Data for the public good

NATIONAL
INFRASTRUCTURE
COMMISSION

Foreword

Advances in technology have always transformed our lives and indeed whole industries such as banking and retail. In the same way, sensors, cloud computing, artificial intelligence and machine learning can transform the way we use and manage our national infrastructure. Government could spend less, whilst delivering benefits to the consumer: lower bills, improved travel times, and reduced disruption from congestion or maintenance work.

The more information we have about the nation's infrastructure, the better we can understand it. Therefore, data is crucial. Data can improve how our infrastructure is built, managed, and eventually decommissioned, and real-time data can inform how our infrastructure is operated on a second-to-second basis.

However, collecting data alone will not improve the nation's infrastructure. The key is to collect high quality data and use it effectively. One path is to set standards for the format of data, enabling high quality data to be easily shared and understood; much that we take for granted today is only possible because of agreed standards, such as bar codes on merchandise which have enabled the automation of checkout systems.

Sharing data can catalyse innovation and improve services. Transport for London (TfL) has made information on London's transport network available to the public, paving the way for the development of apps like Citymapper, which helps people get about the city safely and expediently. But it is important that when information on national infrastructure is shared, this happens with the appropriate security and privacy arrangements.

Our report sets out clear actions in three areas: collecting the right data; setting standards for data; and sharing that data securely. These actions can improve services whilst saving society billions of pounds.

Bearing in mind our 10-30-year timeframe, the report also sets out a roadmap towards a national digital twin: a digital model of our national infrastructure which will be able both to monitor our infrastructure in real-time, and to simulate the impacts of possible events, for example, a natural disaster, or a new train line. A digital twin could help plan and manage the nation's infrastructure more effectively, and will only be possible through the secure sharing of high quality, standardised data across infrastructure.

As Bill Gates put it in his book *The Road Ahead*, "We always overestimate the change that will occur in the next two years and underestimate the change that will occur in the next ten. Don't let yourself be lulled into inaction."

Andrew Adonis
Chair, National Infrastructure Commission

REALISING THE BENEFITS OF A SMART INFRASTRUCTURE THROUGH DATA SHARING

Population growth, economic growth and climate change are straining our infrastructure.

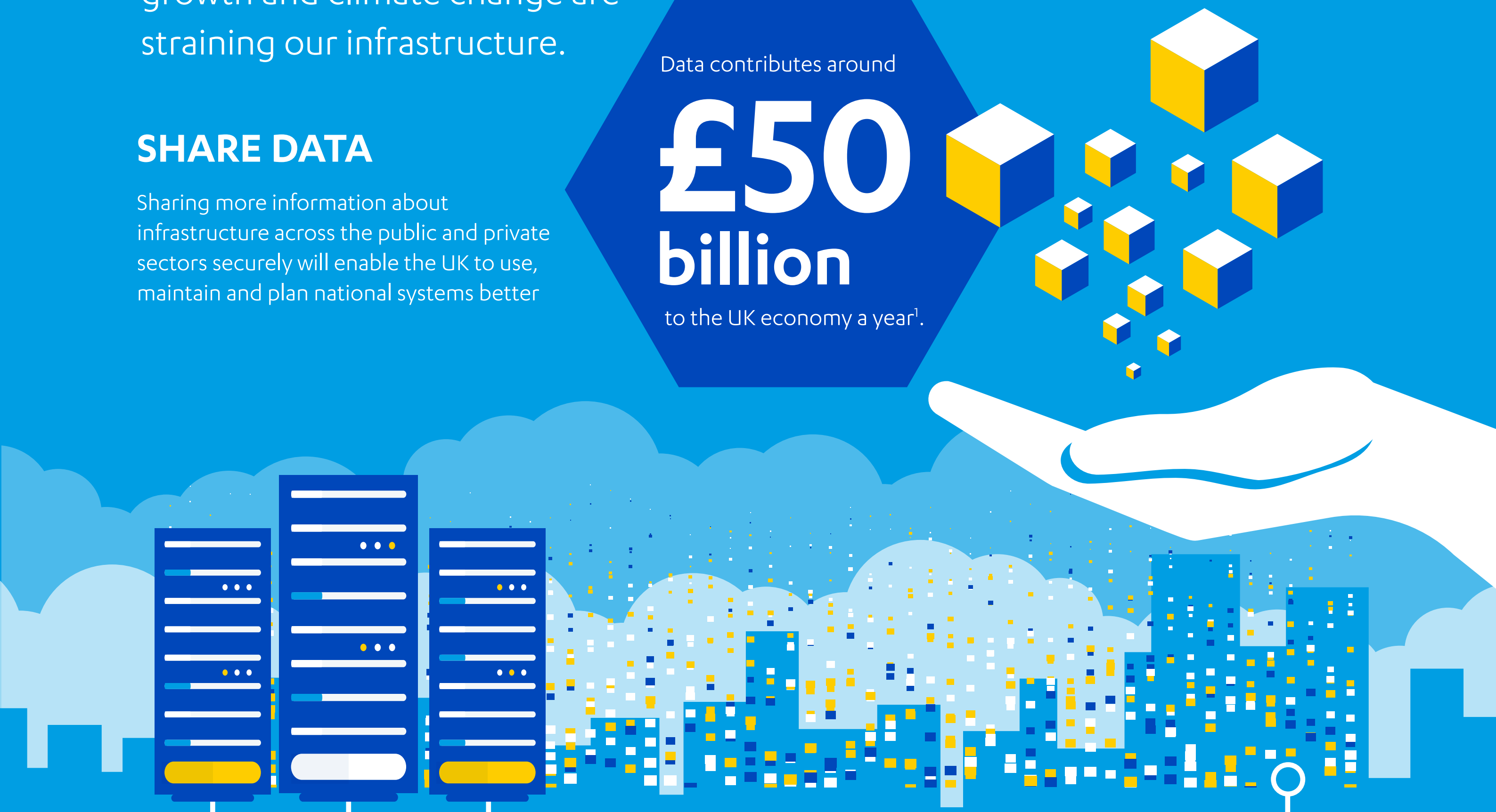
SHARE DATA

Sharing more information about infrastructure across the public and private sectors securely will enable the UK to use, maintain and plan national systems better

Data contributes around

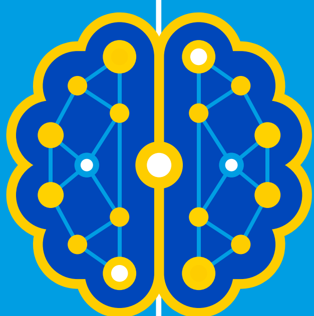
£50 billion

to the UK economy a year[1].



THE TECHNOLOGY IS ALREADY AVAILABLE

Artificial intelligence (AI) and machine learning can help get more from infrastructure, by extracting vast amounts of information

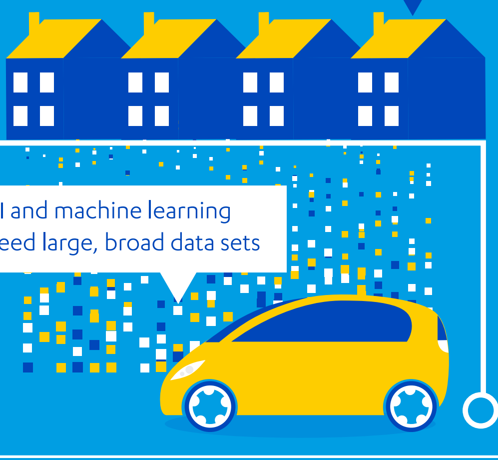


Artificial intelligence could add

10.3%

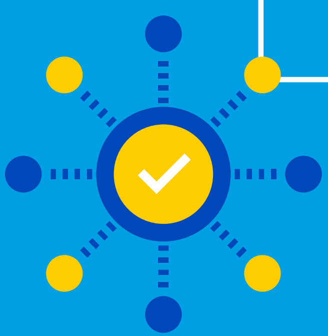
to the UK economy by 2030 – that's up to £2,300 per household

£2,300



A COORDINATED APPROACH

A closed attitude and an array of regulatory, commercial and cultural barriers prevent more effective data sharing



THE IMPACT OF IMPROVED DATA SHARING

1 LOWER CONSUMER BILLS

Greater efficiencies and lower operating costs for suppliers

£1 spent on water smart meters can provide returns of up to

£2.70

2 REDUCED IMPACT ON ENVIRONMENT

Smart meters reduce energy usage, avoid site visits and reduce carbon emissions

3 IMPROVED TRANSPORT

Reduced time spent in traffic and fewer train and bus delays

Digitising asset information of the UK's rail network could save Network Rail up to

£770 million

over the next 8 years

The construction industry saved

£840 million

in 2014 just from using Building Information Modelling

4 SMART CITIES CAN BECOME A REALITY

Faster implementation of Internet of Things, driverless cars and other new technology

£8.9 billion

Direct benefits of UK public sector open data

The Commission recommends:

1. **A Digital Framework** for secure sharing of infrastructure data

2. **A Digital Twin** (computer model) of Britain's Infrastructure, to help plan, predict and understand our assets – a pilot should start in 2018

3. **Coordination** of key players including the Centre for Digital Built Britain through a Digital Framework Task Group

Sources: IDC and Open Evidence, PwC, Network Rail, Deloitte, Mott MacDonald and HM Government

Contents

In Brief

Increasing population, economic growth and climate change are putting significant pressure on infrastructure. To address this, the UK's existing infrastructure needs to become smarter: it needs to work as an optimised system; reducing disruption and congestion.

Artificial intelligence techniques such as machine learning can deliver greater insights into infrastructure assets and systems, enabling greater efficiency. To realise the full benefits, the infrastructure industry needs to adapt to the new world of big data and data analytics and work together.

REGULATORS, NETWORK OPERATORS AND UTILITIES PROVIDERS TO PRIORITISE DATA

Data can provide significant economic benefits. Sharing data, with the appropriate security and privacy arrangements, can catalyse innovation and improve user experience. A Digital Framework Task Group will work across infrastructure sectors to realise the benefits of data.

A DIGITAL FRAMEWORK FOR INFRASTRUCTURE DATA

Under the leadership of the Centre for Digital Built Britain (CDBB) key infrastructure organisations will need to collaborate to develop appropriate data standards, building on existing regulations and guidelines.

COLLABORATION AND DATA SHARING ACROSS THE INFRASTRUCTURE INDUSTRY

The Infrastructure Client Group (ICG) will lead industry engagement in the digital framework, and cultivate a shift towards increased data sharing in the infrastructure industry to encourage collaboration and innovation. The UK's economic regulators may also play an important role in raising the quality and openness of infrastructure data, and should be closely involved.

A ROADMAP TOWARDS A NATIONAL DIGITAL TWIN

High quality, standardised data on all our infrastructure assets, along with the ability to share this securely, will enable the UK's infrastructure to be viewed as an interdependent, dynamic system. CDBB will work with the Alan Turing Institute (ATI) to develop a national digital twin: a virtual model of our national infrastructure which will both monitor infrastructure in real-time and have predictive capability. This will help manage, plan, predict and understand the UK's infrastructure, delivering resilient, responsive, high-performance systems.

Executive summary

Data for the public good

Having more information day-to day helps people make better decisions. Knowing what the weather will do, what time the train will run or when that broadband outage might occur helps minimise wasted time in the day and allows people to work effectively and maximise their leisure time. No one enjoys spending time in a traffic jam because of a road closure they didn't know about or waiting in for the broadband engineer to arrive.

The same is true for how the UK maintains and operates its infrastructure (digital, energy, flood defence, transport, waste and water). Having more information or data about infrastructure assets enables them to be used more productively. Knowing where all the country's infrastructure is and how it is being used will help decision-makers and operators to plan and maintain these crucial national systems better. But simply having the data is not enough. It needs to be shared across the public and private sectors with the appropriate levels of secure access to enable its value to be fully leveraged for public benefit. The UK can get more out of our existing infrastructure through the judicious use of sensors, data and machine learning and by sharing the resulting insights in an appropriate way.

Transport for London has already demonstrated that releasing data to the public can save users time to the economic value of between £15m and £58m per year.[1] Open travel data can support travel apps and real-time alerts to save time, reduce uncertainty and lower information costs, supporting growth in the tech economy and increasing the use of public transport.[2] Ofgem[3] suggest that smart meters will save consumers £47 per household on their annual bills by 2030 through increased awareness of usage, representing potential benefits of £1.3bn and potentially up to £5.7bn,[4] through reduced energy usage, avoided site visits and reduced carbon emissions.

A recent study found[5] that data contributes around £50bn in direct, indirect and induced impacts to the UK economy a year at present. Big data is predicted to contribute more to economic growth in the period from 2012-2025 than the typical contributions from R&D.[6]

In the context of the population growth and climate change challenges identified by the National Infrastructure Commission's National Infrastructure Assessment,[7] data can help extract more value from every unit of infrastructure output. The benefits that accrue from having better data about our infrastructure could be amplified if that data could be shared appropriately. This could mean:

- Lower consumer bills.
- Reduced impact on environment.
- Reduced time spent in traffic.
- Reduced train and bus delays.
- More efficient management of our water resources.
- A faster move to full implementation of new technologies such as CAVs and internet of things (IoT).
- Smart cities in the UK become a reality.
- Greater efficiencies, resilience and lower operating costs of physical infrastructure.

In Autumn Budget 2016, the Chancellor asked the National Infrastructure Commission ('the Commission') **"how can new technologies improve the productivity of infrastructure?".** The Commission has found that technologies which generate and use data such as sensors, machine learning, digital twins and IoT, all features of the 'Smart City', are able to get more out of existing infrastructure.

Data creates value

- **New technologies, such as data capture and processing technologies (sensors, artificial intelligence and digital twins) can generate and manage better quality data about our infrastructure, which can be used to improve the way that assets are planned and maintained.** Data science can help increase the productivity of infrastructure by extracting information from data about infrastructure assets, helping to optimise networks, preventing asset failures, and better targeting maintenance interventions or renewals.

> Mott MacDonald estimates that the total financial and economic cost of rail disruption due to asset failure costs the UK economy £1.3bn – £1.9bn a year. Monitoring technologies can reduce this loss significantly. In the water sector, tools which support leakage identification, such as smart water meters and acoustic noise loggers can provide returns of up to £2.70 per £1 spent due to their ability to help optimise networks, and support better targeted maintenance and renewals.

Data is infrastructure

- Data from infrastructure can be used to generate and send signals and instructions to change the way infrastructure functions, particularly when used as an input into artificial intelligence (AI) models and machine learning. For example, sensors on traffic lights generate data on traffic and pedestrian flow which can then be used to change the timing of the lights. "As data generates the behaviour of infrastructure, it can be said that data is in a sense also a hard infrastructure and that it needs to be maintained and managed through a formal approach, analogous to the way that physical infrastructure itself is managed".[8]

- Data informs how our infrastructure is built, managed and eventually decommissioned, and real-time data can inform how our infrastructure is operated on a second-to-second basis. Data is now as much a critical component of our infrastructure as bricks and mortar. **Data is part of infrastructure** and needs maintenance in the same way that physical infrastructure needs maintenance. It must be updated, housed and made secure. Data must also be supported by physical infrastructure such as data centres, high speed broadband, and widespread high capacity mobile and broadband networks.

Using AI and machine learning can help to extract maximum information from vast amounts of data about infrastructure assets and turn it into useful insights and predictions so that efficient and effective decisions can be made. It is estimated that AI will add up to 10.3 per cent to the UK economy (up to £2,300 per household) by 2030.[9]

Data enables innovation

- **Greater access to open data enables greater innovation.** The Industrial Strategy White Paper[10] sets out the Grand Challenge "we will put the UK at the forefront of the AI and data revolution", and emphasises the gains that AI can bring to the UK economy. The Grand Challenge sets "we will make the UK a global centre for AI and data-driven innovation" as a key priority. The increased sharing of high quality data is crucial to the development of AI and what it can achieve for infrastructure and consumers.

Transport System Catapult estimate that improved data sharing could lead to incremental benefits through mobility solutions of around £15bn in value, by 2025.[11]

The direct economic benefits of UK public sector open data have been estimated as £8.9bn in 2016,[12] while the total impact could be around four times larger when including indirect and wider benefits.

Doing nothing is a big risk

- **Without coordination and collaboration across the infrastructure network operators, regulators and users,** the **benefits** that could be achieved from data and the application of data science **will not be realised**. Progress towards greater innovation is being hindered by a closed attitude to data across the infrastructure sector, and by an array of regulatory, commercial and cultural barriers. The infrastructure sectors will not coordinate the sharing of data and enable innovation without direction from Government and regulators to ensure that both private and public benefits are fully realised.

- **Without seamless mobile and broadband connectivity the benefits of new technologies won't materialise.** Government needs to take forward the recommendations of the Commission's 5G report and consider the appropriate incentives for investment. The National Infrastructure Assessment[13] will make further recommendations on the combination of fixed, mobile and converged technologies to offer the level of connectivity required to deliver the benefits expected from new technologies such as IoT.

Data needs to be shared safely and securely

- Data about our infrastructure assets needs to shared in way that opens up the benefits yet maintains appropriate levels of privacy and security. Our culture must change from one of closed, siloed thinking to an open, transparent culture of effective data management. For example, private companies could make more data about their infrastructure networks, assets and organisation open whilst still protecting personal data. The model needs to move from keeping all data confidential to minimum levels of commercial confidentiality.
- Regulators should ensure that operators take responsibility for collating this data, verifying its **quality**, making it available to the appropriate parties and using it in a safe and ethical way..
- The Centre for Digital Built Britain (CDBB) and Building Information Modelling (BIM) can provide the foundations for a more consistent and open approach to infrastructure data and underpin a move towards digitisation of the built environment which includes existing as well as new assets.

