



GEMINI

CROSS-SECTOR UK DATA SHARING INFRASTRUCTURE

DATA SHARING INITIATIVES – HIGH-LEVEL LANDSCAPE SNAPSHOT

Digital Twin Hub Data Sharing Working Group

This report draws on inputs from the Digital Twin (DT) Hub's Data Sharing Working Group, the DT Hub Governance and Trust Working Group, DT Hub Board and the Catapult Network.

October 2024



Just as we built roads, highways, and airports in the 20th century, we must now build a digital infrastructure that is open, accessible, and empowers everyone.

Bill Gates



**If you want to go fast, go alone.
If you want to go far go together.**

African proverb



Do it once and share it many times

Data Sharing Working Group



Data Sharing Infrastructure

Executive Summary

Government and industry recognise the need to share data more efficiently and effectively to address the missions of the day.

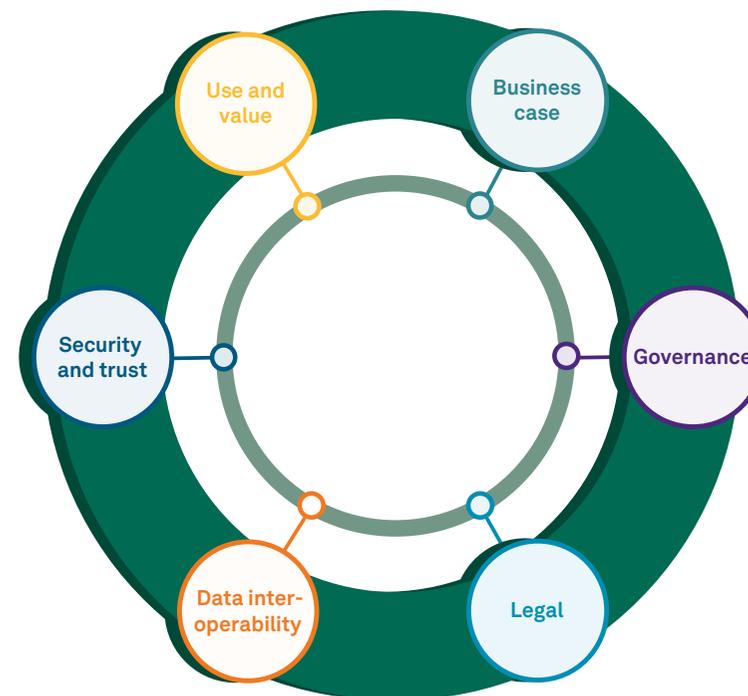
The case for Data Sharing Infrastructure to enable safe, seamless and secure sharing of data across organisations and sectors has been made. Opinions differ on how to achieve this and where government intervention and funding is required. There is, however, unanimous support for greater coordination across data sharing initiatives, now and in the future.

A first step to identifying where to intervene, fund and coordinate, is to understand what's already happening. This snapshot report is based on the expert knowledge from within the Digital Twin (DT) Hub's Data Sharing Working Group, which forms the kernel for a future Data Sharing Hub.

This report calls urgently for greater coordination across the governance, legal, data interoperability and security and trust components of Data Sharing Infrastructure.

The recommendation is not for centralised, top-down guidance without flexibility, but for collaboratively driven support and consensus for common yet adaptable governance, legal, data interoperability and security and trust solutions which can be applied across sectors to minimise the cost and risk of sharing data.

Data Sharing Infrastructure



What is Data Sharing Infrastructure?

Data Sharing Infrastructure is infrastructure as it enables the safe, secure and seamless sharing of data across organisations and sectors. It is infrastructure because it enables common access to shared resources on a secure basis, in a way that is essential to functioning of society.

The National Data Strategy published in 2020 recognises the importance of data as vital to a future functioning economy (Mission 1: Unlocking the value of data across the economy)¹ and the role of data as the essential ingredient to technologies such as AI and digital twins. The [Industrial Strategy Green Paper](#)² recognises “a huge opportunity for the UK to use its data more strategically, driving innovation and economic growth” and also that the UK is lagging behind in terms of achieving productivity growth through sharing and using data.

