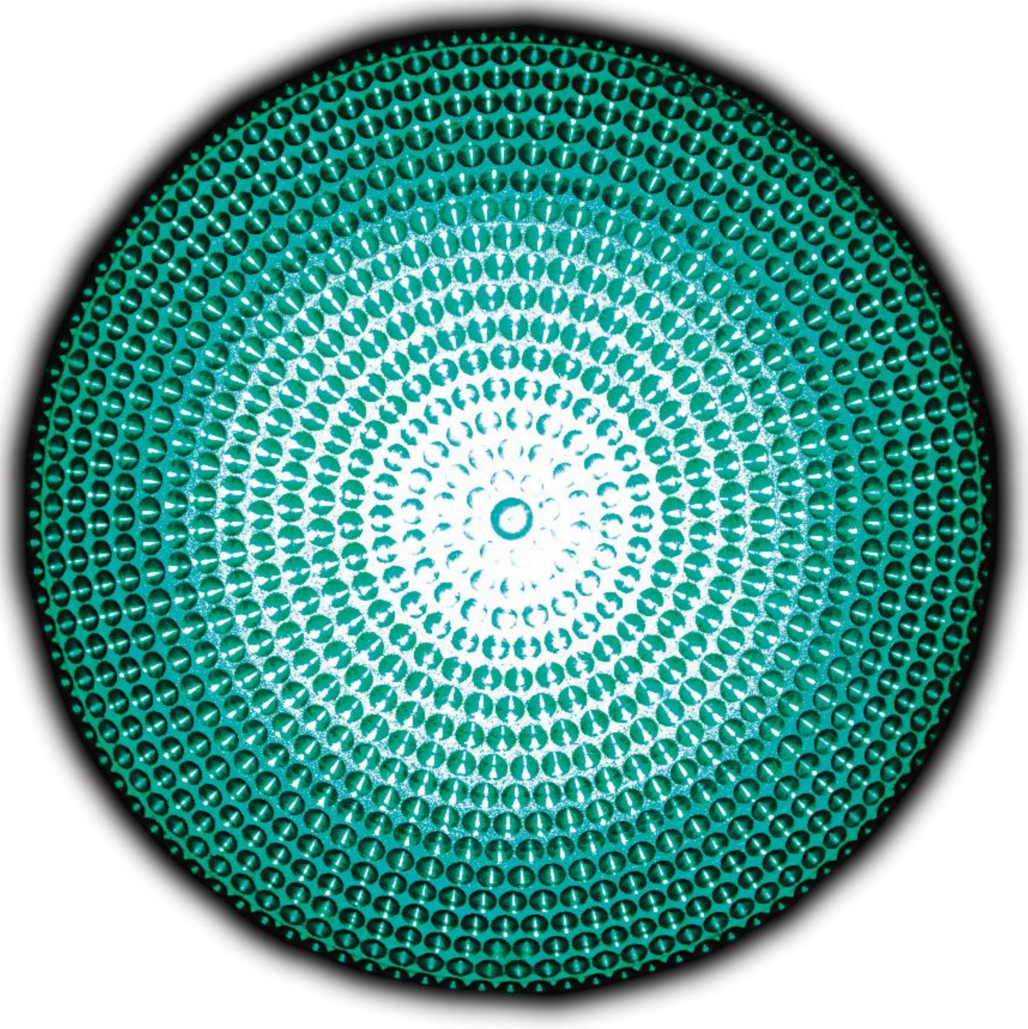
Deloitte.



New Technologies Case Study: Data Sharing in Infrastructure
A final report for the National Infrastructure Commission

Deloitte Confidential: Public Sector – For Approved External Use
November 2017

Important notice from Deloitte

This final report (the "Final Report") has been prepared by Deloitte LLP ("Deloitte") for the National Infrastructure Commission (NIC) in accordance with the contract with them dated 19th September 2017 ("the Contract") and on the basis of the scope and limitations set out below.

The Final Report has been prepared solely for the purposes of analysing barriers to data sharing in the UK's infrastructure sectors, as set out in the Contract. It should not be used for any other purpose or in any other context, and Deloitte accepts no responsibility for its use in either regard, including its use by NIC for decision making or reporting to third parties.

The Final Report is provided exclusively for NIC's use under the terms of the Contract. No party other than NIC is entitled to rely on the Final Report for any purpose whatsoever and Deloitte accepts no responsibility or liability or duty of care to any party other than NIC in respect of the Final Report or any of its contents.

The information contained in the Final Report has been obtained from NIC and third party sources that are clearly referenced in the appropriate sections of the Final Report. Deloitte has neither sought to corroborate this information nor to review its overall reasonableness. Further, any results from the analysis contained in the Final Report are reliant on the information available at the time of writing the Final Report and should not be relied upon in subsequent periods.

All copyright and other proprietary rights in the Final Report remain the property of Deloitte LLP and any rights not expressly granted in these terms or in the Contract are reserved.

Any decision to invest, conduct business, enter or exit the markets considered in the Final Report should be made solely on independent advice and no information in the Final Report should be relied upon in any way by any third party. This Final Report and its contents do not constitute financial or other professional advice, and specific advice should be sought about your specific circumstances. In particular, the Final Report does not constitute a recommendation or endorsement by Deloitte to invest or participate in, exit, or otherwise use any of the markets or companies referred to in it. To the fullest extent possible, both Deloitte and NIC disclaim any liability arising out of the use (or non-use) of the Final Report and its contents, including any action or decision taken as a result of such use (or non-use).

Contents

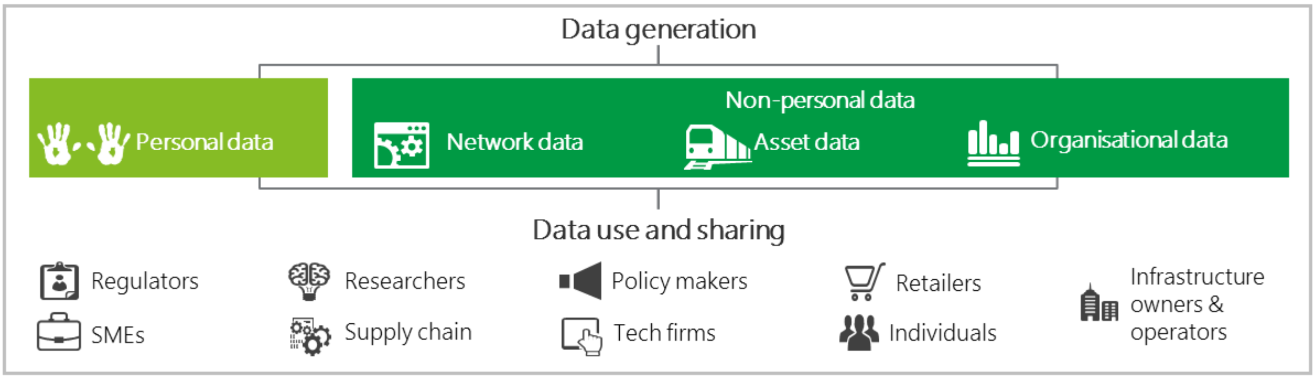
Glossary

Term	Definition
Application Programming Interface (API)	A set of functions and procedures that allow for the creation of applications that access the features or data of an operating system, application, or other service.
Artificial Intelligence (AI)	A broad field of science which aims to employ computers to do tasks that would normally require human intelligence. In this context, human intelligence includes the ability to plan, reason, learn, communicate and build some perception of knowledge.
Augmented Reality (AR)	A technology that superimposes information, such as a digital image, onto a user's field of view as they perceive the real world. The information typically informs the user about an object or place at which they are looking.
Blockchain	A type of distributed ledger technology, where the growing list of records ('blocks') are linked by cryptography. Allows the sharing of data over the internet, without the need for an intermediary.
Building Information Modelling (BIM)	In construction, BIM gives a digital representation of both the physical and functional characteristics of a given asset that allows for a shared knowledge resource of all data on a construction project. BIM allows for a real-time view of behaviour and performance, supporting the exchange of information from the actual asset to its digital twin, and facilitating performance-optimising adjustments from the digital twin back to the asset.
Data sharing	The making available of data by an organisation that originally created or collected the data with other organisations, individuals and public bodies that seek to use or re-use it for a variety of purposes. Data sharing can be unilateral or multilateral. It may take the form of an exchange of data or the creation of a centralised repository or 'pool' of data.
Demand-side response (DSR)	In energy, demand-side response services allow energy customers (potentially consumers and businesses) to intelligently adjust their energy use in real-time. For instance, this could be in combination with time-varying tariffs.
Distributed denial-of-service (DDoS) attacks	A type of cyber-attack that can shut down targeted servers and infiltrate corporate networks by overwhelming them with traffic from multiple compromised computer systems.
Distributed ledger	An asset database that can be shared ('distributed') across a network of multiple users, all of which have their own identical copy of the ledger. The assets can be financial, physical, legal or electronic, and any updates are transparent and are distributed within minutes to each copy of the ledger.
General Data Protection Regulation (GDPR)	Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. It regulates the

Term	Definition
	handling of EU citizens' personal data by public and private sector organisations. It will come into force in the UK on 25 May 2018.
Information Commissioner's Office (ICO)	The UK's independent authority set up to uphold information rights in the public interest, promoting openness by public bodies and data privacy for individuals.
Infrastructure	Generally, infrastructure refers to the physical and organisational networks, structures and facilities that are fundamental to the operation of an economy. For the purpose of this report, infrastructure is confined to the coverage of the National Infrastructure Assessment sectors: transport, energy, water and wastewater, digital communication, solid waste and flood risk management.
Internet of Things (IoT)	Describes the interconnection of a network of physical objects, which are embedded with unique identifiers and software that allows the collection and exchange of data.
Machine learning	An artificial intelligence capability where computer systems improve their performance by exposure to data without the need to follow programmed instructions.
National Infrastructure Commission (NIC)	Works with HM Treasury to provide the Government with impartial, expert advice on major long-term infrastructure challenges.
Open data	Open data is data that anyone can access, use and share. For data to be considered 'open', it must be; accessible, which usually means published on the web, available in a machine-readable format and have a licence that permits anyone to access, use and share it - commercially and non-commercially.
Personal data	Data from which a person can be identified, including data that can be combined with other information to identify a person.
Smart city	A city that optimises use of its infrastructure and resources using a network of connected physical objects, which collect data on infrastructure use using smart objects.
Smart grid	In electricity, the smart grid uses digital communications and embedded devices to connect the existing electricity infrastructure to a network. This network connectivity allows for the collection and subsequent use of data to manage demand and supply.
Smart meters	Unless otherwise specified, smart meters in this report refers to the smart meters that are being rolled out by the Government since 2011 with an objective to cover all homes by 2020. They digitally capture and send consumption data (potentially in half-hourly intervals where the user consents) and offer in home displays.
Virtual Reality (VR)	Technology that allows users to perceive and interact with a simulated environment, which may be realistic or fictional. Virtual reality is primarily experienced through sight and sound (for example using VR headsets).

Executive summary

Data is generated across the infrastructure sectors and used by a variety of stakeholders



Data generation

Personal data

Non-personal data

Network data

Asset data

Organisational data

Data use and sharing

Regulators

SMEs

Researchers

Supply chain

Policy makers

Tech firms

Retailers

Individuals

Infrastructure owners & operators

Promoting greater data sharing in infrastructure could realise annual benefits in the order of ***£15 billion****, resulting from…*

Improved efficiencies

Lower costs, better capacity management, lower emissions, increased outputs

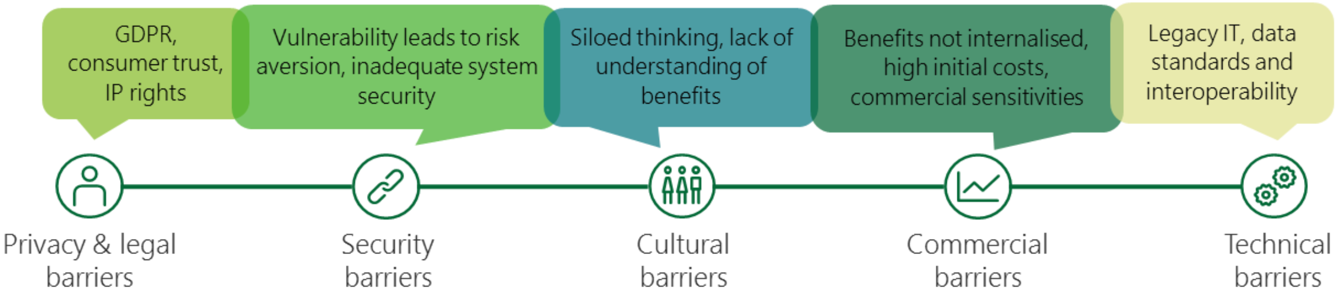
Increased competition and innovation

Effective competition and market entry, consumer savings, development of innovative services

Network planning and resilience

More accurate resilience models, reduced frequency, duration and impact of disruptive events

However, these benefits depend on current ***barriers to data sharing*** *being addressed*



GDPR, consumer trust, IP rights

Privacy & legal barriers

Vulnerability leads to risk aversion, inadequate system security

Security barriers

Siloed thinking, lack of understanding of benefits

Cultural barriers

Benefits not internalised, high initial costs, commercial sensitivities

Commercial barriers

Legacy IT, data standards and interoperability

Technical barriers

Reducing these barriers requires ***action*** *from Government, as well as a collaborative and committed approach from public bodies, infrastructure players, academia and third party data users, such as…*

New industry-led data sharing groups

Industry-led groups in each sector, facilitated by Government to identify challenges & use cases, working collaboratively towards on open standards, privacy and security solutions

Creating an infrastructure data framework

An overarching set of principles to promote best practice and provide clarity on data ownership, standards, contracting, compliance with security requirements, etc

Active role for Government in regulated industries

A regulatory impetus may be necessary to fully address cultural and commercial barriers, ensuring that regulatory frameworks provide appropriate incentives for data sharing

Support to Local Authorities

Challenges in accessing local Government data could be addressed by funding, supported by clearly identified use cases and analysis of benefits

The benefits of data in infrastructure

The economic potential of data is widely recognised. In 2016 the aggregate economic benefit from data was estimated to exceed £50 billion in the UK.[1] As the number of connected devices continues to grow, so will the quantity of data they generate, further increasing the expected value of data and its benefits to society.

This holds true in the infrastructure sectors,[2] where by 2024 it is forecast that there will be almost 100 million Internet of Things (IoT) connections.[3] Benefits arising from data use and re-use in the infrastructure sectors include:

- Improving efficiency, by giving infrastructure owners, operators and users information to make more informed decisions and support coordination between different parties. In the energy sector, data sharing can be an important component to enable new solutions such as demand-side response and intelligent storage that help to manage demand and supply fluctuations, with potential benefits in the order of £2 billion per year.[4] In the transport sector, data-driven solutions such as digital signalling and smart traffic management show potential to increase rail capacity and reduce road congestion by 10-40% in some cases.[5] In construction, data sharing as part of Building Information Modelling (BIM) can lead to cost savings in the region of 20-30%.[6]

- Informing better infrastructure planning and strengthening resilience by providing a holistic, real-time understanding of infrastructure assets and how they are used, facilitating long-term planning, predictive maintenance and the management of disruption. For example in the water sector, data from smart meters can give infrastructure operators a more holistic view of the infrastructure in order to understand weaknesses and existing leaks. In energy, visualisations are helping to provide improved situational awareness across the grid by making it easier to draw actionable insights from large, complex datasets.

- Increasing competition and innovation by providing third-parties and potential new market entrants with the data required to develop new services and apps. Already in the transport sector, open data made available by TfL has been accessed by over 10,000 developers and powers around 600 apps used by 42% of Londoners,[7] generating annual benefits of up to £130m.[8]

In the longer-term, greater data sharing across infrastructure sectors has the potential to amplify the above benefits beyond current levels and create new benefits. For example, in the transport segment, more data sharing could enable the deployment of 'Intelligent Mobility' solutions that cut journey times, pollution and optimise infrastructure planning and management, with estimated benefits of £15 billion over the period to 2025.[9] Similarly, more data sharing across all of the infrastructure sectors could inform the development of large-scale digital twins, modelling critical infrastructure networks and assets under different scenarios to inform planning decisions and prototype new solutions accurately. Although trials of digital twins are taking place at city-level to explore this concept, the full potential benefits are not yet known.

Current barriers to data sharing in infrastructure

The value that is delivered from infrastructure data is dependent on the extent to which it is shared with stakeholders in the wider economy. A qualitative analysis of data sharing in the sector and discussions with industry and academic

[1] IDC and Open Evidence for the European Commission (2017), 'European Data Market Study'
[2] Throughout this report, 'infrastructure sectors' refers to the transport, energy, water and wastewater, digital communication, solid waste and flood risk management sectors, as per the NICs remit.
[3] Cambridge Consultants for Ofcom (2017), 'Review of latest developments in the Internet of Things'
[4] Carbon Trust and Imperial College London (2016), 'An analysis of electricity system flexibility in Great Britain'
[5] Venture Beat (https://venturebeat.com/2017/05/05/how-chinas-meshing-ride-sharing-data-with-smart-traffic-signals-to-ease-road-congestion/, accessed October 2017); Network Rail (2015), 'Wessex Route Study'
[6] Building Information Modelling (BIM) Task Group (http://www.bimtaskgroup.org/bim-faqs/, accessed October 2017)
[7] TfL Open Data Policy (https://tfl.gov.uk/info-for/open-data-users/open-data-policy)
[8] TfL (https://tfl.gov.uk/info-for/media/press-releases/2017/october/tfl-s-free-open-data-boosts-london-s-economy)
[9] Transport Systems Catapult (2017), 'The case for Government involvement to incentivise data sharing in the UK Intelligent Mobility Sector'

stakeholders has shown that there currently exist a number of barriers which limit data sharing. Barriers relating to privacy, security and organisational culture are reported to be particularly significant.