Approaches must ensure security and privacy

- Security and privacy are key and must be prioritised. Sharing data for the public good means that some datasets are public, whilst others will only be available to certain parties and some specific data will have the highest levels of security protection. Levels of access to data must be allocated in the context of an agreed approach to risk management.
- Data protection is fundamental to the development and successful deployment of smart city models and functions.

Digital twins can help manage data about infrastructure and establish a leadership position for the UK

- A digital twin is a computer model which mirrors and simulates an asset or a system of assets and their surrounding environment. Digital twin models can help organise data and pull it into interoperable formats so that it can be used to optimise infrastructure use. Digital twins can also share this data, with defined levels of access, to inform better decisions about which future infrastructure to build and how to manage current and future infrastructure.

GE define a digital twin as "a dynamic digital representation of an industrial asset, that enables companies to better understand and predict the performance of their machines and find new revenue streams, and change the way their business operates".

- The UK with its research capability across the university network is already making progress in this area along with the national institute for data science, the Alan Turing Institute (ATI), and is well placed to be a world leader in the development of digital twin infrastructure models.

Virtual Singapore is a model of the built environment used to optimise current use and planning of infrastructure. The UK has a world leading data science research capability in the Alan Turing Institute and private sector expertise across the UK.[14] The UK has the resources to build digital twin models of the most advanced kind.

The UK needs a Digital Framework for Infrastructure Data and a digital twin pilot project

- To make this happen the country needs a digital framework for sharing data about our infrastructure assets. This would be a national resource and must be as open as possible, while still addressing security risks and concerns.
- A digital twin pilot project would provide the opportunity to demonstrate the benefits that could be achieved from transforming data about infrastructure assets into a shared interoperable format and the gains that can be made from having a greater understanding of the interdependencies of our infrastructure system.

If action is taken to share data in this way, the productivity of the nation's existing and future infrastructure could be improved and the potential cost and risk of future infrastructure breakdown and attack could be minimised. **In short, this could increase the performance, efficiency and resilience of the national infrastructure system.**

The **principles** for smarter data sharing:

- Data quality, consistency and availability should be given a high priority through a Digital Framework for Infrastructure Data. This framework should be integrated with the development of data trusts proposed by the independent AI Review[15] to improve trust and ease around sharing data by working closely with the Centre for Data Ethics and Innovation.

- Non-personal data (network, asset and organisational data) should be treated as open wherever possible, assuming there are no security issues around the release of the data.
- There should be a presumption of minimum levels of commercial confidentiality (e.g. through BIM contracting and regulatory participation).
- Data security and trust should be prioritised in line with recommendations of the AI Review[16] and the Royal Society and British Academy's report on *Data management and use: Governance in the 21st Century*.
- Regulators should play a role in improving data collection, enforcement and cross-sector innovation. The Commission will look at the regulatory challenges to data sharing as part of its ongoing work.

To make this a reality, the Commission makes the following **recommendations**:

1. **1. The Government should task the Centre for Digital Built Britain (CDBB) with the establishment of a digital framework for infrastructure data, drawing together key organisations and existing initiatives both large scale (BIM) and smaller scale:**

    a. A Digital Framework Task Group for infrastructure data should be established with a high-profile chair who can act as a national champion for this agenda.

    b. CDBB should set out a roadmap to a digital framework to develop standards and formats for collating and sharing data.

    c. Key organisations which should be involved in the Task Group and in developing the framework include the Alan Turing Institute (ATI), Infrastructure Client Group (ICG), Construction Leadership Council, Infrastructure and Projects Authority (IPA), Cambridge Centre for Smart Infrastructure and Construction, Project 13, Office for National Statistics, Ordnance Survey, Open Data Institute and the recently announced Geospatial Commission and Centre for Data Ethics and Innovation. Wider consultation input should also be sought from the digital twin working group referenced in section 5.

    d. On standards development, in order to understand the existing standards landscape thoroughly at both a national and supranational level and to ensure agility, CDBB should consult extensively with industry on current behaviours and future requirements for how different infrastructure sectors and sub-sectors use data. CDBB may wish to commission an external standards organisation to conduct some of this work on their behalf.

e. IPA and ICG should be engaged with closely when designing final data standards and performance measures. The ICG amongst others will also be important to consult in respect of appropriate safeguards for commercial confidentiality, with the aim of keeping these to a minimum level wherever possible.

f. There should also be close collaboration with the Centre for the Protection of National Infrastructure (CPNI) and the National Cyber Security Centre (NCSC) on security requirements and levels of access and to develop standards for security, consistent with the objectives of the framework and an agreed approach to risk management.

g. In order to support effective implementation, CDBB should also lead a scoping exercise for a framework to assess industry progress in adopting and using the framework and to measure the level of innovation achieved by industry.

h. CDBB should complete these actions and provide a public report on progress by 1 September 2018, setting out their recommendations for next steps. The NIC would assess this as part of its wider role in monitoring progress against its recommendations.

2. The Infrastructure Client Group (ICG) and the Digital Framework Task Group Chair should lead industry engagement in the framework and cultivate a shift towards minimum levels of commercial confidentiality.

a. ICG should report to CDBB on current industry compliance with minimum levels of commercial confidentiality agreed with CDBB in recommendation 1.

b. ICG should work collaboratively with industry and the Digital Framework Task Group to identify opportunities to make data available and reduce the unnecessary use of commercial confidentiality (e.g. through reviewing and revising existing digital contracts), and should set out an agreed plan with milestones towards achieving the proposed shift.

c. ICG should report on progress in reducing the application of commercial confidentiality to infrastructure data by December 2018.

3. The Digital Framework Task Group (see recommendation 1) should work with the UK Regulators Network and relevant Government departments to review and, where possible, strengthen the role of economic regulators in improving the quality and openness of infrastructure data. This should include:

a. Participation by the UK Regulators Network in the formulation of the digital framework set up by CDBB to ensure that it is effectively aligned with regulatory work on innovation and data.

b. Assessment of the potential role of regulators and of possible barriers within current regulatory frameworks regarding:

- Ensuring compliance by regulated network operators and utilities with the national framework and adherence to data collection standards and formats.
- Sharing of data to inform better understanding of asset performance and user experience.
- Sharing of data across infrastructure sectors and the value chain to enable greater innovation in the development of new technologies and data management focused on better asset management and increased productivity.

c. Support for CDBB's engagement with network operators and utilities around the provision of data of verified quality for the development of a national infrastructure digital twin over the long term.

d. Identification of relevant areas for further trials or studies to enable regulators, and regulated industries, to understand and demonstrate how monitoring technologies and data can support cost-effective maintenance decisions and proactive asset management, working with relevant research organisations.

4. **CDBB should collaborate with the Alan Turing Institute (ATI) and the UK Infrastructure Transitions Research Consortium (ITRC) in pioneering digital twin models with predictive capability in the UK. This initiative should draw upon the AI expertise concentrated in the UK across universities and the public and private sectors.**

a. CDBB should work with BEIS and other potential funders to take forward a digital twin pilot project to explore and experiment with the benefits of building a digital twin of a specific geographical area. CDBB, ATI and ITRC should draw upon input from the digital twin working group identified in section 5.

b. A project review of the pilot digital twin should be completed by October 2018. This should consider the lessons learnt for any future development of larger-scale or more complex digital twins, and the most effective institutional structures to support continuing progress in this area.

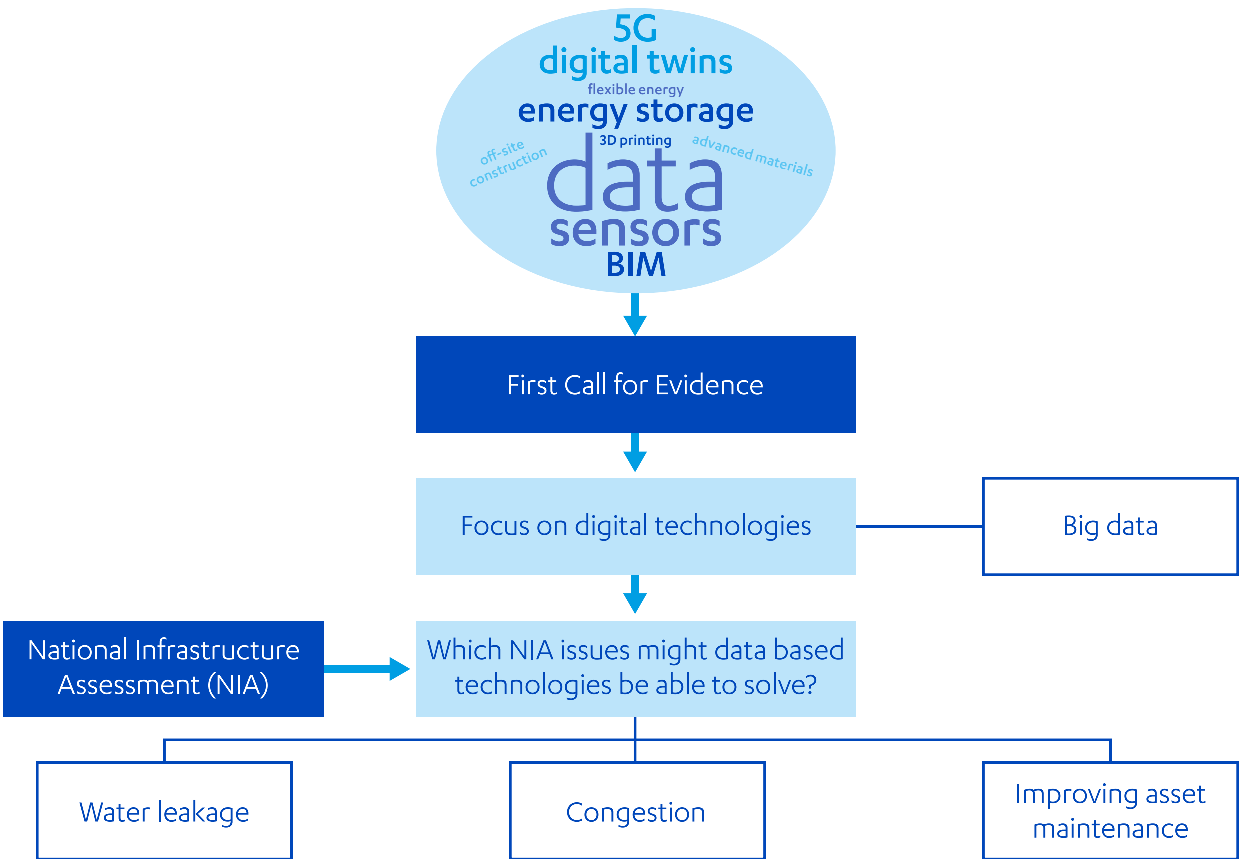
Data for the public good

Introduction and background

In his Autumn Statement 2016, the Chancellor asked the National Infrastructure Commission ('the Commission'), "How can new technologies improve the productivity of infrastructure?" and to undertake a study which has led to this report. The Commission's central finding is that the country can get more out of its existing infrastructure through the judicious use of sensors, data and machine learning but it will not be possible to fully realise the benefits unless the resulting insights are shared in a more open yet secure way.

Focus on data

A public call for evidence and the work so far on the National Infrastructure Assessment led to a focus on digital technologies. The Commission set out four case study areas to evaluate the potential for digital technologies to improve the way infrastructure is used and managed: use of new technologies to reduce leakage and increase efficiencies in the **water sector**, new technology to address **congestion** and maximise traffic flow on roads, **big data** and **asset maintenance**.



As set out in the terms of reference[17] to the study, emerging technologies have the potential to radically improve the way the UK manages infrastructure. The UK has a well-developed infrastructure which was mostly built before the advent of the technologies considered in this report. The Commission has therefore focused on how these technologies can be used to get more out of existing infrastructure and has found that there is limited understanding about the way the UK's infrastructure works as a system. Data and the new technologies that analyse data will be key to getting more out of existing infrastructure. Areas like sensors, digitalisation, the internet of things, big data, and AI can all create opportunities for improving the way the UK operates infrastructure, maintains existing assets, and enhances the capacity and resilience of its networks.* **However, none of this is currently done in a coordinated way;** use of new technologies to improve infrastructure productivity is piecemeal and not as collaborative as it could be. A recent study by IDC and Open Evidence showed that new data driven technologies currently contribute significant benefits to the UK economy of up to £50bn[18] a year of which a significant portion accrues to the infrastructure sector. If data were more effectively shared, the benefits to the public and society that could arise from improved coordination in the use of new technologies applied to infrastructure would be far greater than with the current approach.

Data can be collected, verified and assimilated into models of infrastructure systems. A model of UK infrastructure which captures in real-time the functioning and layout of infrastructure would enable optimisation of use and maintenance of transport, water, energy and communications networks. A model which brings data about these assets together provides a system level view that can be used for planning and simulations and can be used to increase the resilience of UK infrastructure. This type of model is referred to throughout this report as a "**digital twin**".

The key new technologies

The Commission considered a long list of technologies from blockchain to virtual reality[19] New (and in some cases not so new but not widely used) technologies such as sensors, AI and machine learning enable the collection, collation and processing of data. Data science can help the nation increase the productivity of infrastructure by extracting information from data about infrastructure assets. This in turn can enable decision-makers to understand more about what infrastructure exists, where it is, how it interacts, how it is used and crucially how the system as a whole can be made more efficient. Implementation is key and it is important to ensure that available new technologies are being used.

* The Commission is also interested in how the increasing proliferation of digital systems in infrastructure could present new vulnerabilities if not addressed now at the outset of their introduction. This is being explored through the National Infrastructure Assessment and is the basis of a commissioned report by Arup and University College London.

New technologies – definitions

A **digital twin** is a digital model: a dynamic representation of a system which mimics its real-world behaviour. This will typically be a real-time, updated collection of data, models, algorithms or analysis. A digital twin is a virtual representation of a physical object or system across its lifecycle, using real-time data to enable understanding, learning and reasoning.

Big data is a term which is used to cover to processing of data which 'traditional' analysis approaches are not suitable for, due to the data volume, type of or frequency of update. It also encompasses modern data science techniques such as machine learning, which are used to extract value (usable information) from this data. All references to data in this study should be considered as 'big data', with allied processing also being characteristic of this.

Onward use of data is required to extract the full value from a digital twin. This value extraction could use statistical approaches, physically representative models or machine learning driven approaches. These approaches can also be used in conjunction with each other. An example could be using a statistical approach to characterise train speed, which may be used with the train power and weight to calculate expected track wear. A machine learning algorithm could then be used to generate a tool to make track renewal choices to accelerate track renewal and reduce costs. Users may also benefit as the same data can provide real-time estimates of arrival times, adding to the benefits from reduced system costs.

Artificial Intelligence (AI) is an umbrella term for the science of making machines smart. This includes technologies that can perform complex tasks traditionally considered to be out of reach of computer programs, such as visual perception, speech recognition, natural language processing, reasoning, learning from data and solving all manner of optimisation problems.[20] AI holds great promise for improving the management of complex infrastructure systems.

Machine learning (ML) is a branch of AI technology that allows computers to perform specific tasks intelligently by learning from examples, data and experience. Machine learning is closely related to the fields of statistics and data science and at its most basic level involves computers processing a large amount of data to predict outcomes.[21] Due to technical developments in the field, increased availability of data, and growing computing power, the capabilities of machine learning have advanced significantly. Through an expanding ability to extract insights from ever-larger volumes of data, machine learning could increase productivity and provide more effective public services.

Through this study and through stakeholder consultation, two new technologies have come to the fore as offering potential significant productivity gains to infrastructure – **sensors and a digital twin** – but without high quality, shareable data the full gains won't be realised.

Digital twin – practical uses

Rolls-Royce, The Norwegian University of Technology Science (NTNU), research organisation SINTEF Ocean, and classification society DNV GL have signed a memorandum of understanding (MoU) with the aim of creating an open source digital platform for use in the development of new ships.

The platform will allow the creation of so called 'digital twins'. A digital twin is a digital copy of a real ship, including its systems, that synthesises the information available about the ship in a digital world. This allows any aspect of an asset to be explored through a digital interface, creating a virtual test bench to assess the safety and performance of a vessel and its systems, both before its construction and through its lifecycle.[22]

The purpose of this report is to send a clear message to Government and the private sector: the UK needs to open up infrastructure data and make full use of data science and machine learning to get more out of existing infrastructure and to make the right decisions about future infrastructure. The opportunity for the UK to develop a digital twin of the national infrastructure system offers the potential to increase both the efficiency and resilience of infrastructure. The work required is challenging and complex but the emphasis placed on the development of AI through the Industrial Strategy White paper and the AI Sector Deal[23] demonstrates the Government's commitment to progression in this field.

The requirement for seamless mobile and broadband connectivity is a key consideration. Government needs to take forward the recommendations of the National Infrastructure Commission's 5G report and consider the appropriate incentives for investment. The National Infrastructure Assessment will make further recommendations concerning the combination of fixed, mobile and converged technologies to offer the level of connectivity required to deliver the benefits from new technologies such as the Internet of Things (IoT). **The opportunity to use new technologies to get more out of existing infrastructure assets will be constrained unless the UK has seamless, high capacity mobile and broadband coverage in both urban and rural areas.**

1. Data creates value

1. Data creates value: leveraging data to get more out of infrastructure

What is data?

Data generated and collected in the infrastructure sectors includes personal and non-personal data. Infrastructure data can be classed as personal, network, asset and organisational data.

