/pex/carnegie_uk_trust/2018/08/03110116/Online-Data-Privacy-from-Attitudes-to-Action-CUKT.pdf

[218] Ipsos Mori – World Economic Forum (2019). Global Citizens & Data Privacy. https://www.ipsos.com/sites/default/files/ct/news/documents/2019-01/ipsos-wef_-_global_consumer_views_on_data_privacy_-_2019-01-25-final.pptx_lecture_seule_0.pdf

[219] CIGI-Ipsos (2019). 2019 CIGI-Ipsos Global Survey on Internet Security and Trust. www.cigionline.org/internet-survey-2019

[220] GfK (2017). Press release: More people firmly agree with sharing personal data in return for rewards, than firmly disagree. *Growth from Knowledge.* https://www.gfk.com/insights/press-release/more-people-firmly-agree-with-sharing-personal-data-in-return-for-rewards-than-firmly-disagree/

[221] Hawkins, A. (2017, May 24). Chinese Citizens Want the Government to Rank Them. *Foreign Policy.* https://foreignpolicy.com/2017/05/24/chinese-citizens-want-the-government-to-rank-them/

[222] Kostka, G. (2018, September 17). China's social credit systems are highly popular – for now. *Mercator Institute for China Studies (Merics).* https://www.merics.org/en/blog/chinas-social-credit-systems-are-highly-popular-now

[223] Ding, J. (2019) ChinAI #61: A Backlash to Social Credit Blacklists? *ChinAI*. https://chinai.substack.com/p/chinai-61-a-backlash-to-social-credit

[224] Jing, M. (2018, Jan 5). China consumer group accuses Baidu of snooping on users of its smartphone apps. *South China Morning Post.* https://www.scmp.com/tech/china-tech/article/2127045/baidu-sued-china-consumer-watchdog-snooping-users-its-smartphone

[225] Athey, S., Catalini, C. & Tucker, C. (2017). The digital privacy paradox: Small money, small costs, small talk. *National Bureau of Economic Research*, *working paper 23488.* https://www.nber.org/papers/w23488

[226] Draper, N. and Turow, J. (2019). The corporate cultivation of digital resignation, *New Media & Society*. https://journals.sagepub.com/doi/abs/10.1177/1461444819833331?journalCode=nmsa

[227] Rose, E. (2005). Data Users versus Data Subjects: Are Consumers Willing to Pay for Property Rights to Personal Information? *Proceedings of the 38th Annual Hawaii International Conference on System Sciences* 180c. IEEE. https://doi.org/10.1109/HICSS.2005.184

[228] Egelman, S., Felt, A.P. & Wagner, D. (2013) Choice architecture and smartphone privacy: there's a price for that, *The Economics of Information Security and Privacy,* 211-236. https://doi.org/10.1007/978-3-642-39498-0_10

[229] Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information systems research, 22*(2), 254-268. https://doi.org/10.1287/isre.1090.0260

[230] Carrascal, J. P., Riederer, C., Erramilli, V., Cherubini, M., & de Oliveira, R. (2013, May). Your browsing behavior for a big mac: Economics of personal information online. *In Proceedings of the 22nd international conference on World Wide Web,* 189-200. https://doi.org/10.1145/2488388.2488406

[231] Hann, I. H., Hui, K. L., Lee, S. Y. T., & Png, I. P. (2007). Overcoming online information privacy concerns: An information-processing theory approach. *Journal of Management Information Systems, 24*(2), 13-42. https://doi.org/10.2753/MIS0742-1222240202

[232] Schreiner, M., & Hess, T. (2013). On the willingness to pay for privacy as a freemium model: First empirical evidence. *ECIS 2013 Research in Progress. 30.* http://aisel.aisnet.org/ecis2013_rip/30

[233] Acquisti, A., John, L. K. & Loewenstein G. (2013). What is privacy worth? *The Journal of Legal Studies*, *42*(2), 249-274. https://doi.org/10.1086/671754

[234] Ofcom (2016). Adults' media use and attitudes: Report 2016. https://www.ofcom.org.uk/__data/assets/pdf_file/0026/80828/2016-adults-media-use-and-attitudes.pdf

[235] European Commission, (2015). Special Eurobarometer 431: Data protection. https://doi.org/10.2838/552336. https://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_431_en.pdf

[236] Miller, C., Kitcher, H., Perera, K., Abiola, A., (2020) People, Power and Technology: The 2020 Digital Attitudes Report. London: Doteveryone. https://doteveryone.org.uk/report/peoplepowertech2020

[237] Melumad, S., & Meyer, R. (2020). Full Disclosure: How Smartphones Enhance Consumer Self-Disclosure. *Journal of Marketing*, *84*(3), 28-45.https://doi.org/10.1177/0022242920912732

[238] Gramlich, J. (2018). Defending against terrorism has remained a top policy priority for Americans since 9/11. *Pew Research Center.* https://www.pewresearch.org/fact-tank/2018/09/11/defending-against-terrorism-has-remained-a-top-policy-priority-for-americans-since-9-11/ [Accessed April 2020]

[239] Bakir, V., Cable, J., Dencik, L., Hintz, A. & McStay, A. (2015). Public Feeling on Privacy, Security and Surveillance: A Report by DATA-PSST and DCSS. *ESRC, Bangor University, and Cardiff University.* http://orca.cf.ac.uk/87335/1/Public-Feeling-on-Privacy-Security-Surveillance-DATAPSST-DCSS-Nov2015.pdf

[240] Ipsos Mori (2016). Global Trends 2016. https://www.ipsos.com/sites/default/files/2017-05/global_trends.pdf

[241] Samson, R., Gibbon, K. and Scott, A. (2019, September). *About data about us.* Open Data Institute, Luminate and RSA. https://www.thersa.org/discover/publications-and-articles/reports/data-about-us

[242] Understanding Patient Data. *How do people feel about the use of data?* https://understandingpatientdata.org.uk/how-do-people-feel-about-use-data, and references therein. [Accessed August 2020]

[243] Skatova, A., McDonald, R. L., Ma, S., & Maple, C. (2019). Unpacking Privacy: Willingness to pay to protect personal data. *PsyArXiv Preprint.* https://psyarxiv.com/ahwe4

[244] Open Data Institute (2020, March 11). Sharing data to create value in the private sector. *Open Data Institute.* https://theodi.org/article/report-sharing-data-to-create-value-in-the-private-sector/

[245] UK Department for Business, Energy and Industrial Strategy and Department for Digital, Culture, Media and Sport. Smart data: putting consumers in control of their data and enabling innovation, *gov.uk.* https://www.gov.uk/government/consultations/smart-data-putting-consumers-in-control-of-their-data-and-enabling-innovation

[246] Information Commissioner's Office (2020). Investigation into data analytics for political purposes. https://ico.org.uk/action-weve-taken/investigation-into-data-analytics-for-political-purposes/

[247] O'Flaherty, K. (2019). Apple issues new blow to Facebook and Google with this bold privacy move. *Forbes.* https://www.forbes.com/sites/kateoflahertyuk/2019/11/06/apple-issues-new-blow-to-facebook-and-google-with-this-privacy-move/#59da96ce481d

[248] Competition and Markets Authority (2020). Online platforms and digital advertising market study. https://www.gov.uk/cma-cases/online-platforms-and-digital-advertising-market-study

[249] Internet Health Report (2019). The Attention Merchants: Google, Facebook and Baidu. *Mozilla*. https://internethealthreport.org/2019/how-the-biggest-internet-companies-make-money/

[250] Chan, C. (2018, July 12). Outgrowing Advertising: Multimodal Business Models as a Product Strategy. *Andreessen Horowitz.* https://a16z.com/2018/12/07/when-advertising-isnt-enough-multimodal-business-models-product-strategy/

[251] Kiniulis, M. (2020, July 28). List of Largest Internet Companies in the World. https://www.markinblog.com/largest-internet-companies/

[252] Meeker, M. (2019). Internet Trends 2019, *Bond Capital.* https://www.bondcap.com/report/itr19/

[253] European Commission (2020). Entrepreneurship and Small and medium-sized enterprises (SMEs). https://ec.europa.eu/growth/smes_en

[254] IDC & the Lisbon Council (2019). D2.4 Second report on facts and figures: Updating the European Data Market Monitoring Tool. http://datalandscape.eu/sites/default/files/report/EDM_D2.4_2ndReport-FactsFigures_26032019.pdf

[255] CB Insights (2020). The Global Unicorn Club: Current Private Companies Valued At $1B+. https://www.cbinsights.com/research-unicorn-companies [updated 17 July 2020]

[256] Hall, W., & Pesenti, J. (2017). Growing the artificial intelligence industry in the UK, gov.uk. https://www.gov.uk/government/publications/growing-the-artificial-intelligence-industry-in-the-uk

[257] Kahn, J. (2018, August 16). Why Can't Europe Do Tech? *Bloomberg Businessweek.* https://www.bloomberg.com/news/features/2018-08-16/inside-europe-s-struggle-to-build-a-truly-global-tech-giant

[258] UN (2019). Data economy: radical transformation or dystopia? *Frontier Technology Quarterly.* https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/FTQ_1_Jan_2019.pdf

[259] European Commission (2019, April 8). Artificial intelligence: Commission takes forward its work on ethics guidelines. https://ec.europa.eu/commission/presscorner/detail/en/IP_19_1893