If we don't share data safely and securely, we will fail to address the missions of the day, for example, climate change. As set out in [Data Sharing Infrastructure – Digital Twin Hub](#), “Data Sharing Infrastructure is infrastructure and needs to be developed for the benefit of all. This means that Data Sharing Infrastructure needs funding and regulation, in the same way as for example, transport infrastructure. Public and private sector data both need to be shared using this infrastructure but according to a common set of rules that keep us all safe.”

If we want to assess progress in data sharing, we first need to find a way to describe different technical and non-technical components of Data Sharing Infrastructure. Data Sharing Infrastructure is not just a single piece of software, it is a combination of software, hardware, legal and cultural arrangements. Just as we can describe the barriers to sharing data as: technical, legal, cultural, financial and economic – so are the solutions.

Data Sharing Infrastructure is part of national Cyber Physical Infrastructure (CPI) and this report builds upon the work set out in the government consultation describing the enablers of CPI³ and the Catapult consortium report⁴.

Landscape framework and snapshot

This landscape snapshot presents a methodology for understanding Data Sharing Infrastructure

It outlines the DT Hub Data Sharing Working Group framework for describing components of Data Sharing Infrastructure.

The Data Sharing Working Group was formed in 2022 and meets every week to discuss, review, identify and refine best practice. It is made up of people from different industries, government departments and universities.

This Data Sharing Working Group framework can be used to examine in greater detail, data sharing use cases and initiatives as part of a wider landscape review. Further work to understand the maturity of different elements of Data Sharing Infrastructure could build upon such a landscape review.

This report provides a framework and a snapshot of DSI initiatives and recommendations.



Conceptual components of Data Sharing Infrastructure

Discussions over the period 2022-2024 in the Data Sharing Working Group have addressed the question, “How can we describe Data Sharing Infrastructure?”

The data sharing stakeholder community is very broad and diverse with a mix of technical and non-technical skills. Language and terminology have become a significant barrier to making progress in sharing data because we don’t know if we’re talking about the same thing.

The Data Sharing Working Group has been working to describe the components of Data Sharing Infrastructure, to identify what has already been achieved and what needs to be carried out. It took both a bottom-up approach to examining the detailed components of the Data Sharing Infrastructure and a top-down approach looking at major categories.

The Data Sharing Working Group recognises that there is no definitive answer, as perspective on data sharing will impact how each person or organisation describes it.

The work of the EU Data Spaces Support Centre (EU DSSC) considers the building blocks of data spaces. This report considers the concept of data spaces as aligned to the concept of a Data Sharing Infrastructure. The EU Data Spaces Support Centre describes the building blocks in the diagram below.



Organisational and Business Building Blocks

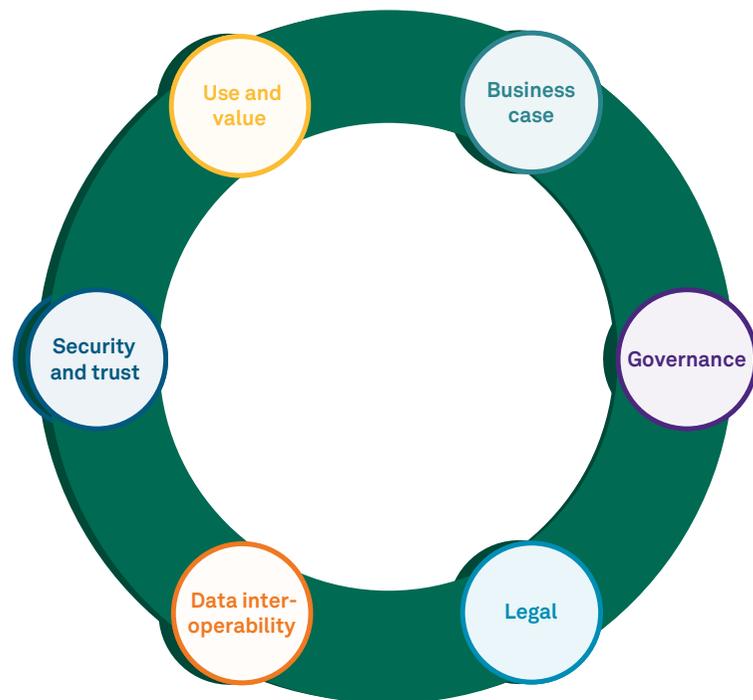


Technical Building Blocks



Discussions with stakeholders have shown that where an organisation, whether in the public sector or private sector, faces the need to share data to solve problems, often across sectors and across supply chains, that organisation will have to go through the process of addressing these six areas in its own bespoke way. These six areas are aligned to the six building blocks in the EU DSSC diagram.