Types of infrastructure data

Personal data	Non-personal data		
	Network data	Asset data	Organisational data
• Consumption • Geospatial location • Payment information • Socio-demographics	• Operational • System resilience • Schedules • Traffic flows • Emissions	• Condition • Licences • Housing • Road • Location	• Financial • Performance • Internal processes

Source: Deloitte

There is overlap between personal and non-personal data in infrastructure. For example, personal location data translated into a traffic minimisation app, or energy consumption data from smart meters and IoT devices being used for the management of smart grids. Personal and network data can be converted to non-personal data, either anonymised or pseudonymised in order to share it, as well as non-personal data collected to reveal personal details so the appropriate privacy protections are required. These approaches may enable wider data sharing whilst ensuring privacy and legal requirements are met. For example, aggregated and otherwise anonymised data from smart meters, with regulatory approval, may be used by energy network operators to enable more efficient and smarter infrastructure management.

Representative data science and supporting technologies are identified below to illustrate some of the current and potential future use of data in infrastructure:

Data collection technologies	Data management technologies	Data analysis and visualisation technologies
Sensors – for example fibre optic, wireless, acoustic noise loggers and smart meters	Building Information Modelling (BIM)	Artificial Intelligence (AI)
Imagery and Measurement – for example, satellite and drone	Digital twin	Machine Learning
CCTV	Geographic Information Systems	Predictive Analytics
Supervisory Control and Data Acquisition (SCADA) & Internet of Things (IoT) data collection	Semantic Web & Resource Description Framework (RDF)	Physical Modelling

What these technologies can do

The nation's infrastructure is experiencing a digital revolution. Some, but not all, infrastructure operators are applying digital technologies to their physical assets, transforming them into smart infrastructure assets which can provide real-time data on performance, condition, and use, as well as a range of other information, unlocking a range of benefits across the infrastructure lifecycle.

The real benefits of this transformation across public and private infrastructure are not being realised, however, because there is limited coordination across organisations and sectors, leading to significant variability in the quality, format and availability of the data that is generated and how it is accessed and stored.

The central benefit of data generated by smart infrastructure lies in its potential to **optimise the performance** of infrastructure assets. By having such data, which first depends on an understanding of what assets exist, the country could be better positioned to make improved decisions about how to use and operate assets and networks. This data can be used together with analytics technologies such as machine learning techniques to realise more efficient ways of operating infrastructure, improving capacity, or supporting better demand management.

In addition to improving how infrastructure operates on a day-to-day basis, more data on how assets are performing, their condition and use, enables **better economic design and asset management**, supports improved decision making around aspects such as **maintenance**, and provides insight into the wider **systems** in which infrastructure operates,

"Getting the most out of our existing assets, networks and systems means improving, for example, availability, capacity and resilience; while reducing congestion, carbon intensity, emissions and running costs, and the need for new build"

Infrastructure and Projects Authority report on Transforming Infrastructure Performance (2017)[24]

"Before visualisation software, designers would have to go [to] the railway and plot the location of the new signals and map how the signals could be viewed by train drivers... Using virtual reality modelling, most of this work can be done from a desktop"

Liverpool City Region Combined Authority call for evidence response both above and below ground. These elements can improve the UK's ability to prolong the life of its assets, prevent critical failures, and realise the greatest possible value from its infrastructure.

Smart infrastructure – providing us with better quality data on assets

The Cambridge Centre for Smart Infrastructure and Construction (CSIC) identified that smart infrastructure is a global opportunity worth somewhere between £2 trillion and £4.8 trillion[25] and the UK is leading the way in many areas, such as building information modelling (BIM), and developing a digital railway. Highlights include:

- Network Rail has a plan to invest approximately £190m between 2019 and 2024 in asset monitoring technologies, analysis platforms and information management systems, which is anticipated to deliver gross benefits of £270m over the next ten years, in the form of asset reliability and maintenance operational efficiencies. In addition, a further £190m may be invested to enable efficiencies in capital expenditure and capacity from 2024 onwards, but this will dependent on the production and agreement of robust business cases.
- Ofgem has made £70m available each year for distribution network operators to compete for funding to support projects that help operators understand what they need to do to address environmental issues, support cost reductions and ensure security of supply, with a particular focus on new technologies.
- Water companies are undertaking substantial investment in transforming their networks for the future. Thames Water, for example, is investing £300m over the next fifteen years to install smart water meters in three million homes throughout London and parts of the South-East. In this region, such technology has demonstrated potential to reduce consumption by 12 per cent in the average home, and support faster identification of leaks.[26]
- Highways England has committed to a £150m innovation fund to 2021 which will include trials of connected and autonomous vehicle technologies, provision of better information to customers and improved management of its network and infrastructure assets.[27]

However, as a rapidly evolving field, the case for investing in smart infrastructure is still being demonstrated across sectors. CSIC recently noted that the infrastructure industry "puts a value on physical assets … but not yet on digital assets"[28] which collect data and support a range of benefits. As such there is a need to share insights and learnings across sectors to better understand the benefits and application of smart infrastructure in a variety of contexts.

In the international context, Singapore, Hong Kong, South Korea and China are making significant advancements in the use of smart infrastructure and the development of smart cities.

Case Study: China and AI[29]

China has developed a strategy to become the world leader in AI by 2030, aiming to build a domestic industry worth almost $150bn. Current infrastructure technologies in development include:

- Meituan's O2O Real-Time Logistic Dispatch System uses AI on the supply and delivery chain, employing big data analysis to improve delivery route efficiency.
- Baidu's Apollo will launch an open, complete and reliable software platform for autonomous driving. The company plans to introduce fully autonomous driving capabilities on highways and urban roads by 2020.
- AliBaba's cloud solutions provider, AliCloud, uses a computational system called Apsara. Some intelligent products by AliCloud include 'City's Intelligent Eye' a 90.46 per cent accurate vehicle detection system; the 'ET City Brain for Hangzhou' which helps police to respond to traffic collisions much faster (3 mins compared to 15 mins); and the 'ET Environment Brain' which can intelligently monitor pollution of water, air and soil in Jiangsu province, providing warnings to local Governments. It is expected to be helpful in disaster forecasting, extreme weather warning and environmental protection.

1. Better quality data on infrastructure can help to optimise its use:

The opportunities for using data to support improved efficiencies through demand and capacity management are clearly demonstrated in the transport sector. Providing users with increased information on transport choices, for example through smartphone applications such as Citymapper, can help to maximise network efficiencies and realise additional capacity, particularly if integrated with data-driven solutions such as smart traffic management or digital signalling. In China, for example, a major ride-hailing app (Didi), as a commercial operator, set up a platform to share data from smart traffic signals with traffic authorities which helped to optimise traffic light changes in real time and ease congestion by an estimated 11 per cent in the city of Jinan.[30] In the UK rail sector, 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 per cent may be achievable, at a lower cost than through conventional approaches.[31]

Case Study: M42 Motorway Active Traffic Management

Active Traffic Management (ATM) was launched as a pilot scheme on the M42 operating between junction 3a and 7 near Birmingham with mandatory variable speed limits, hard shoulder running, better driver information signs and a new incident management system. The monitoring technology allows operators to direct and control the flow of traffic, and open and close any lane to traffic to help manage congestion or an incident. The approach has reduced journey times by 25 per cent, accidents by 50 per cent, pollution by 10 per cent, and fuel consumption by 4 per cent – at only 20 per cent of the cost of widening the road.[32]

2. AI and infrastructure - applying machine learning techniques can amplify the benefits

The traffic lights in Milton Keynes are the same as any other traffic lights but what makes them 'smart' is the system that supports them - the sensors gathering data from the roads, the application of machine learning to the data collected, and the development of a platform to provide journey information to pedestrians and road users to influence their behaviour. This exemplifies how AI can improve the way infrastructure is used, for instance targeting reduced congestion and journey times.

Case Study: Smart traffic lights in Milton Keynes:

Around 2,500 cameras are being deployed across a 50-square mile area of the city, and once fully deployed (March 2018 target date) should provide more detailed traffic information than is currently available. Any personal information is deleted by the camera, with all analysis conducted on anonymised data. The sensor package has been designed as a low-cost unit that has low power requirements and can be retrofitted to existing roadside infrastructure. A neural network will be applied to the data collected, to predict the location of vehicles, congestion and journey times across all possible routes for a given arrival time (i.e. proactive, not limited to reactive measures based on real-time traffic information).

The next phase of work concerns the communication of this analysis, through integration of this system with traffic control centres, connected & autonomous vehicles, and a redevelopment of the Milton Keynes smart cities data platform. This project should deliver short-term benefits for the management of the road network, and is also designed to support road network optimisation in the long-term through the application of AI.

Machine learning can be applied directly to the way infrastructure is used to increase efficiency.

Case Study: Using AI to increase the efficiency of Google's data centres

DeepMind applied its machine learning to Google's data centres, reducing the amount of energy used for cooling by 40 per cent. Servers generate heat, and as a result, a lot of power is consumed to keep Google's data centres cool and to keep servers running. The cooling system is complex and dynamic; each data centre requires a range of equipment and has over one thousand sensors – all generating a lot of data. It must also be sensitive to internal and external changes, such as the weather. Furthermore, every data centre has its own unique architecture and environment.

To respond to this challenge, DeepMind developed a general intelligence framework to understand the data centre's interactions, and they applied machine learning to make recommendations for operating Google's data centres more efficiently. DeepMind trained neural networks on historical information on the data centre cooling systems. When the model was applied to a live data centre, it consistently achieved a 40 per cent reduction in the amount of energy used for cooling, which equates to a 15 per cent reduction in overall power usage effectiveness, after accounting for electrical losses and other non-cooling inefficiencies.

Because the model is general purpose, it is possible that it could also be applied to increase energy efficiency in other complex industrial systems.

AI has emerged as a key technology that could support improvements in infrastructure and industrial performance. Consistent with the vision set out in the Industrial Strategy White Paper[33] and the AI Review,[34] AI is a key pillar driving growth in the UK economy.

> *"There is a vast range of potential benefits from further uptake of machine learning across industry sectors and the UK is well placed to take a leading role in the future development of machine learning. Ensuring the best possible environment for the safe and rapid deployment of machine learning will be essential for enhancing the UK's economic growth, wellbeing, and security, and for unlocking the value of 'big data'"*
>
> Royal Society call for evidence response

> *""...supercomputing and advanced whole system modelling should be prioritised to drive simulation; and artificial intelligence to underpin optimised planning, management and operation of the whole railway system and support faster, more effective decision-making"*
>
> Network Rail/Rail Delivery Group call for evidence response

3. Insights from data about infrastructure assets can support better targeted maintenance: reducing lifetime costs, prolonging the life of assets and reducing disruption

The UK has an ageing infrastructure asset base, and measures to maintain and prolong the life of assets are important in ensuring the long-term resilience of infrastructure. Improved data on the location, condition and performance of assets can support better decision making, and enable infrastructure operators to be proactive rather than reactive when it comes to maintenance, inspection and refurbishment programmes, driving down lifetime costs of asset management and potentially prolonging the life of assets. More cost-efficient approaches to managing public services and monitoring the performance of contractors can be of particular benefit to resource-tight local authorities.[35]

Case Study: CSIC and Network Rail: Understanding the structural response of masonry arch bridges through sensing at Marsh Lane

Working with Network Rail, CSIC has been monitoring an operational Victorian masonry arch viaduct in Leeds for the past two years. The objective of the pilot study has been to better understand the dynamic response (movement under a changing load) of the viaduct to rail traffic and identify the mechanisms that drive its degradation. A combination of technologies was used to monitor the structure, including fibre optic sensors and laser scanners. Data collected provided a new understanding of how arches in the viaduct react to rail traffic, and new evidence that can explain existing damage, as well as identify parts of the structure where damage is likely to concentrate in the future. Data similar to that collected as a consequence of this pilot has potential to inform better targeted future maintenance and renewal techniques, as well as improved management of assets known to have defects across Network Rail's portfolio.

The Commission sought analysis from Mott MacDonald on the Marsh Lane pilot, to understand the benefits of monitoring structural assets. The analysis found that monitoring technologies, similar to those applied at Marsh Lane, represent a significant opportunity to support improved asset management, particularly in terms of maintenance, renewals and avoiding unforeseen asset failures. In addition they found that the largest benefits are likely to be obtained from combinations of multiple monitoring technologies, together with investments to improve asset information management and develop other components of smart infrastructure systems such as decision support tools.

Given the early stage of the Marsh Lane pilot, it was difficult to quantify the benefits of having additional data to inform maintenance decisions for structural assets. The analysis did find, however, that improved information would have supported more focused, less expensive maintenance interventions at Marsh Lane in the past, especially if combined with decision support tools such as predictive maintenance models.

In terms of renewals, the analysis assumed that there was strong potential for monitoring technology to allow Network Rail to extend the life of assets. A scenario was developed to test the case for smart monitoring to enable a deferral of bridge renewals. The analysis determined that the technology would pay for itself if 20 per cent of bridges scheduled for renewal could be delayed by five years, based on the assumption that technology rollout costs were £20,000 per structure. Benefits significantly increase if renewals can be deferred longer or the cost of this monitoring technology reduces. This is a significant opportunity given Network Rail's ageing asset base, with many Victorian-era bridges soon due for renewal.

The Marsh Lane analysis shows that newer and more innovative applications of monitoring technologies are expensive. Recent applications of monitoring technology have therefore generally targeted assets which have known problems (i.e. reactively) rather than proactively to predict future issues and inform maintenance. Research undertaken by Deloitte found that a proactive rather than reactive approach to maintenance increases equipment uptime by 10-20 per cent while reducing overall maintenance costs by 5-10 per cent.[36] Better maintenance that may result from more widespread monitoring could reduce disruption for passengers, as well as overall maintenance costs for industry.

Current approaches are missing a significant opportunity, and infrastructure operators and regulators need to accommodate innovation and research in this area, test technologies and ensure the realisation of benefits from improved data. This could be demonstrated through further studies in collaboration with the various sectors and researchers. In the case of Network Rail, future pilot projects should target a wider range of structural assets, and focus on developing technology that is more easily deployable, in order to understand how to improve maintenance processes.

The Commission looked beyond structural assets to examine how data is improving maintenance processes in the UK and in Hong Kong. This could also be subject to further research or trials throughout the UK. Robust ex-post evaluation of this work would help to identify and build use cases for these new technologies.

Case study: Transport for London – real-time monitoring of Tube train condition

Transport for London (TfL) is currently developing a new tool that will analyse data feeds from Tube trains to provide maintenance staff with live information about the condition of a train.

Using the tool, staff can analyse the data and identify where faults exist or might be developing and remedy them before they cause service issues. For example, by monitoring the real-time performance of the time it takes for a train door to open, it is possible to identify if a fault is likely. Pre-emptive maintenance work can then be undertaken to ensure that a fault – which might cause the train to be taken out of service, causing delays and disruption – does not develop.

The tool has strong potential to make maintenance planning more efficient and prevent costly faults leading to service delays from occurring. The in-house capability will also help TfL save money by reducing third-party spend – currently around £46m over five years on an external maintenance support contract.

Case Study: Use of artificial intelligence to deploy engineers on the Hong Kong metro

The MTR, the owner and operator of Hong Kong's metro system began using AI to plan and schedule maintenance works in 2004. This is recognised as one of the most advanced metro systems in the world; it monitors the entire subway line to determine critical maintenance tasks and directs the 10,000-strong workforce by priority using AI. Sensors along the tracks, switches and signals generate real-time data to assist in informed decision-making.

The rollout of the technology has supported better targeted maintenance and scheduling from a workforce planning perspective. At an initial cost of $0.7m in 2004, the tool is delivering annual savings of US$0.8m per year, due to the removal of time spent preparing work schedules, and allowing resources to be targeted more efficiently.

4. Data identifies where the country's infrastructure assets are

One of the biggest maintenance and renewal challenges facing infrastructure operators is understanding where assets are physically located and knowing what condition they are in. This is a particular problem for underground assets. The lack of data on the location of assets underground costs the utilities sector £150m annually in London alone due to accidental strikes during excavation. This complicates the design and construction process unnecessarily.[37]

Advances in detection technologies, combined with improved data management and collection, has significantly improved the ability of infrastructure operators to locate and better understand underground asset condition. International examples of improving the accuracy of underground asset location data reveal a return on investment as high as £21 for every £1 spent.[38]

When combined with asset condition monitoring technologies, improved asset mapping can support better management of underground networks – such as earlier detection and prediction of problems occurring underground such as water leakage.

Case study: Affinity Water using permanent acoustic logging technology to detect water leakage to inform maintenance decisions

Since the mid-1990's, the total volume of water leaking from networks has reduced by 37 per cent,[39] however innovative approaches are now required to continue these positive gains. Affinity Water has deployed 20,000 acoustic loggers across the 25 per cent of its network most prone to leakage between January and May 2017. The loggers, which are distributed at 200-400 metre intervals, detect noises caused by leaks and allow users to pinpoint approximate locations to enable engineers to act on the data to repair assets, improving the efficiency of leakage detection, maintenance and renewals.

The Commission sought independent analysis to understand what kind of benefits could be achieved by rolling out acoustic logging technology throughout the UK, testing benefits in six separate regions. The analysis found that benefits vary considerably by region, however up to £2 per £1[40] spent can be achieved in some cases, particularly at locations with high marginal cost of supply due to the technology's potential to support faster, more accurate leakage detection. At other locations, even after accounting for a future reduction in cost and potential benefits, the analysis found the technology not to be viable.