[260] Martin, N., Matt, C., Niebel, C. & Blind, K. (2019). How Data Protection Regulation Affects Startup Innovation. *Information System Frontiers, 21,* 1307-1324. https://doi.org/10.1007/s10796-019-09974-2

[261] Internet Health Report (2019). How do the biggest internet companies make money? *Mozilla.* https://internethealthreport.org/2019/how-the-biggest-internet-companies-make-money/

[262] Reiff, N. (2019, December 9). How Amazon Makes Money. *Investopedia. i*https://www.investopedia.com/how-amazon-makes-money-4587523

[263] e.g. statcounter (2020). Search Engine Market Share Worldwide. *GlobalStats.* https://gs.statcounter.com/search-engine-market-share; statcounter (2020) Social Media Stats Worldwide. *GlobalStats.* https://gs.statcounter.com/social-media-stats

[264] Gurman, M. (2019, March 23). Apple Reinvention as Services Company Starts for Real Monday. *Bloomberg.* https://www.bloomberg.com/news/articles/2019-03-23/apple-s-reinvention-as-a-services-company-starts-for-real-monday

[265] Libra (2020). Welcome to the Libra project. *Libra Association.* https://libra.org/en-US/

[266] Hartmans, A. & Meisenzahl, M. (2010, February 12). All the companies and divisions under Google's parent company, Alphabet, which just made yet another shake-up to its structure. *Business Insider.* https://www.businessinsider.com/alphabet-google-company-list-2017-4?r=US&IR=T

[267] Mazzucato, M. & Semieniuk (2017). Public financing of innovation: new questions. *Oxford Review of Economic Policy, 33*(1), 24-48. https://discovery.ucl.ac.uk/id/eprint/1553082/1/Mazzucato_grw036.pdf

[268] Shapiro, C. (2018). Antitrust in a Time of Populism. *International Journal of Industrial Organization*, *61*, 714-748. https://www.sciencedirect.com/science/article/abs/pii/S0167718718300031

[269]Noordyke, M. (2020). Big Tech's shift to Privacy. *International Association of Privacy Professionals.* https://iapp.org/resources/article/big-techs-shift-to-privacy-2/

[270] Schaake, M. (2020). Big Tech companies want to act like governments. *Financial Times.* https://www.ft.com/content/36f838c0-53c5-11ea-a1ef-da1721a0541e

[271] Fortune (2019). Global 500 Rankings. *Fortune Magazine.* https://fortune.com/global500/2019/search/?hqcountry=China

[272] OECD (2020). Gross Domestic spending on R&D. OECD data. https://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm

[273] Heggestuen, J. (2014, February 11). Alipay Overtakes PayPal As the Largest Mobile Payments Platform In The World. *Business Insider.* https://www.businessinsider.com/alipay-overtakes-paypal-as-the-largest-mobile-payments-platform-in-the-world-2014-2?r=US&IR=T

[274] Chorzempa, M. (2018, April 26). How China Leapfrogged Ahead of the United States in the Fintech Race. *Peterson Institute for International Economics.* https://www.piie.com/blogs/china-economic-watch/how-china-leapfrogged-ahead-united-states-fintech-race

[275] Yang, Y. Lockett, H. (2019, November 25). What is China's digital currency plan? *Financial Times.* https://www.ft.com/content/e3f9c3c2-0aaf-11ea-bb52-34c8d9dc6d84

[276] Shih, G. (2014, October 16). Alibaba affiliate Alipay rebranded Ant in new financial services push. *Reuters.* https://www.reuters.com/article/us-china-alibaba-idUSKCN0I50KJ20141016

[277] Shu, C. (2015, January 28). Data from Alibaba's E-Commerce Sites Is Now Powering A Credit-Scoring Service. *TechCrunch.* https://techcrunch.com/2015/01/27/data-from-alibabas-e-commerce-sites-is-now-powering-a-credit-scoring-service/

[278] Mullen, J. & Wang, S. (2017, September 1). Pay with your face at this KFC in China. *CNN Business.* https://money.cnn.com/2017/09/01/technology/china-alipay-kfc-facial-recognition/index.html

[279] South China Morning Post (2019). China Internet Report 2019. https://www.scmp.com/china-internet-report

[280] Candelon, F., Yang, F. & Wu, D. (2019, May 22). Are China's Digital Companies Ready to Go Global? *BCG Henderson Institute.* https://www.bcg.com/en-gb/publications/2019/china-digital-companies-ready-go-global.aspx

[281] Geromel, R. (2019, June 17). As Tech Cold War Looms, Chinese Internet Giants Like Alibaba and Tencent Tackle Emerging Markets. *Forbes.* https://www.forbes.com/sites/ricardogeromel/2019/06/17/as-tech-cold-war-looms-chinese-internet-giants-like-alibaba-and-tencent-tackle-emerging-markets/#17962cb25ee0

[282] Apple (2019). App Store Top Free Charts. https://apps.apple.com/story/id1484100916?ign-itsct=BestOfApps_SC09_PT006_WW%2F&ign-itscg=10000

[283] US. Federal Trade Commission (2019, February 27). Video Social Networking App Musical.ly Agrees to Settle FTC Allegations That it Violated Children's Privacy Law. *Federal Trade Commission.* https://www.ftc.gov/news-events/press-releases/2019/02/video-social-networking-app-musically-agrees-settle-ftc

[284] Grieger, G. (2016) Why China's public procurement is an EU issue. *European Parliamentary Research Service.* http://www.europarl.europa.eu/RegData/etudes/ATAG/2016/593571/EPRS_ATA(2016)593571_EN.pdf

[285] Donnan, S. & Leonard, J. (2019). U.S. Blacklists Eight Chinese Tech Companies on Rights Violations. *Bloomberg.* https://www.bloomberg.com/news/articles/2019-10-07/u-s-blacklists-eight-chinese-companies-including-hikvision-k1gvpq77

[286] Acxiom (2020). Effective consumer segmentation made easy. https://www.acxiom.co.uk/what-we-do/consumer-segmentation-personicx/

[287] Rieke, A., Yu, H. Robinson, D. & von Hoboken, J. (2016). Data Brokers in an open society. *Upturn.* https://www.opensocietyfoundations.org/uploads/42d529c7-a351-412e-a065-53770cf1d35e/data-brokers-in-an-open-society-20161121.pdf

[288] US. Federal Trade Commission (2014, May). Data Brokers: A Call for Transparency and Accountability: A Report of the Federal Trade Commission. https://www.ftc.gov/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014

[289] Vermont Office of the Attorney General (2018, December 11). Guidance on Vermont's Act 171 of 2018: Data Broker Regulation. https://ago.vermont.gov/wp-content/uploads/2018/12/2018-12-11-VT-Data-Broker-Regulation-Guidance.pdf

[290] Privacy International (2018, November 8). Privacy International files complaints against seven companies for wide-scale and systematic infringements of data protection law. https://privacyinternational.org/press-release/2424/press-release-privacy-international-files-complaints-against-seven-companies

[291] Information Commissioner's Office (2018). Investigation into data analytics for political purposes summary report. https://ico.org.uk/media/action-weve-taken/2260270/executive-summary.pdf

[292] Information Commissioner's Office (2019, April 12). Bounty UK fined £400,000 for sharing personal data unlawfully. https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/04/bounty-uk-fined-400-000-for-sharing-personal-data-unlawfully/

[293] Tham, E. (2018, August 23). Data dump: China sees surge in personal information up for sale. *Reuters.* https://uk.reuters.com/article/us-china-dataprivacy/data-dump-china-sees-surge-in-personal-information-up-for-sale-idUKKCN1L80IW

[294] Ram, A. & Murgia, M. (2019). Data brokers: regulators try to rein in the 'privacy deathstars'. *Financial Times.* https://www.ft.com/content/f1590694-fe68-11e8-aebf-99e208d3e521

[295] European Data Protection Board (2019, May 22). 1 year GDPR – taking stock. https://edpb.europa.eu/news/news/2019/1-year-gdpr-taking-stock_en

[296] Privacy Affairs (2020). GDPR Fines Tracker and Statistics. https://www.privacyaffairs.com/gdpr-fines/. [Accessed August 2020, last updated 20 July 2020].