Components of Data Sharing Infrastructure



The diagram shows that data sharing use cases or initiatives need to address each of the six areas (or building blocks) to share data.

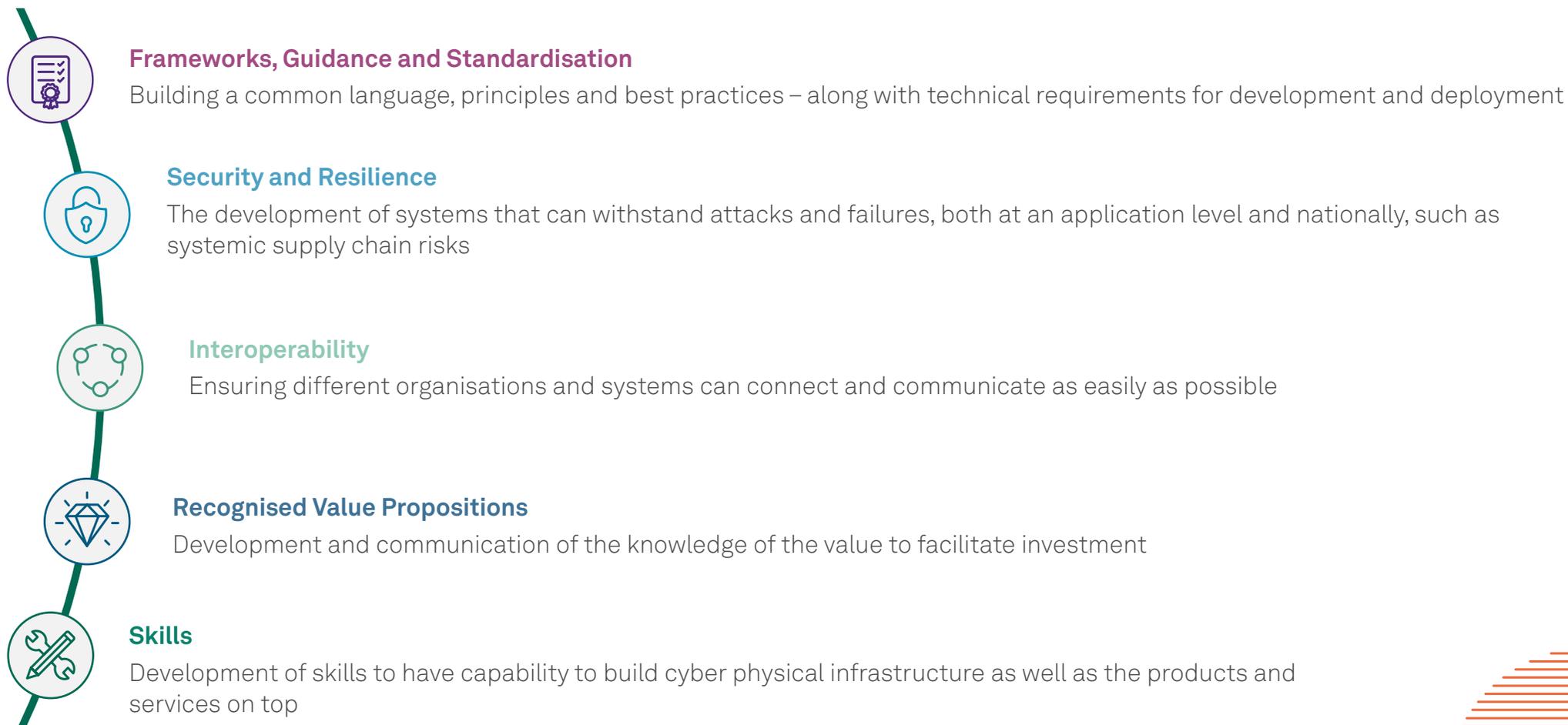
- **Business case.** The starting point, as previous studies have demonstrated is often with the use case and business case where a clear purpose and vision for data sharing is required.
- **Governance.** Initiatives need to have some form of organisational governance and data governance (rules for sharing the data).
- **Legal.** Initiatives need some form of a data licence and to adhere to relevant legislation.
- **Data interoperability.** Initiatives need to consider the use of a data model and the appropriate data structure to ensure the data is ready and able to be shared. This component is relevant to data preparation nodes and data sharing mechanisms⁶.
- **Security and trust.** These are key considerations and whilst this diagram shows a circular and sequential nature to the key components of a Data Sharing Infrastructure, many of these components must be considered from the outset.
- **Use and value.** Users will be able to share and use the data to derive value.