This demonstrates a wider conundrum faced by infrastructure operators when deploying innovative technologies. Benefits achieved in one region cannot be automatically assumed in another, and as such the uptake and rollout of technology needs to be considered on a local basis.[41]

5. Data can help to prevent, manage and mitigate asset failures and disasters

Asset failure significantly impacts infrastructure, which in turn disrupts the economy, endangers people's health and safety, and incurs environmental consequences. Mott MacDonald found that in the rail sector, asset failures and other incidents related to rail infrastructure have a total financial and economic cost of £1.3bn – £1.9bn, including fare losses and the cost of passengers' time on delayed or cancelled trains. Improved data on asset performance and degradation, also known as 'health monitoring', is important in preventing or minimising the impact of failures.

Case study: benefits of monitoring low-clearance bridges susceptible to road vehicle strikes

Low-clearance rail bridges are highly susceptible to failure due to the impact of road vehicle strikes. There are some 1750 strikes on average per year across the country, costing over £80m. The introduction of monitoring technologies at susceptible locations can help prevent strikes from occurring or minimise impacts. For example, CCTV monitoring can enable Network Rail to remotely assess the impact of a strike ahead of an engineering inspection crew arriving, allowing for rail services to continue if the safety risk is identified as minimal (providing a financial and economic benefit of £24m[42] over the next 20 years at today's prices), whilst more advanced forms of monitoring can detect the height of approaching road vehicles (particularly HGVs and buses) and trigger a warning signal to the driver (providing a benefit of £306m[43] over the next 20 years at today's prices). Future monitoring technologies have the potential to connected with vehicles, autonomously stopping them in the event of a threat of strike.

In addition to preventing asset failure, rapid deployment of monitoring technologies can provide insights about structural integrity and condition and help support a quick return to service. For example, the Forth Road Bridge in Scotland was brought back into service two weeks earlier than expected, following a crack being detected under the carriageway, due to data provided on its structural health by monitoring technologies.

Case study: Rapid deployment of monitoring technologies to support the reopening of the Forth Road Bridge

In December 2015, the Forth Road Bridge, one of Scotland's busiest crossings, was forced to close after the unexpected discovery of defective steelwork.

Initial estimates suggested the bridge would be closed for over a month whilst undergoing full inspections, testing, and repairs. The fast deployment of innovative sensors provided crucial diagnostic monitoring and engineering information to the bridge operator, and this supported a decision to open the bridge two weeks earlier than anticipated.[44]

Notwithstanding the wider economic impacts, the closure of the bridge is estimated to have cost the freight industry over £40m, with costs of £0.6m per day being incurred due to additional travel time and fuel requirements for the 10,500 HGVs which use the bridge daily.[45] In a separate survey, 60 per cent of businesses in the Fife region reported profits decreasing, like-for-like with the same period the previous year, due to the impacts of the bridge closure on connectivity.[46]

Since the bridge was closed, however, a sophisticated structural monitoring system has now been applied to parts of the Forth Road Bridge at risk of similar failure, at a cost of £1m to £1.5m.[47] This is a fraction of the cost of the impact of the bridge's closure on the wider economy.

A Scottish Parliament Committee tasked with reviewing the Forth Road Bridge incident found that the operator had not installed monitoring technologies to the component of the bridge that failed due to difficulties in justifying such investment when no indication of the potential for failure is apparent. This approach was in line with what is currently considered best practice in industry. If monitoring technologies had been previously installed on the bridge to detect such structural deficiencies in advance of closure, however, preventative maintenance could have been carried out either to prevent closure of the bridge or to further minimise the amount of time the bridge was out of action. **It is difficult to estimate the benefit from preventing a bridge falling down or closing for a period of time but it is an important element of the benefits that increased quality of data about infrastructure assets from monitoring technologies will deliver.**

Like the Forth Road Bridge example, asset failures in the water sector can have similar economic consequences, causing disruption to people and businesses. In London for example, there was an increase in late 2016 in the frequency of major urban flooding events due to trunk mains failures, a consequence of an ageing network and poor information on priorities for maintenance and renewal.[48] The local water company for the region, Thames Water, has had to increase its budget directed at repairing and monitoring such assets by an additional £97m (beyond the £147m already committed as part of the price control period).[49] One of the priorities of this investment, beyond repairing assets, will include gathering improved data on the condition of the existing network, and developing a risk model to predict the potential for failure, allowing for better targeted maintenance. This provides an effective interim measure ahead of the need for full scale mains replacement.

6. Aggregated data helps inform a system level view of infrastructure

Smart water meters are able to communicate water usage remotely. They have the potential to provide hourly data, instead of just two or so readings a year for conventional meters. This allows leaks within individual properties to be identified and addressed. Smart water meters can also allow householders to be more aware of their water usage by providing information on water consumption. Whilst both smart meters and automated meter readers largely eliminate the need for manual readings, the former provides richer data, allowing customers to review their water consumption in real-time through status displays. Smart water meters could also support seasonal or variable tariffs as seen in Australia, with prices adjusted to reflect how much water is being used or to respond in times of shortages. However, there appears to be little support for these within the industry or public at present.

Since 2013, Thames Water has been rolling out compulsory smart water meters to its customers in and around London. To date, Thames Water have installed over 200,000 smart water meters. These smart meters will support and enable Thames Water to protect water supplies for the future, by helping Thames Water to better manage water demand, as metered customers typically use 12 per cent[50] less water supply. Smart meters will also help Thames Water in identifying and fixing leaks more efficiently.

Case Study: Anglian Water – smarter ways to measure water

Anglian Water (AW) is currently undergoing a major digital transformation. One element of this is a smart water metering trial, which is being rolled out in their Innovation Shop Window in Newmarket. By November 2017, AW had installed 5,800 smart meters to customers, accompanied by a new online portal that allows customers to view and understand their water consumption habits.

AW estimate that compulsory smart meter roll out across their region (the East of England) would reduce customer service pipe leakage by 48 per cent[51] overall over a 10-year period, saving around 18.9m litres of water each day by 2030. This could help reduce the need to build new infrastructure such as new reservoirs, or to abstract water from local rivers and it thereby helps minimise environmental impacts.

There is also evidence that smart water metering can help to reduce consumer water consumption. Customers that were previously on a standard meter and moved to a smart meter showed a 2.2 per cent reduction in consumption. For those customers who were previously on an unmeasured supply and moved to a smart meter, their water consumption reduced by 16 per cent. AW estimate that if they were to replace all metered properties with a smart water meter, then the saving of water for their region would equate to over 20m litres per day by 2030.

The smart water metering trial, albeit early in its journey, has so far demonstrated a number of key benefits which could potentially support AW in meeting a number of its core objectives: by better engaging customers, delivering more for less, enhancing resilience, improving environmental sustainability and affordability, and reducing the need for new capital infrastructure.

The interim National Infrastructure Assessment concluded that the country needs more resilience against drought. There is estimated to be a 1 in 10 chance of a drought in the UK over next 25 years, but there is low awareness around the issue. Climate change, a growing population and higher environmental standards are increasing pressures, with the South East of England set to face significant supply deficits by 2050. Despite the challenge, there has been limited progress in even 'low regrets' actions to make the most of existing infrastructure, such as leakage reduction and water metering. The final National Infrastructure Assessment, due summer 2018, will consider recommendations on the scale and types of investments needed to ensure that the country is resilient to drought. To inform this work, the Commission is carrying out a review of existing studies, as well as producing independent evidence. This includes assessing the impact of different metering policies, taking into account costs and benefits of a more widespread and consistent roll-out of smart water meters. The analysis will also consider a new water supply infrastructure that is likely to be needed even assuming ambitious demand management and deployment of new technology.

Separate analysis looking at several water companies, requested by the Commission, identified that in some cases, **returns up to £2.70 for every £1 spent**[52] can be gained through smart water metering[53] due to reduced consumption, faster identification of leaks, lower meter reading costs and better understanding of network behaviour.

Smart water metering has the potential to offer net savings of approximately £600m – £1.5bn.[54] However, potential returns are subject to significant variation across water companies, as well as uncertainty around the future underlying cost of smart metering technology. In addition, innovation in smart metering approaches, including the use of 5G, IoT networks and new battery technology, may provide a significant shift in underlying cost, increasing the potential returns. Benefits which current approaches struggle to appraise (system-wide benefits from calm networks, customer valuation for improved visibility of water consumption or wider exploitation of data) may also strengthen this business case. The development of further, system and infrastructure level, appraisal approaches may be necessary to understand these benefits.

These findings correlate with analysis undertaken in the energy sector. The Department for Business, Energy & Industrial Strategy (BEIS) has forecast that the roll-out of smart meters will reduce domestic consumer bills by approximately £47 per household in 2030.[55] At a national level, the roll-out of smart meters are projected to deliver net benefits of around £5.7bn through reduced energy usage, avoided site inspections and reduced carbon emissions.[56, 57]

> ***"Having smart and connected meters at each demand point allows a much more focussed effort on demand management and leakage management"***
>
> WSP call for evidence response

Network operators and utilities see benefits in having better information about their networks and operations. As such, they are moving in the direction of using data and digital technologies at a system wide level. The problem is that this is currently uncoordinated between players within and across industries.

> ***"...new lower-cost sensors which are IoT enabled provide new and innovative ways of obtaining network coverage and generating data on asset condition, use and utilisation"***
>
> Transport Systems Catapult call for evidence response

> ***"...costly investment in new civil engineering infrastructure is unlikely in general to offer value for money, as a means of tackling congestion [on road networks]."***
>
> ***"The application of digital technologies is likely to be more cost-effective – what might be termed the 'digital roadway', by analogy with the digital railway"***
>
> University College London call for evidence response

Case Study: Network Rail – ORBIS Programme

The Offering Rail Better Information Services (ORBIS) is a seven-year, £330m programme. It is the first real Digital Asset Management programme of its type, a whole-system intelligence-led approach that puts asset information, knowledge and intelligence in the hands of engineers and planners. It is designed to improve the way Network Rail captures, stores and exploits asset information. It has realised £216m to date and has been a core part in developing Network Rail's investment plans for its next Control Period (2019-2024).

Since 2012 it has delivered a number of ground-breaking digital tools and solutions that are changing how data is used and viewed by seeing it as an asset in its own right and putting it at the heart of the railway.

Key outputs from the ORBIS programme to date have included 1) Geo-RINM Viewer (GRV) – which uses data from a range of sources, including images from existing master assets and aerial surveys, to give a clear picture of the entire railway network and how it relates to the wider environment. 2) Integrated Network Model (INM) – a digital milestone achieved in September 2017 when this was introduced as Network Rail's new and improved Track Asset Register. INM delivers a geospatial picture of the track asset with up-to-date data, and displays the location of assets in the real world. With its logical view of the entire rail network, INM is transforming the way track and infrastructure data is accessed. 3) Decision Support Tools (DST) – a number of DSTs have been developed by the programme. This includes a Track DST which is helping track teams to predict where and when Track Speed Restrictions could be imposed on the railway network. This is significantly reducing TSRs across the rail network.

Some cross-sector coordination is being achieved through Building Information Modelling (BIM). Since 2016, the Government has mandated BIM Level 2 for all public sector projects, and this process has been supported by the Government-backed BIM Taskforce.

To enable this, BIM tools have been developed which can exchange project and spatial data in common file formats. BIM Level 2 has improved delivery across many projects resulting in construction cost savings of £840m in 2013/14.[58]

Case Study: HS2 are relying on BIM to provide an efficient exchange of information, and to provide a precursor to a 'digital era'

HS2 have been relying on a Building Information Modelling (BIM) platform to transform from a traditional document management approach to a real-time data driven environment. This has enabled efficient and accurate exchange of information between all parties involved in current work. In the future HS2 see the development of a digital representation of the railway with BIM technologies as a key element in optimising processes that will also link to other information (such as schedule or survey data). They have identified that Health and Safety will be improved with this technology, through both decreased site and manufacturing time as well as better identification of risks. The further application of BIM is therefore seen as a key component in the generation of railway for the digital era.

BIM Level 2 currently focusses on the design and initial delivery of an asset. BIM approaches are being extended through the life of an asset.[59] These extensions will provide a more interoperable system which can then support additional capabilities including paperless contracts and enhanced information discovery (through the use of semantic web concepts – aligning physical objects to online records), supporting the full lifecycle of assets.

BIM is evolving to support maintenance and operations as well initial construction, becoming a live data repository which will be maintained post-construction. This will enable the creation of a digital picture of the way infrastructure functions, which will be able to assimilate data from smart meters, large scale network models, building information models, and other external elements to a cohesive whole. Such models can provide insight into the way that infrastructure as a whole depends upon where and how people live.

The convergence of BIM and geospatial data is part of this evolution. "If BIM is about the purposeful management of information throughout the project life cycle – for infrastructure as well as buildings – then geospatial data will become a significant aspect of that, particularly as we move towards Digital Built Britain and the management of entire estates and portfolios", according to Dr Anne Kemp, ICE.[60]

2. Data is infrastructure

2. Data is infrastructure

Data informs how infrastructure is built, managed and eventually decommissioned, and real-time data can inform how infrastructure is operated on a second-to-second basis. Data is now as much a critical component of national infrastructure as steel, bricks and mortar. **Data is part of infrastructure** and needs maintenance in the same way that physical infrastructure needs maintenance. Data needs to be updated, housed and made secure. Data must also be supported by physical infrastructure such as data centres, high speed broadband, and widespread high capacity mobile and broadband networks.

The UK Digital Strategy[61] sets out a plan to strengthen national data infrastructure: the assets, technology, processes and organisations that not only create data, but open it up and allow it to be shared.[62] This includes working with the Open Data Institute to create an environment to open up customers' data across more sectors through the use of Application Programming Interfaces (APIs), and a firm commitment to implementing the EU's General Data Protection Regulation (GDPR) by May 2018 to ensure that businesses and services can continue to operate across international borders. This report builds on the recommendations in the UK Digital Strategy and the recommendations of the National Infrastructure Commission's report on 5G. Seamless mobile and broadband connectivity it required to support the use and sharing of data.[63]

Peter Kawalek's report 'Data as Infrastructure' sets out that the realisation of the value of data across the public realm depends on the use and governance of vast amounts of data.[64] The report explores the possible role of Government in data governance, ranging from Government as a provider of data to Government as a 'smart system'.[65]

Smart Dubai

Smart Dubai is a city-wide Government initiative to 'deliver world-class smart services and infrastructure to create happiness'. As Dubai's infrastructure is rapidly evolving to cater for the next 30m visitors by 2020, the city aims to bake smart design into the fabric of the city through new smart initiatives. Focal areas include citywide connectivity and access, sensing and actuation to drive efficiency, data orchestration and analysis for real-time intelligence, smart service delivery apps, and a centralised monitoring and management layer. This delivers many benefits, but may come at the expense of issues such as privacy, ownership and flexibility.

The Government's strategy for data governance will need to be shaped by its vision for the future, alongside relevant current thinking.[66] Identification of the problems and complexities associated with infrastructure will help to develop the most appropriate form of data governance, as data and technology are used to address these challenges. The strategy will also need to be joined up across both the public and private owners and operators of infrastructure, and ideally consider ways to encourage private entities to develop solutions to infrastructure challenges (as in the case of Busan Green u-City in South Korea outlined in the box).[67]

"Governance mechanisms need to evolve to recognise data as a valued asset"

Ordnance Survey call for evidence response

"The creation and maintenance of an interoperable data asset is key to unlocking the potential of our existing infrastructure, and this in turn requires the potential of data itself needs to be unlocked"

Ordnance Survey call for evidence response

Case Study: Busan Green u-City in South Korea

Busan is the second largest city in South Korea. The u-City programme aims to make Busan a 'Smart City', integrating next-generation technology with the city's built environment. The programme focusses on collecting and sharing harmonised and interoperable municipal data: an Integrated Operation Centre has been developed to integrate both public and private city data; this data is then available to different stakeholders through a shared platform, the 'blueprint ecosystem'. The blueprint ecosystem also supports an application development platform called the Busan Mobile Application Centre (BMAC). By making the municipal data, and the BMAC system, available to both public and private stakeholders, Busan has provided support for private entities to attempt to solve urban challenges. Private entities are now providing a range of services that contribute to the well-being of citizens and sustainable management of Busan. Some of the services provided by private entities include digital signage and smart bus shelters or energy management services for commercial and residential businesses.

3. Sharing data

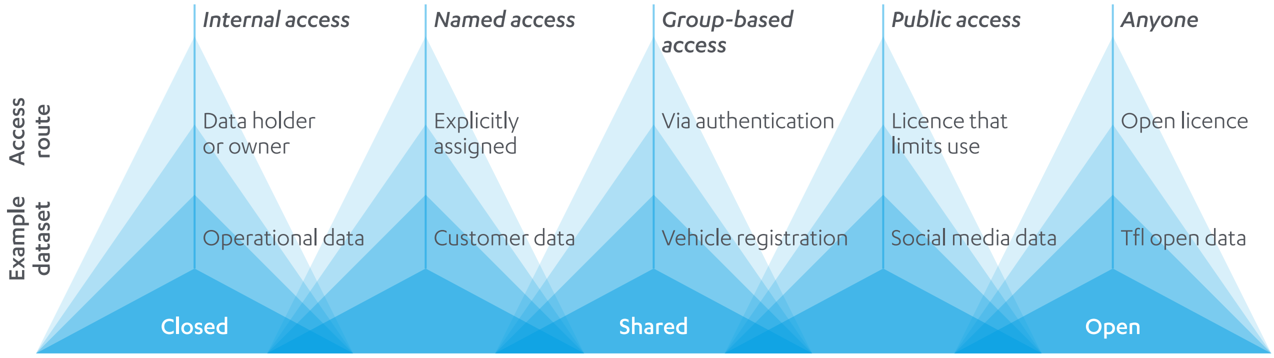
3. Sharing data

There are unrealised benefits from sharing data within and across sectors (energy, digital, transport, water, waste and flood defence).