[297] European Commission (2020, June 24). Commission report: EU data protection rules empower citizens and are fit for the digital age. https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1163

[298] Vinocur, N. (2019, December). 'We have a huge problem': European regulator despairs over lack of enforcement. *Politico.* https://www.politico.eu/article/we-have-a-huge-problem-european-regulator-despairs-over-lack-of-enforcement/

[299] Data Transfer Project (2020). About us. https://datatransferproject.dev/

[300] Cyphers, B. and O'Brien, D. (2018, July 24). Facing Facebook: Data Portability and Interoperability Are Anti-Monopoly Medicine. *Electronic Frontier Foundation.* https://www.eff.org/deeplinks/2018/07/facing-facebook-data-portability-and-interoperability-are-anti-monopoly-medicine

[301] Grelf, B. (2018, October 10). Study: Google Is the Biggest Beneficiary of the GDPR. *Cliqz.* https://cliqz.com/en/magazine/study-google-is-the-biggest-beneficiary-of-the-gdpr

[302] Urban, T., Tatang, D., Degeling, M., Holz, T., and Pohlmann, N. (2020). Measuring the Impact of the GDPR on Data Sharing in Ad Networks. *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security (ASIA CCS '20), October 5–9, 2020, Taipei, Taiwan.* https://doi.org/10.1145/3320269.3372194

[303] Information Commissioner's Office (2019). Update report into adtech and real time bidding. https://ico.org.uk/media/about-the-ico/documents/2615156/adtech-real-time-bidding-report-201906.pdf

[304] Information Commissioner's Office (2020, May 7). ICO statement on Adtech work. https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2020/05/ico-statement-on-adtech-work/

[305] Ren, J., Dubois, D.J., Choffnes, D., Mandalari, A.M., Kolcun, R. & Haddadi, H. (2019). Information exposure from consumer IOT devices: A multidimensional, network-informed measurement approach. *Proceedings of the Internet Measurement Conference,* 267-279. https://moniotrlab.ccis.neu.edu/wp-content/uploads/2019/09/ren-imc19.pdf

[306] World Bank (2016). World Development Report 2016: Digital Dividends. *World Bank*, https://doi.org/10.1596/978-1-4648-0671-1 http://documents.worldbank.org/curated/en/896971468194972881/pdf/102725-PUB-Replacement-PUBLIC.pdf

[307] Ferracane, M. F. & van der Marel, E. (2018). Do Data Policy Restrictions Inhibit Trade in Services? *European Centre for International Political Economy.* https://ecipe.org/wp-content/uploads/2018/10/Do-Data-Policy-Restrictions-Inhibit-Trade-in-Services-final.pdf

[308] Ferracane, M. F., van der Marel, E. & Kren, J. (2018, October). The Cost of Data Protectionism. *European Centre for International Political Economy, blog.* http://ecipe.org/blog/the-cost-of-data-protectionism/

[309] Ferracane, M. F., Kren, J, & van der Marel, E. (2018, October). Do Data Policy Restrictions Impact the Productivity Performance of Firms and Industries? *European Centre for International Political Economy.* https://ecipe.org/wp-content/uploads/2018/10/Do-Data-Policy-Restrictions-Impact-the-Productivity-Performance-of-Firms-and-Industries-final.pdf

[310] Lewis, J. (2018). Economic Impact of Cybercrime – No Slowing Down. *McAfee.*https://www.mcafee.com/enterprise/en-us/assets/reports/restricted/rp-economic-impact-cybercrime.pdf

[311] Phillips, J. (2019, April). One year on, what has been the impact of GDPR on data security? *Intelligent Cisco.* https://www.intelligentciso.com/2019/04/16/one-year-on-what-has-been-the-impact-of-gdpr-on-data-security/

[312] Goldfarb, A., Gans, J., & Agrawal, A. (2019). The Economics of Artificial Intelligence: An Agenda. *University of Chicago Press*. https://static1.squarespace.com/static/5a9ef8f689c1729a3dba4225/t/5d15399dc7e0e300014ebc8e/1561672094284/The+Economics+of+Artificial+Intelligence+-+Chapter+19.pdf

[313] OECD. Working Party on International Trade in Goods and Services Statistics (2019). OECD-WTO Handbook on Measuring Digital Trade. *OECD Statistics and Data Directorate, Committee On Statistics And Statistical Policy.* http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/CSSP/WPTGS(2019)4&docLanguage=En

[314] Open Data Institute (2019). What are the links between data infrastructure and trade competitiveness? *Open Data Institute.* https://theodi.org/article/what-are-the-links-between-data-infrastructure-and-trade-competitiveness/

[315] US, California. Office of the Attorney General (2020). California Consumer Privacy Act (CCPA): Proposed Regulations Package Submitted to OAL. https://oag.ca.gov/privacy/ccpa

[316] Pereira, F. (2019, April 25). Brazil – Is the new Brazilian data protection law a cut-and-paste of the European GDPR. *Lus Laboris Lawers.* https://www.lexology.com/library/detail.aspx?g=307adbe3-69df-49d7-9bb0-8539de93a29b

[317] UNCTAD (2016). Data protection regulations and international data flows: Implications for trade and development. *United Nations Conference on Trade and Development.* https://unctad.org/en/PublicationsLibrary/dtlstict2016d1_en.pdf

[318] International Monetary Fund (2017). World Economic Outlook: Gaining Momentum? Chapter 2: Roads less traveled: growth in emerging market and developing economies in a complicated external environment. https://www.imf.org/en/Publications/WEO/Issues/2017/04/04/world-economic-outlook-april-2017#Chapter%202

[319] OECD (2019). Real GDP forecast (indicator). https://doi.org/10.1787/1f84150b-en

[320] UK. Ministry of Defence (2018). Global Strategic Trends: The Future Starts Today. 6th Edition. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/771309/Global_Strategic_Trends_-_The_Future_Starts_Today.pdf

[321] UNCTAD. (2018). Key Statistics and Trends in International Trade 2018. *United Nations Conference on Trade and Development.* https://unctad.org/en/PublicationsLibrary/ditctab2019d2_en.pdf

[322] World Trade Organisation (2019). World Trade Statistical Review 2019., https://www.wto.org/english/res_e/statis_e/wts2019_e/wts2019_e.pdf

[323] OECD (2019). OECD Economic Outlook, Volume 2019 Issue 1. https://doi.org/10.1787/b2e897b0-en.

[324] ITU (2018). *Measuring the Information Society Report 2018: Volume 1*. http://handle.itu.int/11.1002/pub/8114a552-en. [Updated 2018 statistics from www.itu.int]

[325] The World Bank (2017). World Development Indicators. https://openknowledge.worldbank.org/handle/10986/26447

[326] United Nations, Department of Economic and Social Affairs, Population Division (2019). Probabilistic Population Projections based on the World Population Prospects 2019: http://population.un.org/wpp/

[327] GSMA Intelligence (2019). The Mobile Economy Sub-Saharan Africa 2019. https://www.gsmaintelligence.com/research/?file=36b5ca079193fa82332d09063d3595b5&download

[328] Facebook. (2018). Facebook Q4 2018 Results. https://s21.q4cdn.com/399680738/files/doc_financials/2018/Q4/Q4-2018-Earnings-Presentation.pdf

[329] Weber, S. (2017). Data, development, and growth. Business and Politics, 19(3), 397-423. doi:10.1017/bap.2017.3

[330] ITU. (2020). *Country Classifications*. International Telecommunications Union. https://www.itu.int/en/ITU-D/Statistics/Pages/definitions/regions.aspx [Accessed April 2020]

[331] McGowan, K., Vora, P., Homer, M. & Dolan J. (2018). Personal data empowerment: restoring power to the people in a digital age. *Pathways for Prosperity Commission.* https://pathwayscommission.bsg.ox.ac.uk/Personal-data-empowerment-paper

[332] ICANN (2016, October 1). Stewardship of IANA Functions Transitions to Global Internet Community as Contract with U.S. Government Ends. https://www.icann.org/news/announcement-2016-10-01-en

[333] Balmforth, T. & Kolomychenko, M. (2019, February 12). Russian lawmakers back bill on 'sovereign' Internet. *Reuters.* https://www.reuters.com/article/us-russia-internet/russian-lawmakers-back-bill-on-sovereign-internet-idUSKCN1Q11RJ

[334] Russian Federation (2014). Federal Law No. 242-FZ of July 21, 2014 on Amending Some Legislative Acts of the Russian Federation in as Much as It Concerns Updating the Procedure for Personal Data Processing in Information-Telecommunication Networks (with Amendments and Additions. https://pd.rkn.gov.ru/authority/p146/p191/

[335] Republic of India. (2018). The Personal Data Protection Bill, 2018. https://meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill%2C2018_0.pdf

[336] Goel., V. (2018). India Pushes Back Against Tech 'Colonization' by Internet Giants. *The New York Times.* https://www.nytimes.com/2018/08/31/technology/india-technology-american-giants.html

[337] Srivastava, A. (2019, July 14). India must reclaim its lost digital space. *The Hindu Businessline.* https://www.thehindubusinessline.com/opinion/india-must-reclaim-its-lost-digital-space/article27941890.ece

[338] European Data Protection Board (2016). Guidelines 3/2018 on the territorial scope of the GDPR (Article 3) - Version for public consultation. https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_3_2018_territorial_scope_en.pdf

[339] UNCTAD (2020). Data Protection and Privacy Legislation Worldwide. *United Nations Conference on Trade and Development.* https://unctad.org/en/Pages/DTL/STI_and_ICTs/ICT4D-Legislation/eCom-Data-Protection-Laws.aspx [Accessed August 2020]

[340] CNIL (2019). Data Protection Around the World. Commission Nationale de l'Informatique et des Libertés https://www.cnil.fr/en/data-protection-around-the-world

[341] Bradford, A. (2020). How Europe Rules the Digital Economy. *Project Syndicate Special Edition Magazine, Spring 2020: Beyond the Techlash.* https://www.project-syndicate.org/onpoint/brussels-effect-digital-economy-by-anu-bradford-2020-04

[342] Gamble, C. & Gao, J. (2018, August 17). Safety-first AI for autonomous data centre cooling and industrial control. *Deepmind.* https://deepmind.com/blog/article/safety-first-ai-autonomous-data-centre-cooling-and-industrial-control