- Privacy and legal barriers around the sharing and use of personal data have been cited as a key reason why organisations may be unable or reluctant to share the data they hold with others. With the incoming General Data Protection Regulation (GDPR), some stakeholders noted a lack of clarity – particularly in the short-term – around the lawfulness of data sharing and, specifically, whether and in what circumstances consumer consent is required. Coupled with consumer reluctance to give consent (where required) for non-essential use of data, or for new, unexpected uses of their personal data which were not notified to the consumer when their data were first collected, this can inhibit data sharing. Personal data collected via, for example, smart meters, connected vehicles or smart tickets, may be patchy or unavailable for the development of innovative services.

  Non-personal data is generally less affected by legal barriers, though infrastructure contracts may restrict the scope for data sharing where these have not evolved to recognise the importance of data and where underlying cultural and commercial barriers are reflected in the terms used.

- Security barriers refer to concerns that sharing data may lead to security breaches, data losses and, in extreme instances, high-impact cyber-attacks. As data sharing grows and new technologies drive a more interconnected infrastructure ecosystem, the range of potential threats is reported to be expanding. Organisations' security fears are heightened further by large potential fines under GDPR and Network and Information Systems (NIS) Directive rules, though the reputational and commercial repercussions of a breach may yet be greater. These fears are thought to be discouraging greater data sharing, especially in instances where the benefit case is uncertain or not fully understood.

- Commercial barriers refer to data not being shared because the costs of sharing are perceived to be greater than the expected benefits. The value of data may be poorly understood or sharing data may be perceived to entail a loss of competitive advantage, while sharing data may entail financial costs if IT systems and data management practices have not been set up appropriately. Commercial barriers are higher where data sharing and collaborative approaches are not the norm, as there may be a 'free rider' problem where data shared by one firm is used by other parties who do not reciprocate by sharing their own data.

- Cultural barriers refer to habits, policies and attitudes within organisations that oppose data sharing. These exist both within the private and public sectors and can reflect trust issues and a belief that data should only be shared on a 'need to know' basis, as well as a lack of understanding and focus on the potential benefits of data and new technology. This type of organisational culture appears to be deep-rooted within several large infrastructure players who may be relatively slow to adapt to new trends. Where cultural barriers exist these are likely to exacerbate other barriers, as organisations may fail to recognise commercial benefits from data sharing and take an excessively cautious view of privacy and security issues.

- Technical barriers refer to challenges in sharing the data from a technical perspective, such as the data not being in the right format (for example due to a lack of common standards), being stored only in legacy systems that are not built for sharing or the data not being digitised at all. Linked to commercial barriers, if the value and benefit of sharing data is not understood, then there will be no incentive to invest in new systems to share data.

Analysis of existing studies suggests that, by tackling these barriers, increased data sharing in the future could lead to annual benefits from data in the order of £15bn across the UK's infrastructure sectors, compared to current levels of around £8bn.[10] On the other hand, where these barriers are not addressed, data would remain siloed within individual organisations, with fragmented data-driven innovations or initiatives taking place only among a restricted set of parties

[10] Deloitte analysis of Cebr (2016), 'The Value of Big Data and the Internet of Things to the UK Economy'; McKinsey Global Institute (2013), 'Open data: unlocking innovation and performance with liquid innovation'. See Section 4.5 for more details.

who can access the relevant data. As such, the UK's infrastructure sectors would risk lagging behind other countries, with infrastructure productivity and innovation potentially being impaired.

Potential remedies: overcoming barriers to data sharing in infrastructure

A balance needs to be achieved between tacking cultural, commercial and technical barriers to realise the benefits from data sharing while simultaneously ensuring that security and privacy risks are appropriately addressed. This study identifies key areas where Government can take steps to facilitate market-led solutions to improve data sharing, whilst remaining mindful of the above trade-offs and the need for a measured and targeted approach. In the case of commercial and cultural barriers, these may be more difficult to address through an industry-led approach alone; therefore an active role for Government and regulators is also considered.

The suggested measures are intended to complement one another, rather than being seen as alternative options, and should also complement existing initiatives around using procurement to compel data sharing, ongoing publication of open data by the public and private sector and specific industry activities to share data.

New industry-led data sharing groups

Government (potentially via regulators, the NIC or other public bodies) can consider facilitating the creation of industry-led groups in different infrastructure sectors to tackle particular challenges around data sharing, including:

- Articulating economic, social and environmental challenges specific to the sector and what data and data sharing is required to address them;
- Providing use cases and guidance on the (monetary) value of data, its ROI and the benefits of sharing;
- Considering the implications of GDPR and developing a common approach to complying with it whilst still sharing data, to foster clarity, certainty and consumer trust;
- Considering the security issues that arise within the sector, promoting greater awareness and understanding of these and working collaboratively with external organisations, such as NCSC and the Cyber Security Information Sharing Partnership (CiSP) towards solutions, such as developing secure gateways to share data;
- Developing harmonised open standards that can be applied across the sector; and
- Promoting the development of open APIs and more data being available as open by default.

The membership of these groups would be jointly decided between industry and Government, but is likely to include sector bodies, a representative sample of infrastructure operators, SMEs involved in the infrastructure sectors, regulators, consumer groups and the relevant central and devolved departments.

An infrastructure data framework

A key gap identified by many stakeholders was the absence of an overarching set of principles that provided guidance and clarity on issues such as data ownership, what constitutes data, what might be interpreted as personal and non-personal, ensuring security by design, and so forth. While such a framework cannot ever be considered definitive, a common set of principles applicable across the whole sector (which can be customised) can be used as a starting point for subsequent data sharing, building on the work by industry groups to providing overall guidance.

The principles of the framework could cover areas including:

- Best practice guidance for organisations to carry out an internal audit of their data, classifying different types and identifying data that can be shared, either as open data or with restrictions.

- Best practice guidance for data quality and formatting for different categories of data.
- Approaches to specifying contracts that give appropriate emphasis to data requirements, clarity around responsibilities and liabilities related to data, and ensure there is scope for data to be used and re-used.
- Approaches to data anonymization and aggregation so that confidential data may become shareable.
- Steps to deal with grey areas around data ownership, data and IP, personal and non-personal data, etc.
- Appropriate security measures for data sharing in infrastructure, building on the Government's '10 steps guidance' and NIS Directive principles to build awareness and understanding among infrastructure players, setting out explicitly how best practice in cybersecurity can be achieved by infrastructure organisations.

This framework would benefit from leadership by a public body with an invested interest in each industry, which would be complementary to the work carried out by industry-led groups. Inputs should be sought from industry and academia, and facilitated by public bodies such as regulators and NIC.

Active role for Government and regulators in regulated industries

Past experience shows that where Government plays an active role in stimulating data sharing, rapid and significant changes can be brought about. For example, widespread open data initiatives have helped reduce cultural and commercial barriers to data sharing among public bodies.

In the regulated infrastructure sectors, a similar misalignment of incentives to data sharing often exists. The commercial incentive to work towards innovative data-driven solutions and make the necessary investments is at times insufficient. A historic focus on engineering-based solutions may mean that a cultural shift may be needed to enable greater focus on data and new technology. Stakeholder discussions suggest that a full reduction of these barriers is unlikely to be achieved through industry-led initiatives alone.

Against this backdrop, there appears to be scope for Government and regulators to a greater regulatory impetus to promote data sharing. Any regulatory-led action, such as adjustments to regulatory frameworks, guidance and specific targets for data sharing, would work in parallel with industry-led data sharing groups and the development of a broader infrastructure data framework, to maximise the likelihood of a pervasive shift towards greater data sharing across stakeholders. While recognising the potential benefits of data sharing, any intervention would need to carefully consider the costs involved and how these may be dealt with as part of regulatory frameworks and incentive mechanisms.

Support to Local Authorities

Some local authorities have made progress in sharing infrastructure data – for example the Thermal Harrow open data initiative helps to identify heat loss from buildings;[11] Data Mill North is the first platform to bring together open data from different sectors of cities and promote sharing and re-use, with over 90 transport datasets available.[12] However, stakeholders and previous studies highlight challenges in accessing data from local Government, particularly transport-related data including traffic monitoring data, cycle counts, road closures and diversions.[13]

The reported reason for this was often a lack of funding to either share the data or maintain its quality and integrity. While recognising fiscal constraints, the return from making this data available more widely as open data is significant. Local authorities could benefit from Government and industry support via challenge funds that local authorities and

[11] Harrow Council (http://www.harrow.gov.uk/info/100006/environment/1514/thermal_harrow, accessed October 2017)

[12] Leeds City Council (http://www.leeds.gov.uk/opendata/Pages/Data%20Mill%20North.aspx, accessed October 2017)

[13] See for example Ricardo Energy and Environment for the Department for Transport (2017), 'Scoping Study into Deriving Transport Benefits from Big Data and the Internet of Things in Smart Cities'; Transport Systems Catapult (2017), 'The case for Government involvement to incentivise data sharing in the UK Intelligent Mobility Sector'; DotEcon (March 2015), 'Independent evaluation of the OFT's 2006 market study into the Commercial Use of Public Information (CUPI)'

data users or re-users can bid for, to support the opening up of specific datasets for wider use. The work of the industry group would feed into this remedy by identifying specific datasets and use cases that are particularly valuable across the infrastructure sectors.

Each of these remedies will need to be mindful of the potential risks from more data being shared in infrastructure, balancing several important trade-offs around consent for personal data, rising security risks from more connected systems and the return on investment from improving data sharing systems when the future benefit is uncertain or unknown.

1 Introduction

This report focuses on identifying barriers to data sharing in the UK infrastructure sectors and remedies to address these.

1.1 Context and background

The National Infrastructure Commission (NIC) has been established to provide an independent, credible voice on infrastructure policy and strategy for the long-term. It seeks to support sustainable economic growth across all regions of the UK, improve competitiveness and the quality of life. It focuses on economic infrastructure covering the transport, energy, digital communications, water, solid waste and flood defence sectors.

In November 2016, the Government asked the NIC to conduct a new study on how technology can improve infrastructure productivity. In particular, it was tasked with exploring the application of technologies such as digitalisation, big data, artificial intelligence (AI) and the internet of things (IoT) to the infrastructure sectors and how Government could support their deployment. A common component across all these technologies is their use of data.

1.2 Scope of this project

To support the publication of its study, NIC has commissioned Deloitte to focus specifically on data sharing in the infrastructure sectors. Recognising that data and its availability across a wide variety of stakeholders are critical to the deployment and development of these technologies, this study seeks to examine:

- The current barriers to data sharing in the infrastructure sectors;
- The potential benefits of increased data sharing in the infrastructure sectors; and
- The role Government can play in facilitating greater data sharing.

The main focus of the analysis has been on the energy and transport sectors as the largest sectors, although the findings are applicable across other infrastructure sectors. Particular focus has been paid to the role of the digital twin as an emerging technology, showcasing the potential of data sharing to transform the infrastructure sectors.

This study has been conducted over a period of eight weeks during September and October 2017 and has involved a literature review, interviews with a range of Deloitte industry experts, external stakeholders and cross-sector roundtables.[14] In its analysis of legal issues, Deloitte has received expert support from Berwin Leighton Paisner. Stakeholders were selected to ensure that a broad range of views and evidence would be collected from organisations involved in different infrastructure sectors, as well as academics and institutions that have written on the subject of data sharing, and stakeholders that were suggested by internal Deloitte experts and the NIC itself. Hypotheses have been formed and validated with industry stakeholders, academics, public bodies and other data experts. Nevertheless, within each of the infrastructure sectors the number of stakeholders was limited by time constraints; therefore, the findings do not necessarily represent the full range of views across all sectors.

Deloitte would like to thank all stakeholders who participated in interviews and roundtables, as well as staff at the NIC for their input to the study.

[14] External stakeholders consulted for this study are listed in the Appendix.

2 The data landscape in infrastructure

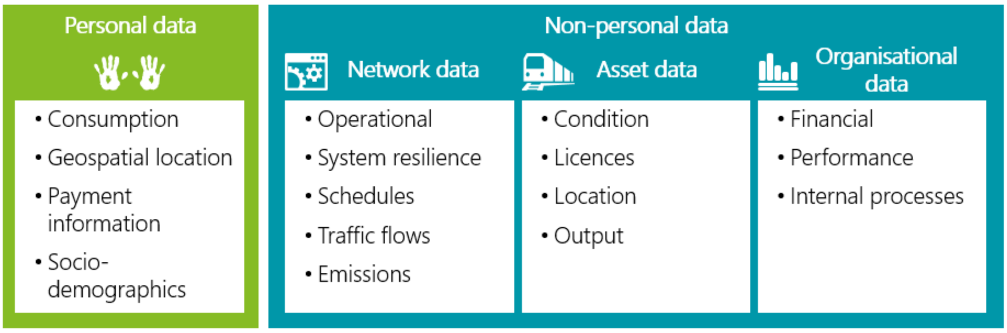
From smart meters and smart tickets to traffic flows and the location of water pipes, the infrastructure sectors hold and generate significant flows of data. This data is generated not just by the infrastructure assets and their owners and operators, but importantly the users as well. The data covered is of a personal, network, asset and organisational nature, and is used and re-used by a variety of stakeholders.

2.1 Data in the infrastructure sector

Infrastructure data is generated by users of infrastructure, the networks and their assets as well as the wider environment (for example weather conditions). The ubiquity of the 'internet of things', sensing technologies and smart phones coupled with the 'velocity and verbosity' of the data collected mean the volume of this data being generated and subsequently collected will be significant – potentially more than any other part of the economy. For example, a report for Ofcom forecast that by 2024 there will be 55 million IoT connections in the Automotive sector, 36 million in Utilities, 6 million as part of Smart Cities and 2 million in Construction.[15]

Broadly speaking, the main types of data being generated and collected in the infrastructure sectors include personal data and non-personal data.

Figure 1: Types of infrastructure data



Personal data

- Consumption
- Geospatial location
- Payment information
- Socio-demographics

Non-personal data

Network data

- Operational
- System resilience
- Schedules
- Traffic flows
- Emissions

Asset data

- Condition
- Licences
- Location
- Output

Organisational data

- Financial
- Performance
- Internal processes

Source: Deloitte

These data types are not necessarily mutually exclusive. For example, data from smart meters or IoT devices may come to be seen as network data for the management of smart grids, but may also be personal data if it is identifiable at the household or individual level.

Where personal data starts and stops is sometimes contested and open to interpretation. The disclosure of personal data in an infrastructure sector data sharing initiative is not necessarily a requirement and personal data may be converted to non-personal data through:

- Aggregation, where a combined dataset is produced with no individual-level information.

[15] Cambridge Consultants for Ofcom (2017), 'Review of latest developments in the Internet of Things'

- Anonymisation, where individual-level data is retained, but with potential identifiers removed (such as name or home address).

However, these data types are not necessarily constant; as the capability to mosaic and combine different datasets increases, so too does the possibility of making individuals associated with the data identifiable, even from a dataset that was initially anonymised or aggregated. For example, location data, identification numbers or online identifiers, such as IP addresses, cookies and RFID tags, could provide ways to make data personally identifiable.[16]

The boundaries between different types of non-personal data may equally be difficult to ascertain. For example, data across several assets may be aggregated to form datasets that give insights across an entire network; this is the case where sensors are used to track the location of individual vehicles such as buses, but data is also aggregated to better understand how the network as a whole operates.[17]

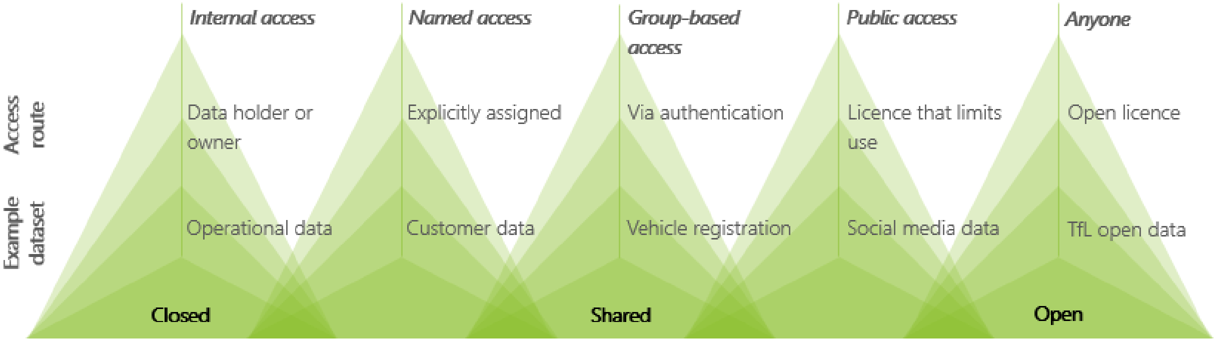
The classification of data may also vary across other important dimensions. For example, whether or not data is commercially sensitive or not may affect the scope for data sharing. There can also be a separate distinction between static and dynamic data, such as real-time data, with different potential uses based on data sharing.