Data sharing does not mean all data is open

At present, data sharing is well established across some pockets of infrastructure but not all. It is important to distinguish between data that is freely shared as 'open data', and other data which may be shared with conditions or under licence or contract (which may or may not involve a fee). Some data, such as sensitive and operationally critical data from power plants, may be shared only through secure mechanisms.

The data spectrum in infrastructure



Source: ODI (example datasets adapted by Deloitte)

The benefits of sharing data

The lack of coordination of data within infrastructure sectors and across sectors means that the UK is losing out on a large chunk of potential benefits from sharing data. Careful management, proactive industry leadership and effective systems to address concerns around privacy and security could enable greater data sharing, leading to significantly increased benefits for industry and society.

> ***"The benefits to having open data in the modern age are unprecedented, especially where they impact public services. Open data and accessible APIs can lead to greater public awareness and engagement with infrastructure, new services, greater safety and gains in efficiency. It also opens the sector to greater innovation from data science firms, especially the UK's wealth of start-ups and SMEs in this space"***
>
> Alan Turing Institute call for evidence response

Case Study: The benefits of opening up Transport for London's data

Transport for London (TfL) releases a significant amount of data – such as timetables, service status and disruption information – in an open format for anyone to use, free of charge. This allows thousands of developers and partners to bring new products and services to market more quickly, and therefore extends the reach of TfL's own information channels within stations, at bus stops and online. At last count, more than 600 apps were being powered specifically using the TfL open data feeds, used by 42 per cent of Londoners.

TfL has already demonstrated that releasing data to the public can save users time to the economic value of between £15m and £58m per year.[68] Recent analysis by Deloitte found that the provision of transport information through travel apps and real-time alerts is saving £70m-£95m per year in time, reduced uncertainty and lower information costs. Further, release of open data by TfL has supported the growth of London's tech economy to the value of £14m annually in gross value added (GVA) and over 700 jobs, whilst also unlocking new revenue and savings opportunities and new ways of working at TfL, including a £20m increase in bus usage, as customers are more aware of service opportunities.[69, 70] This data has been used by a range of apps, from early stage start-ups to global leading technology platforms, saving time and reducing stress. There are currently 13,700 registered users of TfL's open data.

Case Study: Greater London Authority: London Infrastructure Mapping Application

The London Infrastructure Mapping Application is a new platform that allows utilities, boroughs, the Mayor of London and TfL to share information relevant to infrastructure investment and planning. By bringing together a range of data, the application facilitates improved collaboration between actors, joined-up approaches to construction and design, and better identification of future demand and capacity constraints. Information is visualised spatially through a bespoke mapping application that has been developed in consultation with users. Early evidence has found that the tool supports better alignment of investment – unlocking housing growth and reducing disruption throughout London – by allowing projects such as roadworks to be timed better, and saving costs through joined-up approaches to construction (e.g. joint ducting of utility cables and pipes). Analysis by the GLA found that the application eliminates one per cent of congestion caused by long-term works on the TfL road network over a single year, meaning £4.2m could be saved in avoided delays. This significantly exceeds the cost of developing the tool to date.

Case Study: Benefits of open data for the Environment Agency

In April 2018, the Environment Agency (EA) will be removing all charges levied for the use of its data. Some of their data will be published as open whilst other data sets, such as those which use data from third parties, will not. The EA will also design new datasets to be open by design where possible.

The opening of the EA's Risk of Flooding from Rivers and Seas dataset has increased public awareness and knowledge of flooding. Applications based on the data have helped raise awareness about environmental issues and flood risks, helping the EA achieve its core objectives. The data was also published to provide direct value benefits, for example as inputs to improve insurance models for flooding. Furthermore, the EA's open and collaborative approach has fostered stronger relationships with external stakeholders. In addition, data can be corrected or improved on the basis of user feedback, which provides the EA with a free quality assurance resource, ultimately resulting in higher quality published data.

Case Study: Benefits of open data for the public

Archaeologists are now able to use the EA's LIDAR data to locate Roman roads. Organisations such as Surfers Against Sewage use the open data to conduct scientific analysis and trend monitoring, to organise and campaign for remedial action and to hold water companies and the environmental agencies to account. ImageCat Ltd use EA flood risk data to help build forecasting tools for infrastructure and insurance sectors.

A broad stakeholder review and a second call for evidence highlighted the proliferation of data sharing initiatives but also demonstrated that these initiatives are largely sector focused, whereas the benefits would be greater if the data sharing were extended across the infrastructure sectors. For example, there is no centralised database of all underground assets so that when work needs to be done underground it can be unclear which existing assets will be affected. If this information were shared across sectors to select parties with authorised access, the sharing of this data would reduce the cost of underground work relating to the maintenance of existing assets and the construction of future assets. Similarly, if a utility in one sector brings its assets up to a certain level of protection against natural hazards, such as flooding, that information needs to be shared across sectors, so that the other utilities in the same geographical area can work to the same standard.

It is difficult to put a value on this kind of benefit but this is a live example of where money is not being spent efficiently, duplicating or working with incomplete data sets, thereby increasing the cost of infrastructure.

How to enable greater sharing of data

There are different classes of data relating to infrastructure: consumer data and non-consumer data (asset, network, organisational and environmental data). Any data sharing activity must take full account of the important security, privacy and legal considerations in this area, and it should never be the case that all data is be expected to be shared completely openly. Different levels of access for different parties will be appropriate. For example, only authorised personnel would need data pertaining to the location of strategic infrastructure hubs. Levels of access must be clearly defined and will require ongoing maintenance as uses and requirements evolve. Appropriate measures, such as the use of anonymization systems, will be necessary to protect fully any personal data.

Security

Security measures are required to mitigate risks from data loss on system operation. As data about national infrastructure becomes more interconnected it is essential to consider the security risk from terrorist threats, malicious activity and cybercrime. Where the intention is to share data for the public good, security measures need to be considered in the context of the appropriate risk management strategy and can enable greater sharing of data. Security measures need to be developed as part of an infrastructure data sharing strategy from the outset under the supervision of the Centre for the Protection of National Infrastructure (CPNI) and the National Cyber Security Centre (NCSC).

Legal considerations

Data can be shared within the boundaries of competition law, intellectual property rights and existing contracts. Where there is a public good argument for sharing that data on a basis that would be considered unlawful, then there may be a case for reconsidering the appropriate legislation and contractual arrangement. For example, the introduction of GDPR changes the way personal data can be used to provide protection to consumers whilst enabling sharing of data.

The NIS Directive (Directive on Security of Networks and Information Systems) provides legal measures to boost the overall level of cybersecurity ensuring: "a culture of security across sectors which are vital for our economy and society and moreover rely heavily on ICTs, such as energy, transport, water, banking, financial market infrastructures, healthcare and digital infrastructure. Businesses in these sectors that are identified by the Member States as operators of essential services will have to take appropriate security measures and to notify serious incidents to the relevant national authority. Also key digital service providers (search engines, cloud computing services and online marketplaces) will have to comply with the security and notification requirements under the new Directive".[71]

Privacy

The protection of personal data is governed by the General Data Protection Regulation (GDPR), which strengthens individuals' rights over the processing of their personal data. The UK Government is supportive of GDPR and the Commission does not make any further recommendations regarding personal data beyond what is considered under GDPR. Similarly, the Commission does not make any further recommendations with respect to the Data Protection Bill, which will make provisions for the regulation of the processing of information relating to individuals.

The treatment of personal data is an issue of particular sensitivity, and strong systems would be needed to ensure that any such data is fully protected. Nonetheless, it may be possible that some personal data relating to the use of infrastructure and location, when suitably treated and anonymised, may be shared in a legally compliant way alongside network, asset and organisational data. The Commission recognises that protecting personal data should be considered a key public policy objective and that effective data protection safeguards are a cornerstone of a vibrant digital economy. Data protection by design and data protection impact assessments can help ensure privacy and innovation and provide a strong basis for data to be used anonymously.

Case Study: General Data Protection Regulation (GDPR)

The GDPR 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. 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 Commissioner's Office (ICO), or a different ground relied upon, so as to comply with the GDPR.

The GDPR also provides individuals with enhanced rights in relation to their data and data portability. They may restrict or object to specific uses of data and request to have their data erased, with some exceptions.

Addressing security, legal and privacy considerations will be crucial to enabling increased sharing of data in the infrastructure sector, but will not be sufficient to ensure progress. There are also **technical, commercial and cultural barriers and challenges** to sharing data which need to be overcome to harness the greater benefits from sharing data about infrastructure.

Technical barriers occur where data is in different, incompatible formats, limiting its availability, quality or interoperability. Inadequate standards or poorly constructed systems present a costly but surmountable barrier to greater sharing of data for the public good. Coordinating interoperable formats for data and data access can reduce costs of data management for individual companies over time and induce greater benefits from the ability to share models and data in an interoperable format.

Modern history provides plenty of examples of how the move from fragmented regulations to uniform standards can unlock significant economic benefits. The introduction of the standardised shipping container in the post-war period helped to lower the price of shipping goods drastically, facilitating increased global trade. Similarly, the mass adoption of the barcode was important in allowing large retailers to scale up their services. As economist Tim Harford reflects: "Part of the difficulty was getting everyone to move forward on a system which didn't really work without a critical mass of adopters".[72] Prior to this, World War II provided an imperative to reduce barriers to exchanging advanced technological equipment within allied countries, helping to drive the standardisation of nuts and bolts, and laying the foundations for international standards.[73]

As institutions and companies take time to recognise the benefits of sharing data, with the appropriate data protection in place, the UK risks incurring greater costs by delaying the inevitable need for coordination. As companies continue to develop their own individual data strategies and move in the direction of developing digital twins, it will become harder to coordinate data and models without common standards and formats. It is less costly to adopt common standards proactively rather than to harmonise retrospectively. Now is the time to act.

Commercial barriers arise where organisations believe that data sharing may harm their commercial model or result in additional costs. Companies treat internal data as confidential on the assumption that sharing data may damage their competitive position and ultimately revenues. Even where a company recognises the benefit from sharing data, the company may be prevented from doing so by confidentiality requirements. For example, infrastructure companies who do not compete could share supply chain data to reduce costs and delays.

> ***"It is vital that we remove barriers to innovation and ensure that data is used in a way that is both safe and fair to individuals"***
>
> Industrial Strategy White Paper[lxxiv]

Driving national benefits

Communication of data to and from smart meters is being managed across Great Britain, by the Data and Communications Company (DCC), an independent licensed entity regulated by Ofgem. This approach enables short term and long term benefits in terms of efficiency, interoperability and the simplification and improvement of industry processes that will improve consumer experience (e.g. change energy supplier) and be a catalyst for a smarter, more resource efficient energy system. For example, network operators with approval from Ofgem will be able to use aggregated/anonymised data from smart meters to make more targeted investment decisions.

An example of where natural business cases could provide a barrier to data sharing is smart traffic management. The application of machine learning techniques to traffic data could provide insights to improve road capacity, emergency incident response and network-level resilience. However, this depends on the combination of 'big' datasets to produce a very large volume of relevant data for analysis.

At present, responsibility for managing the road network lies with local authorities, which also bear responsibility for all data relevant to their administrative area. This could put cross-boundary approaches to data collection and sharing (which could be more effective and less expensive) at odds with conventional business cases which are focused on value to the authority making the case. Understanding value capture from transport data at a national and network level will likely form part of the Department for Transport's (DfT) ongoing work in this area. The need for commercial entities and public authorities to share data in the transport space is becoming more pressing.

Cultural barriers persistently restrict the sharing of data

Cultural barriers create a 'closed mentality', resulting in siloed thinking. This may be exacerbated due to a lack of understanding from network operators, utilities, regulators and consumers regarding the value of sharing data and opportunities to do so. A risk-averse attitude towards sharing data, stemming from understandable concerns about the negative security, privacy and commercial implications of sharing data, may limit the extent to which parties are willing to share data. As set out above, however, if these privacy and security issues are taken fully into account from the start, they can be addressed in a manner which still enables significant benefits from enhanced data sharing.

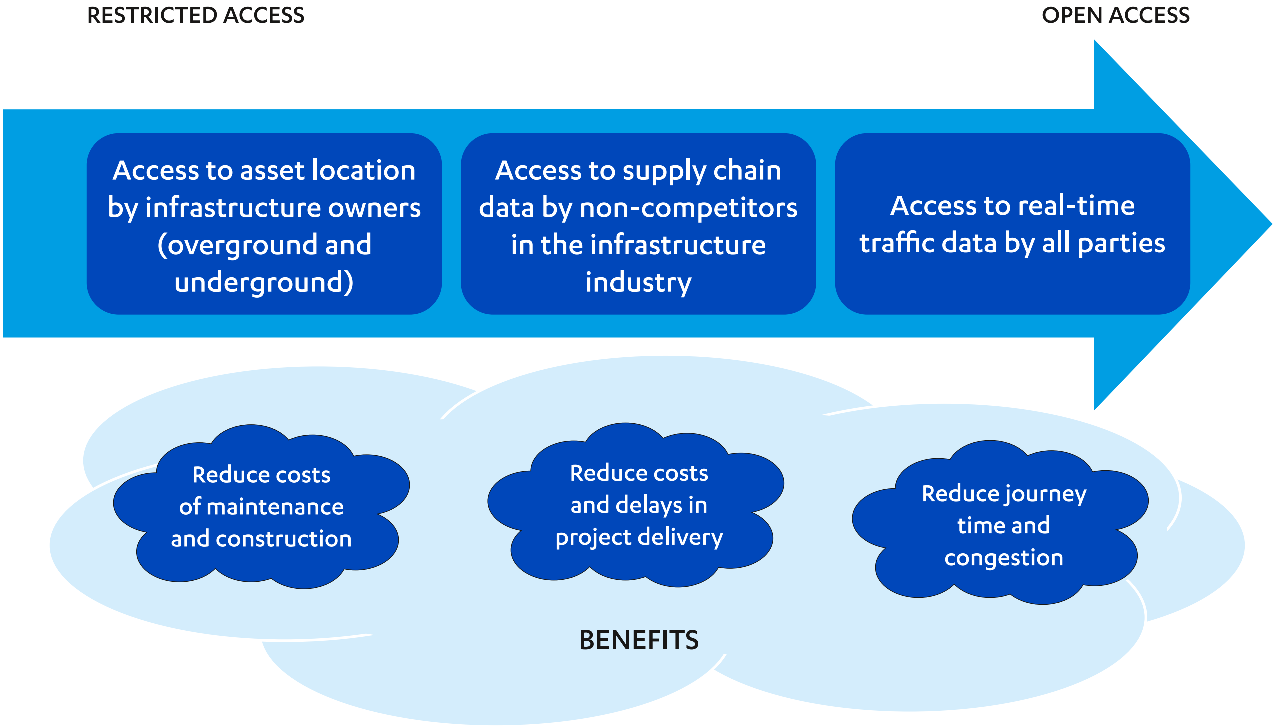
Cultural barriers in the rail industry

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.[75] As a result, many technological advances have been missed.

Cultural and commercial barriers are putting the UK behind

An infrastructure network operator or utility may treat much of its data as commercially confidential unless forced to reveal that data by the sectoral regulator. Companies do not fully realise the value of their data because confidentiality is a default setting. Network operators, utilities and regulators are resistant to greater sharing of data and argue that the benefits must be demonstrated before more data is shared. This demonstrates that the benefits are not well understood. Furthermore, this approach makes it impossible to take account of benefits beyond those which can be identified through a proof of concept. New and unpredictable uses of shared data will generate external benefits which cannot currently be incorporated into a formal business case. The original case for opening up TfL's data, for instance, would not have incorporated the full range benefits from the applications that have arisen since.

Where the costs of collecting, processing and sharing data are greater than the estimated private benefit, a company will not rationally act to share that data. This means that when such sharing would also have great public benefit, society may lose out. By refusing to share data, a private company or organisation keeps control of that data as it grows. As set out in this report, as the volume of data increases and machine learning techniques are applied, the quality of the data improves and so becomes more valuable. Thus there are increasing returns to data, which if retained in the private sphere, will remain as narrow returns to the private company rather than wider returns to the economy as a whole.



Data sharing examples and benefits:

The potential directions in which infrastructure data may be shared in the future, in an open data or controlled manner, are diverse. The examples below illustrate potentially beneficial links:

- Utilities may be able to open access to suitably packaged records of planned renewal, upgrade and maintenance activities. This would provide a real-time portal for consumers to minimise disruption to their planned travel.
- The sharing of asset shape and location data between sectors would allow simulation of future schemes to reduce spatial collisions. This will enable better design of systems in increasingly tight spatial areas, e.g. cities.
- Controlled interfaces may be used to enable data sharing between organisations in sectors, which may need to share detailed technical and operational data. This will reduce repeated and costly manual actions to access data, allowing more real-time automation.
- Wider industry users may need to access detailed asset data for engineering analysis and system maintenance. By having customisable controls, an infrastructure data framework may be able to grant access to potentially sensitive data for known users.
- Opening access to curated views of infrastructure data (which could include operational history, low detail spatial information, aggregated usage and indicative asset types) will support innovation, facilitating the development of technologies which will be applicable across sectors.

Now is the time for the UK to start to fully explore the potential from reducing commercial confidentiality requirements. Greater availability of data will enable the application of machine learning, an area where the UK plays a leading role.