Data sharing initiatives that address more than one use case will have different emphases on different areas and it is the purpose of this review to identify where progress is being made and where greater emphasis and resources are required to help build out Data Sharing Infrastructure.

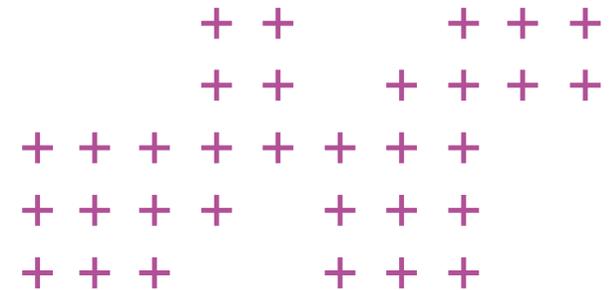
It is logical that the components of Data Sharing Infrastructure relate to the enablers of Cyber-Physical Infrastructure, illustrated.

In the UK, government and industry are thinking about these components and enablers in the same way as in the EU Data Spaces work, but the UK requires proper funding to take a consistent approach to actively advance the development of Data Sharing Infrastructure. The UK is at risk of being left behind.

Cyber-Physical Infrastructure Enablers



Data Sharing Infrastructure Components	Cyber-Physical Infrastructure Enablers
Business case	Recognised value propositions
Governance	Frameworks, guidance and standardisation
Legal	Frameworks, guidance and standardisation
Data interoperability	Interoperability
Security and Trust	Security and resilience
Use and value	Skills Recognised value propositions



Data sharing landscape

Appendix 1 sets out a number of identified data sharing use cases and initiatives. It is a small sample of the wider world of data sharing initiatives and is not intended as a representation of a complete list.

The reader is invited to submit information on data sharing use cases or initiatives here: [Data Sharing Landscape Review](#).

Whilst there may be thousands of data sharing initiatives in the UK in 2024, they are fragmented and without shared infrastructure each must develop bespoke solutions in isolation making benefits realisation more costly and resulting in low levels of interoperability.

A Data Sharing Infrastructure would enable the development of unlimited data sharing use cases in the future at a lower cost without continually needing to reinvent the wheel for each initiative.

We need a way to discover and find existing use cases so that the market can continue to develop new use cases and build upon what is already available and accessible. This is not to suggest that all data sharing use cases and initiatives across public and private sectors should be publicly accessible but that they should be findable by the right people.

A secure catalogue of data sharing use cases and initiatives should be part of Data Sharing Infrastructure. This would form part of a shared knowledge base within Data Sharing Infrastructure and could be part of a National Data Library. In the Apollo Protocol⁷, this is called a Body of Knowledge, based

on a distributed and federated approach so that initiatives are discoverable and there is collective understanding of where this can be achieved and shared between organisations.

The use cases and initiatives range across sectors of the economy; environment, health, finance, economic infrastructure and social infrastructure. There has been significant progress across banking, energy, water, transport and defence together with some infrastructure cross-sector initiatives.

Without a complete catalogue of use cases and initiatives, it is not possible to map out what each initiative is using for governance, legal, data interoperability and security and trust arrangements. However, this report sets out a framework to map out and compare data sharing use cases and initiatives and could be used as a basis for monitoring and evaluation in the future.

It is based on work in the EU Data Spaces Support Centre and on extensive discussions in the Data Sharing Working Group. Digitalising the use of such a framework could capture use cases and initiatives on a more automatic and efficient basis. Digitalising this type of governance process incurs a cost but will enable Data Sharing Infrastructure to be both governable and more widely used.

The table below sets out a snapshot of data sharing initiatives using this framework, showing that by using these categories it is possible to describe different data sharing initiatives alongside each other in summarised form.

The detail in this table comes from the Data Sharing Working Group and is open to further refinement. A full landscape review could take this approach and use it as the basis of a survey questionnaire.