2.2 Data sharing in the infrastructure sector

Based on stakeholder discussions, 'data sharing' is defined as the sharing of data between organisations that originally created or collected the data with others that seek to use or re-use it for a variety of purposes.

There is a spectrum of data sharing – on the one hand, data can be freely shared as 'open data', whilst on the other hand, it may be shared with conditions or under licence or contract (which may or may not involve a fee). It should be stressed that there will be some instances of data that will not be suitable for sharing under any mechanism, such as sensitive and operationally critical operating data from power plants. This spectrum is show in Figure 2.

Figure 2: The data spectrum in infrastructure



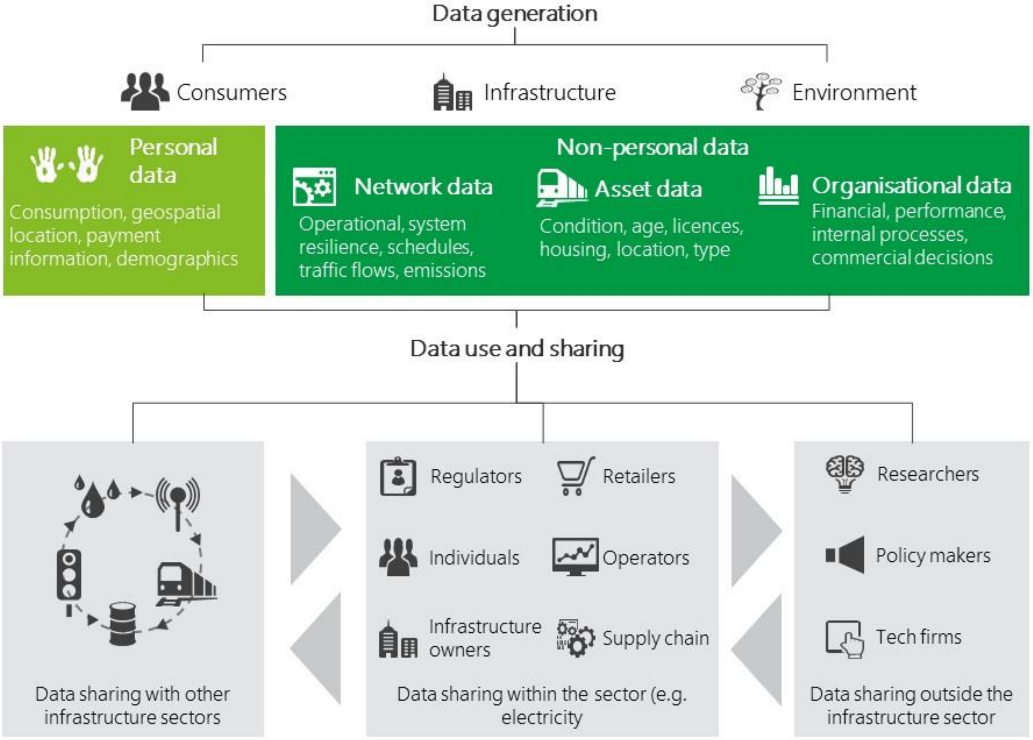
Source: ODI (example datasets adapted by Deloitte)

Data sharing is already well established across pockets of the infrastructure sector. Data is used and shared by a variety of organisations and individuals, including infrastructure owners and operators, regulators, retailers, researchers, third parties and private citizens.

[16] These possibilities are recognised in article 4(1) and recital 30, GDPR.

[17] See for example Houses of Parliament (2014), 'Big and Open Data in Transport'

Figure 3: Data value chain in infrastructure


Data generation
Consumers
Infrastructure
Environment
Personal data
Consumption, geospatial location, payment information, demographics
Non-personal data
Network data
Operational, system resilience, schedules, traffic flows, emissions
Asset data
Condition, age, licences, housing, location, type
Organisational data
Financial, performance, internal processes, commercial decisions
Data use and sharing
Data sharing with other infrastructure sectors
Regulators
Retailers
Individuals
Operators
Infrastructure owners
Supply chain
Data sharing within the sector (e.g. electricity
Researchers
Policy makers
Tech firms
Data sharing outside the infrastructure sector

Source: Deloitte

There are several examples of current data sharing initiatives in the infrastructure sectors that seek to promote and ease data sharing.

Box 1: Selected examples of data sharing in the infrastructure sectors

- Data is often shared between different parties in the supply chain as part of infrastructure construction projects, including as part of BIM level 2 projects which provide a common environment for data sharing. Under this standard, "*design information is shared collaboratively using a common file format, enabling data to be combined to create (and check) a federated BIM model*". However, adoption of this standard remains limited (see Box 14), as does sharing of data outside of the project supply chain (see Box 11).
- The Infrastructure Transitions Research Consortium (ITRC) recently launched the Data Analytics Facility for National Infrastructure (DAFNI), which aims to build "*a national database for visualisation and analysis*" that can act as "*a shared, secure system for academic research and a resource for businesses, innovators and policy-makers*". DAFNI is backed by £8m of funding and builds on the ITRC's prior experience of building a National Infrastructure Systems Model (NISMOD), which is a database with over 400 layers, but DAFNI aims to provide greater breadth and detail, "*to represent individual buildings and to develop plausible connectivity networks*".[18]
- Roadworks data shared by different infrastructure players is aggregated on a centralised, publicly accessible portal by Elgin, collating data shared by 170 Local Highways Authorities, metropolitan transport authorities, Network Rail and gas, water, electricity and telecom companies.[19]
- Bodies such as TfL, Rail Delivery Group and Environment Agency have released open data feeds or open APIs. For example, TfL releases live open data on different modes of transport, including through a unified API;[20] Rail Delivery Group offers Fares,

[18] ITRC (2017, http://www.itrc.org.uk/wp-content/PDFs/DAFNI-launch-notes.pdf, accessed October 2017)
[19] See https://www.elgin.org.uk/products/roadworksorg
[20] See https://tfl.gov.uk/info-for/open-data-users/

Routes and Timetable data, as well as links to data feeds from National Rail;[21] the Environment Agency has released more than 1,000 datasets as open data, including national flood risk datasets.[22]

- Water and energy regulators are working to explore better uses of data and information sharing to identify and assist vulnerable customers.[23]

Nonetheless, despite these initiatives, stakeholders reported the scope for much more data sharing. The Royal Academy of Engineering has noted that "*while Government has led the way* [with data sharing], *much potentially valuable data remains locked away in corporate silos or within sectors*".[24] Even in the case of the public sector, there is potential for improvement, for example by Local Authorities.[25] The reasons for why data sharing is not happening more frequently or to greater levels is discussed in the next Chapter.

[21] See https://www.raildeliverygroup.com/our-services/rail-data.html

[22] See https://data.gov.uk/publisher/environment-agency

[23] UKRN (2017, http://www.ukrn.org.uk/news/better-use-of-data-and-information-sharing-to-identify-customers-in-vulnerable-situations-august-project-update/, accessed October 2017)

[24] Royal Academy of Engineering and the Institute of Engineering and Technology (2015), 'Connecting data: driving productivity and innovation'

[25] As noted for example in Transport Systems Catapult (2017), 'The case for Government involvement to incentivise data sharing in the UK Intelligent Mobility Sector' and DotEcon (March 2015), 'Independent evaluation of the OFT's 2006 market study into the Commercial Use of Public Information (CUPI)'

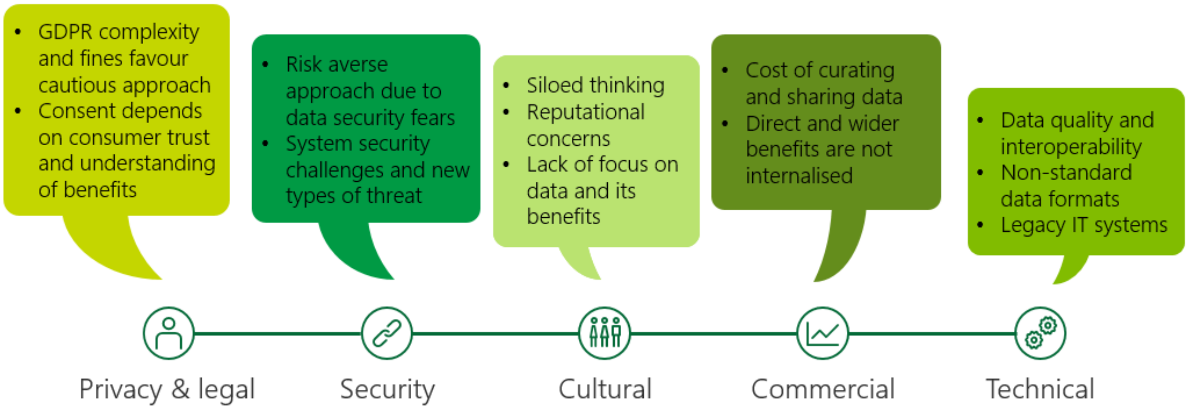
3 Identified barriers to data sharing

Analysis and discussions with industry stakeholders confirm that a number of barriers to data sharing currently exist. Many of these barriers are interrelated and, in some cases, reinforce one another. Cumulatively, the impact is that data sharing is not occurring at the optimal level for the UK to take advantage of new technologies in infrastructure.

3.1 Overview of barriers

Stakeholders have identified the following barriers to data sharing in the infrastructure sectors.

Figure 4: Barriers to sharing data in infrastructure sectors



Privacy & legal
- GDPR complexity and fines favour cautious approach
- Consent depends on consumer trust and understanding of benefits

Security
- Risk averse approach due to data security fears
- System security challenges and new types of threat

Cultural
- Siloed thinking
- Reputational concerns
- Lack of focus on data and its benefits

Commercial
- Cost of curating and sharing data
- Direct and wider benefits are not internalised

Technical
- Data quality and interoperability
- Non-standard data formats
- Legacy IT systems

Source: Deloitte

These barriers do not exist independently – many are interrelated and may exacerbate each other. For example, a suspicious culture among consumers may mean that they do not give consent for the sharing of personal data due to privacy reasons, and an organisational culture may place little emphasis on exploring the commercial potential from data sharing. However, this interrelatedness also means that addressing certain key barriers could have a 'halo effect' in subsequently reducing other data sharing barriers.

The following sections considers each barrier in turn.

3.2 Privacy and legal barriers

Privacy concerns and legal restrictions can apply to personal data, while other legal barriers can affect non-personal data. These issues are considered in turn below.

Personal data

Across the infrastructure sector, personal data is being collected via IoT devices, payments and consumption data and geospatial tracking. The collection, storage and use of this personal data can trigger a number of actual or perceived barriers that potentially reduce data sharing. In particular, certain laws and regulations may restrict the scope for sharing personal data through strict sharing conditions, measures to safeguard sensitive data or by preventing data from being shared outright. Where sharing is permissible, these laws and regulations may increase the costs to sharing data or dissuade organisations from sharing personal data for non-essential purposes for fear of contravening regulations. A key regulation in this respect is the incoming General Data Protection Regulation (GDPR), due to come into force on 25 May 2018, which was raised by nearly all industry stakeholders.[26]

Box 2: Summary of key GDPR components

- The GDPR updates the prior EU data protection directive (95/46/EC) and regulates the handling ('processing') of any form of personal data. In the infrastructure sectors this can include customer or passenger data linked to specific individuals or households, as well as data from smart meters, smart tickets or connected vehicles.
- The GDPR strengthens requirements around consent, where this is used as the lawful basis for the use of personal data. Consent must be given freely and actively (i.e. through 'opt-in' rather than 'opt-out'), and be specific, informed and unambiguous.
- Data reliant on consent which was held and used prior to the GDPR may need to be 're-consented' according to guidance from the Information Commissioners Office (ICO), or a different ground relied upon, so as to comply with the GDPR. This can apply even in the case where there is no anticipated change in the use of the data.
- Revised privacy notices will be required in order to meet enhanced transparency requirements in the GDPR, such as the right to inform individuals about their right to object to processing or to complain to the ICO.
- The GDPR also provides individuals with enhanced rights in relation to their data. They may:
  - restrict or object to specific uses of data (with exemptions from the right to object applying in the case of processing for scientific or historical research purposes or statistical purposes);
  - request for their data to be provided in a machine readable format or transferred between organisations; and
  - request to have their data erased.
- Organisations are obliged to report any data security breaches and the maximum fines for breaches are increased significantly, up to a maximum of 4% of worldwide turnover or €20 million, whichever is higher.

The new provisions in the GDPR around consent illustrate a complex trade-off: the regulation aims to safeguard personal privacy as the use of digital services and IoT devices increases, but strengthening requirements around consent may make it less likely that consumers give consent for their data to be shared and used, even in cases where it would benefit them. The risk of placing reliance on consent is that "*if consumers are expected to engage in detail with technical matters for which they lack the time or understanding, this can lead to uninformed consent ('choice fatigue', which defeats the purpose of the policy) or to an uninformed refusal to give consent (which may be a loss to everyone)*".[27]

Applied to the infrastructure sector, there was broad consensus among stakeholders that the GDPR will act as a barrier to data sharing, with the increased potential for fines encouraging a more risk-averse approach. For instance, failure to conduct a Data Protection Impact Assessment (DPIA) in the case where the type of data processing is likely to result in a high risk to the rights and freedoms of the individual could result in financial penalties. In itself, the obstacle of carrying out a DPIA could also act as a further factor to deter organisations from sharing data.

However, there were differing views as to the overall impact of the GDPR on data sharing in the medium-to-long run. Specifically, some stakeholders argued that in the short-run the regulation would have a large 'chilling effect' on data sharing, but this could only be temporary until uncertainty around how the GDPR is to be applied is resolved. Already, some organisations are attempting to address compliance with other standards and norms (such as ISO 27001[28]) at the

[26] Regulation (EU) 2016/679. Note that the UK Government has clarified that this will be implemented in UK law and remain in place following the UK's withdrawal from the EU. The Data Protection Bill (HL Bill 66), currently going through Parliament, confirms this.

[27] Analysys Mason (2014), 'Data-driven innovation in Japan: supporting economic transformation'

[28] A family of standards that helps organisation keep information assets secure by providing requirements for an information security management system (https://www.iso.org/isoiec-27001-information-security.html)

same time as addressing GDPR compliance and gaining a holistic understanding of the types of personal (and non-personal) data held, which can reduce these barriers in the longer term.

In contrast, other stakeholders believe the GDPR represents a fundamental philosophical shift away from the previous Data Protection Act 1998, with stronger requirements for consent materially limiting data sharing on an ongoing basis. It was noted that current guidance by regulators, such as the ICO, was taking a conservative interpretation.

Although the GDPR has not yet come into force, several stakeholders across the infrastructure sectors cited examples of how the regulation is already reducing data sharing or making it more difficult.

Box 3: Examples of personal data in the infrastructure sectors

- In the energy sector, smart meter data is subject to the Data Access and Privacy Framework, which is broadly consistent with the GDPR requirements. Half-hourly data, which industry stakeholders and experts see as valuable to ensure the long term viability of the energy grid, is only shared if customers explicitly opt-in. Thus far, there has been consumer resistance to installation and limited opt-in consent, leading to 'patchy' half-hourly data.
- Similar restrictions could prevent beneficial uses of other types of infrastructure data, such as from smart tickets or connected vehicles. For example, the EU Intelligent Transport Systems platform considers vehicle-to-vehicle and infrastructure-to-vehicle communications as personal data. This could restrict the potential uses and sharing of this data, potentially slowing the development of innovative third-party services based on this data.
- Stakeholders have raised examples where privacy concerns have obstructed benevolent uses of data, for example to contact vulnerable customers who could benefit from improved insulation in their homes.[29] Similarly, in the transport sectors stakeholders suggested that privacy restrictions and fears may be preventing innovative uses of smart ticket data, for example to provide real-time notifications to passengers.

It is worth noting that consent is not the sole legal basis for sharing personal data in the infrastructure sectors. Data sharing and processing can be legitimised where it is necessary for the performance of a task that is in the public interest,[30] or where necessary for the purposes of the legitimate interests pursued by the party disclosing the data, or the party receiving it, as balanced against the rights, freedoms and interests of the individuals.

In addition to the GDPR, the new ePrivacy Regulation[31] may affect IoT and machine-to-machine communications. The regulation includes additional requirements regarding unsolicited marketing, cookies, confidentiality and the processing of communication data and metadata which may raise the costs of data sharing. However, these restrictions appear to be creating relatively little concern for the infrastructure sectors relative to the GDPR.

Non-personal data

The infrastructure sectors also generate a wealth of non-personal data, for instance on asset locations, asset condition and network flows. Non-personal data may be subject to concerns about potential breaches of competition law, where data sharing is perceived as collusive or anticompetitive, could dampen incentives to share, for example in relation to price or volume data. This issue has not generally been raised by the majority of stakeholders and may only take effect in a limited number of cases.

Where data sharing is subject to complex or onerous contractual terms, these may create a significant burden for organisations accessing the data, particularly SMEs. This has previously been raised as an issue in relation to some data held by public bodies, though barriers appear to have been reduced through the adoption of the Open Government

[29] For an example of a situation where a refusal to share data gas account data has been linked to the deaths of vulnerable customers, http://news.bbc.co.uk/1/hi/england/london/3342059.stm

[30] These grounds correspond to Articles 6.1(f) and 6.1(e) of the GDPR.