A new approach would require companies to make the case for keeping certain datasets confidential in order to minimise levels of commercial confidentiality and maximise the amount of open data and therefore the potential for insight and innovation. As set out above, there are good reasons for keeping some data confidential, and the crucial point is that **greater benefits from sharing data can be enabled from allowing different levels of access to different types of data.**

The Government is making progress in sharing more data

The recent **AI Review**[76] makes recommendations to address the barriers to sharing data and ultimately to facilitate the development of AI in the UK. The recommendations, which the National Infrastructure Commission fully endorse, include:

- Development of data trusts, to improve trust and ease around sharing data.
- Making more research data machine readable.
- Supporting text and data mining as a standard and essential tool for research.
- A programme to support public sector use of AI.

This report sees the development of data trusts, which would need to ensure that personal data is processed in accordance with the GDPR, as providing greater reassurance to parties sharing data, and as conducive to reducing the requirements for commercial confidentiality and enabling greater sharing of data.

AI Review:

A recent independent review work on the growth of AI technologies, undertaken by Professor Dame Wendy Hall and Jérôme Pesenti,[77] has identified £630bn of benefits from the application of AI technologies in the UK (up to 2035). A number of recommendations are made to help ensure that the opportunities identified in AI don't turn into barriers:

- Efforts to improve coordination: Data trusts to be established to improve data sharing, supporting the exploitation of AI driven processes, making more data machine readable, ensuring that IP transfer is possible, and supporting widespread text and data mining. This will help to break down the technical, commercial/legal and cultural barriers.
- Ensuring support for public sector use of AI: Establishment of a council to support this and AI more generally, which engages with the Industrial Strategy Challenge Fund (ISCF) and Small Business Research Initiative (SBRI). This will help to reduce cultural and legal barriers to innovation.

The report also recommended supporting opportunities for boosting the UK's capability further in terms of skills with funding allocated to university level qualifications in AI. AI is a key tool in realising infrastructure efficiencies, and as such a vibrant UK AI community will support future infrastructure.

The government will create a **new Centre for Data Ethics and Innovation** to enable and ensure safe, ethical and ground-breaking innovation in AI and data-driven technologies.[78] This world-first advisory body will work with Government, regulators and industry to lay the foundations for AI adoption. The Government will invest over £75m to take forward key recommendations of the independent AI review, including exploratory work to facilitate data access through 'data trusts'. The Government will create new AI fellowships, and initially fund 450 PhD researchers, to secure the UK's leading position in the global AI market.

The Government will be investing £40m per year for the next two years to establish a **Geospatial Commission** to maximise the value of all UK Government data linked to location.[79] Its first task will be to work with Government and the Ordnance Survey by May 2018 to open up freely OS MasterMap data, in particular for the benefit of UK-based small businesses.

The new Commission will draw together HM Land Registry, the Ordnance Survey, the British Geological Survey, the Valuation Office Agency, the UK Hydrographic Office and the Coal Authority with a view to:

- Improving the access to, links between, and quality of their data.
- Looking at making more geospatial data available for free and without restriction.
- Setting regulation and policy in relation to geospatial data created by the public sector.
- Holding individual bodies to account for delivery against the geospatial strategy.
- Providing strategic oversight and direction across Whitehall and public bodies who operate in this area.

The Department for Transport is advancing work to understand how data sharing and standards should develop within the context of a roadmap for the future traffic management of UK roads. This includes reviewing the continued likely effectiveness of Urban Traffic Management Control (UTMC), a UK initiative which has helped develop standards and supported Local Authorities and suppliers as they integrate them into their traffic management systems. DfT's roadmap will ensure that standards-setting procedures are fit to meet the challenge of connected vehicles and next-generation intelligent traffic systems within an open data platform.

"If data sharing processes were digitised and automated, this could raise utility companies profits by 20-30%, and increase staff productivity by 15%"[lxxx]

McKinsey & Company

Government-led initiatives such as the Geospatial Commission, the Centre for Data Ethics and Innovation and initiatives at departmental level, for example at the Department for Transport, need to be coordinated and joined up to exploit the full benefits from sharing data.

4. A coordinated approach

4. A coordinated approach is required to enable benefits from data sharing: a digital framework for data on infrastructure

The UK needs a digital framework for data on infrastructure to harness the benefits from sharing better quality information about its infrastructure; how it is used, maintained and planned. A digital framework will enable better understanding of interdependencies between infrastructure sectors and help to break down silos. As private companies increasingly recognise the benefits of sharing performance data (anonymised in some contexts) with competitors and regulators, opening up more information to customers and in some cases selling specific datasets to realise the true value of data, it is likely that network operators and utilities will look to drive the increased sharing of data about infrastructure. But this is not yet happening across infrastructure sectors in a coordinated way. There is an opportunity right now to set a framework that will establish common standards of interoperability for sharing data about infrastructure so that the UK can realise the benefits sooner and at a lower cost.

Case Study: A New Partnering Arrangement at Southern Water

Southern Water (SW) are interested in piloting a new data platform, and a new approach to building a platform, that will integrate and leverage expertise from a number of partner organisations across utilities, Government and the technology sector. They are working collaboratively in developing a new partner framework that can rapidly deliver innovation, technology and data. This new partnering framework will involve each partner will contributing funding. This mechanism and approach will provide the partners with several benefits:

- Firstly, and most importantly, it will provide a more open and transparent monitoring of the environment for the regulators of the water sector. This will enable the regulators, with partners, to create their own bespoke dashboards and reports, help reduce regulatory reporting burdens, and provide greater and more assured access to data.
- Secondly, the data platform will provide an 'open data' forum, which provides public access to non-commercial/sensitive datasets. In this way, SW see the platform as a tool both for customer engagement and to facilitate innovation for the sector by making datasets available to innovators, suppliers and academics to develop and test their products or hypothesis.
- Lastly, this platform and partnering approach will offer the opportunity to be curious, to be agile and to innovate. The aim is not to 'solutionise' but to explore opportunities in a low risk, low cost approach.

Although SW are leading on the development of this platform and approach, the SW vision is that this becomes a water sector-wide platform. This could allow all water companies and other data suppliers to participate, share data and, importantly, share benefits.

The initiative driven by Southern Water, which is currently early stage, provides an excellent example of a utility driving data sharing and innovation. Similarly, Future Cities Catapult, British Geological Survey (BGS) and the Ordnance Survey (OS) have been working together to gain a clearer understanding of where assets are underground through Project Iceberg.

Case Study: Collaboration to improve efficiency – Project Iceberg

Project Iceberg is an exploratory project being undertaken by Future Cities Catapult, British Geological Survey (BGS) and the Ordnance Survey (OS). The remit of the project is to understand what lies beneath the subsurface, who holds the data about it, who accesses it and how a single digital subsurface platform could drive efficiencies for the built sector, reduce traffic congestion and improve health & safety. Ground risks are listed as one of the main causes of project delay (50 per cent) and, according to TfL, road works account for 38 per cent of the most serious and severe traffic disruptions across London, at a total cost of £752m. To date there have also been approximately 12 deaths and 600 injuries due to strikes. This suggests a case for creating a shared geo-spatial framework for sharing underground infrastructure asset information across sectors. In order to do this, Project Iceberg have identified a need to develop a national standard that addresses commercial sensitivities and security risks concerning subsurface data sharing, which will ultimately underpin and support an integrated data framework. This framework will encourage the sharing of interoperable data that can support multiple infrastructure benefits and efficiencies.

This kind of initiative needs coordination under a digital framework for data on infrastructure with the many other initiatives across the infrastructure sector and the work of the new Geospatial Commission and the prospective data trusts through the Centre for Data Ethics and Innovation. The Geospatial Commission provides an excellent example of Government setting direction in an area where there was enthusiasm but an absence of coordinated effort. There are many initiatives looking to coordinate data in a previously uncoordinated space. But these initiatives themselves require a framework to ensure consistency. There are many initiatives which look to coordinate data to improve the way infrastructure is used, maintained and planned but the missing element is a digital framework to bring all these current and future initiatives under a single umbrella. The Commission therefore makes the recommendation to establish a digital framework led by the Centre for Digital Built Britain (CDBB) to coordinate inputs from a range of stakeholders.

"There are inconsistencies across the quality and format of data, as well as poor data sharing across the utilities industry, with no official bespoke standard to guide practice" [lxxxi]

Institution of Civil Engineers

"This is not about Government investing in technology, but in setting and maintaining the standard" [lxxxii]

Turner & Townsend call for evidence response

Recommendation 1: The Government should task the Centre for Digital Built Britain (CDBB) with the establishment of a digital framework for infrastructure data, drawing together key organisations and existing initiatives both large scale (BIM) and smaller scale:

a. A Digital Framework Task Group for infrastructure data should be established with a high-profile chair who can act as a national champion for this agenda.

b. CDBB should set out a roadmap to a digital framework to develop standards and formats for collating and sharing data.

c. Key organisations which should be involved in the Task Group and in developing the framework include the Alan Turing Institute, Infrastructure Client Group, Construction Leadership Council, Infrastructure and Projects Authority, Cambridge Centre for Smart Infrastructure and Construction, Project 13, Office for National Statistics, Ordnance Survey, Open Data Institute and the recently announced Geospatial Commission and Centre for Data Ethics and Innovation. Wider consultation input should also be sought from the digital twin working group referenced in section 5.

d. On standards development, in order to understand the existing standards landscape thoroughly at both a national and supranational level and to ensure agility, CDBB should consult extensively with industry on current behaviours and future requirements for how different infrastructure sectors and sub-sectors use data. CDBB may wish to commission an external standards organisation to conduct some of this work on their behalf.

e. IPA and ICG should be engaged closely when designing final data standards and performance measures. The ICG amongst others will also be important to consult with respect to appropriate safeguards for commercial confidentiality.

f. There should also be close collaboration with CPNI and NCSC on security requirements and levels of access and to develop standards for security, consistent with the objectives of the framework and an agreed approach to risk management.

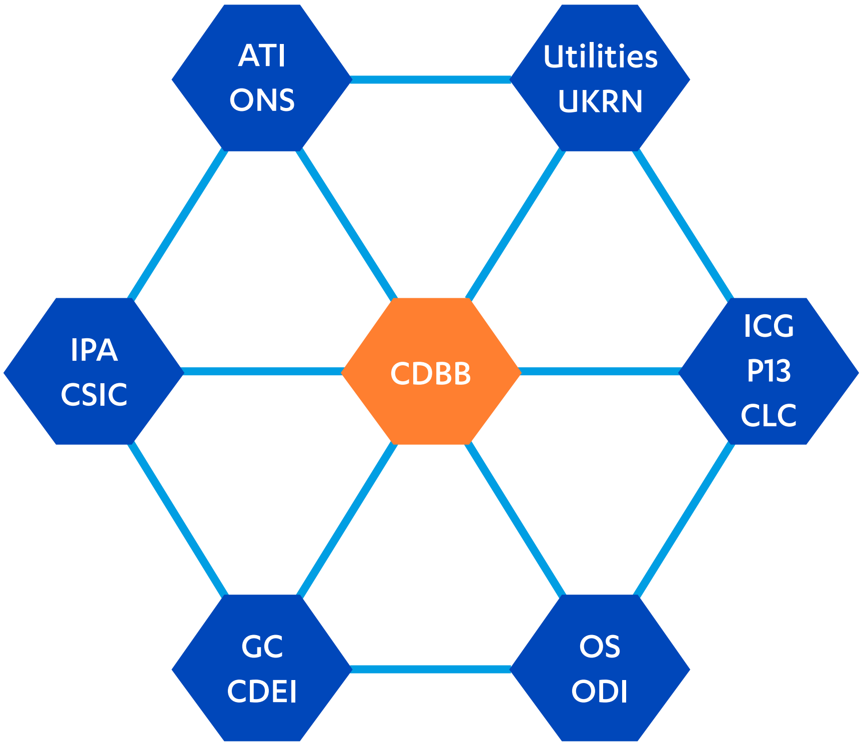
g. In order to support effective implementation, CDBB should also lead a scoping exercise for a framework to assess industry progress in adopting and using the framework and to measure innovation.

h. CDBB should complete these actions and provide a public report on progress by 1 September 2018, setting out their recommendations for next steps. The National Infrastructure Commission would assess this as part of its wider role in monitoring progress against its recommendations.

> *"Any physical infrastructure that is built has to be planned for and must conform to standards...[we] need to think of data in exactly the same way"*
>
> Chartered Institution of Highways & Transportation call for evidence response

> *"The development and use of voluntary consensus industry standards has an important role to play in leveraging rapid innovation by bringing stakeholders together, sharing expertise and establishing good practice.*
>
> ....
>
> *The development of industry capability through knowledge exchange and dissemination of good practice is a clear benefit"*
>
> British Standards Institution call for evidence response

Digital Framework



Institutions integral to the digital framework (as an evolving list)	
Centre for Digital-Built Britain (CDBB)	Utilities and network operators
Cambridge Centre for Smart Infrastructure and Construction (CSIC)	Regulators
Infrastructure Client Group (ICG)	Open Data Institute (ODI)
Project 13	Office for National Statistics (ONS)
Infrastructure Projects Authority (IPA)	Ordnance Survey (OS)
Construction Leadership Council (CLC)	Geospatial Commission (GC)
Alan Turing Institute (ATI)	Centre for Data Ethics and Innovation (CDEI)

There is interest and enthusiasm from the private and public sectors to coordinate and collaborate. To make this happen successfully, engagement from the network operators, utilities and regulators is key. The output of this framework should ensure an agile approach to standard setting, based on input from these key parties. CDBB may wish to commission an external standards organisation to conduct some of this work on their behalf. The framework will drive the requirement for rigorous data curation. The Digital Framework Task Group should set out the requirements for data custodians responsible and accountable for the maintenance of the common data environment.

Coordination must be designed such that the generality associated with a national framework does not lose the efficacy of detail of the most individual, tailored data project. The requirement for an overarching framework to pull all the initiatives together should not result in a shift towards generality. For example, the conceptual development of a digital twin model may require skills and expertise drawn from geography and related fields, dealing with how infrastructure and the built and natural environment all fit together but more sector specific skills such as structural rail engineering and AI will still be required to input to the model and to assess the validity of outcomes.

Industry engagement

A national champion, working as the Digital Framework Task Group Chair, is needed to drive this digital framework forward and coordinate across sectors. There is a common thread of willingness to share ideas and data which is hampered by commercial confidentiality. As explained in the previous section, the current default is to keep all data commercially confidential rather than exploring the potential benefits of sharing and in some cases charging for it. A move towards minimum levels of commercial confidentiality requires a cultural shift to be led by the national champion working across industry with the ICG. The Commission therefore recommends that ICG and the national champion lead industry engagement in and alignment with the framework and minimum levels of commercial confidentiality.

Recommendation 2: The Infrastructure Client Group (ICG) and the Digital Framework Task Group Chair should lead industry engagement in the framework and cultivate a shift towards minimum levels of commercial confidentiality.

a. ICG should report to CDBB on current industry compliance with minimum levels of commercial confidentiality agreed with CDBB in recommendation 1.

b. ICG should work collaboratively with industry and the Digital Framework Task Group to identify opportunities to make data available and reduce the unnecessary use of commercial confidentiality (e.g. through reviewing and revising existing digital contracts), and should set out an agreed plan with milestones towards achieving the proposed shift.

c. ICG should report on progress in reducing the application of commercial confidentiality to infrastructure data by December 2018.

Role of regulators

The economic regulators are in a unique position to understand what network operators, utilities and consumers can get out of greater sharing of data. The perspective held by a regulator should enable oversight across an entire industry. But it does not enable oversight across the entire infrastructure of the UK. A regulator's concern for consumer interest covers only the consumers within that sector. For example, Ofgem's principal objective is to protect the interests of current and future electricity and gas consumers. If an electricity or gas network company cuts into a water pipe supplying that town, Ofwat must get involved to protect the interests of water customers. If there were a regulatory mechanism to ensure that data about where all the infrastructure networks are is shared with other network operators and utilities (at an authorised level of access, taking into account security considerations) then a cross-sector system level view could be taken which could avoid the energy company cutting into the water pipe and cutting off water supply and delaying the connection of additional electricity resource.

Regulators need to take a more cross sector approach to encourage open data within and across industries and enable greater innovation. This will help to multiply the benefits of having better data about national infrastructure. The Industrial Strategy Challenge Fund[83] will establish a £10m Regulators' Pioneer Fund to support UK regulators to develop innovative approaches to emerging technologies. This fund is not allocated to any specific sector and therefore regulators can pool efforts to harness the benefits of new technologies.

Utilities and network operators need to drive this framework as the data owners and service providers to the benefit of consumers. The active participation of utilities and network operators is crucial to ensure the framework serves the purpose of facilitating data sharing to the ultimate benefit of consumers. To do this it needs to be flexible and responsive to the changing needs of industry and consumers. Therefore the Commission recommends that the Digital Framework Task Group work with the UK Regulators Network as set out in Recommendation 3.

> ***"The regulatory framework needs to encourage innovation with long-term benefits, rather than focus on the current regulatory period"***
>
> United Utilities call for evidence response

> ***"One outcome of the Northumbrian Water Group Innovation Festival in July 2017 was consensus that a common infrastructure map which facilitates the wider sharing of underground asset information could greatly contribute to greater resilience and enhanced industry innovation"***
>
> Ordnance Survey call for evidence response

Recommendation 3: The Digital Framework Task Group (see recommendation 1) should work with the UK Regulators Network and relevant Government departments to review and, where possible, strengthen the role of economic regulators in driving up the quality and openness of infrastructure data. This should include:

a. Participation by the UK Regulators Network in the formulation of the digital framework set up by CDBB to ensure that it is effectively aligned with regulatory work on innovation and data.

b. Assessment of the potential role of regulators and of possible barriers within current regulatory frameworks regarding:

- Ensuring compliance by regulated network operators and utilities with the national framework and adherence to data collection standards and formats.
- Sharing of data to inform better understanding of asset performance and user experience.
- Sharing of data across infrastructure sectors and the value chain to enable greater innovation in the development of new technologies and data management focused on better asset management and increased productivity.

c. Support for CDBB's engagement with network operators and utilities around the provision of data of verified quality to input to the development of a national infrastructure digital twin over the long term.

d. Identification of relevant areas for further trials or studies to enable regulators, and regulated industries, to understand and demonstrate how monitoring technologies and data can support cost-effective maintenance decisions and proactive asset management, working with relevant research organisations.