[343] Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., ... & Luccioni, A. (2019). Tackling climate change with machine learning. *arXiv preprint* https://arxiv.org/pdf/1906.05433.pdf

[344] The Shift Project, (2019). Lean ICT: Towards Digital Sobriety. https://theshiftproject.org/wp-content/uploads/2019/03/Lean-ICT-Report_The-Shift-Project_2019.pdf

[345] Belkhir, L., & A. Elmeligi (2018). Assessing ICT global emissions footprint: Trends to 2040 & recommendations. *Journal of Cleaner Production, 177, 448-463.* https://www.sciencedirect.com/science/article/pii/S095965261733233X

[346] Climatescope (2019). Climatescope 2019: Key Findings. *Bloomberg NEF.* http://global-climatescope.org/key-findings

[347] IEA (2020, June). Data centres and data transmission networks. *International Energy Agency.* https://www.iea.org/reports/data-centres-and-data-transmission-networks [Accessed August 2020]

[348]Shehabi, A., Smith, S., Sartor, D., Brown, R., Herrlin, M. Koomey, J., Masanet, E., Horner, N., Azevedo, I., Linter, W. (2016, June). United States Data Centre Energy Usage Report. Ernest Orlando Lawrence Berkley National Laboratory. l-1005775_v2.pdf

[349] Gossart, C. (2015). Rebound effects and ICT: a review of the literature. In *ICT innovations for sustainability*. 435-448. https://link.springer.com/chapter/10.1007/978-3-319-09228-7_26

[350] Evans, D. (2011, April). *The Internet of Things: How the Next Evolution of the Internet is Changing Everything.* Cisco. https://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf

[351] Cisco. (2020, March 9). *Cisco Annual Internet Report (2018–2023) White Paper.* https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html

[352] Dragland, A. (2013, May 22). Big Data, for better or worse: 90% of world's data generated over last two years. *Science Daily.* https://web.archive.org/web/20130604154149/http:/www-01.ibm.com/software/data/bigdata/

[353] Data aggregated from Cisco Annual Internet Reports from 2008-2017. The most recent annual report at time of publication can be found at: Cisco (2020, March 9). *Cisco Annual Internet Report (2018–2023) White Paper.* https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

[354] Mohajeri Moghaddam, H., Acar, G., Burgess, B., Mathur, A., Huang, D. Y., Feamster, N., ... & Narayanan, A. (2019, November). Watching you watch: The tracking ecosystem of over-the-top tv streaming devices. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security.*131-147. https://tv-watches-you.princeton.edu/tv-tracking-acm-ccs19.pdf

[355] Van Elegem, L. (2020, February 2020). The most innovative Smart City projects in the world. *Nexxworks Blog.* https://nexxworks.com/blog/the-most-innovative-smart-city-projects-in-the-world

[356] Mulligan, C. E. A. (2014). The Impact of Datafication on Strategic Landscapes. *Ericsson and Imperial College Business School.* http://www.ericsson.com/res/docs/2014/the-impact-of-datafication-on-strategic-landscapes.pdf

[357] Drew, L. (2019, July 24). The ethics of brain–computer interfaces. *Nature, 517,* s19-s21 https://www.nature.com/articles/d41586-019-02214-2

[358] Singer, N. (2018, July 9). Facebook's Push for Facial Recognition Prompts Privacy Alarms. *The New York Times.* https://www.nytimes.com/2018/07/09/technology/facebook-facial-recognition-privacy.html

[359] NHS (2019). *The NHS Long Term Plan.* https://www.longtermplan.nhs.uk/

[360] Maple, C. (2017). Security and privacy in the internet of things. *Journal of Cyber Policy* 2.2: 155-184. https://doi.org/10.1080/23738871.2017.1366536

[361] De Montjoye, Y. A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the crowd. The privacy bounds of human mobility. *Scientific reports, 3,* 1376. https://doi.org/10.1038/srep01376

[362] Narayanan, A. & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. *IEEE Trans. Secur. Priv. 8*, 111–125 https://doi.org/10.1109/SP.2008.33

[363] L. Sweeney. (2002). k-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 10*(5), 557-570. https://epic.org/privacy/reidentification/Sweeney_Article.pdf

[364] Garfinknel, S., Abowd, J.M., Martindale, C. (2018). Understanding Database Reconstruction Attacks on Public Data. *Security, 16*(5). https://queue.acm.org/detail.cfm?id=3295691

[365] Rocher, L., Hendrickx, J.M. & de Montjoye, Y. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. *Nature Communications* 10, 3069 https://doi.org/10.1038/s41467-019-10933-3https://doi.org/10.1038/s41467-019-10933-3

[366] Government Statistical Service Quality Centre (2018). Privacy and data confidentiality methods: A National Statistician's Quality Review. *Government Statistical Service.* https://gss.civilservice.gov.uk/policy-store/privacy-and-data-confidentiality-methods-a-national-statisticians-quality-review-nsqr/

[367] Page, H., Cabot, C. & Nissim, K. (2018). Differential privacy: an introduction for statistical agencies. *Government Statistical Service.* https://gss.civilservice.gov.uk/wp-content/uploads/2018/12/12-12-18_FINAL_Privitar_Kobbi_Nissim_article.pdf

[368] Brynjolfsson, E., Collis, A. & Eggers, F. (2019). Using massive online choice experiments to measure changes in well-being. Proceedings of the National Academy of Sciences, 116(15), 7250-7255. https://doi.org/10.1073/pnas.1815663116

[369] Allcott, H., Braghieri, L., Eichmeyer, S., Gentzkow, M. (2020) The Welfare Effects of Social Media. *American Economic Review, 110*(3), 629-676, https://doi.org/10.1257/aer.20190658

[370] ODI (2019) Open Banking, Preparing for lift off: Purpose, progress and potential. *Open Banking Implementation Entity*. https://www.openbanking.org.uk/wp-content/uploads/open-banking-report-150719.pdf

[371] Stats NZ (2018, July 2). Integrated Data Infrastructure. https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/

[372] UK Statistics Authority (2020). Better Use of Data: Statistics and Research. https://www.statisticsauthority.gov.uk/about-the-authority/better-useofdata-statistics-and-research/

[373] ESRC (2020). Administrative Data Research UK. *UKRI.* https://esrc.ukri.org/research/our-research/administrative-data-research-uk/

[374] Consumer Data Research Centre (2020). *University of Leeds.* https://www.cdrc.ac.uk/

[375] e-estonia (2020) we have built a digital society and we can show you how. *Enterprise Estonia.* https://e-estonia.com/

[376] Davies, G. (2019, June 21). *Challenges in using data across government.* National Audit Office. https://www.nao.org.uk/report/challenges-in-using-data-across-government/

[377] Centre for Data Ethics and Innovation (2020, July 20). *Addressing trust in public sector data use,* gov.uk. https://www.gov.uk/government/publications/cdei-publishes-its-first-report-on-public-sector-data-sharing/addressing-trust-in-public-sector-data-use#fn:6

[378] The Centre of Excellence for Information Sharing (2016, July 6). *Protecting vulnerable children and families: information sharing,* gov.uk. https://www.gov.uk/government/publications/protecting-vulnerable-children-and-families-information-sharing

[379] Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., ... & Ledsam, J. R. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *The Lancet Digital Health*, *1*(6), e271-e297. https://doi.org/10.1016/S2589-7500(19)30123-2

[380] Cystic Fibrosis Trust. *UK Cystic Fibrosis Registry.* https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry

[381] Graetz, N., Friedman, J., Osgood-Zimmerman, A., Burstein, R., Biehl, M. H., Shields, C., ... & Reiner, R. C. (2018). Mapping local variation in educational attainment across Africa. *Nature, 555*(7694), 48-53. https://doi.org/10.1038/nature25761

[382] United Nations Global Pulse (2013, October) *Mobile Phone Network Data for Development.* http://www.unglobalpulse.org/sites/default/files/Mobile%20Data%20for%20Development%20Primer_Oct2013.pdf

[383] MIT (2020) atlas of inequality. https://inequality.media.mit.edu/

[384] UN Global Pulse (2020) About UN Global Pulse. *United Nations.* https://www.unglobalpulse.org/about-new

[385] Identity Theft Resource Center. (2018). *End of year data breach report.* https://www.idtheftcenter.org/wp-content/uploads/2019/02/ITRC_2018-End-of-Year-Aftermath_FINAL_V2_combinedWEB.pdf

[386] Information is Beautiful (2020*).* World's biggest data breaches and hacks. https://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/

[387] Bradshaw, S. & Howard, P.N. (2019). The global disinformation order: 2019 global inventory of organised social media manipulation. *Oxford Internet Institute, Computational Propaganda Research Project.* https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf

[388] Konger, C. (2019, August 19). Facebook and Twitter Say China Is Spreading Disinformation in Hong Kong. *New York Times.* https://www.nytimes.com/2019/08/19/technology/hong-kong-protests-china-disinformation-facebook-twitter.html

[389] Chala, E. (2018). Leaked Documents Show That Ethiopia's Ruling Elites Are Hiring Social Media Trolls (And Watching Porn). *Global Voices.* https://globalvoices.org/2018/01/20/leaked-documents-show-that-ethiopias-ruling-elites-are-hiring-social-media-trolls-and-watching-porn/