[31] Procedure 2017/0003/COD

Licence for some datasets and the simplification of commercial licences.[32] As a case in point, licence terms for Ordnance Survey have become more accessible and open data initiatives such as the OS OpenSpace API encourage use of Ordnance Survey data without administrative or legal barriers.[33]

Today, legal issues appear somewhat more likely to arise in relation to data sharing as part of major capital projects. The adoption of more collaborative BIM approaches, with greater data sharing, raises several legal considerations.[34]

- Intellectual property: Data shared may incorporate elements of IP, and as more parties contribute to the project it is important that individual IP rights continue to be recognised. In general, UK IP law is capable of recognising the rights of parties making a distinct contribution (including recognition of pre-existing 'background IP rights'), but complex situations may arise where joint contributions are made.

  For example, where the end product of data sharing has been contributed to by more than one organisation, a situation of joint ownership may arise. Unless there are specific contractual provisions, the rights of such joint owners may not be clear and future use of the data could require the consent of all contributors. Issues may also arise where an IP licence is obtained in relation to data use on a project and expires before project members have ceased using the data.

- Data reliance and liability: Data sharing may create new classes of possible liability, as more parties use and rely on data that could include errors made by other parties. As collaboration increases, any change in one piece of information will impact several other parts of the model.

- Confidentiality: Where information shared is considered to be confidential, appropriate non-disclosure clauses may be required. In some cases, provisions may be required to allow confidential data to be held outside of the shared environment.[35]

These issues may require construction contracts to change, though stakeholders and existing evidence suggest that solutions should generally be manageable within current legal frameworks. Stakeholders point to the NEC suite of standard form contracts and the most recent NEC3 suite as evidence that contracting has already evolved to allow more collaborative and open ways of working.[36] Consistent with this, the BIM Industry Working Group found in 2011 that "*little change is required in the fundamental building blocks of copyright law, contracts or insurance to facilitate working at Level 2 of BIM maturity*", with greater collaboration and data sharing.[37]

Various organisations have published protocols, templates and guidance to overcome contractual issues, including:

- The Construction Industry Council's BIM Protocol, which establishes the contractual and legal framework for the use of BIM in eight clauses and is supported by NEC guidance on how to use this protocol with NEC3 contracts.[38]

- The British Standards Institute has published a specification for information management in a BIM context, setting out how to share information with common practices, standards and software, as well as other related specifications available as part of the BIM Level 2 suite of documents.[39]

[32] DotEcon (March 2015), 'Independent evaluation of the OFT's 2006 market study into the Commercial Use of Public Information (CUPI)'.
[33] Ordnance Survey (https://www.ordnancesurvey.co.uk/business-and-government/products/opendata-products.html, accessed November 2017)
[34] See for example NBS (2012), 'BIM: mapping out the legal issues'
[35] BIM Working Party (2011), 'A report for the Government Construction Client Group'
[36] For more details see Out-Law.com (2011), 'Standard Form Contracts: NEC'
[37] BIM Working Party (2011), 'A report for the Government Construction Client Group'
[38] CIC (2013), 'BIM Protocol'; NEC3 (2014), 'How to use BIM with NEC3 contracts'.
[39] See http://bim-level2.org/en/standards/downloads/

- Other templates and guidances for use in different contexts have been released by the Joint Contracts Tribunal,[40] Project Partnering Contracts and Alliance,[41] and Chartered Institute of Building.[42]

Overall, where contracts fail to adapt to allow greater scope for data sharing, discussions indicate that this is most likely to reflect an underlying cultural resistance to data sharing, as discussed in section 3.4, or commercial reasons for withholding data, as discussed in section 3.5, rather than any specific contractual or legal barriers.

Summary

Restrictions on the collection, sharing and use of personal data – which are enhanced by the GDPR – appear likely to have a material impact on data sharing in the infrastructure sectors, where personal data is increasingly collected from smartphones and IoT devices. While in principle the rules are not designed to restrict benign uses of data, there appears to be a genuine risk that consumers will opt to not share data (whether by refusing consent or by exercising rights to object or opt-out) where this could have personal and wider benefits. There is a culture among consumers that often opposes any non-essential sharing of data with organisations, reflecting a lack of trust and difficulties in understanding data uses and potential benefits, particularly where these are of a relatively technical or complex nature.

Legal issues related to non-personal data do not generally appear to be creating barriers, though currently contracts often place little emphasis on the importance of data or allow limited scope for sharing, which reflects cultural and commercial barriers discussed in later sections.

3.3 Security barriers

Fears around data security could discourage organisations from sharing data, while the security of infrastructure systems themselves is potentially affected by increased data sharing and new technologies. These issues are considered in turn below.

Data security

A key concern around the sharing of data across infrastructure is that it could fall into the hands of users who seek to use it maliciously. This applies to non-personal data as well as any personal data that is not treated as open data, for example due to commercial sensitivities. Data may be compromised by:

- Accidental breaches are possible where users or systems fail to comply with applicable data security restrictions or protocols. Examples of such breaches can include instances of data being sent to incorrect recipients, failure to redact data, accidental publishing or unencrypted transmission of restricted data, or saving data in an unsecure location.

- Malicious breaches or attacks occur where there is deliberate action by one or more individuals, for instance -

    - External or unauthorised users may be able to access and extract data (data exfiltration), either by hacking data storage (for example through malware or credential theft through keystroke logging) or intercepting data transmissions (for example through a man-in-the-middle attack).

    - Internal, authorised users may also be responsible for this type of breach, for example by deliberately leaking data to other parties.

[40] JCT (2011), 'Public sector supplement: Fair payment, transparency and Building Information Modelling'.
[41] See http://ppc2000.co.uk/
[42] CIOB (2013), 'CIO Contract for use with complex projects'

The likelihood of any kind of breach occurring can depend on a system's security design and configuration. Where systems are not set up with adequate measures in place there is greater scope for breaches, for instance as a result of flawed cryptography, weak passwords, outdated software or use of default security settings.

When data is shared externally, there can be a heightened risk of security breaches, particularly if it is difficult to verify the security measures taken by third parties. There have been instances of such breaches taking place in the telecommunications sector. For example, the security of customer data that Verizon had shared with a third-party partner was recently compromised due to inadequate security settings used by the third-party.[43] Similarly, data on T-Mobile US customers held by Experian was breached in 2015.[44]

The risk may increase where bespoke data exchange platforms or centralised data hubs that various parties may access are created. In such cases, any security breach – including by insiders – could provide access to vast quantities of data. An example of this is the healthcare sector, where extensive data sharing is commonplace and exchange platforms are used, but security breaches have been prevalent both in the UK[45] and internationally.[46]

While the infrastructure sectors are currently not among the 'worst offenders' with regard to cybersecurity incidents – the ICO identified 216 such incidents from January to June 2017, of which only 4 pertained to the transport and utilities sectors[47] – stakeholders noted that fears around security lapses may discourage data sharing. This may be particularly relevant for dynamic or real-time data, for example collected from IoT sensors, which involve constant or repeated data flows that may be relatively difficult to secure compared to static data.

The potential financial and reputational repercussions of data security issues create incentives to take a risk-averse approach to data sharing, potentially reinforcing a closed culture (see section 3.4 below). Stakeholders suggest that this effect is observed across sectors and also in the public sector, where security may impose an additional layer of cost and administrative burden for data sharing; this may be a barrier for Local Authorities in particular.[48]

Box 4: Examples of potential repercussions from data security breaches

- In the telecoms sector, a cyber-attack leading to the compromise of more than 155,000 TalkTalk customers' personal data resulted in a fine of £400,000 from the data protection regulator, the ICO, as well as the loss of an estimated 100,000 customers and additional costs of circa £60m to settle damages and improve data security.
- The hacking of 1.5 billion Yahoo users' data caused major reputational damage and a fall in its share price of 7%.[49]

Data security concerns apply equally to customers, when sharing their data with organisations. In particular, as more consumer IoT devices become available across the infrastructure sectors, users' willingness to adopt the devices and opt-in to data sharing may be heavily influenced by data security perceptions. Such security fears have been reported as a significant factor leading to consumer resistance towards the smart meter rollout, preventing more energy consumption data from being shared,[50] though stakeholders noted that data security fears among consumers are a cross-sector phenomenon.

[43] The Verge (2017, https://www.theverge.com/2017/7/12/15962520/verizon-nice-systems-data-breach-exposes-millions-customer-records)

[44] Financial Times (2015, https://www.ft.com/content/226e970e-6901-11e5-97d0-1456a776a4f5)

[45] The Telegraph (2017, http://www.telegraph.co.uk/news/2017/03/17/security-breach-fears-26-million-nhs-patients/)

[46] Brookings (2016), 'Hackers, phishers and disappearing thumb drives: Lessons learned from major health care data breaches'

[47] ICO (2017, https://ico.org.uk/action-weve-taken/data-security-incident-trends/, accessed October 2017)

[48] Law Commission (2014), 'Data sharing between public bodies: a scoping report'

[49] Catapult (2017), 'The case for Government involvement to incentivise data sharing in the UK mobility sector. Briefing paper.'

[50] The Times (2017, https://www.thetimes.co.uk/article/millions-of-homeowners-reject-smart-meters-over-hacking-fear-rhhm98ps2, accessed October 2017)

System security

Aside from data security breaches, the infrastructure sectors are potentially susceptible to system security breaches that could result in wide-reaching impacts for infrastructure users, as well as owners and operators. While these are not new threats and are not solely brought about by data sharing, stakeholders have suggested that increasingly connected physical and digital systems affected by new technology and big data can increase the scope for infrastructure security breaches, which risks discouraging greater integration and data sharing.

Legacy industrial control systems, such as those used in the energy and water sectors, may already have vulnerabilities to cyber-attacks. Media reports and industry papers suggest that these systems have been the victim of hacks,[51] which are becoming more frequent,[52] with the energy sector in particular being targeted (see Box 5).[53] These threats could increase as smart technologies and IoT adoption evolve: more data being exchanged and accessed from a variety of endpoints (PCs, smartphones, etc.) will lead to a broadly connected ecosystem of physical and digital systems that will inherently face vulnerabilities.

A key concern is that there could be a heightened risk of low-frequency, high-impact incidents, such as the cyber-attacks in the Ukraine that caused widespread power outages.[54] These risks may manifest in various forms of malicious attacks, such as distributed denial-of-service (DDoS) attacks, malware, ransomware, data tampering, false data injection or any other intervention which prevents systems from functioning normally.

Box 5: System security issues in the energy sector

- New technologies in the energy sector have the potential to create a more complex interconnected ecosystem, with the increasing role played by renewables (including small generators), DSR, storage, smart appliances and other innovations.
- An MIT study finds that "*the growing complexity and interconnectedness of electric grids is increasing the number of potential targets and vulnerabilities*", as the number of points where an unauthorised used may try to enter the system or extract data is increasing. Similarly, a study for the European Parliament notes that "*a massively expanding 'attack surface' now forms the operational foundation of the energy ecosystem. As the energy system is also fundamentally interconnected with every other critical infrastructure network, the cyber security threat to the energy sector impacts every aspect of our modern society*".[55]
- While it is new technologies mainly driving these developments, the data sharing aspect that comes with a more interconnected ecosystem is significant. Concerns are that information flows in advanced energy grids could be manipulated to cause malfunctions or outages,[56] a risk that may increase as more and more data flows between different users, generators and organisations at different points of the grid.
- The European Parliament study finds that good progress is being made to address these challenges, but advises that limited coordination and collaboration between organisations risks leading to outcomes that are not comprehensive. A separate EU study substantiates this, focusing on the importance of effective sharing of information about cyber security incidents in order to combat evolving threats.[57]

While the energy sector appears particularly affected, all critical infrastructure could be vulnerable to these types of threats. For example, the increasing range of digital technologies used in transport for navigation, tracking, signalling and other purposes "*are often interconnected through networks and remote access terminals, which may allow malicious actors easier access to key nodes*".[58] A recent cyber-attack on shipping company Maersk "*disrupted data-reliant processes*

[51] See, for example, Thales (2013), 'Cyber Security for SCADA Systems'
[52] Houses of Parliament (2017), 'Cyber Security of UK Infrastructure'
[53] The Telegraph (2017, http://www.telegraph.co.uk/technology/2017/07/18/hackers-targeting-uk-energy-grid-gchq-warns/, accessed October 2017)
[54] US Department of Homeland Security (2016), 'Cyber-Attack Against Ukrainian Critical Infrastructure'
[55] European Parliament Directorate-General for Internal Policies (2016), 'Cyber Strategy for the Energy Sector'
[56] Accenture Consulting (2017), 'Outsmarting Grid Security Threats'
[57] European Union Agency For Network And Information Security (2016), 'Report on cyber security information sharing in the energy sector'
[58] US Department of Homeland Security and Department of Transportation (2015), 'Transportation systems sector-specific plan'

such as creating arrival notices and obtaining customs clearance",[59] created congestion at some ports and overall caused estimated damages of up to $300 million. Similarly, telecom networks may be a significant target for cyber-attacks; in one overseas case, critical services were even shut down as a result of what proved to be a false claim of an attack, while the matter was investigated.[60]

Complementary to the forthcoming GDPR legislation, the imminent implementation of the EU Network and Information Systems (NIS) Directive[61] regulates cyber security with the aim of ensuring that essential networks and infrastructure services are adequately protected from system breaches. The focus of the NIS Directive is to improve cooperation and incident reporting between Member States and the sharing of cyber security related national strategies. The legislation establishes requirements for Operators of Essential Services (including services in energy, transport, water and digital infrastructure) to take appropriate measures to prevent and minimise the impact of incidents affecting the security of network and information systems used for the provision of essential services.

The penalties to be imposed under the NIS Directive and additional reporting obligations for non-personal data breaches could act as a deterrent to increased data sharing. However, stakeholder discussions suggest that the NIS Directive holds a lower profile than the GDPR, with relatively limited awareness and understanding of its contents.

Box 6: UK implementation of the Networks and Information Systems (NIS) Directive

The NIS Directive will require the UK's essential infrastructure operators to abide by four proposed high-level principles, each of which entails more specific requirements.

- Organisational structures, policies and processes to govern the security of network and information systems: This encompasses requirements that cover appropriate governance policies, risk management, asset management and supply chain risks.

- Proportionate security measures to protect from cyber-attacks or system failures: This includes measures to prevent unauthorised access to data and ensure that systems are appropriately configured and resilient, as well as appropriate policies, processes and staff training.

- Capabilities to ensure effective defences and to detect cyber security events: This requires ongoing monitoring of security statuses and effectiveness of security measures, with any anomalies being detected.

- Capabilities to minimise the impact of any cyber security incidents: Incident management processes should be clear and well tested, with any incident being analysed to improve resilience in the future.

The principles complement the GDPR measures concerning personal data security and breaches are subject to similar potential penalties as under the GDPR, potentially reaching £17 million or 4% of global annual turnover.

Source: Department for Digital, Culture, Media & Sport (2017), 'Security of Network and Information Systems: Public Consultation'

The principles set out by the NIS Directive are particularly important given the increasing role played by data and new technologies, which could make infrastructure systems vulnerable to new types of attack. Where IoT, Machine Learning, AI and automation are used, emerging threats are possible even from authorised users who may be able to affect expected outcomes by manipulating data flows and compromising the data integrity. Where the data is used for modelling or decision-making, including as part of automated processes, this could lead to adverse outcomes.

[59] Reuters (2017), 'Maersk says too early to predict financial impact of cyber-attack'

[60] Deloitte, Global Cyber Executive Briefing: Telecommunications case studies,

[61] Directive (EU) 2016/1148 concerning measures for a high common level of security of network and information systems across the Union. It is also referred to as the Cybersecurity Directive. Implementation by Member States is required to take effect from 10 May 2018.

Box 7: Example of security threats from authorised users

- Machine learning and AI technology could be compromised by being fed incorrect data that is 'authorised'. For instance, consider an autonomous car that relies on external sensors and data from other vehicles. The functioning of the car could be compromised if another authorised road user tampers with the vehicle-to-vehicle transmissions from their own car, or otherwise interferes with the road environment so that the required data cannot be read.
- This differs from hacking as it is the data input stream that is being compromised, potentially leading to adverse outcomes without any breach of secure systems or data. As the number of data sources increases and data sharing increases, this risk could become more prominent.

Overall, there is a lack of consensus among stakeholders on whether fears around data and system security breaches constitute a major barrier to data sharing specifically, or whether the issue should be seen more broadly as an ongoing challenge that industries face as data and technology become more prominent. While some stakeholders expressed a degree of scepticism about security-by-design[62] and whether it can feasibly mitigate all risks, other experts noted that there has always been an 'arms race' between data controllers and hackers in developing greater sophistication, which continues to be the case at present. Indeed, a number of the stakeholders highlighted new technologies that could be used to support security-by-design to mitigate these risks in infrastructure.