Initiative	Business case	Governance	Legal	Data interoperability	Security and Trust	Use and Value
Apollo Protocol	An initiative to unlock the benefits of connected digital twins delivering a mechanism for formalising communication between sectors	Executive Board and Advisory Board supporting Apollo Forums	Open initiative currently chaired by The Alan Turing Institute and the Advanced Manufacturing Research Centre	Focused on information management practices	Focus on security and privacy	Articulating the value of connected digital twins across domains for sharing ideas and proven best practices
CReDo	Climate change adaptation digital twin platform to improve system-wide resilience across infrastructure networks. Current use cases include strategic resilience planning use case. Business case to reduce cost of disruptions, regulatory use case	Led by Connected Places Catapult, alongside partners Anglian Water, BT Group and UK Power Networks, funded by Innovate UK, Ofgem SIF, Ofwat Breakthrough	Bespoke data licence based on energy data exploration licence and Partnership Agreement	Application-level ontologies were developed to create a knowledge graph that integrates data across energy, water, telecoms	Use of central secure host originally and shift to distributed architecture	£4.4bn Net Present Value of CReDo cross-sector digital twin for extreme heat and flooding scenarios at Great Britain level to 2080 (£2.6bn to 2055) ⁸
Earth Observation (EO) Data Hub	Increase use of EO Data to new markets and improve ease of access to all EO data sources. Minimise duplication and transfer of large datasets	EO Data Hub Board, EO Data Hub end user and stakeholder forum	Licensing: Open access for open data. Individual data licences between supplier and end user	Suppliers using industry standards for EO data and geospatial data sharing	OAuth (Open Authorisation Standard) used across distributed architecture. Supplier-end user access to be enabled	Focus on climate and land use change monitoring. Open-source code, offering both open and commercial data
Energy Data Sharing Infrastructure	Net zero and resilience, multiple use cases	Ofgem consulting on interim Data Sharing Infrastructure coordinator, Ofgem Data Best Practice. NESO pilot (requested by DESNZ) under Virtual Energy System programme governance	For pilot and minimum viable product use case leveraging existing energy grid codes (license obligation to share data with NESO). In long run, legal framework to leverage Trust Framework	Pilot and minimum viable product are exchanging data using Common Information Model. Review of standards planned	Trust Framework to be developed as part of Data Sharing Infrastructure pilot and minimum viable product	Use across energy system



Initiative	Business case	Governance	Legal	Data interoperability	Security and Trust	Use and Value
Interoperability Flagship	Reduction in time to map data models between systems. Use case: carbon measurement and reporting within construction	Led by Digital Catapult		Automated semantic mapping between different but related data models in different systems. Translation from data generator systems into an adaptive universal language and translation into a consumer data model	Provide an indication of data gaps, completeness and confidence across use cases	Creating carbon reporting data models
NDTP⁹	Develop standards, frameworks, guidelines, methodologies, and tools that are foundational to a functioning market in digital twins	HMG Programme funded and managed by DBT, with cross-Whitehall contributions (e.g., DfT, MoD)	Open-source software (OSS) available to all under an Apache 2.0 License. Multiple MoUs across government and industry (e.g., NESO, Stream, MoD etc.)	Developing open-source software, called the Integration Architecture (IA), to serve as a distributed Data Sharing Infrastructure. This enables seamless data exchange and supports interoperability by mapping data to a common 4D ontology, known as the Information Exchange Standard (IES) ¹⁰ (IES). IES spans multiple sectors and gaining widespread adoption.	Developed using Secure by Design principles (also covered in Secure by Design Policy , authored by CDDO and Cabinet Office). Is a Zero Trust Architecture as defined by NCSC . Employs ABAC (attribute-based access control) and currently working on developing PBAC (policy-based access control)	Multi-million-pound open-source software, accessible to all, and already adopted by NESO and MOD. Includes four (4) data-exploitation demonstrators, such as IRIS, which accelerates the ECO4 eligibility and identification process by over 20%
NUAR¹¹	Digital map of underground pipes and cables for authorised users in England, Wales, NI, accelerating data sharing process from six days (currently) down to 60 seconds (with NUAR) ¹²	Geospatial Commission funded the development of NUAR. OS as future operator of the service while Government remain accountable for the register and policy oversight	Minimum viable product is initially accessible to asset owners who operate in England, Wales or Northern Ireland and have signed the NUAR Data Distribution Agreement and agreed a Data Ingestion Specification	Harmonised data model. Use of OGC MUDDI schema for harmonisation. Centralised data sharing architecture in pilot phase	Role based access – based upon role and approved purposes. All access is audited and sensitive assets have ‘enhanced’ security measures attached	NUAR is envisaged to deliver over £400m per year of economic growth through increased efficiency, reduced accidental damage and reduced disruption for citizens and businesses. The service is transitioning to Public beta in Spring 2025
Rail Data Marketplace	The central platform for finding and sharing data across the rail sector	DfT funded. Cross-industry steering group. Delivered and run by Rail Delivery Group	License builder or default provided https://raildata.org.uk/helpAndInformation/policies/data-sharing-agreement	Supports APIs, flat files and pub/sub technologies. Common Meta data (RSSB standard)	Full role-based access. Full security model followed	Currently 149 data products published across: Open data, shared (limited use) and monetised. Aim to be commercially neutral
Stream	Data sharing across water sector. Use case agnostic	Independent governance framework including partners, observers and advisors, Data triage and identify rules to put in place to govern who can use the data for what purpose	Overarching legal framework with scheme specific terms for use cases where data sharing is a factor. For open data, agreement over standardisation on open data terms/licensing	Standards defined and documented for what and how the data needs to be shared – agreed among members of the scheme. Adoption of existing standards where these exist is a key principle	Technical enactment of the rules including validation/verification of publisher, user and data as appropriate	Stream’s role is to enable the creation of value and to capture and share/spotlight stories of value generation to demonstrate ROI
4D SIG	A network of information management enthusiasts exploring the organisational, technical and theoretical challenges within and between information systems	Independent with a leadership team led by Dame Wendy Hall	Network managed out of the University of Southampton	Focuses on 4-dimensionality (4D) that is a key characteristic of the foundation data model used by the National Digital Twin Programme	Dedicated working group investigating future security models utilising this approach	Dedicated working group investigating the value in 4D approaches sharing examples where possible