The Commission makes these recommendations in full acknowledgement of the complexity of delivery. Coordination is not currently happening because it is not easy. The challenge to draw together the key players set out in the framework in order to achieve greater sharing of infrastructure data is complex and requires a willingness to collaborate through a new cross sector approach.

> ***"As an industry we have the opportunity to define a consistent standard for sharing of supply chain data. This will enable effective and efficient access to data to make a step change in benefits delivered for both clients and the government"***
>
> Heathrow Expansion

5. Digital twin

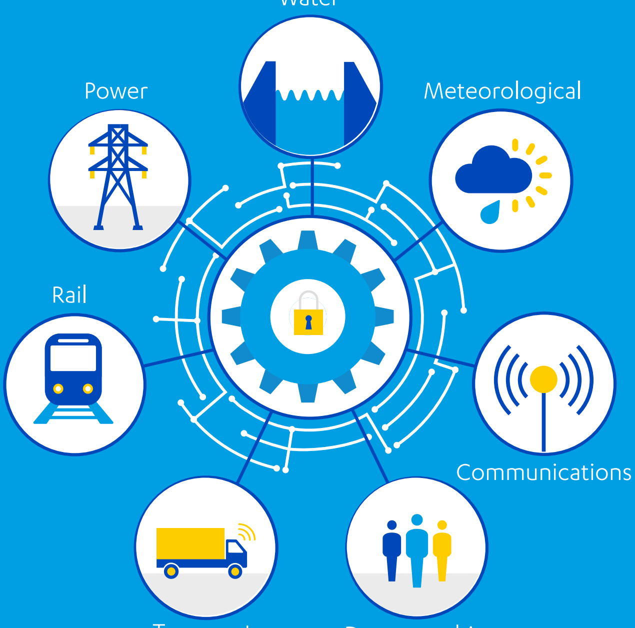
DIGITAL TWIN: A PIONEERING APPROACH

WHAT IS A DIGITAL TWIN?

A digital twin of the UK would bring together data in a computer model to simulate the entire infrastructure system

A digital twin pilot would demonstrate the benefits from sharing infrastructure data

Examples of data sources



Machine learning, big data and other technologies process data and produce insights

Security is at its core: Not all information should be in the public domain

A digital twin that unifies these separate systems can answer questions such as:

Is it possible to:

- Avoid building a new hospital carpark by managing appointment times and traffic flows?
- Reduce energy consumption by 10% per person over six months?
- Assess the impact of closing a road in the event of a water leak?

WHY DOES THE UK NEED A DIGITAL TWIN?

A national digital twin would enable the UK to develop a richer understanding of the way our infrastructure works and optimise it, so government and industry can make more informed decisions about the future.

- Optimise use of resources such as energy and water
- Reduce disruption and delay for transport and ease traffic flow
- Increase resilience in the face of terrorist attacks
- Boost quality of life for UK citizens
- Improve responsiveness in natural disasters

AN OPPORTUNITY TO LEAD

The UK has world leading data science research capability and AI expertise. This gives us the opportunity to be at the forefront of digital twin technology.

5. A digital twin model to manage, plan, predict and demonstrate

There is a concentration of data science and AI expertise in the UK. The AI Review sets out an approach to nurturing this capability to the best advantage of the country. The Industrial Strategy White Paper[84] establishes the Grand Challenge "we will put the UK at the forefront of the AI and data revolution", and emphasises the gains that AI can bring to the UK economy. The Grand Challenge sets "we will make the UK a global centre for AI and data-driven innovation" as a key priority.

The UK has a particular opportunity to demonstrate global leadership in digital twin technology, through the development of a national digital twin of its infrastructure, aided by the big data and machine learning approaches in which it has significant strengths. Such a digital twin would bring together individual infrastructure models capturing data on national infrastructure and the interdependencies between infrastructure systems, supporting the development within the infrastructure sector of a data-driven economy. Advanced approaches including data science, machine learning and predictive analytics will contribute to each model of an infrastructure asset, network or system. This modelling, analysis and simulation will power predictive asset maintenance, support planning decisions and enable performance optimisation. For example, a digital twin model could:

- Integrate and prioritise the maintenance needs of transport networks (for example by road and rail) from a spatial perspective to minimise disruption to consumers.
- Overlay planned infrastructure on existing infrastructure to show interdependencies and evaluate the optimal timing of investment (the GLA London Infrastructure mapping application provides an example of this kind of tool).
- Identify efficiencies in energy use which can be tested across different sectors through simulations.

Through the input of verified data to these digital twin models it will be possible to develop a richer understanding of the way the infrastructure system works. Applying deep learning techniques may offer extended predictive capabilities. In the future, it may be possible to ask questions of the model like:

If the population of London were to increase by 50 per cent by 2050 how do we change the way we use transport?

How can we reduce energy consumption by 25 per cent per person by 2050?

How do we avoid building another carpark in the city centre? How can we reoptimize the traffic flow and reduce congestion?

A digital twin model can provide insights beyond what is currently seen with existing infrastructure models and can be used as a tool to aid decision making. This twin will capture data on infrastructure, which can be augmented with models, predictive asset management approaches and advanced data science to optimise the performance of infrastructure systems.

> *"The Digital Twin allows you to have access to all archive and real-time data associated with the national asset within a single environment, opening-up access to that data and insight across the asset owners' business"*
>
> Transport Systems Catapult call for evidence response

Case Study: Virtual Singapore:

Virtual Singapore, expected to be deployed in 2018, will be a platform providing a dynamic, 3D virtual model (a digital twin) of the urban areas of Singapore. It will incorporate static and real-time data on e.g. climate, demographics, terrain attributes, energy consumption or building elevation. It will also contain semantic information such as the composition and materials of buildings. It has been funded by Singapore's National Research Foundation, and is being delivered in conjunction with a French company, Dassault Systèmes.

Virtual Singapore is expected to deliver benefits to the quality of life of Singapore's citizens through encouraging collaboration across all stakeholders (public/private/academic) with an interest in the city's built environment. It will enable more efficient and integrated urban planning, as well as response preparation for extreme scenarios. The project faces complex challenges around data security, change management, and data standards.

A digital twin pilot project

The Commission recommends that Government support a digital twin pilot to test out the benefits of a digital twin approach. A dedicated proof of concept project will explore the potential benefits from developing digital twin models of infrastructure and will set out some of the challenges and security risks involved in doing so. The Commission proposes a digital twin pilot project to help provide additional evidence, and support the future approach to digital twin technology. This pilot project should aim to:

- Demonstrate how a digital twin may be used to make better decisions in the management of national infrastructure.
- Identify the most plausible organisational, regulatory and technical approaches required to facilitate an effective digital twin for the UK.
- Continue to engage all public and private sector stakeholders. This should build on the engagement undertaken in this study, acknowledging that broad acceptance of this technology is the key to future success.

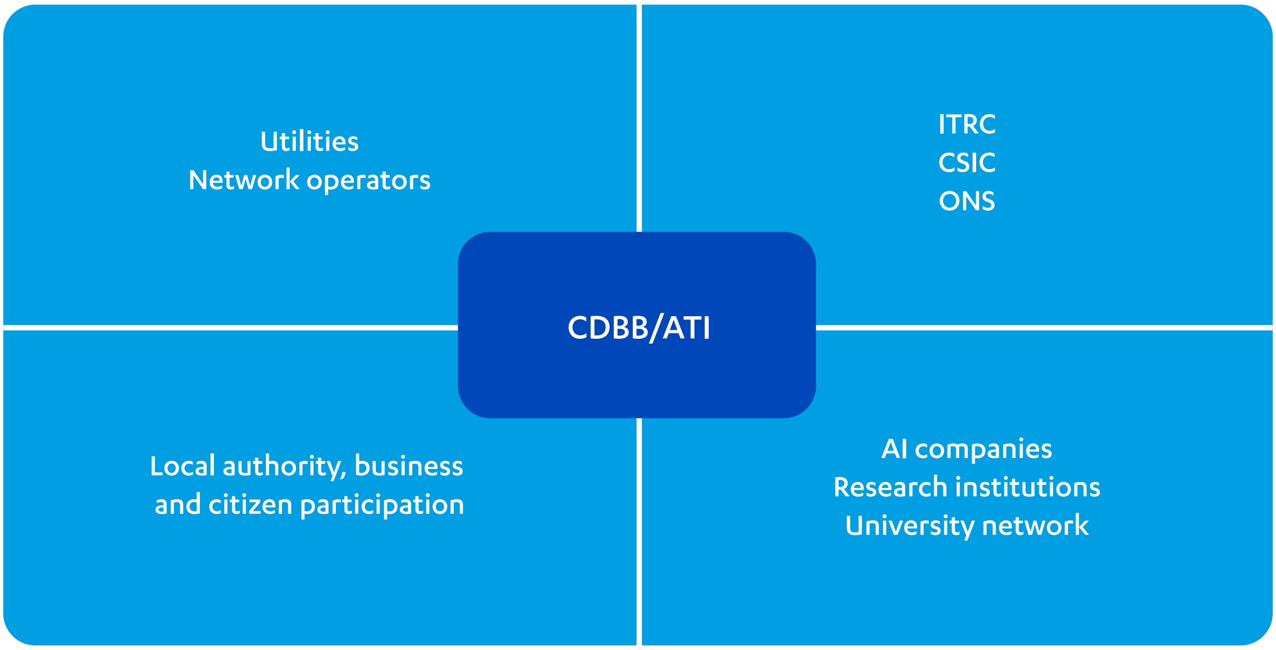
The scope of a national infrastructure digital twin is extremely broad, covering multiple sectors, systems and users and it will require substantial inputs from infrastructure experts, data scientists and supporting systems. In order to narrow the scope, an effective pilot study will focus only on infrastructure in one specific geographic region and would apply machine learning to provide answers to a specific infrastructure problem, for example:

- How far could energy consumption can be minimised? Could it be reduced by 10 per cent per person over 6 months?
- What reduction in average car journey time can be achieved?
- What efficiency be gained by coordinating street works?

Through the development of the pilot it will be important to consider security aspects and how to engage CPNI and NCSC in risk assessment. The pilot should consider the requirement for a data protection impact assessment if the use of personal data is involved. It will be necessary to consider how the pilot study sits within the existing body of data about infrastructure which can be used to inform system effect and how the model can be scaled up to a national version through a collection of regional models. The Commission recommends that the pilot study run over a period of six months and a detailed appraisal be undertaken to evaluate the impact such digital twins can have along with identifying the best organisational, commercial and technical approach to move the pilot to full scale deployment. The pilot study should also lay the foundations for assessing the cost of developing a national digital twin and making further recommendations to Government.

The Commission has identified a number of key stakeholders whose collaboration would be required in developing these models through a digital twin working group.

Digital twin pilot project working group



This is a unique opportunity for the public, private and academic realms to collaborate on a project which could put the UK at the forefront of digital twin technology and make use of our world-class machine learning talent. The task of modelling existing infrastructure in the UK is arguably more difficult than in other countries with newer infrastructure systems, yet ultimately all national digital twins will need to encompass existing infrastructure and the UK has the combination of academic and private sector data science and AI research and development capability to lead in this area. There may also be the potential for synergies with ongoing public sector initiatives, such as the Electronic Property Information Mapping Service (e-PIMS).

> ***"A national digital twin is important to provide modelling and forecasting to an ageing UK infrastructure. It should provide a platform for using data and data science to validate and reinforce existing mathematical models of complex engineering systems and assist in the development of new models"***
>
> Alan Turing Institute call for evidence response

High quality data is a key input to a digital twin and the Commission recommends the use of real data to address existing challenges in the development of the digital twin pilot.

Exeter City Futures: accessing and sharing data

A range of projects are being coordinated under the Exeter City Futures banner differentiated by 1) a scientific focus on whole-system models across a small geography and core population of only 125,000; 2) the use of new mechanisms to share personal and commercial data for innovation; 3) the leveraging of data from existing connected devices to provide immediate scale and replicability; 4) a technology-agnostic approach focussed on development of cost-effective solutions to high-value problems. The ongoing projects aim to unlock data silos and pool city data from IoT devices, create trust and support for data sharing by integrating privacy and security by design, and providing the procurement, collaborative innovation and funding mechanisms to accelerate commercial solutions to city-wide problems to help save money for the public sector and citizens.

A successful digital twin pilot project will demonstrate the value of better coordinating and sharing data about infrastructure assets. A digital twin will be able to offer solutions to difficult decisions arising from population growth, congestion, climate change and the development of new technologies. Collaboration across academia, the technology industry, network operators, utilities, relevant public sector bodies and consumers will be key to making a digital twin pilot project work.

Recommendation 4: CDBB should collaborate with the Alan Turing Institute (ATI) and the UK Infrastructure Transitions Research Consortium (ITRC) in pioneering digital twin models with predictive capability in the UK. This initiative should draw upon the AI expertise concentrated in the UK across universities and the public and private sectors.

a. CDBB should work with BEIS and other potential funders to take forward a digital twin pilot project to explore and experiment with the benefits of building a digital twin of a specific geographical area. CDBB, ATI and ITRC should draw upon input from the digital twin working group identified in section 5.

b. A project review of the pilot digital twin should be completed by October 2018. This should consider the lessons learnt for any future development of larger-scale or more complex digital twins, and the most effective institutional structures to support continuing progress in this area.

Definitions

In other usage there may be differing meanings from those used in this study, these are acknowledged where there may be significant risk of confusion. Where precise technical definitions are not material for the interpretation of this study they have not been provided.

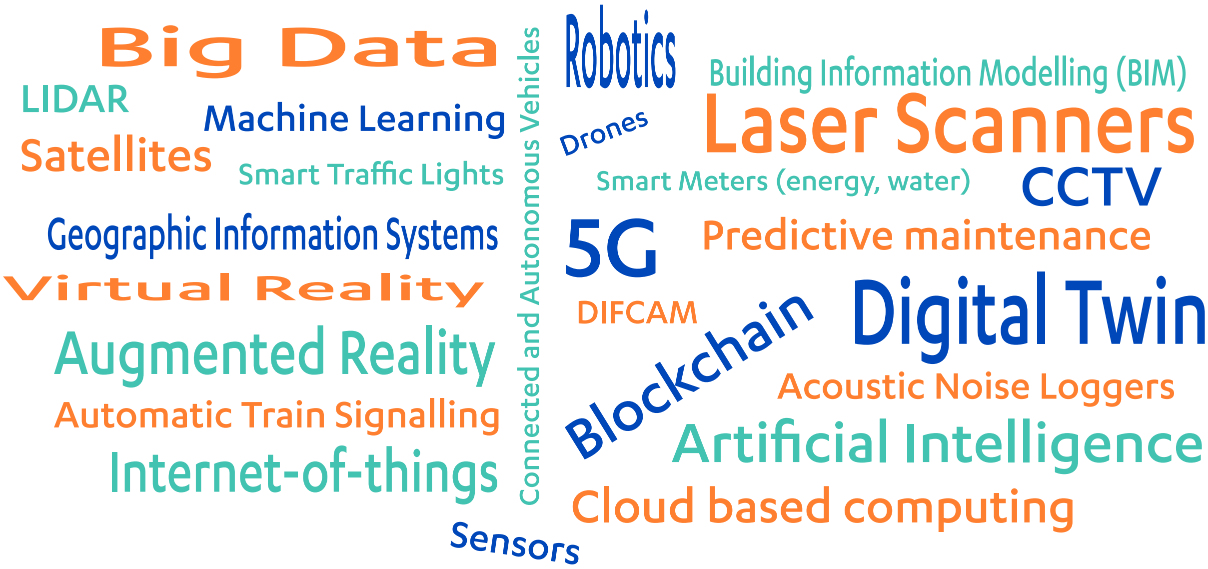
Term	Definition/Meaning
Acoustic Logger/Logging[85]	A system which monitors noise from flow in a pipe, at intervals along its length, to infer the location of a leak.
Algorithm	A series of steps which may be used to perform a specific activity, typically used by a computer.
Analysis/Analytics	A process used to gain an understanding from an item, which may be specifically applied to data.
Artificial Intelligence (AI)	An umbrella term for the science of making machines smart. This includes technologies that can perform complex tasks traditionally considered to be out of reach of computer programs, such as visual perception, speech recognition, natural language processing, reasoning, learning from data and solving all manner of optimisation problems.[86]
Asset Management[87] (Reactive/Proactive)	Practices associated with ensuring that assets remain usable. This may respond to the state of the asset (reactive) or take action in light of the expected future state of the asset (proactive).
Assets	Any item which has economic value; all Infrastructure may be considered to be an asset.
Autonomy/Automation	The ability to make decisions independent of human input. typically applying artificial intelligence approaches.
Big Data[88]	Data sets which are large enough to preclude the use of 'traditional' processing techniques.
Building Information Modelling (BIM)	The use of digital representations of a system, to support construction and ongoing asset management.[89]
Cloud Solutions	The use of computational resource as shared infrastructure, typically providing ubiquitous access and economies of scale.
Critical Failures/Asset Failure	When an asset can no longer meet its primary purpose.
Cyber-Attack	The targeting of a digital system for a malicious purpose.
Data/Open Data	A set of values that represent some aspect of a system or process. Open Data is the concept that data should be freely available for all to use.
Data Feeds	Connections, typically computerised and to other systems, which provide data.
Data Mining	The exploration of patterns in data sets, often done using machine learning.
Data Processing Technologies	Generically refers to all big data, machine learning, data science, cloud, modelling, predictive analytics or other methods that take input data and act on it to derive insight.
Data Science/Data Scientist	The inter-disciplinary field, or member of this field, which plans data capture, organises large data sets and extracts insight from data.
Data security	Data protection measures to ensure that data is available for appropriate use.