Box 8: Examples of potential security solutions driven by new technologies

- Machine learning and AI may increasingly be valuable to detect tampered or invalid data, which can give additional protection to systems from attacks based on data manipulation. For example, machine learning algorithms can automatically learn relationships between various parameters – such as the pH and chlorine levels of water – and warn if parameters deviate from expected values.[63]
- Distributed ledger technology such as blockchain may also be used to develop new tools for cybersecurity. This technology involves a decentralised system for storing digital data, which eliminates the need for intermediaries and may protect against the risks of data security breaches when data is held centrally. The potential of this technology has been recognised in previous Government studies: "*The opportunity is for distributed ledger technologies to provide much greater security for data than is available in current databases, but this is not a given. This is another area where much research and development is needed as part of the development of the technology*"[64]
  - Specific applications may include uses to "*prevent fraudulent activities through consensus mechanisms, and detect data tampering based on its underlying characteristics of immutability, transparency, auditability, data encryption & operational resilience (including no single point of failure)*".[65]
  - In the infrastructure sectors, blockchain has been considered or trialled for the purpose of facilitating secure transactions for small generators[66] or to deliver secure IoT connectivity.[67] The US is currently implementing blockchain technology to develop a cyber resilient energy transmission network.[68]

  Stakeholders and experts note that future applications of distributed ledger technology remain uncertain and it should not be expected to be a panacea for all cybersecurity issues.[69] Nevertheless, it is increasingly recognised as a valuable technology for controlling secure access to shared databases, especially where a history of transactions or interactions is needed, which may apply to organisational and market data across the infrastructure sectors.

[62] Article 25 of the GDPR requires a controller to implement the principles.

[63] Water Online (https://www.wateronline.com/doc/industrial-internet-of-things-iot-identifying-the-vulnerabilities-of-field-devices-0001, accessed October 2017)

[64] UK Government Chief Scientific Adviser (2016), 'Distributed Ledger Technology: beyond block chain'.

[65] Deloitte (2017), 'Blockchain & Cyber Security: Let's Discuss'

[66] Electron (http://www.electron.org.uk/blog.html, accessed October 2017)

[67] Monitor @ Deloitte (2016), 'How blockchain can impact the telecommunications industry and its relevance to the C-Suite'

[68] Guardtime (https://guardtime.com/blog/us-department-of-energy-contracts-guardtime-pnnl-siemens-and-industry-partners-to-develop-blockchain-cybersecurity-technology-for-distributed-energy-resources, accessed October 2017)

[69] See for example Deloitte (2017), 'Blockchain & Cyber Security: Let's Discuss'

Summary

As infrastructure systems and data sources become more connected, there is a material risk of increased vulnerability to data security breaches and system resilience issues. Therefore, any measures to promote data sharing should ensure that organisations have sufficient tools and incentives to pursue adequate security solutions to mitigate these risks.

This may involve setting clear standards around minimum requirements (for example regarding firewalls, intrusion detection systems, encryption and access authorisation) and promoting a collaborative and coordinated approach across organisations and sectors, as opposed to fragmented or ad hoc approaches that increase the risk of potential security lapses. Information Sharing and Analysis Centres (ISACs), such as the UK's Cyber Security Information Sharing Partnership, can support this. By collecting and disseminating information on security threats and hazards amongst members, ISACs enable collaboration on common cyber security issues.[70] This type of approach, with a strong commitment and participation by infrastructure players, appears vital in the light of evolving cyber threats as data and technologies play an increasing role.

3.4 Cultural barriers

Cultural norms, attitudes and habits can be an important barrier to data sharing, which may exacerbate other types of barriers. Stakeholders noted that several organisations have a relatively closed mentality that opposes sharing data on principle, privileging access on a 'need to know' basis only even if the data is neither personal nor sensitive. Many infrastructure players are large organisations, with long-established cultures and a deep-rooted focus on engineering; stakeholders noted this meant they may lack commitment to exploring potential benefits offered by sharing network or asset data and developing new uses of data based on digital technologies. For example, a recent study found that "*the transport industry has a particularly conservative and siloed culture in both the private and public sector*".[71]

Organisational culture may result in siloed data, reflecting that different business units operate in isolation without considering the scope for data sharing. This might create barriers in two directions. First, data holders might not share data as they may not recognise its relevance elsewhere. Second, those without data may not request data to be shared as they are not aware of its existence. Across the sectors considered, stakeholders have suggested there is a tendency to focus only on individual data requirements, rather than establishing a wider data sharing culture.

Box 9: Data silos in the transport industry

- In the transport sector overall, there is evidence from existing studies of siloed internal structures and a lack of holistic thinking across modes of transport.[72]
- This is illustrated by the rail industry. Some bodies such as the Rail Delivery Group make some types of data freely accessible, such as train arrival times and expected delays. However, a recent study based on a wide spectrum of rail industry stakeholder views found evidence of data being fragmented and siloed.[73] As a result, many technological advances are said to have been missed as the pool of people with the data processing knowledge is limited.
- Another recent study looking at smart cities finds that barriers include geographical silos, where information only exists at regional level across different local authorities and operators, and modal silos, including in relation to automated cycle count data, urban traffic management and historic ticketing data.[74]

[70] ENISA (2016), 'Report on Cyber Security Information Sharing in the Energy Sector'

[71] Ricardo Energy and Environment for the Department for Transport (2017), 'Scoping Study into Deriving Transport Benefits from Big Data and the Internet of Things in Smart Cities'

[72] Transport Systems Catapult (2017), 'The case for Government involvement to incentivise data sharing in the UK Intelligent Mobility Sector'

[73] Hacktrain (2016), 'B.A.R.R.I.E.R.S Report'

[74] Ricardo Energy and Environment for the Department for Transport (2017), 'Scoping Study into Deriving Transport Benefits from Big Data and the Internet of Things in Smart Cities'

A culture that opposes data sharing may be sustained by underlying trust issues and reputational concerns. Data holders may be concerned that shared data could have an adverse impact on their organisation; this could be due to data quality issues, different intended uses of the data, or data misuse by third parties. This barrier has been found to apply to some public bodies such as local authorities, where fear of scrutiny to the opening up of data has deterred data sharing.[75] Stakeholders suggest that an improved understanding of new third-party business models, use cases and benefits should help in this respect, for example through initiatives such as 'hackathons'.

Trust and reputation concerns can also create make customers reluctant to share data with organisations. A culture of suspicion and a lack of trust in infrastructure organisations could stem from a poor reputation of these organisations from the perspective of customers.

Box 10: Trust and reputation issues in the energy industry

- The reputation of the energy sector has suffered in recent years and in 2016 it was reportedly less trusted than other sectors including technology, food and beverages, consumer packaged goods, telecommunications, pharmaceuticals and automotive.[76]
- Perceptions of trustworthiness and service quality may impact consumers' willingness to share data. An estimated 43% of consumers see no service improvement as a result of sharing data with utility companies and around one in four felt that their data was used to extract more money from them.[77]

Where data sharing today is limited by cultural barriers, the curation of data within organisations may suffer as the positive value of sharing data is not fully understood. As a result, investments to 'futureproof' datasets and make them accessible for later use are not made. Stakeholders involved in infrastructure construction projects commented that data archiving does not typically give consideration to future use by other parties, such as asset managers. Similarly, tender documents and contracts may give limited emphasis to data, with stakeholders suggesting that for many organisations it is not typical to consider potential future efficiencies that could be driven by data over the asset's life. Where this is the case, infrastructure operators or managers may face higher costs through duplication of data collection, or by having to resolve issues related to incomplete, incompatible or low quality data sets.

Box 11: Costs of duplication in infrastructure construction

- Stakeholders noted that valuable data collected as part of the construction process is often difficult or impossible to access subsequently by asset managers or operators, leading to duplication and inefficiency. A recent study states that "*technical standards have a vital part to play in ensuring that technologies used in infrastructure fulfil basic standards in terms of interoperability and performance – irrespective of the supplier. They also go some way in minimising the tendency to re-invent the wheel every time a new project is rolled out*."[78]
- Building Information Modelling (BIM) standards aim to address this in the long run, by providing asset data that can be accessed by different parties throughout the lifecycle of the asset.[79]

Stakeholders noted that cultural barriers can affect public and private sector organisations alike. At a central Government level, much public sector data has now been published, including infrastructure data on flood risk and public transport. Nevertheless, barriers in the public sector do remain in particular at the local level, as highlighted in

[75] DotEcon (March 2015), 'Independent evaluation of the OFT's 2006 market study into the Commercial Use of Public Information (CUPI)'
[76] Edelman (2016), 'Trust Barometer 2016: Building Trust in the Energy Market'
[77] Fujitsu (2013), 'Power to the people: data security and trust in the utilities sector'
[78] Pinsent Masons (2017), 'The evolution of Infratech: How technology is shaping the future of infrastructure'
[79] HM Government (2015), 'Digital Built Britain: Level 3 Building Information Modelling – Strategic Plan'

stakeholder discussions and recent reports,[80] including a study for the DfT which finds evidence of a silos-driven organisational culture where Local Authorities focus only on activity within their own boundaries.[81]

Summary

Siloed mentalities and limited attempts to curate data have a profound impact on data sharing in infrastructure. Stakeholder discussions have suggested that this is partly due to a limited understanding and appreciation of the potential benefits. In addition, the deep-rooted nature of organisational cultures in infrastructure sectors and the multitude of organisations that may need to coordinate to share data further deter action.

3.5 Commercial barriers

Stakeholders have noted that commercial barriers can further dampen incentives to share data between organisations. The costs and commercial risk of sharing data – perceived or actual – may be high, meaning that significant potential benefits may be needed to justify investing in data sharing. Coupled with a low-awareness of the 'return' from sharing data, this may add to the effects of cultural barriers.

This barrier also touches on the debate over what constitutes 'data': some stakeholders have noted that a number of datasets in the infrastructure pertaining to operations could be better thought of as Intellectual Property, for example in relation to major infrastructure construction projects. The way this data has been collected, organised and stored could contain institutional knowledge, which may be commercially sensitive. Related to this, previous studies have attempted to distinguish 'unrefined' or raw data from 'refined' data which has been manipulated or combined with other datasets in order to add value.[82]

The direct costs of sharing data could involve negotiating licensing agreements, investing in or upgrading IT infrastructure and data sharing platforms, and the ongoing management of the data sharing process. These costs may be substantial in the infrastructure sectors. For example, local authorities have reportedly avoided publishing data from traffic monitoring systems due to the costs associated with data extraction and cleaning, whilst operating under tight fiscal constraints.[83]

Government may be able to intervene to provide funding for public bodies to release more data – as in the case of the Environment Agency's flood risk data, for example[84] – but private sector organisations will need to justify this expenditure with an expected return on investment. In the regulated industries, stakeholders mentioned some uncertainty about whether IT investments to support data sharing could be fully recouped, though there was some consensus that regulatory frameworks are moving to place greater emphasis on innovation which would allow for certain investments around data and data sharing to be counted against the regulated asset base.

Box 12: Costs of sharing smart meter data

- The roll-out of smart meters and sharing of their data nationwide came with limited commercial incentive for private energy providers. As a result, and in recognition of the sizeable long-term benefits that are estimated to be realised through the

[80] DotEcon for the CMA (2015), 'Independent evaluation of the OFT's 2006 market study into the Commercial Use of Public Information (CUPI)'
[81] Ricardo Energy and Environment for the Department for Transport (2017), 'Scoping Study into Deriving Transport Benefits from Big Data and the Internet of Things in Smart Cities'
[82] Select Committee on Communities and Local Government Committee (2008, https://publications.parliament.uk/pa/cm200708/cmselect/cmcomloc/268/268we05.htm, accessed October 2017)
[83] Transport Systems Catapult (2017), 'The case for Government involvement to incentivise data sharing in the UK Intelligent Mobility Sector'
[84] Government Digital Service, (2014, https://data.blog.gov.uk/2014/06/26/funding-agreed-for-important-new-open-data-projects/)

widespread use of smart meters, the Department of Business, Energy and Industrial Strategy (BEIS, formerly DECC) established Smart DCC.

- The data and communications infrastructure of Smart DCC is in place to enable the secure sharing of smart meter data with authorised users. Smart DCC's costs are projected to exceed £400m by 2020.[85]

Other more specific costs or risks may depend on the type of data involved. For example, stakeholders noted that telecom operators may be reluctant to share information about buried assets due to the risk of copper theft.

In addition to direct costs, a possible indirect cost of sharing data could be perceived loss of competitive advantage. Proprietary data could be seen valuable to the organisation internally, particularly where the boundaries between data and knowledge are not clearly defined, resulting in fears that any form of data sharing could erode an organisation's competitive advantage or benefit its competitors. However, there may be ways to mitigate commercial concerns to enable data to be shared.

Box 13: Ways of sharing commercially sensitive data

- On Elgin's roadworks data portal, planned works in the near-term, around 12-18 weeks, are publicly available, while longer-term plans, which are considered commercially sensitive, are shared only on a restricted version of the platform with authorised users.
- In another example, oil and gas operators were able to share commercially sensitive performance data on a data platform which displayed information with a neutral party in anonymised form. This allowed benchmarking to help identify potential improvements in operations and maintenance. [86]

Given the potential costs involved, a sufficient return on investments in data sharing may depend on an industry-wide commitment. Otherwise, 'free-riders' will benefit without sharing data. For instance, high-performing firms may be reluctant to share data that allows other organisations to improve their performance – for example, this has been previously identified as an issue for fleet operators in the transport and logistics sector[87] – while low-performing firms may equally be reluctant to share data that could attract higher-performing entrants to the market. Stakeholders have expressed the view that procurement models in the infrastructure sectors discourage data sharing for such reasons – for example when bidding for rail franchises, incumbents have an interest in retaining as much market and operational data as possible that may give useful insights when preparing bids at the re-tender stage. Similarly in construction, a recent study has found that "*there's often little commercial incentive to share knowledge – whatever the wider benefits might be*", as data could reveal knowledge gained during a project, which is seen as a source of competitive edge.[88]

Summary

Examples of commercial barriers are seen to exist in the infrastructure sectors. However, workarounds have been demonstrated with positive results, while the removal of other barriers will improve the business case for sharing data and lessen the impact of commercial barriers. In particular, a cultural shift toward greater awareness and understanding of the potential long-term value of data sharing, combined with a collaborative approach to exploring the potential of data and new technologies, could be effective in reducing commercial barriers.

[85] Smart DCC (https://www.smartdcc.co.uk/about-dcc/business-plan/, accessed October 2017)
[86] Wood Group Tech tool, cited in RAEng (2015), 'Connecting data: driving productivity and innovation'
[87] McKinsey (2013), 'Open data: Unlocking innovation and performance with liquid information'
[88] Pinsent Mason (2017), 'The evolution of Infratech: How technology is shaping the future of infrastructure'

3.6 Technical barriers

Technical barriers to data sharing may manifest in the form of inadequate data availability, quality and interoperability. Non-standardised datasets with data quality issues restrict interoperability, with some stakeholders believing this is a pervasive problem affecting various private and public sector infrastructure players who have not adopted basic core data principles.

These issues appear to be closely related to other barriers: the previous sections suggest that cases of inadequate data formats and standards in both the public and private sectors may reflect a culture that places little emphasis on the importance of curating data, as well as a lack of commercial incentive to make necessary investments to overcome technical issues. Discussions with stakeholders in the rail, energy and construction sectors have suggested substantial data still exists in analogue form or in outdated IT systems, making it difficult to access or extract.

The more specific issue of standardisation of data formats, and the governance arrangements around these, have been identified as important by a number of stakeholders in different contexts. While there are generally benefits of common data standards to facilitate data sharing, setting these standards would benefit from considering the level of maturity of the relevant technologies and intended data uses. In the early stages where data is collected and manipulated, the data curators and users will require some degree of flexibility on standards. Later on, more extensive common standards become more important as common use and re-use increases.

Box 14: Data standards in construction

- BIM level 2 provides for a common data environment with set technical standards that facilitate data sharing and collaborative working.
- However, while BIM level 2 is now mandated for Government procured projects, much of the industry still operates at level 1 with limited data sharing.[89] BIM level 2 maturity is expected to increase over time and there are plans to develop BIM level 3 standards by 2025, which are expected to be developed to allow data sharing across a wider range of players and throughout the lifecycle of assets.[90]

These issues are well-recognised, and there are a number of initiatives in the infrastructure sectors to provide common data standards and platforms have been launched to address technical barriers and aid the development of data sharing, each in response to specific needs, objectives or use cases.

Box 15: Recent initiatives for common data standards in the infrastructure sectors

- The British Standards Institute has developed four sets of standards so far to aid with the development of smart cities, including guides to establishing models for data interoperability and decision-making frameworks for data sharing.[91] This includes common definitions for data structures and types, and high-level guidance on establishing a data sharing culture, identifying benefits of data sharing, managing the anonymization of sensitive data, and dealing with privacy and security issues.
- The global association of mobile operators, GSMA, has launched an 'IoT Big Data API Directory', which provides harmonised data sets from a number of sources worldwide. This approach is designed to support a common approach to data sharing that will support IoT development and applications such as smart cities. Operators such as China Mobile, Orange and Telefónica are already using this to share harmonised IoT data.

[89] Pinsent Mason (2017), 'The evolution of Infratech: How technology is shaping the future of infrastructure'

[90] HM Government (2015), 'Digital Built Britain: Level 3 Building Information Modelling – Strategic Plan'

[91] Catapult (http://futurecities.catapult.org.uk/project/cities-standards-institute/, accessed October 2017))

- In the energy sector, a data exchange platform between national and regional grid operators at European level is being established to stimulate innovation and the development of new digital tools to manage electricity flows, under a common framework.[92]
- In the water sector, water companies have been working to bring business customer data up to the minimum standards set by the Market Operator Services Limited,[93] to improve data quality.[94]

Efforts to establish common data standards are ongoing in many instances and some stakeholders suggested that greater certainty could be provided, quicker. For example, in the energy sector there do not yet seem to be agreed standards for smart meter data that is provided to customers in machine readable format upon request (as per GDPR rules). Customers may opt to share this data with other suppliers or third parties, but a lack of standardisation risks being a barrier to effective sharing. In this area, stakeholders noted the 'My Energy Data' work being carried out at European level to explore "*a possible industrial initiative on a common format for energy data interchange*",[95] but there appears to be some remaining uncertainty over progress and implementation – as compared for example to the 'Green Button' initiative in the US which already provides a standardised way for customers to get their energy data.