[390] Kosinski, M., Stillwell, D. & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*(15), 5802-5805. https://www.pnas.org/content/pnas/110/15/5802.full.pdf

[391] Matz, S.C., Kosinski, M., Nave, G. & Stillwell, D.J. (2017) Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences, 114*(48), 12714-12719. https://www.pnas.org/content/pnas/114/48/12714.full.pdf

[392] Ribeiro, F. N., Saha, K., Babaei, M., Henrique, L., Messias, J., Benevenuto, F., ... & Redmiles, E. M. (2019, January). On microtargeting socially divisive ads: A case study of Russia-linked ad campaigns on facebook. In *Proceedings of the Conference on Fairness, Accountability, and Transparency,* 140-149, https://arxiv.org/pdf/1808.09218.pdf

[393] Gerber, A.S., Huber, G.A., Doherty, D., Dowling, C.M., Panagopoulos, C. (2013). Big Five Personality Traits and Responses to Persuasive Appeals: Results from Voter Turnout Experiments. *Politcal Behaviour 35***,** 687–728. https://doi.org/10.1007/s11109-012-9216-y

[394] Kalla, J. L., & Broockman, D. E. (2018). The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. American Political Science Review, 112(1), 148-166. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3042867

[395] Guess, A., Lyons, B., Montgomery, J. M., Nyhan, B., & Reifler, J. (2018). Fake news, Facebook ads, and misperceptions: Assessing information quality in the 2018 US midterm election campaign. Hanover: Dartmouth College. http://www.dartmouth.edu/~nyhan/fake-news-2018.pdf

[396] Allcott, H. & Gentzkow, M. (2017) Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives, 31*(2), 211-236. https://web.stanford.edu/~gentzkow/research/fakenews.pdf

[397] Pennycook, G., Cannon, T., Rand, D.G. (2018). Prior Exposure Increases Perceived Accuracy of Fake News. *Journal of Experimental Psychology: General, 147*(12), 1865-1880, https://doi.org/10.1037/xge0000465

[398] Kurtzleben, D. (2018, April 11) Did Fake News on Facebook Help Elect Trump? Here's What We Know. *npr.* https://www.npr.org/2018/04/11/601323233/6-facts-we-know-about-fake-news-in-the-2016-election?t=1570109118015

[399] Lazer, D.M.J., Baum, M.A.., Benkler, Y., Berinskey, A.J., Greenhill, K.M., Menczer, F., Metzger, M.J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S.A., Sunstein, C.R., Thorson, E.A., Watts, D.J. & Zittrain, J.L. (2018, March 9) The Science of fake news. *Science, 359*(6380), 1094-1096. https://scholar.harvard.edu/files/mbaum/files/science_of_fake_news.pdf

[400] Darktrace (2018). The Next Paradigm Shift: AI-Driven Cyber-Attacks. https://www.maxwara.ee/sites/default/files/the_next_paradigm_shift_-_ai_driven_cyber_attacks.pdf

[401] Centre for Data Ethics and Innovation (2020) *CDEI Review of online targeting,* gov.uk. https://www.gov.uk/government/publications/cdei-review-of-online-targeting

[402] Hitaj, B., Gasti, P., Ateniese, G., & Perez-Cruz, F. (2019, June). Passgan: A deep learning approach for password guessing. *International Conference on Applied Cryptography and Network Security,* 217-237 https://arxiv.org/pdf/1709.00440.pdf

[403] Centre for Data Ethics and Innovation (2019, September 12) *Snapshot Paper - Deepfakes and Audiovisual Disinformation.* https://www.gov.uk/government/publications/cdei-publishes-its-first-series-of-three-snapshot-papers-ethical-issues-in-ai/snapshot-paper-deepfakes-and-audiovisual-disinformation

[404] Mirsky, Y., Mahler, T., Shelef, I. & Elovici, Y. (2019). CT-GAN: Malicious tampering of 3D medical imagery using deep learning. *28th {USENIX} Security Symposium,* 461-478, https://arxiv.org/abs/1901.03597

[405] Fried, O., Tewari, A., Zollhöfer, M., Finkelstein, A., Shechtman, E., Goldman, D. B., ... & Agrawala, M. (2019). Text-based editing of talking-head video. *ACM Transactions on Graphics (TOG), 38*(4), 1-14. https://www.ohadf.com/projects/text-based-editing/

[406] Zakharov, E., Shysheya, A., Burkov, E., & Lempitsky, V. (2019). Few-shot adversarial learning of realistic neural talking head models. *Proceedings of the IEEE International Conference on Computer Vision,* 9459-9468, https://arxiv.org/pdf/1905.08233.pdf

[407] See e.g. Heaven, D. (2019, October 9) Why deep-learning AIs are so easy to fool. *Nature, 574(7777), 163-166.* https://www.nature.com/articles/d41586-019-03013-5

[408] Brundage, M. et al. (2018). The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. https://maliciousaireport.com/

[409] Gu, T., Liu, K., Dolan-Gavitt, B. & Garg, S. (2019). Badnets: Evaluating backdooring attacks on deep neural networks. *IEEE Access*, *7*, 47230-47244. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8685687

[410] Tramèr, F., Kurakin, A., Papernot, N., Goodfellow, I., Boneh, D., & McDaniel, P. (2017). Ensemble adversarial training: Attacks and defenses. *arXiv preprint.* https://arxiv.org/abs/1705.07204

[411] Panel for the Future of Science and Technology European Science-Media Hub (2019) Automated tackling of disinformation. *European Parliamentary Research Service.* https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624278/EPRS_STU(2019)624278_EN.pdf

[412] Charlton, E. (2019, May 21). How Finland is fighting fake news - in the classroom. *World Economic Forum.* https://www.weforum.org/agenda/2019/05/how-finland-is-fighting-fake-news-in-the-classroom/

[413] Data aggregated from Identity Theft Resource Center Annual Breach Reports, which can be found at: Identity Theft Resource Center (2020). *Data Breaches.* https://www.idtheftcenter.org/data-breaches/

[414] Amodei, D. & Hernandez, D. (2018, May 16) AI and Compute. *Open AI.* https://openai.com/blog/ai-and-compute/

[415] Sun, C., Shrivastava, A., Singh, S., & Gupta, A. (2017). Revisiting unreasonable effectiveness of data in deep learning era. *Proceedings of the IEEE international conference on computer vision,* 843-852. https://arxiv.org/abs/1707.02968

[416] Popejoy, A. B., & Fullerton, S. M. (2016). Genomics is failing on diversity. *Nature News, 538*(7624), 161. https://www.nature.com/news/genomics-is-failing-on-diversity-1.20759

[417] Sirugo, G., Williams, S. M., & Tishkoff, S. A. (2019). The missing diversity in human genetic studies. *Cell, 177*(1), 26-31. https://doi.org/10.1016/j.cell.2019.02.048

[418] Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Conference on fairness, accountability and transparency,* 77-91. http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

[419] Grother, P., Ngan, M., & Hanaoka, K. (2019). *Face Recognition Vendor Test (FVRT): Part 3, Demographic Effects.* National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf

[420] Doshi, T (2018, September 6) Introducing the Inclusive Images Competition. *Google AI Blog*. https://ai.googleblog.com/2018/09/introducing-inclusive-images-competition.html

[421] Smith, J.R. (2019, January 29) IBM Research Releases 'Diversity in Faces' Dataset to Advance Study of Fairness in Facial Recognition Systems. *IBM Research Blog.* https://www.ibm.com/blogs/research/2019/01/diversity-in-faces/

[422] Chin, C. & Ashok, B. (2019, November 18) Highlights: Addressing fairness in the context of artificial intelligence. *Brookings.* https://www.brookings.edu/blog/techtank/2019/11/18/highlights-addressing-fairness-in-the-context-of-artificial-intelligence/

[423] Lee, N.T., Resnick, P. & Barton, G. (2019, May 22) Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. *Brookings.* https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/

[424] Conger, K., Fausset, R., & Kovaleski, S. F. (2019, May 14). San Francisco Bans Facial Recognition Technology. *The New York Times.* https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html

[425] Rawat, W. & Wang, Z. (2017) Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Computation, 29*(9), 2352-2449, https://doi.org/10.1162/neco_a_00990

[426] Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and brain sciences, 40.* https://arxiv.org/abs/1604.00289

[427] Lake, B. M., R. Salakhutdinov, and J. B. Tenenbaum. (2019). The Omniglot challenge: a 3-year progress report. *Current Opinion in Behavioral Sciences 29: 97-104.* https://arxiv.org/pdf/1902.03477.pdf

[428] Royal Society (2017, April). *Machine learning: the power and promise of computers that learn by example.* royalsociety.org/machine-learning

[429] Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science advances*, *4*(1), eaao5580. https://doi.org/10.1126/sciadv.aao5580

[430] Jung, J., Concannon, C., Shroff, R., Goel, S., & Goldstein, D. G. (2017). Simple rules for complex decisions. http://dx.doi.org/10.2139/ssrn.2919024

[431] Mittal, S. (2019). A Survey on optimized implementation of deep learning models on the NVIDIA Jetson platform. *Journal of Systems Architecture (97) 428-442,* https://www.researchgate.net/publication/329802520_A_Survey_on_Optimized_Implementation_of_Deep_Learning_Models_on_the_NVIDIA_Jetson_Platform