Term	Definition/Meaning
Data-driven economy	A digitally connected economy that realises significant value from connected, large-scale data that can be rapidly analysed by technology to generate insights and innovation.
Deep Learning	An approach to machine learning that uses architectures consisting of several layers of artificial 'neurons'. This is inspired by the human brain and has recently led to many AI breakthroughs.
Diagnostic Monitoring	Monitoring which is tailored to identify the cause of a particular behaviour
Digital Signalling	Smart Infrastructure concepts applied to rail networks, using digital systems to control train moments.
Digital System	A system which is formed of a series of computers (or other digital electronic circuits).
Digital Twin	A representation of a system which mimics its real-world behaviour (and, in some cases, the surrounding environment). This may typically be a real-time updated collection of data, models, algorithms or analysis.
Dynamic Response	The motion of an object which results from a changing force.
Engineering Information	Data which may be useful to understand the performance of a system.
Fault (Condition/State)	When a system has a part or sub-system which is not fully functional.
Fibre Optic Sensors	A sensor system which uses optical fibres, typically with many individual sensors along a fibre.
Framework/Digital Framework	An infrastructure information management landscape and data environment which will adequately define parties, processes, information and technology; a reference point to facilitate data use in line with security, legal, commercial, privacy and other relevant concerns.
General Data Protection Regulation (GDPR)	A regulation that lays down rules relating to the protection of individuals' personal data.[90]
Geographic Information Systems (GIS)	A database or other software system which is designed to interact with spatial data, typically to support wider understanding of a system or object.
Health Monitoring (Structural)	The use of monitoring for asset management purposes, which may be specifically linked to the characterisation of the response of an object to applied forces.
Information security	The protection of information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction in order to provide confidentiality, integrity, and availability.[91]
Intellectual Property (IP)	Intellect which may be of value to an organisation, often protected by law.
Internet of Things (IoT)[92]	The concept that objects may be embedded with data connectivity to allow them to be monitored or controlled remotely (typically across the internet).
Key Performance Indicators (KPIs)	Data which is used to track how a system is behaving, typically in a contractual manner and with reference to targets.
Laser Scanners/Scanning	The use of LIDAR approaches with a moveable beam, in order to capture the shape of an object.
Light Detection and Ranging (LIDAR)	A system using light to measure distance, which may be used to build up an understanding of the size or motion of an object.
Machine Learning (ML)	A branch of artificial intelligence that allows computer programs to learn directly from examples, data and experience, rather than following hard-coded rules.[93]
Maintenance/Predictive Maintenance	Activities which ensure that an asset continues to be of use. Predictive maintenance will use predictive analytics to trigger these activities.

Term	Definition/Meaning
Model/Digital Model/Digital Representation/Network Model/ System Model	A computerised representation of a system, which is typically simplified through the use of assumptions.
Modelling Technology	The tools and techniques used to construct or analyse models.
Monitoring/Monitoring Technology	The use of observations (which may be from sensors or other digital systems) to keep records of behaviour.
Multi-Modal Transport	The use of many types of infrastructure (for example roads and rail) to move people or goods.
Neural Network	Machine learning architectures, inspired by structures in the brain, consisting of several layers of artificial 'neurons' (see 'Deep learning').
Open Data	Open data and content can be freely used, modified, and shared by anyone for any purpose.[94]
Personal Data	Data which may be attributed to an individual. Any information which relates to an identified or identifiable individual. Article 4 of GDPR defines personal data as the following: "means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".[95]
Physical Model/Modelling	The use of a model which considers the real-world behaviour of an object or system.
Predictive Analytics	Using analysis of current or historical data to show how an asset may behave in the future.
Pseudonymisation	Defined by Article 4 of GDPR: "means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person".
Real-Time System/ Live Information	The ability to process, or simply receive, data which is immediate enough for a computer (or person) to keep up with the current state of an object or system.
Renewals	Replacement of assets.
Resource Description Framework (RDF)	An approach for representing and linking to data about objects captured on the internet.
Risk Model	A representation of a system which is constructed to understand how likely/ frequent future events are expected to be, potentially also covering their severity.
Semantic Web	Concepts which promote common data formats on the internet.
Sensors	A device which detects events or changes in its environment.
Simulation	The use of a model to understand the future behaviour of a system.
Smart Grids	Electrical distribution networks that harness smart infrastructure concepts.
Smart Infrastructure	Infrastructure which includes a variety of elements which can sense or control (cause change).
Smart Meters/Metering	A digital device (or use of this device) such that utility consumption can be recorded, and potentially transmitted, frequently. It may also be able to control connected systems.

Term	Definition/Meaning
Smart Traffic Management/ Active Traffic Management (ATM)	Road networks that harness smart infrastructure concepts.
Spatial or Geospatial Data/Information[96]	Any data with a direct or indirect reference to a specific location or geographical area.
Structural Integrity	The ability of an object to adequately resist degradation due to applied forces.
Supervisory Control and Data Acquisition (SCADA)	The use of computers and remote monitoring to allow the state of a system to be understood.
System/Network	A set of items which work together as part of a whole.

List of Technologies considered



Acknowledgements: List of organisations consulted during the study

3D Repo
Affinity Water
AI Foundry
AMCL
Analysys Mason
Anglian Water
ARM
Arqiva
Arup
Association for Project Management
Atkins
BAM Nuttall
BEIS
Bristol Water
British Geological Survey
British Standards Institute
Buckinghamshire County Council
CAA
Cabinet Office
Cambridge Civil Engineering and Centre for Smart Infrastructure and Construction
Capgemini
Centre for Digital Built Britain
Centre for Transport Studies, UCL
Centrica
CGI IT UK Ltd
CH2M
CIHT
Cisco Systems Ltd
City Science
Civil Engineering Contractors Association
Costain
Council of Science and Technology
CPNI
Cranfield University
DCMS
decisionLab
DeepMind
DEFRA
Deloitte
DevicePilot
DFT
Digital Catapult
DIT
Doosan Babcock
Ecotricity
Ellen MacArthur Foundation
Energy Networks Association
Energy Systems Catapult
Energy UK
Environment Agency
Essex County Council
Exeter City Futures
EY
Flexilink
ForeScout Technologies
Frazer-Nash Consultancy
Fujitsu
Future Cities Catapult
Gaist Solutions Ltd.
Gardiner
GCHQ/NCSC
Geode Networks
GO Science
Grantham Institute
Guide Dogs
High Value Manufacturing Catapult
Highways England
HMT
IBM
Infrastructure and Projects Authority
Innovate UK
Institute for Mechanical Engineers
Institution of Civil Engineers
Intelsat
Invoke Capital
Iotic Labs
Itron
Jaguar Land Rover
Joint Radio Company Ltd
KPMG
Liverpool City Region Combined Authority
Living PlanIT
Loughbrough University
Macquarie Group
Materials for life
Met Office
Microsoft
Mind Foundry
Mole Solutions
Mott MacDonald
National Composites Centre
National Energy Action
National Grid
NESTA
Network Rail/ Network Rail SDG
Newcastle University
Northumbrian Water
Ofcom

Offshore Renewables Energy Catapult
Ofgem
Ofwat
Open Data Institute
Ordnance Survey
ORR
Oxford Centre for Industrial & Applied Maths
PETRAS IoT Research Hub
Pinsent Masons
Project 13
Revaluetech
Royal Academy of Engineering
Royal Society
Sandbag
Satellite Applications Catapult
SciSys
Scottish Power
Sensat
Sensus UK
SGN
Siemens
Skanska
Southern Water
Tarmac
Tech City
techUK
Thales
The Alan Turing Institute
The Infrastructure Client Group
Three
Transport for West Midlands
Transport Systems Catapult
Transport Technology Forum
Turner & Townsend
UK Space Agency
UKCloud Ltd
UKCRIC
UKPN
UKRN
United Utilities
University of Exeter
Vivacity Labs
Water UK
Waterwise
Web Sciences Institute/DCMS AI Review
Welsh Water
WSP
Yorkshire Water

Endnotes

1 BEIS (2013) Market Assessment of Public Sector Information **https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf**

2 The release of open data by TfL has supported the growth of London's Tech economy to the value of £14m pa in GVA and over 700 jobs and has led to a £20m increase in bus usage. Deloitte (2017) 'Assessing the value of TfL's open data and digital partnerships' accessed **http://content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf**

3 Ofgem (2017) **https://www.ofgem.gov.uk/publications-and-updates/infographic-energy-network**

4 BEIS (2016) Smart Meter Roll Out: Cost Benefit Analysis **https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/567167OFFSEN_2016_smart_meters_cost-benefit-update_Part_I_FINAL_VERSION.PDF**

5 IDC and Open Evidence for the European Commission (2017), 'European Data Market Study'.

6 Goodridge and Haskel, Imperial College London Business School (2015) How does big data affect GDP? Theory and evidence for the UK. **http://spiral.imperial.ac.uk/bitstream/10044/1/25156/2/Goodridge_2015_06.pdf**

7 National Infrastructure Commission (2016), **The impact of population change and demography on future infrastructure demand**, December 2016; National Infrastructure Commission (2017), **The impact of the environment and climate change on future infrastructure supply and demand**, June 2017.

8 P Kawalek "Data as Infrastructure" forthcoming December 2017.

9 PwC "The economic impacts of AI on the UK economy" June 2017.

10 **https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/664563/industrial-strategy-white-paper-web-ready-version.pdf**

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

12 Deloitte (2017), New Technologies Case Study: Data Sharing in Infrastructure – A final report for the National Infrastructure Commission, November 2017 p.37.

13 To be published in summer 2018.

14 Non-exhaustive list of DeepMind, Improbable, MindFoundry, Sensat.

15 Department for Business, Energy & Industrial Strategy (2017) **Growing the artificial intelligence industry in the UK**

16 •••ref•••

17 HM Treasury (2016) **Terms of Reference – New Technology Study**.

18 IDC and Open Evidence for the European Commission (2017), **European Data Market Study**.

19 The long list is included at the end of this report.

20 Royal Society (2017), **Machine learning: the power and promise of computers that learn by example**, April 2017.

21 Royal Society (2017), **Machine learning: the power and promise of computers that learn by example**, April 2017.

22 **https://www.rolls-royce.com/products-and-services/marine/discover/2017/simulation-platform.aspx**

23 **https://www.gov.uk/government/publications/industrial-strategy-sector-deals/introduction-to-sector-deals**

24 Infrastructure and Projects Authority (2017), **Transforming Infrastructure Performance** p.38.

25 Cambridge Centre for Smart Infrastructure and Construction (2016) '**Smart Infrastructure: getting more from strategic assets**'.

26 Utility Week (2017) Article: **Smart Meters Could Revolutionise Water Sector**.

27 Highways England (2016) **Innovation, Technology and Research Strategy**.

28 Cambridge Centre for Smart Infrastructure and Construction (2017) '**Transforming infrastructure through smarter information – Annual Review 2017**'.

29 Goldman Sachs Equity Research (2017), China's Rise in Artificial Intelligence – The New New China, 31 August 2017.

30 Sawers, P (2017) **How China's meshing ride-sharing data with smart traffic lights to ease road congestion**.

31 Network Rail (2015), 'Wessex Route Study'.

32 McKinsey (2013) '**Infrastructure Productivity: How to save $1 trillion per year**'.

33 HM Government (2017) **Industrial Strategy: Building a Britain fit for the future**

34 Department for Business, Energy & Industrial Strategy (2017) **Growing the artificial intelligence industry in the UK.**

35 Institution of Engineering and Technology (2017), **Local Authority Guide to Emerging Transport Technology 2017-18** p.13.

36 Deloitte (2017) **Predictive Maintenance Taking pro-active measures based on advanced data analytics to predict and avoid machine failure**.

37 Greater London Authority (2014) GLA Networked Utilities: Summary of Stakeholder Workshop.

38 Zeiss, G. (2014) Accelerating world wide initiatives to map underground utilities.

39 Ofwat (2017) **Delivering Water 2020: Consulting on our methodology for the 2019 price review**.

40 Mott MacDonald (2017) Value Analysis: Better Asset Management – final report.

41 Mott MacDonald (2017) Value Analysis: Better Asset Management – final report.

42 Mott MacDonald (2017) Value Analysis: Better Asset Management – final report.

43 Mott MacDonald (2017) Value Analysis: Better Asset Management – final report.

44 Environmental Engineering (2016) **Structural Monitoring Helps Re-Open Forth Road Bridge**.

45 Road Haulage Association (2015) **Continued closure of the Forth Road Bridge to HGV traffic will cost the industry an additional £40m says RHA**.

46 Carnegie UK Trust (2016) **Breaking the Link: Fife and the Forth Road Bridge closure**.

47 The Scottish Parliament, Infrastructure and Capital Investment Committee (2016) **Inquiry into the circumstances surrounding the closure of the Forth Road Bridge**.

48 Ofwat (2017) **Ofwat statement on Thames Water leakage commitments**, accessed online December 2017.

49 Thames Water (2017) Trunk Mains: Strategic Review.

50 Thames Water (2017) **What is a Smart Meter**?

51 Servelec Technologies (2017): Newmarket Metering Trial Consumption Analysis – Interim Report.

52 Mott MacDonald (2017) Value Analysis: Better Asset Management – final report.

53 Mott MacDonald (2017) Value Analysis: Better Asset Management – final report.

54 Mott MacDonald (2017) Value Analysis: Better Asset Management – final report.

55 Department for Business, Energy & Industrial Strategy (2016) **Smart Meter Roll Out: Cost Benefit Analysis**.

56 Ofgem (2017) **Infographic: Energy Network**.

57 Department for Business, Energy & Industrial Strategy (2016) **Smart Meter Roll Out: Cost Benefit Analysis**.

58 HM Government (2015) **Digital Built Britain Level 3 Building Information Modelling – Strategic Plan**.

59 PAS 1192-3:2014, "Specification for information management for the operational phase of assets using building information modelling".
60 NBS (2016) **National BIM Report 2016**.
61 Department for Digital, Culture, Media & Sport (2017) **UK Digital Strategy 2017**.
62 Department for Digital, Culture, Media & Sport (2017) **Policy Paper 7: Data – unlocking the power of data in the UK economy and improving public confidence in its use**.
63 National Infrastructure Commission (2016) **Connected Future**.
64 Kawalek "Data as Infrastructure".
65 Smart Dubai (2017) **Foundation Dimensions**.
66 PAS 185 (2017) "Smart Cities: Specification for establishing and implementing a security-minded approach".
67 Ernst & Young (2017): Study on international best practices in using technology to improve productivity.
68 BEIS (2013), Market Assessment of Public Sector Information **https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf**
69 Deloitte (2017) **Assessing the value of TfL's open data and digital partnerships**.
70 Transport for London (2017) advice provided to the Commission.
71 **https://ec.europa.eu/digital-single-market/en/network-and-information-security-nis-directive**
72 Tim Harford (2017), Fifty things that made the modern economy, Little, Brown – Chapters 17 and 18, direct quote from p.108.
73 Steven M. Spivak and F. Cecil Brenner (2001), Standardization Essentials: Principles and Practice p.13.
74 **https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/664563/industrial-strategy-white-paper-web-ready-version.pdf**
75 Hacktrain (2016), 'B.A.R.R.I.E.R.S Report'.
76 Department for Business, Energy & Industrial Strategy (2017) **Growing the artificial intelligence industry in the UK**.
77 Department for Business, Energy & Industrial Strategy (2017) **Growing the artificial intelligence industry in the UK**.
78 HM Treasury (2017) **Autumn Budget 2017**.
79 HM Treasury (2017) **Chancellor to unlock hidden value of government data** **https://www.gov.uk/government/news/chancellor-to-unlock-hidden-value-of-government-data**.
80 McKinsey & Company (2016) **The digital utility: New opportunities and challenges**.
81 ICE (2015) **Improving visibility and resilience of our buried services: creating an industry standard**.
82 Turner and Townsend response to NIC New Technology Second Call for Evidence, September 2017
83 **https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/664563/industrial-strategy-white-paper-web-ready-version.pdf**
84 **https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/664563/industrial-strategy-white-paper-web-ready-version.pdf**
85 This term has a specific meaning in this study, which relates to the use of acoustic logging in the water sector.
86 Royal Society (2017), **Machine learning: the power and promise of computers that learn by example**, April 2017.
87 Asset management may also be used to specifically refer to the management of financial assets. This is not the primary meaning in this study.
88 Big Data may also be used in relation to more specific concepts or approaches, a general meaning has been considered in this study.
89 See PAS 1192-2:2013 and PAS 1192-3:2014.
90 Regulation (EU) 2016/679, being implemented in May 2018.
91 NIST SP800-53.
92 IoT may also be used in relation to more specific concepts or approaches, a general meaning has been considered in this study.
93 Royal Society (2017), **Machine learning: the power and promise of computers that learn by example**, April 2017.
94 **http://opendefinition.org/**
95 **http://www.privacy-regulation.eu/en/4.htm**
96 This definition may also be restricted to data or information used by dedicated Geographic Information Systems. A broader meaning has been considered in this study, which is consistent with the INSPIRE Directive, **https://data.gov.uk/location/inspire**.