Summary

Technical barriers in the form of inadequate data standards appear to be a significant barrier in selected cases, though important steps are being taken to provide overarching frameworks and standards as part of BIM and sector-specific initiatives. While further work may be required, technical barriers are generally more likely to be reduced if organisational culture shifts to place greater emphasis on the importance of data sharing, and if commercial barriers are reduced to allow investment in addressing technical issues

[92] Euractiv (2017), 'European power grid operators cook up 'App Store' for smart grids'
[93] The Market Operator Services Limited (MOSL) is a private company funded by water company members that is responsible for the effective and efficient operation of the water retail market.
[94] Ofwat (2017), 'Unlocking the value in customer data: a report for water companies in England and Wales'
[95] European Smart Grids Task Force (2016), 'Expert Group 1 – Standards and Interoperability; My Energy Data'

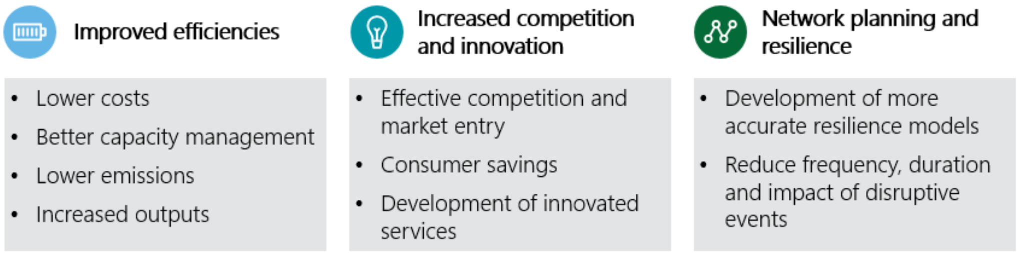
4 The potential benefits of greater data sharing

Addressing the barriers to data sharing in the infrastructure sectors through targeted, balanced measures can yield significant benefits to the UK economy, as long as safeguards are in place to manage potential risks. Areas of impact include efficiency savings, enhanced competition, innovations in products and services and greater system-wide resilience and capacity.

4.1 Overview of benefits

The economic benefits from sharing data are well-known. Data can facilitate efficiencies, competition, innovation, resilience and better use of networks, which in turn can generate jobs, support exports and reduce harmful emissions.

Figure 5: Benefits of data sharing



Improved efficiencies

- Lower costs
- Better capacity management
- Lower emissions
- Increased outputs

Increased competition and innovation

- Effective competition and market entry
- Consumer savings
- Development of innovated services

Network planning and resilience

- Development of more accurate resilience models
- Reduce frequency, duration and impact of disruptive events

Source: Deloitte

Given the size and breadth of the infrastructure sector, the potential impact of these benefits is likely to be significant, reaching across the economy and affecting consumers and organisations alike, but there are also risks from increased sharing.

4.2 Improved efficiency

Increasing volumes of infrastructure data amplify the potential for infrastructure owners and operators to improve their understanding of their customers, assets and networks. This understanding can then be used to support the development of more efficient solutions in terms of lower costs, better capacity management and lower emissions. For example, potential cost savings from use of BIM level 2 standards in construction – which facilitate data sharing across the supply chain in a common environment – are reportedly in the region of 20-30%.[96]

Increased efficiency can take the form of improved matching of demand and supply. For example, in the energy sector, Government and wider stakeholders have recognised that data sharing may be crucial to enable demand-side response (DSR), where consumer demand can adjust to changes in price, contributing to overall system efficiency.

[96] Building Information Modelling (BIM) Task Group (http://www.bimtaskgroup.org/bim-faqs/, accessed October 2017)

Box 16: Data sharing and electricity network efficiency

- DSR has been used by Ofgem since 2015 as part of their flexibility programme for commercial and industrial electricity customers. A 2016 survey by Ofgem suggested that up to 3GW of untapped demand reduction flexibility potential in this customers base could be realised using effective DSR,[97] while among residential customers DSR has yet to be widely implemented and adopted.
- In 2016, it was estimated that flexibility technologies such as DSR and batteries could give net benefits of £1.4-£2.4 billion per year by 2030, though stakeholders note that the rapid rate of technological progress may meant that benefits are achievable much faster.[98] However, widespread uptake of DSR remains highly uncertain as it relies on consumer willingness to share data with and potentially give up control of their energy supply to suppliers and other third parties. A recent study suggests there may be significant costs in "*driving behavioural change in consumers, marketing campaigns for acceptance, contract design, incentive structures to encourage adoption*".[99]
- Other solutions, where consumers agree to share data and give consent to suppliers to control aspects of their energy usage, could include distributed storage, for example where smart charging of electric car batteries helps to balance the grid. This technology is being introduced by Ovo Energy in the UK.[100]

Data sharing could also increase efficiency by driving solutions that increase capacity. Sharing data such as scheduling or operations data can raise awareness of spare capacity across networks or potential alternative flow patterns and routes. This can lead to reduced congestion and overcrowding. For example, China has merged data from a major ride-hailing app with data from smart traffic signals to optimise traffic light changes in real time and ease congestion by an estimated 11%.[101]

Stakeholders noted that the 'Digital Railway' offers benefits such as improved performance and increased capacity through implementation of digital signalling, data links between trackside equipment and trains, data-drive traffic management systems and driver dashboards. However, this has yet to be implemented on a large scale in the UK.

Box 17: Data sharing and rail network capacity

- Rail capacity improvements may be crucial in meeting growing passenger volumes. Data-driven solutions such as digital signalling could enable trains to run closer together while preserving safety. On some routes, capacity increases in the order of 40% may be achievable, at a lower cost than through conventional approaches.[102]
- Currently some collaboration built on data sharing is taking place in this space, primarily in the form of small-scale trials.[103] Stakeholders noted that other countries, such as Japan, have already implemented advanced solutions.
- A key barrier in this area appears to be commercial, as the relevant technologies require substantial investment. Some stakeholders noted that individual organisations may have relatively weak incentives to pursue innovation due to franchising and contract durations that favour a short-term mentality.

More integrated and comprehensive datasets of infrastructure sectors could also help to give regulators a holistic overview of markets to inform any interventions. Machine learning is currently used by the Financial Conduct Authority

[97] Ofgem (2016), 'Industrial & Commercial demand-side response in GB: barriers and potential'

[98] Carbon Trust and Imperial College London (2016), 'An analysis of electricity system flexibility in Great Britain'

[99] Carbon Trust and Imperial College London (2016), 'An analysis of electricity system flexibility in Great Britain'

[100] Ovo Energy (2017), 'OVO Energy launches EV Everywhere, offering energy customers free membership to the largest nationwide electric vehicle charging network'.

[101] Venture Beat (https://venturebeat.com/2017/05/05/how-chinas-meshing-ride-sharing-data-with-smart-traffic-signals-to-ease-road-congestion/, accessed October 2017)

[102] Network Rail (2015), 'Wessex Route Study'

[103] Coventry University (http://www.coventry.ac.uk/research/research-directories/research-news/2017/paving-the-way-for-digital-railway-in-the-uk/, accessed October 2017)

to identify groups of consumers who may be at risk of consumer detriment and other regulators could increasingly adopt similar data-driven approaches.

4.3 Increased competition and innovation

Increased data sharing would promote a transparent environment that could be more conducive to effective competition and market entry. This has been recognised by RAND Europe in a study for Digital Catapult: "*Entry barriers such as closed datasets or data dispersed across multiple sources can create additional costs for firms and prevent their entry into new markets*."[104] Stakeholders have reported that a lack of visibility of market data may prevent new market entry in bidding processes in the infrastructure sectors, in particular in relation to rail franchises.

Box 18: Data sharing and competition

Measures to stimulate competition by increasing data sharing and transparency have been used in different sectors.

- In the banking sector, data on the past financial performance of SMEs is typically held by each businesses' current bank and not widely shared, so alternative providers have a disadvantage when assessing creditworthiness.[105] To address this, the Government mandated some institutions to share data with all providers (with the SME's consent).[106]
- In the electricity retail sector, a CMA investigation found that competition would be enhanced by sharing customer data for those who have been on a basic 'standard variable tariff' for over three years. Ofgem is implementing a customer database where data may be easily shared with price comparison websites to find a better deal.[107]
- At the wholesale level, data sharing may again be important to foster competition. An MIT study notes that data platforms or hubs are becoming more important to facilitate a level playing field for new agents, such as aggregators, DSR providers, small renewable generators, batteries and energy control devices: "*timely and non-discriminatory access to data on network conditions and operation and planning decisions, as well as information on network customers, could be an important facilitator for competition*".[108]

Where data sharing stimulates competition, the outcomes may include consumer savings through lower prices and the development of innovative services. The connection between data and innovation is widely recognised; the OECD defines data-driven innovation as "*a disruptive new source of growth that could transform all sectors in the economy*".[109] Increased sharing of data can lead to new opportunities for retailers, operators, planners, researchers, developers and start-ups to innovate. Third parties may devise new applications or new approaches to infrastructure management and service delivers, including using technologies such as AR, VR or AI.

Box 19: Innovative services currently enabled by data

Examples of innovative services currently enabled by shared data exist across the infrastructure sectors.

- In the transport sector, open data made available by TfL has been accessed by over 10,000 developers and powers around 600 apps used by 42% of Londoners,[110] generating annual benefits of up to £130m.[111]

[104] Digital Catapult Centre

[105] HM Treasury (https://www.gov.uk/Government/consultations/competition-in-banking-improving-access-to-sme-credit-data/competition-in-banking-improving-access-to-sme-credit-data, accessed October 2017)

[106] HM Treasury (https://www.gov.uk/Government/news/boost-for-small-businesses-seeking-finance-thanks-to-Government-data-sharing-scheme, accessed October 2017)

[107] Ofgem (https://www.ofgem.gov.uk/consumers/household-gas-and-electricity-guide/how-switch-energy-supplier-and-shop-better-deal/ofgem-energy-customer-database-service, accessed October 2017)

[108] MIT Energy Initiative (2016), 'Utility of the Future'

[109] OECD (2015), 'Data-Driven Innovation: Big Data for Growth and Well-Being'.

[110] TfL, Open Data Policy (https://tfl.gov.uk/info-for/open-data-users/open-data-policy, accessed October 2017)

[111] TfL (https://tfl.gov.uk/info-for/media/press-releases/2017/october/tfl-s-free-open-data-boosts-london-s-economy, accessed October 2017)

- In the energy sector, smart meter data is used to power new services, such as dashboards, usage feedback and switching services. For example, Labrador allows users to share their data and automatically switch to the cheapest suppliers as tariffs change over time, or receive tailored recommendations.[112]
- In the flood risk management sector, several flood alert apps have been built by third parties based on open data released by the Environment Agency and other data made available by third parties, such as Flood Network. Other types of apps include one that uses this data in combination with other datasets to predict the location of road accidents caused by flooding.[113] This could help reduce flood damage costs, which are estimated to exceed £1.1 billion annually in the UK.[114]

In transport, stakeholders and third-party evidence[115] have indicated that further consumer-facing innovation might be achievable from greater data sharing. The rail industry is reportedly working on giving passengers better information about train location by linking datasets;[116] but stakeholders mentioned that the industry is 'behind' on data sharing, which is corroborated by some existing third-party studies.[117] The Bus Services Bill is expected to compel bus operators to share data that can power innovative transport apps.[118] Further innovation could be possible through sharing smart ticket data, though privacy barriers and concerns may limit this.

In energy, new technologies such as IoT and automation are creating scope for innovation. Government estimates suggest that smart meters will save consumers £47 on their annual bills by 2030 through increased awareness of usage, representing potential benefits of up £1.3 billion respectively.[119] However, studies indicate that the extent of savings realised will vary depending on the services and apps available to consumers to manage and understand their usage: granular real-time feedback from smart appliances could eventually have the greatest impact.[120]

Stakeholders suggest that achieving these benefits may rely on third-party innovators accessing shared data. Currently, researchers, SMEs and entrepreneurs are reported to face difficulties in accessing data, though some initiatives are taking place to alleviate this. A five year project funded by the Engineering and Physical Sciences Research Council aims "*to develop a Smart Meter Research Portal (SMRP) to provide vital access to energy data for the UK research community*".[121]

Box 20: Estonia case study – innovation in the energy sector

Estonia has a data sharing platform (Estfeed[122]) that gives energy suppliers and third-parties secure and controlled access to consumption data from smart meters, as well as other data such as weather data and electricity prices. Customers can use the portal to view their data, manage their consent options for each party requesting access to their data. Meanwhile, a testing and development environment is provided to third-parties to create apps for consumers using the data.

So far, the services offered on this platform include apps to:[123]

- Connect small electricity producers and consumers, potentially leading to better matching of supply and demand at the local level;
- Help users calculate potential savings from installing different types of renewable energy solutions (such as solar panels);

[112] The Labrador (https://www.thelabrador.co.uk/about/, accessed October 2017)
[113] ODI (https://theodi.org/ea-going-open-how-the-data-is-used, accessed October 2016)
[114] House of Commons Library (2017), 'Flood risk management and funding'.
[115] Hacktrain (2016), 'B.A.R.R.I.E.R.S Report'; Transport Systems Catapult (2017), 'The case for Government involvement to incentivise data sharing in the UK Intelligent Mobility Sector'
[116] Rail Delivery Group (2016), 'Our Customers Our People: A Railway for the Digital Age'
[117] Hacktrain (2016), 'B.A.R.R.I.E.R.S Report'
[118] House of Commons Transport Committee (2016), 'Bus Services Bill, Eighth Repot of Session 2016-17'
[119] Ofgem (https://www.ofgem.gov.uk/publications-and-updates/infographic-energy-network, accessed October 2017); Deloitte analysis
[120] Mission:Data Coalition (2016), 'Got Data? The Value of Energy Data Access to Consumers'
[121] UCL Energy Institute (https://www.ucl.ac.uk/bartlett/energy/smart-meter-research-portal-smrp, accessed October 2017)
[122] Elering (https://elering.ee/en/smart-grid-development, accessed October 2017)
[123] Elering (2016), 'Smart Grid data sharing platform Estfeed'.

- Identify households with cases of lost heat and inefficiently oversized circuit breakers so that these can be addressed to improve efficiency; and
- 'Aggregate' consumers to negotiate better tariffs, leading to consumer savings.

The former Prime Minister of Estonia suggested that implementing this technology widely would help Europe save up to €100 billion per year.[124]

The Estfeed platform is built on Estonia's nationwide data sharing platform, as discussed in Chapter 6.

4.4 Infrastructure planning and resilience

Making data available more widely can allow it to be used in resilience models to test system-wide shocks and responses under different scenarios, as well as a variety of other applications. This is recognised in the Government's Digital Built Britain strategy with the objective of "*interoperable sharing of information at key stages*", which "*will also be extended across market sectors to enable the cross asset view of a Smart City or Smart Grid*".[125] Further potential benefits may be achievable through more sophisticated uses of digital twin-type technology, as discussed in Chapter 5.

Sharing more data could allow infrastructure planners and operators to reduce the frequency, duration or impact of disruptive events, either by developing their own solutions or using third-party services built on shared data. Stakeholder views suggest that:

- Sharing data in real-time between different parties may enable new traffic management solutions that minimise the likelihood and impact of delays. In the West Midlands, local authorities and public transport bodies are collaborating to identify which datasets to open up to best manage traffic flow in the region.[126]
- Sharing data can encourage innovative third-party solutions. For instance, visualisations could be used to give an improved situational awareness of infrastructure networks that can help with planning and dealing with faults and disruptive events. Such services are being introduced for electricity grids.[127]
- Sharing data on infrastructure condition could allow predictive maintenance to take place, minimising the risk of faults. For example, IoT sensors on trains might one day generate data that infrastructure managers could analyse to identify potential track maintenance; smart meter data allows water companies to monitor network conditions, quickly identify leaks and target repairs.[128]

Box 21: Japan case study – data sharing to improve water network resilience

- In Japan, Metawater collects sensor data monitor the operational status of water treatment infrastructure, including water levels, quality and pressure, daily inspections and crisis response. Metawater covers around 100 water purification plants and shares this information with local Governments and water management enterprises in order to streamline operation and maintenance activities.
- In the future, Metawater plans more extensive data collection and analysis through a platform integrating water treatment operations and pipelines. This is expected to reduce costs further and improve the quality and safety of water treatment and transport.