[432] Top500 (2019) China Extends Lead in Number of TOP500 Supercomputers, US Holds on to Performance Advantage. https://www.top500.org/news/china-extends-lead-in-number-of-top500-supercomputers-us-holds-on-to-performance-advantage/

[433] Hines, J. (2018, June 8). Genomics code exceeds exaops on summit supercomputer. *Oakridge National Laboratory Leadership Computing Facility.* https://www.olcf.ornl.gov/2018/06/08/genomics-code-exceeds-exaops-on-summit-supercomputer/

[434] Oak Ridge National Laboratory (2020). ORNL is in the fight against COVID-19. https://www.ornl.gov/coronavirus

[435] Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., ... & Burkett, B. (2019). Quantum supremacy using a programmable superconducting processor. *Nature, 574*(7779), 505-510. https://www.nature.com/articles/s41586-019-1666-5

[436] Innovate UK (2019), *Quantum Technologies in the USA 2019.* Innovate UK Global Expert Mission. https://admin.ktn-uk.co.uk/app/uploads/2020/03/0183_KTN_USA-QuantumTechnologiesReport_v4.pdf

[437] Hurd, W. (2017). Quantum Computing Is the Next Big Security Risk. *Wired.* https://www.wired.com/story/quantum-computing-is-the-next-big-security-risk/

[438] Davies, M. (2019) Exploring Neuromorphic Computing for AI: Why Spikes? (Part One). *Intel.* https://www.intel.ai/exploring-neuromorphic-computing-for-ai-why-spikes-part-one/#gs.vfxviw

[439] Ray, T. (2019, July 1) Neuromorphic computing finds new life in machine learning. *ZDNet.* https://www.zdnet.com/article/neuromorphic-computing-finds-new-life-in-machine-learning/

[440] Yousefpour, A., Fung, C., Nguyen, T., Kadiyala, K., Jalali, F., Niakanlahiji, A., ... & Jue, J. P. (2019). All one needs to know about fog computing and related edge computing paradigms: A complete survey. *Journal of Systems Architecture, 98,* 289-330. https://arxiv.org/abs/1808.05283

[441] Teerapittayanon, S., McDanel, B., & Kung, H. T. (2017, June). Distributed deep neural networks over the cloud, the edge and end devices. *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS,* 328-339. IEEE. http://www.eecs.harvard.edu/~htk/publication/2017-icdcs-teerapittayanon-mcdanel-kung.pdf

[442] Gaia-X. Gaia-X: A federated data infrastructure for Europe. https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html [Accessed August 2020]

[443] Royal Society (2019) *Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies in data analysis.* https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/privacy-enhancing-technologies-report.pdf

[444] Wood, A., Altman, M., Bembenek, A., Bun, M., Gaboardi, M., Honaker, J., ... & Vadhan, S. (2018). Differential privacy: A primer for a non-technical audience. *Vand. J. Ent. & Tech. L., 21,* 209. https://salil.seas.harvard.edu/files/salil/files/differential_privacy_primer_nontechnical_audience.pdf

[445] National Conference of State Legislatures, (2020, June 4). *Differential Privacy for Census Data Explained.* https://www.ncsl.org/research/redistricting/differential-privacy-for-census-data-explained.aspx

[446] Miller, S. (2020, June 5). *Researchers raise concerns with differential privacy use on census data.* GCN. https://gcn.com/articles/2020/06/05/differential-privacy.aspx
[447] Shaw, D. (2020, February 12). UK's 2021 census could be the last, statistics chief reveals. BBC News. https://www.bbc.co.uk/news/uk-51468919
[448] NHS Digital (2018, June 29) NHS Digital leading the protection of patient data with new patient de-identification solution. https://digital.nhs.uk/news-and-events/latest-news/nhs-digital-leading-the-protection-of-patient-data-with-new-patient-de-identification-solution
[449] Kamm, L., Bogdanov, D., Laur, S., & Vilo, J. (2013). A new way to protect privacy in large-scale genome-wide association studies. *Bioinformatics, 29*(7), 886-893. https://academic.oup.com/bioinformatics/article/29/7/886/253610
[450] McMahan, B. & Ramage, D. (2017, April 6) Federated Learning: Collaborative Machine Learning without Centralized Training Data. *Google AI Blog.* https://ai.googleblog.com/2017/04/federated-learning-collaborative.html
[451] Murphy, H. (2019). How Facebook could target ads in age of encryption. *Financial Times.* https://www.ft.com/content/0181666a-4ad6-11e9-bbc9-6917dce3dc62
[452] Toubiana, V., Narayanan, A., Boneh, D., Nissenbaum, H., & Barocas, S. (2010, March). Adnostic: Privacy preserving targeted advertising. *Proceedings Network and Distributed System Symposium.* https://crypto.stanford.edu/adnostic/adnostic.pdf
[453] Gebhart, G. (2019). A privacy-focused Facebook? We'll believe it when we see it. *Electronic Frontier Foundation*. https://www.eff.org/deeplinks/2019/03/privacy-focused-facebook-well-believe-it-when-we-see-it
[454] McCabe, D., Isaac, M. & Benner, K. (2019). Facebook and Barr escalate standoff over encrypted messages. *The New York Times*. https://www.nytimes.com/2019/12/10/technology/whatsapp-barr-encryption.html
[455] Hub of All Things (2020) Hub of all things: own your own personal data server and private AI. https://www.hubofallthings.com/
[456] digi.me (2020) How is digi.me free for individuals? https://digi.me/business-model/
[457] Brace (2020) You deserve a better internet. https://brave.com/
[458] Basu, T. (2020, January 27). Why private micro-networks could be the future of how we connect. *MIT Technology Review.* https://www.technologyreview.com/s/615094/why-private-micro-networks-could-be-the-future-of-how-we-connect/?utm_source=newsletters&utm_medium=email&utm_campaign=the_download.unpaid.engagement
[459] Valade, P. (2019, August 5) Vision & Trust. *Jumbo Blog.* https://blog.jumboprivacy.com/vision-and-trust.html
[460] Marotta, V., V. Abhishek, and A. Acquisti. (2019). *Online tracking and publishers' revenues: An empirical analysis. Working paper.* https://weis2019.econinfosec.org/wp-content/uploads/sites/6/2019/05/WEIS_2019_paper_38.pdf
[461] BBC (2020, January 15). Cookies Crumbling as Google Phases Them Out. *BBC News*. www.bbc.com/news/technology-51106526
[462] Analysis of data from Ritter, J.R. (2020) Initial public offerings: Technology stock IPOs. *Warrington College of Business, University of Florida.* https://site.warrington.ufl.edu/ritter/files/IPOs2019Tech-Stock.pdf
[463] Fiegerman, S. (2019, May 10) Uber falls more than 7% in disappointing Wall Street debut. *CNN Business.* https://edition.cnn.com/2019/05/10/tech/uber-wall-street-debut/index.html
[464] Ungarino, R. (2019, April 7) Lyft went public at a $24 billion valuation. Here's how that compares to other high-profile tech companies dating back to the dotcom bubble. *Markets Insider.* https://markets.businessinsider.com/news/stocks/lyft-stock-how-valuation-compres-to-other-tech-names-at-ipo-2019-4-1028090958#lyft1
[465] UK. Board of Governors of the Federal Reserve System. (2013) Policy Tools: Open Market Operations Archive. https://www.federalreserve.gov/monetarypolicy/openmarket_archive.htm
[466] Ritter, J.R. (2020, June 23) Initial Public Offerings: Updated Statistics. *Warrington College of Business, University of Florida* https://fas.org/irp/crs/RL31617.pdf
[467] Gurman, M. (2019, March 23) Apple Reinvention as Services Company Starts for Real Monday. *Bloomberg.* https://www.bloomberg.com/news/articles/2019-03-23/apple-s-reinvention-as-a-services-company-starts-for-real-monday
[468] Brochot, G., Brunini, J., Eisma, F., Larsen, R., Lewis, D.J., Zhang, J. (2015). Personal Data Stores. *Study conducted at the Cambridge University Judge Business School.* https://ec.europa.eu/digital-single-market/en/news/study-personal-data-stores-conducted-cambridge-university-judge-business-school
[469] McCullough, B. (2018, December 4) An eye-opening look at the dot-com bubble of 2000 — and how it shapes our lives today. *Ideas.ted.com.* https://ideas.ted.com/an-eye-opening-look-at-the-dot-com-bubble-of-2000-and-how-it-shapes-our-lives-today/
[470] Wigglesworth, R. (2020, May 1). How Big Tech got even bigger in the Covid-19 era. *Financial Times.* https://www.ft.com/content/d2e09235-b28e-438d-9b55-0e6bab7ac8ec
[471] Associated Press (2020, June 3). Zoom booms as teleconferencing company profits from coronavirus crisis. *The Guardian.* https://www.theguardian.com/technology/2020/jun/03/zoom-booms-as-teleconferencing-company-profits-from-coronavirus-crisis