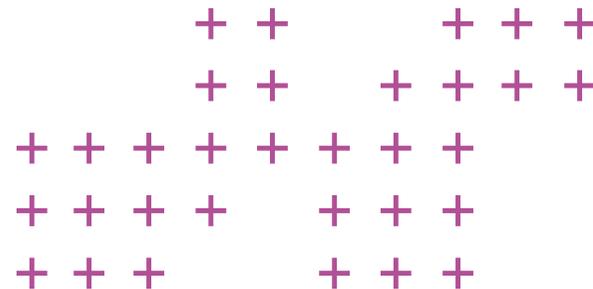
Other data sharing use cases and initiatives

There are a multiplicity of data sharing use cases and initiatives, and no report could hope to capture all. However, digitalising the process of capturing and recording data sharing initiatives could generate up-to-date reports upon request. In the meantime, comparisons can be subjective. Investing in systems to manage complexity and bias is part of getting the benefits from digitalisation.

Open Banking¹³ is a good example of data sharing. It is a system that allows banks and financial institutions to share customer data securely with third-party providers through application programming interfaces (APIs).

This initiative, which came into effect in January 2018, aims to promote competition and innovation in the financial services sector. A key feature of open banking includes consumer consent whereby customers must give explicit consent for their data to be shared, ensuring privacy and control over their information. Access is provided to licensed third-party providers who can then offer services like account aggregation, personalised financial advice, and payment initiation based on the shared data.

Whilst Open Banking is a good example of what is possible, the dataset in scope is narrow compared to the need to make any data set interoperable within a Data Sharing Infrastructure. Therefore, whilst Open Banking is considered a success story and is an example where regulatory compliance is important, it is not wide enough in scope to use as model for Data Sharing Infrastructure.



Lessons learned from selected sectors

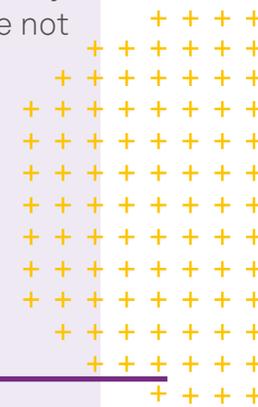
Energy sector data sharing

The development of Data Sharing Infrastructure in the energy sector is being accelerated through the regulator setting direction and industry moving the vision forward.