Source: Analysys Mason (2014), 'Data-driven innovation in Japan: supporting economic transformation'

[124] Netgroup (http://netgroup.ee/project/estfeed/, accessed October 2017)
[125] HM Government (2015), 'Digital Built Britain. Level 3 Building Information Modelling – Strategic Plan'
[126] Department for Transport (2017), Data sharing and collaboration in the West Midlands (https://dftdigital.blog.gov.uk/2017/04/03/roads-data-sharing-in-west-midlands/, accessed October 2017)
[127] See for example Fast Company (https://www.fastcompany.com/40401189/visualizing-the-electric-grid-in-real-time-and-other-world-changing-ideas-in-energy, accessed October 2017)
[128] IBM (2017), The top 5 industrial IoT use cases (https://www.ibm.com/blogs/internet-of-things/top-5-industrial-iot-use-cases/)

Improvements may also be achieved through better coordination between sectors, where data sharing breaks down silos and facilitates more efficient planning and management of the UK's infrastructure. For example:

- ELGIN is a portal that enables roadworks-related data to be shared, providing transparency over planned roadworks and allowing infrastructure players from different sectors to coordinate maintenance activities. As a result, disruption can be reduced. The benefits from an early version of the portal in 2012 were estimated as £25m per year, quantified in terms of efficiency savings and reduced congestion.[129]

- Data from mobile phones may be valuable in combination with data from other sectors; for example Highways England used anonymised mobile phone data to understand travel patterns and use this for transport planning.[130]

Box 22: Singapore case study – data sharing to improve coordination in a 'smart city'

- Singapore's Land Transport Authority (LTA) is building an analytics system which links shared data from various sources, including from the transport sector and telecom sector. This includes data from fare cards, Wi-Fi, CCTV systems and cellular networks.
- This example of data sharing will allow more accurate modelling of commuter flows and planning.
- In the future, private transport data could be integrated into the system.

Source: Singapore Ministry of Communications and Information (2017), 'Towards a smarter, greener, and more inclusive public transport system'

Stakeholders suggest there could be benefits from further data sharing across sectors. Open information on buried assets could support infrastructure planning, or closer coordination between transport operators – who are large buyers of energy – and energy companies could improve planning for future energy demand needs. The latter point is particularly applicable to rail, where Network Rail is the largest non-regulated energy consumer in the UK and has one of the three highest electricity bills in the country.[131]

4.5 Quantifying the size of the prize from greater data sharing

While there have been no dedicated studies on the benefits of data sharing across the full infrastructure sector, the general economic potential of data, including through data-driven technologies and innovation, is well established:

- A study for the European Commission estimates that data contributed €60 billion in direct, indirect and induced impacts to the overall UK economy in 2016.[132]

- The direct economic benefits of UK public sector open data have been estimated as €11.5 billion in 2016,[133] while the total impact could be around three times larger when including indirect and wider benefits.

Analysis of existing studies that consider some of the infrastructure sectors, as set out below, suggests that increased data sharing in the future could lead to annual benefits from data in the order of £15 billion across the UK's infrastructure sector, increasing from current levels of around £8 billion. This may yet underestimate the true potential value, as the full range of benefits from data sharing and future uses of data are not yet known.

[129] Department for Business, Innovation & Skills (2013), 'Market assessment of public sector information'

[130] Arcadis (https://www.arcadis.com/en/global/what-we-do/our-projects/uk/using-mobile-phone-data-for-transport-planning/, accessed October 2017)

[131] CGI Group (2013), Network Rail: Driving the Energy Revolution (https://www.cgi-group.co.uk/case-study/network-rail-driving-the-energy-revolution)

[132] IDC and OpenEvidence for the European Commission (2017), 'European Data Market Study

[133] Capgemini for the European Commission (2015), 'Creating Value Through Open Data'

Over the period 2015-2020, the Centre for Economic and Business Research estimates that big data and IoT will contribute around £19 billion of economic benefits through the telecom sector, £12 billion through energy and utilities and £12 billion through transport and logistics.[134] The majority of the estimated efficiency benefits occur through efficiency gains, though innovation and creation benefits are also captured. Across the entire UK infrastructure sector, this could translate to current annual benefits from data in the order of £8bn.[135]

However, studies have also recognised that increased data sharing could produce economic benefits that exceed current levels. For example, the Transport System Catapult provides an estimate that appears broadly consistent with this, estimating that improved data sharing could lead to incremental benefits through mobility solutions of around £15 billion in value, by 2025.[136]

McKinsey estimates that a more open approach to data across public and private sectors could deliver additional annual benefits of $270 billion globally in transport through infrastructure planning and management, and $150 billion in electricity through optimised investment and operations across generation, transmission and distribution.[137] Across the entire UK infrastructure sector, this scale of impact suggests potential annual benefits in the order of £15bn from increased data sharing.[138]

4.6 Potential risks and trade-offs

Notwithstanding the significant potential benefits from increased data sharing, there may be important risks and trade-offs that any Government policy should consider to avoid harmful or unintended consequences. Any measures to stimulate increased data sharing should be carefully targeted, based on defined objectives and analysis.

Promoting sharing of personal data while respecting privacy

As discussed in section 3.2, a key trade-off exists between encouraging wider sharing and ensuring that consumer preferences are respected. Any sweeping measures that aim to facilitate the sharing of personal data in general might promote use cases that are beneficial to consumers, but would risk allowing greater scope for uses that adversely impact consumers or have an effect on them which they would not reasonably expect and could consider negative.

Within the framework set by the GDPR, consumer consent should be informed and freely given, and as easy to withdraw as to provide. Alternate legal bases to consent (such as legitimate interest) may also be available to support personal data sharing, but depend on ensuring a balance with individuals' rights.

Targeted measures may also be most effective by aiming to educate consumers about the benefits of sharing in the context of key types of data and use cases, as well as providing consumers with clear information about their rights to control how their data is used. This should improve sharing of data in ways that can generate benefits, while minimising the risk of consumers being misled both about the extent of any personal data sharing or its potential to impact them directly, or not, as the case may be.

Implementing wider sharing of data and connected ecosystems while preserving security

As discussed in section 3.3, another trade-off with potentially vast repercussions involves the security implications of increased data sharing. If data is shared more widely and infrastructure systems become connected, it is vital that the

[134] Cebr (2016), 'The Value of Big Data and the Internet of Things to the UK Economy'

[135] Deloitte analysis of Cebr (2016), 'The Value of Big Data and the Internet of Things to the UK Economy', based on average annual benefits estimated for Telecoms, Energy & Utilities and Transport & Logistics, scaled up for the other National Infrastructure Assessment sectors (solid waste management and flood risk management) based on market size.

[136] Transport Systems Catapult (2017), 'The case for Government involvement to incentivise data sharing in the UK Intelligent Mobility Sector'

[137] McKinsey Global Institute (2013), 'Open data: unlocking innovation and performance with liquid innovation'.

[138] Deloitte analysis of McKinsey Global Institute (2013), 'Open data: unlocking innovation and performance with liquid innovation', based on estimates produced for the energy and transport sectors, scaled up for the other four National Infrastructure Assessment sectors (telecoms, water, solid waste management, flood risk management) based on market size.

adequate security measures are simultaneously put in place so as to minimise any increased risk of security issues. Thus, any policy to promote data sharing may face a risk of greater vulnerability, if data sharing is increased without full consideration, implementation and testing of necessary safeguards.

To manage this trade-off it is likely to be important to ensure that suitable cybersecurity expertise and stakeholders are involved in designing and overseeing significant data sharing initiatives, with clear, consistent security requirements being set and enforced. This may require some compromise between pursuing increased data sharing and its potential benefits as quickly as possible, while ensuring that all security issues have been dealt with, including by considering emerging solutions such as blockchain, which may require particularly thorough development and testing.

Maximising the potential benefits while avoiding inefficient investment

As discussed in sections 3.5 and 3.6, there may be significant financial costs involved in increasing data sharing, to invest in new IT infrastructure, data cleaning and extraction, or the development of common data standards, sharing agreements and exchange platforms. If increased data sharing were purely supply-driven, with an objective of making as much data available as possible regardless of the type of data and its potential use cases, inefficient investments could be made in sharing data that is of low value.

Therefore, a focused approach appears necessary, based on a robust analysis of demand for data and the potential benefits achieved in specific cases, as well as supply-side considerations. By aiming to increase data sharing in those cases with the largest achievable benefits, any interventions will minimise the risk of wasteful investment and will also be more likely to overcome cultural and commercial resistance to data sharing.

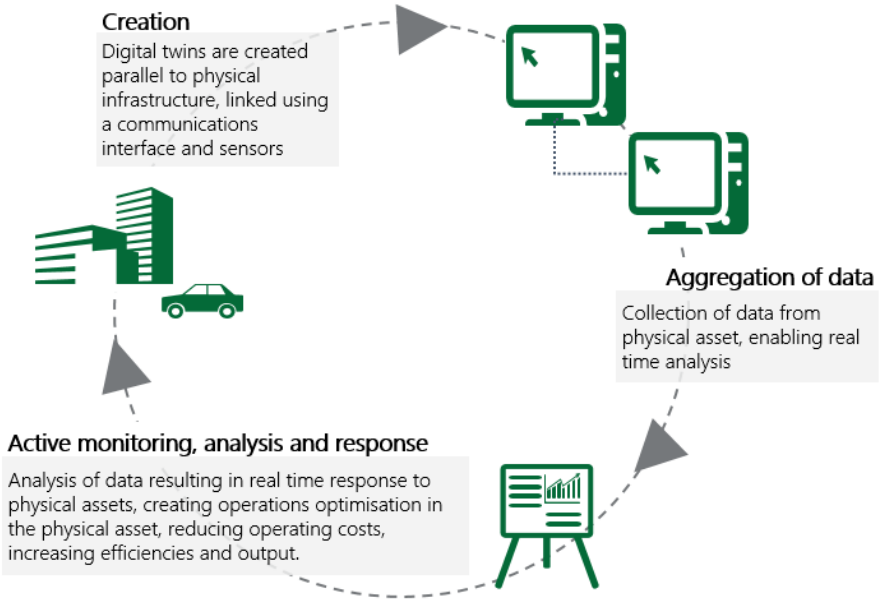
5 Case study: the digital twin

Increased data sharing could allow more advanced digital twin technologies to be used to model the UK's infrastructure in the future. Creating large-scale digital twins or a single national digital twin could generate wide-ranging benefits, facilitating better planning, prototyping and protecting of infrastructure networks, though the significant risks involved may require careful consideration.

5.1 Digital Twin overview

A digital twin is a digital replica of physical assets, processes and systems. The technology is most commonly deployed in the manufacturing sector to test resilience and new processes virtually, though in future it may become possible to create comprehensive digital replicas of UK infrastructure that can be used to analyse usage patterns, conduct scenario planning and test different configurations before deploying changes in the real network.

Figure 6: Digital twin in infrastructure



Creation

Digital twins are created parallel to physical infrastructure, linked using a communications interface and sensors

Aggregation of data

Collection of data from physical asset, enabling real time analysis

Active monitoring, analysis and response

Analysis of data resulting in real time response to physical assets, creating operations optimisation in the physical asset, reducing operating costs, increasing efficiencies and output.

Source: Deloitte

This chapter explores the concepts behind the digital twin and how these may be applied to the infrastructure sectors in the future, as an illustration of the potential benefits driven by data sharing, but also showing how trade-offs and risks discussed in previous chapters may come into play.

5.2 Digital twin applications

The proliferation of data from new sources and development of technologies such as IoT, machine learning, BIM and augmented or virtual reality are enabling organisations to develop more accurate and sophisticated digital models. In some cases this results in a 'digital twin' being created: an accurate digital replica of real-world assets, systems and processes. Unlike other types of models or visualisations, it typically uses real-time or near real-time data from various sources, including IoT sensors, to accurately reflect real-world situations.

There is precedent in the UK and internationally in the creation of digital twins, with most applications taking place on a relatively small scale, though there have been attempts to build digital twins of several assets over large geographic areas. However, successful implementation of larger-scale digital twin projects going forward may depend on effective data sharing between different parties.

Box 23: Examples of digital twin applications and uses

Thus far, digital twins have been primarily used as part of the design and manufacturing process for complex assets such as jet engines. Manufacturers such as General Electric are using digital twins, built from around 100 sensors, to monitor the condition of each engine over time, after it has been built.[139] The digital twin gives an in-depth understanding of each actual engine's use, performance and condition over time, which then informs the development of future designs.

However, larger-scale digital twins have been built at city level in a number of cases:

- Singapore has undertaken a detailed 3D mapping process which may be seen as a digital twin. It will offer a collaborative data platform for consumers, public and private sector organisations and researchers to use.[140]
- Improbable built a full-scale simulation of Manchester to model traffic and population density changes.[141]
- Ordnance Survey is building a smart map of the UK that can be used to optimise plans for 5G rollout.[142]
- In specific sectors, digital twin-type applications are being used or trialled. For example, in the energy sector real-time visualisation software using data from smart meters and operational data have been trialled.[143]

5.3 Potential benefits

Large-scale digital twins could offer a range of benefits. These largely reflect the types of benefits examined in Chapter 5 in terms of improved infrastructure planning, management, efficiency and resilience. Equally, the full range of wider benefits cannot be identified at this stage, as the scope and potential uses of this technology remain somewhat uncertain.

For example, solutions being trialled in the energy sectors indicate that "*collating and visualizing this operational data may provide insights into grid operations, and could minimize outages and cost while increasing safety*".[144] Smart city projects are also working towards creating digital twin-type applications at city level and the British Standards Institute has predicted that "*a single view of city data can highlight improvements and efficiencies and help cities understand how best to improve service delivery*".[145]

As data sharing increases and technologies improve, it may be possible to create digital twins that bring together different assets, systems, processes and networks. This appears to be aligned with the long-term vision for BIM level 3

[139] BBC (http://www.bbc.com/autos/story/20170214-how-jet-engines-are-made, accessed October 2017)

[140] National Research Foundation to the Prime Minister's Office Singapore (https://www.nrf.gov.sg/programmes/virtual-singapore, accessed October 2017)

[141] Wired (http://www.wired.co.uk/article/improbable-quest-to-build-the-matrix, accessed October 2017)

[142] Ordnance Survey (https://www.ordnancesurvey.co.uk/about/news/2016/uk-leading-way-5g-future.html, accessed October 2017)

[143] Pacific Gas and Electric Company (2016), 'Electric Program Investment Charge (EPIC) Final Report'

[144] Pacific Gas and Electric Company (2016), 'Electric Program Investment Charge (EPIC) Final Report'

[145] BSI (2017), 'Smart cities – Guide to establishing a decision-making framework for sharing data and information services'.

set out in the Government's Digital Built Britain strategy, to "*enable the interconnected digital design of different elements in a built environment and will extend BIM into the operation of assets over their lifetime. It will support the accelerated delivery of smart cities, services and grids*".[146] Ultimately, this could culminate in a single national digital twin that captures all of the UK's key infrastructure through a single model, or 'federated models' that connect to one another.

Stakeholders broadly recognised that there are potential benefits from the creation of digital twins or related applications, with exploitation of real-time data and use of visualisations, though there was a mixed response to the concept of a national digital twin. This included some scepticism around the size of potential benefits – which remain very uncertain, as it remains a relatively distant possibility – as well as considerable concerns around the potential costs and risks involved.

5.4 Potential challenges

There appear to be a number of potential challenges and risks associated with creating larger-scale federated or national digital twins encompassing a multitude of infrastructure assets and systems. These provide an illustration of the types of issues discussed in Chapter 3.

- Data security has been raised as a key concern. A national digital twin may require a single organisation to access or act as a gatekeeper to the data, which raises significant data security risks from such quantities of data being held centrally.

- System security is another key issue, given the vulnerabilities already associated with existing industrial control systems, which could increase substantially if different ecosystems became connected to the point of delivering a national digital twin that could influence real-world decisions. This raises questions of accountability, with stakeholder suggestions that individual organisations would be reluctant to take responsibility to deliver this type of project.

- Technical challenges would likely arise through the need to establish IT infrastructure capable of delivering this connected ecosystem and to set harmonised data standards that could make data from different sources interoperable. This would require a clear view of data and output requirements, which would likely be an extremely complex exercise.

- Commercial barriers may be substantial, as stakeholders anticipated large financial costs from addressing these security and technical issues, which infrastructure players would likely be reluctant to undertake. Assuming that IT costs would be funded at least partly by Government, stakeholders expressed a general scepticism over the likelihood of delivering value-for-money for taxpayers from major IT projects.

- Legal issues could arise, as the complexity of logistical and governance arrangements around a digital twin could make issues such as data ownership unclear or contentious, particularly in the case of derived data obtained through algorithmic analysis of the source data.

The nature of these challenges highlights that a collaborative approach across industry and public sector stakeholders would be necessary, with a clear view of uses and benefits to drive the approach to overcoming these challenges.

Summary

A digital twin creates an ecosystem using multiple data sources, which is able to improve efficiencies, reduce silos and enable timely decision making. Without further development of digital twin technology, through a collaborative approach based on data sharing, discussions with stakeholders have suggested there could be valuable insights

[146] Building Information Modelling (BIM) Task Group (http://www.bimtaskgroup.org/, accessed October 2017)

missed and efficiency gains lost. However, the ambition of implementing sophisticated, large-scale digital twins also presents significant practical challenges.

Stakeholder input suggests that any steps to explore this idea would need to build stakeholder buy-in gradually, by first promoting awareness of current digital twin-type applications and potential innovations that may be achievable in the short term, on a relatively small scale. A move toward wider digital twin applications could be informed by trials and proofs-of-concept as these take place internationally as well as in the UK, and justified by a clear assessment of potential benefits and impacts on all stakeholders and consumers.

The ultimate objective of a single digital twin, or a large number of interconnected federated models, presents particular risks, so caution would seem to support initially creating multiple individual digital twins over time and connecting these gradually while the potential security ramifications can be explored and tested further.