[472] Zyskind, G., Nathan, O., Pentland, A. (2015, May). Decentralizing privacy: Using blockchain to protect personal data. *2015 IEEE Security and Privacy Workshops,* 180-184. https://enigma.co/ZNP15.pdf
[473] Vidan, G., & Lehdonvirta, V. (2019). Mine the gap: Bitcoin and the maintenance of trustlessness. *New Media & Society, 21*(1), 42-59. https://journals.sagepub.com/doi/10.1177/1461444818786220
[474] Royal Society, British Academy, & techUK. (2018) Data ownership, rights and controls: Reaching a common Understanding. https://royalsociety.org/~/media/policy/projects/data-governance/data-ownership-rights-and-controls-October-2018.pdf
[475] Erlich, Y., Shor, T., Pe'er, I., & Carmi, S. (2018). Identity inference of genomic data using long-range familial searches. *Science,* 362(6415), 690-694. https://doi.org/10.1126/science.aau4832
[476] Taylor, L., Floridi, L., & Van der Sloot, B. (Eds.). (2016). *Group privacy: New challenges of data technologies* (Vol. 126). Springer.
[477] Mulgan, G., Straub, V. (2019, February 21) The new ecosystem of trust. *Nesta.* https://www.nesta.org.uk/blog/new-ecosystem-trust/
[478] Tennison, J. (2020, February 10). What do we mean by data institutions? *Open Data Institute Blog.* https://theodi.org/article/what-do-we-mean-by-data-institutions/
[479] UK Biobank. https://www.ukbiobank.ac.uk/about-biobank-uk/
[480] MIDATA (2019) Articles of Association of MIDATA Cooperative. *MIDATA.* https://www.midata.coop/wp-content/uploads/2019/08/MIDATA_Statuten_20190626_EN.pdf
[481] Tech Nation (2019, June 4). Introducing the innovative Data Commons for UK Tech. https://technation.io/news/introducing-the-innovative-data-commons-for-uk-tech/
[482] O'Hara, K. (2019) *Ethics, Architecture and Governance for Trustworthy Data Stewardship*. Web Science Institute, University of Southampton. https://eprints.soton.ac.uk/428276/1/WSI_White_Paper_1.pdf
[483] Delacroix, S., & Lawrence, N. D. (2019). Bottom-up data Trusts: disturbing the 'one size fits all' approach to data governance. I*nternational Data Privacy Law, 9*(4), 236-252. https://academic.oup.com/idpl/advance-article/doi/10.1093/idpl/ipz014/5579842#163974452
[484] Lomas, N. (2020, July 20). UK Uber drivers are taking the algorithm to court. *Techcrunch.* https://techcrunch.com/2020/07/20/uk-uber-drivers-are-taking-its-algorithm-to-court/
[485] GovLab (2020) Data Collaboratives. https://datacollaboratives.org/;
[486] ODI (2019, April 15) Data trusts: lessons from three pilots (report). *Open Data Institute.* https://theodi.org/article/odi-data-trusts-report/
[487] British Academy and Royal Society (2017). *Data management and use: Governance in the 21st Century.* British Academy and Royal Society. https://blogs.royalsociety.org/in-verba/2017/06/29/data-management-and-use-governance-in-the-21st-century/
[488] Germany. Data Ethics Commission (2019) *Opinion of the Data Ethics Commission.* Data Ethics Commission of the Federal Government, Federal Ministry of the Interior, Building and Community.https://datenethikkommission.de/wp-content/uploads/191023_DEK_Kurzfassung_en_bf.pdf
[489] Tisne, M. (2018, December 14) It's time for a Bill of Data Rights. *MIT Technology Review.* https://www.technologyreview.com/s/612588/its-time-for-a-bill-of-data-rights/
[490] Tisne, M. (2020). *The Data Delusion: Protecting individual data ins't enough when the harm is collective.* Stanford Cyber Policy Center. https://cyber.fsi.stanford.edu/publication/data-delusion
[491] Mazzucato, M. (2018, June 27) Let's make private data into a public good. *MIT Technology Review.* https://www.technologyreview.com/s/611489/lets-make-private-data-into-a-public-good/
[492] Mejias, U.A. (2019, December 14) Why the Global South should nationalise its data. *Al Jazeera.* https://www.aljazeera.com/indepth/opinion/global-south-nationalise-data-191211082728186.html;
[493] Srivastava, A. (2019, June 14) India must reclaim its lost digital space. *The Hindu Businessline.* https://www.thehindubusinessline.com/opinion/india-must-reclaim-its-lost-digital-space/article27941890.ece
[494] Jones, C. I. & Tonetti, C. (2019). Nonrivalry and the Economics of Data*. National Bureau of Economic Research No. w26260*. https://www.gsb.stanford.edu/faculty-research/working-papers/nonrivalry-economics-data
[495] Ichihashi, S. (2019, June 29) Non-competing Data Intermediaries. *Bank of Canada, Digital Economy and Advanced Analytics Division.* https://shota2.github.io/research/data.pdf
[496] Siciliani, P. & Giovannetti, E. (2019, December 20) Platform competition and incumbency advantage under heterogeneous switching cost — exploring the impact of data portability. Working Paper. *Bank of England.* https://www.bankofengland.co.uk/working-paper/2019/platform-competition-and-incumbency-advantage-under-heterogeneous-switching-cost
[497] Hao, K. (2020, June 12). The two-year fight to stop Amazon from selling face recognition to the police. *MIT Technology Review.* https://www.technologyreview.com/2020/06/12/1003482/amazon-stopped-selling-police-face-recognition-fight/
[498] Budd, J., Miller, B.S., Manning, E.M. *et al.* (2020, August 7). Digital technologies in the public-health response to COVID-19. *Nature Medicine* 26**,** 1183–1192. https://doi.org/10.1038/s41591-020-1011-4
[499] Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dorner, L., Parker, M., Bonsall, D., Fraser, C. (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science, 368*(6491), https://doi.org/10.1126/science.abb6936

[500] Mozur, P., Zhong, R., Krolik, A. (2020, March 1). In Coronavirus Fight, China Gives Citizens a Color Code, With Red Flags. *The New York Times.* https://www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html

[501] Goh, B. (2020, February 26). China rolls out fresh data collection campaign to combat coronavirus. *Reuters.* https://www.reuters.com/article/us-china-health-data-collection/china-rolls-out-fresh-data-collection-campaign-to-combat-coronavirus-idUSKCN20K0LW

[502] The Japan Times (2020, May 13). Green or red light: China coronavirus app is ticket to everywhere. *The Japan Times.* https://www.japantimes.co.jp/news/2020/05/13/asia-pacific/china-coronavirus-app/#.XunZh0XduUk

[503] Gan, N., Culver, D. (2020, April 16). China is fighting the coronavirus with a digital QR code. Here's how it works. *CNN Business.* https://edition.cnn.com/2020/04/15/asia/china-coronavirus-qr-code-intl-hnk/index.html

[504] Davidson, H. (2020, May 26). Chinese city plans to turn coronavirus app into permanent health tracker. *The Guardian.* https://www.theguardian.com/world/2020/may/26/chinese-city-plans-to-turn-coronavirus-app-into-permanent-health-tracker

[505] Prado, S. (2020). Coronavirus: surveillance helps but the programs are hard to stop. *Bloomberg.* https://www.bloomberg.com/news/articles/2020-04-06/coronavirus-surveillance-helps-but-the-programs-are-hard-to-stop

[506] Hern, A. (2020, April 2). Experts warn of privacy risk as US uses GPS to fight coronavirus spread. *The Guardian.* https://www.theguardian.com/technology/2020/apr/02/experts-warn-of-privacy-risk-as-us-uses-gps-to-fight-coronavirus-spread

[507] Fazlioglu, M. (2020, May 1*). Republican senators to introduce the COVID-19 Consumer Data Protection Act.* IAPP*.* https://iapp.org/news/a/republican-senators-to-introduce-the-covid-19-consumer-data-protection-act/

[508] IAPP (2020, May) COVID-19 Consumer Data Protection Act. *IAPP.* https://iapp.org/resources/article/covid-19-consumer-data-protection-act/#:~:text=The%20%E2%80%9CCOVID%2D19%20Consumer%20Data,The%20bill's%20cosponsors%20include%20Sens

[509] Lazzarotti, J.J., Atrakchi, M. (2020, May 13). Federal COVID-19 Consumer Data Protection Bill Introduced. *Jackson Lewis.* https://www.workplaceprivacyreport.com/2020/05/articles/covid-19/federal-covid-19-consumer-data-protection-bill-introduced/?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+WorkplacePrivacyDataManagementSecurityReport+%28Workplace+Privacy%2C+Data+Management+%26+Security+Report%29#page=1

[510] US. Congress (2020, July 5) S.3663 - COVID-19 Consumer Data Protection Act of 2020. https://www.congress.gov/bill/116th-congress/senate-bill/3663

[511] European Data Protection Board (2020, March 19). Statement by the EDPB Chair on the processing of personal data in the context of the COVID-19 outbreak. *European Union.* https://edpb.europa.eu/news/news/2020/statement-edpb-chair-processing-personal-data-context-covid-19-outbreak_en

[512] UK Department for Health and Social Care and The Rt Hon Matt Hancock MP (2020, August 6). Coronavirus (COVID-19): notification to organisations to share information, *gov.uk*. https://www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information#history

[513] Baker, P.C. (2020, March 31). ‘We can’t go back to normal’: how will coronavirus change the world? *The Guardian.* https://www.theguardian.com/world/2020/mar/31/how-will-the-world-emerge-from-the-coronavirus-crisis

[514] Busvine, D. (2020, August 5). Explainer: Europe's coronavirus smartphone contact tracing apps. *Reuters.* https://uk.reuters.com/article/uk-health-coronavirus-europe-tech-explai/explainer-europes-coronavirus-smartphone-contact-tracing-apps-idUKKCN2510N3

[515] Cellan-Jones, R. (2020, April 2). Coronavirus: Privacy in a pandemic. *BBC News.* https://www.bbc.co.uk/news/technology-52135916

[516] Antipolis, S. (2020, June 11). ETSI’s new group on covid-19 tracing apps interoperability moving fast: officials elected and work programme set up. *European Telecommunications Standards Institute.* https://www.etsi.org/committee?id=1780

[517] Greenberg, A. (2020, April 5). Google and Apple Reveal How Covid-19 Alert Apps Might Look. *Wired.* https://www.wired.com/story/apple-google-covid-19-contact-tracing-interface/

[518] Apple (2020, April 10) Apple and Google partner on COVID-19 contact tracing technology. https://www.apple.com/uk/newsroom/2020/04/apple-and-google-partner-on-covid-19-contact-tracing-technology/

[519] Kelion, L. (2020, May 26) Coronavirus: First Google/Apple-based contact-tracing app launched. *BBC News.* https://www.bbc.co.uk/news/technology-52807635

[520] Burgess, M. (2020, June 19). Why the NHS Covid-19 contact tracing app failed. *Wired.* https://www.wired.co.uk/article/nhs-tracing-app-scrapped-apple-google-uk

[521] Sabbagh, D. & Hern, A. (2020, June 18) UK abandons contact-tracing app for Apple and Google model. *The Guardian.* https://www.theguardian.com/world/2020/jun/18/uk-poised-to-abandon-coronavirus-app-in-favour-of-apple-and-google-models

[522] Taylor, J. (2020, May 6) Covidsafe app is not working properly on iPhones, authorities admit. *The Guardian.* https://www.theguardian.com/world/2020/may/06/covidsafe-app-is-not-working-properly-on-iphones-authorities-admit

[523] Kelion, L. (2020, June 18). UK virus-tracing app switches to Apple-Google model. *BBC Technology.* https://www.bbc.co.uk/news/technology-53095336

[524] Kucharski, A.J., Klepac, P., Conlan, A.J.K., Kissler, S.M., Tang, M.L., Fry, H., et al (2020, June 16) Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *The Lancet,* https://doi.org/10.1016/S1473-3099(20)30457-6

[525] Fussell, S., Knight, W. (2020, April 14). The Apple-Google Contact Tracing Plan Won't Stop Covid Alone. *Wired.* https://www.wired.com/story/apple-google-contact-tracing-wont-stop-covid-alone/

[526] Koslowski, M. (2020, May 3). Half of us say we support the COVIDSafe app, but only 16 per cent have downloaded it. *The Sydney Morning Herald.* https://amp.smh.com.au/politics/federal/half-of-us-say-we-support-the-covidsafe-app-but-only-16-per-cent-have-downloaded-it-20200501-p54p53.html

[527] O'Halloran, J. (2020, July 31). UK contact-tracing app to be launched 'shortly' as Northern Ireland deploys service. *ComputerWeekly.com,* https://www.computerweekly.com/news/252486951/UK-contact-tracing-app-to-be-launched-shortly-as-Northern-Ireland-deploys-service

[528] Engineering & Technology (2020, July 21). Ireland donates Covid-19 app code to Linux. *IET.* https://eandt.theiet.org/content/articles/2020/07/ireland-donates-covid-19-app-code-to-linux/

[529] SPI-M-O (2020, April 27) SPI-M-O: Consensus Statement on COVID-19. *UK Government Office for Science.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/888757/S0242_SAGE29_200427_SPI-M-O_Consensus_Statement_Redacted.pdf

[530] NHS (2020) NHS Covid-19 App Ethics Advisory Board Terms of reference. *NHS.* https://www.nhsx.nhs.uk/media/documents/NHS_COVID-19_App_Ethics_Advisory_Board_Terms_of_Reference.pdf

[531] Fahim, K., Kim, M.J., Hendrix, S. (2020, May 2). Cellphone monitoring is spreading with the coronavirus. So is an uneasy tolerance of surveillance. *The Washington Post.* https://www.washingtonpost.com/world/cellphone-monitoring-is-spreading-with-the-coronavirus-so-is-an-uneasy-tolerance-of-surveillance/2020/05/02/56f14466-7b55-11ea-a311-adb1344719a9_story.html

[532] University of Maryland (2020). Disinclined to Download: Americans Not Ready for Virus-Tracking Apps, But Support Business, Immigration Clampdowns, UMD-Post Poll Finds. *University of Maryland, College of Behavioural & Social Sciences.* https://bsos.umd.edu/featured-content/disinclined-download-americans

[533] Asher, S. (2020, July 5) Coronavirus: Why Singapore turned to wearable contact-tracing tech. *BBC News.* https://www.bbc.co.uk/news/technology-53146360

[534] Kim, N. (2020, March 6). 'More scary than coronavirus': South Korea's health alerts expose private lives. *The Guardian.* https://www.theguardian.com/world/2020/mar/06/more-scary-than-coronavirus-south-koreas-health-alerts-expose-private-lives

[535] Tennison, J. (2020, April 2). Why isn't the government publishing more data about coronavirus deaths? *The Guardian.* https://www.theguardian.com/commentisfree/2020/apr/02/government-publish-data-coronavirus-deaths?mc_cid=8c703f5e44&mc_eid=9885091c73

[536] Halbfinger, D.M., Kershner, I., Bergman, R. (2020, March 16). To Track Coronavirus, Israel Moves to Tap Secret Trove of Cellphone Data. *The New York Times.* https://www.nytimes.com/2020/03/16/world/middleeast/israel-coronavirus-cellphone-tracking.html

[537] Privacy International (2020). Apps and Covid-19. *Privacy International.* https://privacyinternational.org/examples/apps-and-covid-19 [Accessed May 2020]

[538] Greenberg, A. (2020, June 5). India's Covid-19 Contact Tracing App Could Leak Patient Locations. *Wired.* https://www.wired.com/story/india-covid-19-contract-tracing-app-patient-location-privacy/

[539] Ranjan, R. (2020, May 11). Madhya Pradesh app to track patients leaks personal data, taken offline. *Hindustan Times.* https://www.hindustantimes.com/india-news/mp-app-to-track-patients-leaks-personal-data-taken-offline/story-WO7ATpaxOMDTsmUxSKduUO.html

[540] Whittaker, Z. (2020, May 2). Security lapse at India's Jio exposed coronavirus symptom checker results. *TechCrunch.* https://techcrunch.com/2020/05/02/jio-coronavirus-security-lapse/amp/?guccounter=1&guce_referrer=aHR0cHM6Ly9wcml2YWN5aW50ZXJuYXRpb25hbC5vcmcvZXhhbXBsZXMvMzgzMS9pbmRpYS1qaW8tc2VjdXJpdHktYnJlYWNoLWV4cG9zZXMtbWlsbGlvbnMtc3ltcHRvbS1jaGVja2VyLWFwcC1sb2dz&guce_referrer_sig=AQAAAAz4zB36lOIilllgJVLMdQKLI07N-OmoOWyEXCEzs2RipTejpneVqA2_1UJhBf1jdV135xAzWb1q0O2WOf4DpOQnzFM1M063sH5w329xAid8dXlab_wuKAvYmIMkdqc9HqOuQfC5Mu4lJjvnTtFOmq8Fl09-2WoFy-CWjaR_ul1W

[541] Jahangir, R. (2020, May 1). "I became a pariah." Coronavirus victims' data is leaked on social media in Pakistan. *codastory.* https://www.codastory.com/authoritarian-tech/pakistan-tech-coronavirus/

[542] Hern, A. (2020, May 27). Qatari contact-tracing app 'put 1m people's sensitive data at risk'. *The Guardian.* https://www.theguardian.com/world/2020/may/27/qatar-contact-tracing-app-1m-people-sensitive-data-at-risk-coronavirus-covid-19

[543] Levy, I. (2020, May 4) The security behind the NHS contact tracing app. *National Cyber Security Centre.* https://www.ncsc.gov.uk/blog-post/security-behind-nhs-contact-tracing-app

[544] Public Health England (2020, March 9). Notifiable diseases and causative organisms: how to report. https://www.gov.uk/guidance/notifiable-diseases-and-causative-organisms-how-to-report#list-of-notifiable-diseases

[545] Public Health England (2020, January 6). Notifications of infectious diseases (NOIDs). https://www.gov.uk/government/collections/notifications-of-infectious-diseases-noids

[546] Amnesty International (2020, April 2). Digital surveillance to fight COVID-19 can only be justified if it respects human rights. *Amnesty International.* https://www.amnesty.org/en/latest/news/2020/04/covid19-digital-surveillance-ngo/

[547] Jahangir, R. (2020, May 1). Pakistan's "patient zero" stigmatized after data leak. *Privacy International.* https://privacyinternational.org/examples/3839/pakistans-patient-zero-stigmatized-after-data-leak