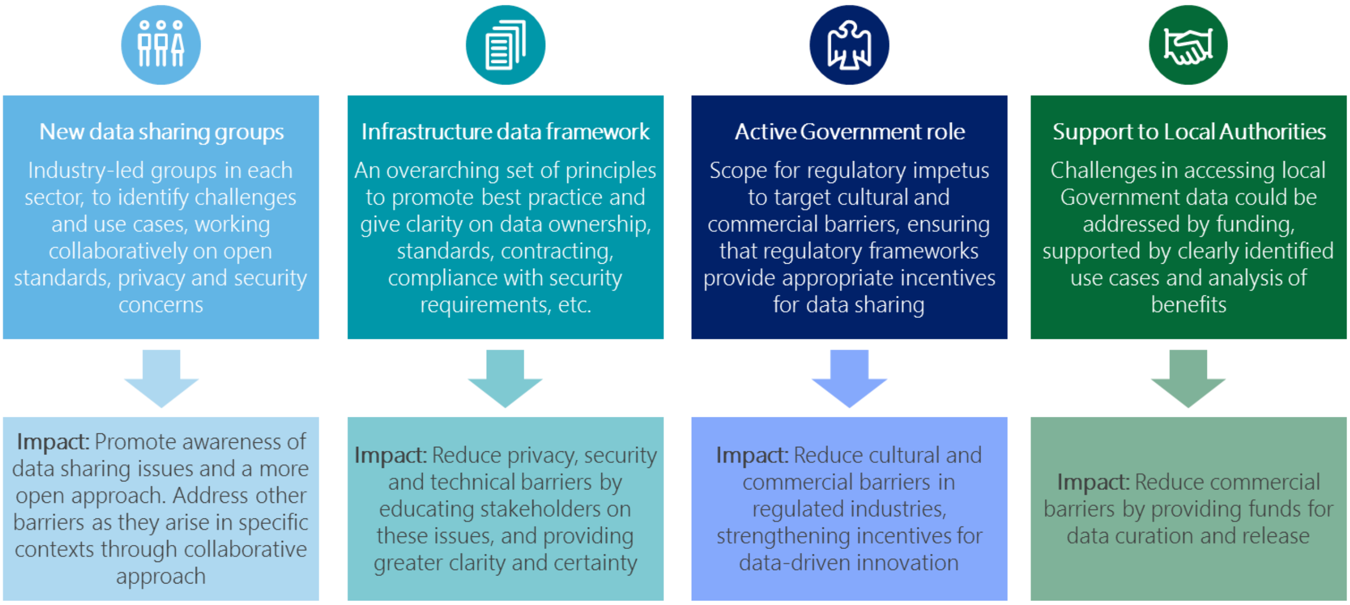
6 Potential remedies

Government can take a key role in working with industry to address many of the identified data barriers.

6.1 Chapter Overview

While industry may be able to work towards solutions to improve data sharing, stakeholders report these may take time to develop and implement and could only represent a partial solution. For this reason, there is a role for Government to address the residual barriers around data sharing. Stakeholder discussions suggest that, due to the large number of stakeholders involved, a clear impetus from central Government or regulators is likely to be important in achieving coordination between organisations.

Figure 7: Summary of potential remedies to promote data sharing



New data sharing groups
Industry-led groups in each sector, to identify challenges and use cases, working collaboratively on open standards, privacy and security concerns

Impact: Promote awareness of data sharing issues and a more open approach. Address other barriers as they arise in specific contexts through collaborative approach

Infrastructure data framework
An overarching set of principles to promote best practice and give clarity on data ownership, standards, contracting, compliance with security requirements, etc.

Impact: Reduce privacy, security and technical barriers by educating stakeholders on these issues, and providing greater clarity and certainty

Active Government role
Scope for regulatory impetus to target cultural and commercial barriers, ensuring that regulatory frameworks provide appropriate incentives for data sharing

Impact: Reduce cultural and commercial barriers in regulated industries, strengthening incentives for data-driven innovation

Support to Local Authorities
Challenges in accessing local Government data could be addressed by funding, supported by clearly identified use cases and analysis of benefits

Impact: Reduce commercial barriers by providing funds for data curation and release

Source: Deloitte

Discussions with stakeholders suggest that any recommendations to Government and industry should be focused and implementable in the short-term. This reflects the sentiment that data sharing must first be improved in individual sectors rather than infrastructure-wide, as well as the fact that some longer-term strategies for data sharing in infrastructure, such as Digital Built Britain, are already in place.

6.2 International precedent for Government supporting data sharing

The UK is recognised as a leading country in open data.[147] Nevertheless, the barriers identified in chapter 3 indicate that there is significant potential to break down remaining data silos within the public sector and to foster a culture of greater data sharing across the private sectors in infrastructure. Examples of steps taken in other countries, including

[147] The World Wide Web Foundation, Open Data Barometer (http://opendatabarometer.org/?_year=2016&indicator=ODB, accessed October 2017)

the creation of new institutions or data exchange platforms, provide relevant precedent when considering possible measures to promote the benefits of data sharing.

In Japan, a Strategic Council for Data Driven Innovation was created in June 2014 by the Ministry of Economy, Trade and Industry. The Council was tasked with encouraging increased sharing and use of data by organisations. It identified key challenges including limited data skills and management principles that led firms to rely exclusively on internal data.[148] The Council's work is focused on areas including fostering collaboration between businesses, improving relationships between consumers and businesses with regard to data sharing, reviewing data protection legislation and cultivating future leaders to lead future innovation activities using data across fields and organisations. [149]

In Singapore, a state-owned enterprise (GovTech) was created in 2016 with the objective of deriving value from data in Singapore. As a relatively new enterprise, there is currently limited evidence about its impact.[150] In the same year, Singapore released MyInfo, an online consent management platform which relies on the national identification system. MyInfo allows citizens to give consent to share data across both public and private entities, choosing which entity receives certain information about them.

In Estonia, data sharing platforms have been central to the country's leadership in implementing e-Government. The objective has been to support data sharing from the bottom-up while centralising activities, common standards and procedures for data exchange.[151] Private and public sector data sharing is built on the secure 'X-Road' network. The network is similar to the UK's Government Secure Intranet, except it may also be used by private sectors such as energy (see Box 20), telecoms and banks. A key feature of the network is that it promotes transparency for consumers, who are easily able to see which X-Road participants hold their data, which participants have the ability to access it, and which participants *have* accessed it.[152] This process is facilitated by blockchain technology, which provides a distributed log of data access for auditing purposes, to monitor who has accessed what information. Similarly to the consent management system in Singapore, this relies on each citizen's national ID.

In Denmark, the Government has embarked on a programme to create a centralised data model and shared platform for distributing Government data, due to be completed in 2018.[153] This aims to address problems including data silos and duplicate information held by different bodies, creating a shared core data set of demographic, geographic and property data. Though this will require an investment of €125 million over seven years, anticipated savings are in the order of €33 million per year for the Government and €66 million per year for the private sector, who will also be able to access the data.[154]

Initiatives are also taking place at city level in Denmark with Copenhagen's City Data Exchange, a new business model that aims to break down data silos.[155] The data exchange integrates data from private companies and public sector open data, with a 'marketplace' that allows data suppliers to find data customers who are interested in their data. This is sustained by an organisation responsible for providing data analytics and support to third-party developers who use the data and to facilitate access to data by businesses and researchers. The City Data Exchange was launched in May 2016 and limited evidence is available thus far regarding its impact. The City of Copenhagen is also a consortium partner in the EU-funded 'SELECT for Cities' project, which aims to develop a data-driven platform for European cities to collaborate in creating, testing and validating new smart city apps and services.

[148] Ministry of Economy, Trade and Industry (http://www.meti.go.jp/english/press/2014/0609_02.html, accessed October 2017)

[149] Ministry of Economy, Trade and Industry (http://www.meti.go.jp/english/press/2014/1105_02.html, accessed October 2017)

[150] Data Futures Partnership New Zealand (2017), 'Exploring different approaches to data sharing'

[151] Data Futures Partnership New Zealand (2017), 'Exploring different approaches to data sharing'

[152] Medium (https://medium.com/sidewalk-talk/how-estonia-became-a-global-model-for-e-Government-c12e5002d818, accessed October 2017)

[153] Data Futures Partnership New Zealand (2017), 'Exploring different approaches to data sharing'

[154] CBO Projects (2015), 'eGovernment: Learning from Denmark and Estonia'

[155] Ricardo Energy and Environment for the Department for Transport (2017), 'Scoping Study into Deriving Transport Benefits from Big Data and the Internet of Things in Smart Cities'

6.3 Potential remedies in the UK

A balance needs to be achieved between tacking cultural, commercial and technical barriers to realise the benefits from data sharing while simultaneously ensuring that security and privacy risks are appropriately addressed. This study identifies key areas where Government can take steps to facilitate market-led solutions to improve data sharing, whilst remaining mindful of the above trade-offs and the need for a measured and targeted approach.

The remedies recognise that the UK is a leader in open data, but that there is evidence suggesting that existing initiatives have not addressed all barriers, particularly at local Government level. In the private sector, stakeholders indicate that existing initiatives such as regulatory sandboxes have had only limited success in promoting a culture shift towards data sharing in infrastructure, indicating that more proactive industry-wide approaches may be required. In the case of commercial and cultural barriers, these may be more difficult to address through an industry-led approach alone; therefore an active role for Government and regulators is also considered.

The suggested measures are intended to complement one another, rather than being seen as alternative options, and should also complement existing initiatives around using procurement to compel data sharing (such as in the Bus Services Bill), ongoing publication of open data by the public and private sector and specific industry activities to share data.

Creating new data sharing industry groups

Industry-led groups in different infrastructure sectors could be facilitated by Government (potentially via regulators, the NIC or other public bodies), to tackle particular challenges around data sharing. These groups could work to ensure that the approach to increasing data sharing is both demand-led – in terms of identifying types of data and sharing arrangements that would be most beneficial for key use cases – and supply-led – in terms of creating greater urgency for data holders to share data where feasible.

It is anticipated that these groups could be involved in:

- Articulating the challenges specific to a sector, and what data and data sharing is required to address them (specifically whether non-personal data can be deployed, thus reducing the regulatory complexity);
- Providing use cases and guidance on the (monetary) value of data, its ROI and the benefits of sharing rather than free riding;
- Considering the implications of the GDPR and developing a common approach to complying with it whilst still sharing data where possible, to foster clarity, certainty and consumer trust;
- Considering the security issues that arise within the sector and working collaboratively with external organisations, such as NCSC and the Cyber Security Information Sharing Partnership (CiSP) towards solutions, such as developing secure gateways to share data;
- Developing harmonised open standards that can be applied across the sector; and
- Promoting the development of open APIs and more data being available as open by default.

Coordination of such an effort could look like Open Banking, the vehicle set up by the Competition and Markets Authority to stimulate competition in the banking sector and encourage innovation made possible by fintech companies. The Open Banking Standard enables third parties to develop mobile and web applications by providing safe access to the data of banking customers, ultimately enabling innovation using data that was limited in the past by the data siloes of legacy banking systems.

The membership of these groups would be jointly decided between industry and Government, but is likely to include sector bodies, a representative sample of suppliers and data users (including SMEs), regulators, consumer groups and the relevant central and devolved departments.

Creating an infrastructure data framework

A key gap identified by many stakeholders was the absence of an overarching set of principles that provided guidance and clarity on issues such as data ownership, what constitutes data, what might be interpreted as personal and non-personal, ensuring security by design, and so forth. While such a framework cannot ever be considered definitive, a common set of principles applicable across the whole sector (which can be customised) can be used as a starting point for subsequent data sharing.

As sector-specific industry groups develop thinking and best practice focused around specific use cases, the framework can build on this by providing overall guidance to be applied across all sectors.

The principles of the framework could cover a range of areas including:

- Best practice guidance for organisations to carry out an internal audit of their data, classifying different types and identifying data that can be shared, either as open data or with restrictions.
- Best practice guidance for data quality and formatting for different categories of data.
- Approaches to specifying contracts that give appropriate emphasis to data requirements, clarity around responsibilities and liabilities related to data, and ensure there is scope for data to be used and re-used.
- Approaches to data anonymization and aggregation so that confidential data may become shareable.
- Steps to deal with grey areas around data ownership, data and IP, personal and non-personal data, etc.
- Appropriate security measures for data sharing in infrastructure, building on the Government's '10 steps guidance' and NIS Directive principles to build awareness and understanding among infrastructure players, setting out more explicitly how best practice in cybersecurity can be achieved in practice by infrastructure organisations.

This recommendation builds on previous work, for example by the British Academy and Royal Society to look at data governance, management and use in the modern economy.[156] Where appropriate, the principles may be supported by examples and templates to promote their application in practice.

This framework would benefit from leadership by a public body with an invested interest in each industry, and would be complementary to the work carried out by the industry-led groups. Inputs should be sought from industry and academia, and facilitated by public bodies such as regulators and NIC.

Active role for Government and regulators in regulated industries

Past experience shows that where Government plays an active role in stimulating data sharing, rapid and significant changes can be brought about. Over recent years, Government open data initiatives have led to large volumes of data being shared across the public sector, with a plethora of innovative use cases and benefits. Without these initiatives, a wealth of data would have remained closed as cultural and commercial barriers prevented a release of datasets that generates wider benefits.

[156] British Academy and Royal Society (2017), 'Data management and use: Governance in the 21st century'

In the regulated infrastructure sectors in particular, a similar misalignment of incentives to data sharing often exists. Though data sharing has been generally recognised by stakeholders to have potentially positive impacts, the commercial incentive to work towards innovative data-driven solutions and make the necessary investments is at times insufficient. A historic focus on engineering-based solutions may mean that a cultural shift may be needed to enable greater focus on data and new technology. Stakeholder discussions suggest that a full reduction of these barriers is unlikely to be achieved through industry-led initiatives alone.

Against this backdrop, there appears to be scope for Government and regulators to a greater regulatory impetus to promote data sharing. Any regulatory-led action, such as adjustments to regulatory frameworks, guidance and specific targets for data sharing, would work in parallel with industry-led data sharing groups and the development of a broader infrastructure data framework, to maximise the likelihood of a pervasive shift towards greater data sharing across stakeholders. Where significant changes are made, appropriate notice or phasing periods can be considered to minimise disruption on the industries.

While recognising the potential benefits of data sharing, any intervention would need to carefully consider the costs involved and how these may be dealt with as part of regulatory frameworks and incentive mechanisms. It may be possible to support an acceleration in investments that enable greater data sharing and the use of associated new technologies, for example through appropriate guidance and rules about costs being factored into regulatory asset bases.

Support to Local Authorities

Some local authorities have made progress in sharing infrastructure data – for example the Thermal Harrow open data initiative helps to identify heat loss from buildings;[157] Data Mill North is the first platform to bring together open data from different sectors of cities and promote sharing and re-use, with over 90 transport datasets available.[158] However, stakeholders and previous studies highlight challenges in accessing data from local Government, particularly transport-related data including traffic monitoring data, cycle counts, road closures and diversions.[159]

The reported reason for this was often a lack of funding to either share the data or maintain its quality and integrity. While recognising fiscal constraints, the return from making this data available more widely as open data is significant. Local authorities could benefit from Government support via funds that local authorities and data users and re-users can bid for to support the opening up of specific datasets for wider use. In addition, communicating clearly the benefits to the local authorities themselves from sharing data would encourage engagement. This could build on previous programmes and local open data projects.[160]

[157] Harrow Council (http://www.harrow.gov.uk/info/100006/environment/1514/thermal_harrow, accessed October 2017)

[158] Leeds City Council (http://www.leeds.gov.uk/opendata/Pages/Data%20Mill%20North.aspx, accessed October 2017)

[159] See for example Ricardo Energy and Environment and the Department for Transport (2017), 'Scoping Study into Deriving Transport Benefits from Big Data and the Internet of Things in Smart Cities'; Transport Systems Catapult (2017), 'The case for Government involvement to incentivise data sharing in the UK Intelligent Mobility Sector'; DotEcon (March 2015), 'Independent evaluation of the OFT's 2006 market study into the Commercial Use of Public Information (CUPI)'.

[160] See for example Local Government Association (https://www.local.gov.uk/our-support/guidance-and-resources/data-and-transparency/making-open-data-work-you-case-studies, accessed October 2017)

Appendix: External stakeholders consulted

Deloitte would like to thank representatives from the following organisations for their contributions to this study.

- Berwin Leighton Paisner
- British Geological Survey
- Cambridge Centre for Smart Infrastructure and Construction
- Elgin
- GCHQ
- GSMA
- Infrastructure Transitions Research Consortium
- Northumbrian Water
- npower
- Open Data Institute
- Ordnance Survey
- Pinsent Masons
- Rail Delivery Group
- Science and Technology Facilities Council
- Smart DCC
- Stagecoach
- TechUK
- Transport for London
- UK Regulators' Network
- A cybersecurity organisation
- A scientific organisation
- Participants at TechUK's Roundtable 'GDPR | Energy & Data Protection in the Connected World' (including consumer groups, energy market organisations, regulators, Government bodies, tech firms)
- Participants at the National Infrastructure Commission's Digital Twin Thinkathon (including from academia and industry)

Deloitte.

Deloitte LLP is a limited liability partnership registered in England and Wales with registered number OC303675 and its registered office at 2 New Street Square, London, EC4A 3BZ, United Kingdom.

Deloitte LLP is the United Kingdom affiliate of Deloitte NWE LLP, a member firm of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"). DTTL and each of its member firms are legally separate and independent entities. DTTL and Deloitte NWE LLP do not provide services to clients. Please see www.deloitte.com/about to learn more about our global network of member firms.