The incorporation of Ofgem’s Data Best Practice and Digitalisation Strategy and Action Plan Guidance into licence obligations in 2021 and 2023 marked a step change in the digitalisation journey of the energy sector. Data Best Practice guidance has resulted in energy networks creating open data portals¹⁴. Ofgem set out its position to take forward an energy Data Sharing Infrastructure in the Future Systems Network Regulation (FSNR) decision of October 2023¹⁵ and has signalled intent to include licence obligations to adhere to the energy Data Sharing Infrastructure in the next price control RIIO 3. Ofgem has also published a consultation on the short-term governance options for energy Data Sharing Infrastructure and is due to consult on the longer-term options in 2025¹⁶. The Department for Energy Security and Net Zero has committed to the National Energy System Operator (NESO) starting work on an outage planning use case Data Sharing Infrastructure pilot in 2024 and to delivering a strategic planning use case Data Sharing Infrastructure minimum viable product in 2025¹⁷.

It should be noted that in addition to the work of NESO, other energy networks are also embarking on data sharing schemes, such as Scottish Power’s Energy System Digital Twin, ENSIGN. Open Energy is an example of a Trust Framework and an energy data catalogue enabling energy data search¹⁸. Ofgem’s proposed interim Data Sharing Infrastructure coordinator will need to develop common rules that apply to data sharing across the energy sector and existing schemes.

Whilst regulatory direction setting and industry action is moving Data Sharing Infrastructure forward within the energy sector, it does not have the remit to progress cross-sector data sharing which is crucial for net zero and resilience use cases which often branch into transport, water, agriculture etc and across supply chains. However, much can be learned from the energy sector and translated across. For example, Data Best Practice is sector agnostic, “It is a principles-based set of guidance on the quality, accuracy and accessibility of data, and its requirements are not specific to the energy sector.”¹⁹



Water sector data sharing

Stream²⁰ is the Data Sharing Infrastructure for the water sector which is opportunity driven rather than regulator driven, although the funding to drive Stream forward has come from Ofwat. Stream is a collaborative initiative involving 16 out of the 18 major UK water companies who all share a vision to unlock the potential of water data to the benefit of customers, society and the environment. Since receiving innovation funding in August 2023, Stream members and partners have established Data Sharing Infrastructure (community + data + processes + technology) to realise their shared vision.

Stream was originally set up in 2020 by water companies who were driven by need and opportunity. It was two years before the Stream collaboration was able to obtain funding, and in this period one water company opted to pursue an independent path outside of the Stream initiative. The introduction of innovation funding from Ofwat since 2022 has accelerated and enabled the development of shared open data across the participating water companies.

Stream is currently examining and pursuing shared data use cases. Stream is looking for further funding to enable maturation of the water sector Data Sharing Infrastructure to a point where it is self-sustaining across participating water companies.

Stream has focused on the operational pillars of:

- 1) Use cases and data sets
- 2) Ecosystem
- 3) Operations
- 4) Technology

Stream has had to create its own governance across its members and rules for sharing data. Stream is currently exploring the opportunities to implement a Trust Framework to enable the secure sharing of shared data (eg not open data). Identity management has been identified as a key issue in shared data. For example, understanding who the consumer of the data is, for what purpose the data will be used, understanding who owns or has rights to the data, who has the responsibility for managing it and understanding the data quality.

Cross-sector data sharing

In the absence of common frameworks or guidance or a usable cross-sector data standard, projects like CReDo²¹, the Climate Resilience Demonstrator, have had to design its own governance, legal arrangements, data interoperability and security and trust mechanisms. CReDo works with confidential asset data which must be kept secure according to the conditions of its current bespoke data licence. When the data licence for CReDo's pilot phase was originally drafted, based on a data exploration licence, the business case was not fully understood because the project needed to share data to inform and develop the business case. It showed that allowing for a pilot phase is helpful as it gives scope to redraft the data licence once the business case is fully developed and understood by legal representatives.

Application-level ontologies were developed to create a knowledge graph that integrates data across energy, water and telecoms sectors. CReDo is evolving from a centralised to a distributed data sharing architecture to enable the scaling of participants and use cases. The evolution of CReDo from demonstrator to decision optimiser will require CReDo to develop and maintain trust across all partners.

Extending and scaling CReDo to other sectors, data sets, use cases and participants requires careful consideration of scalable Trust Framework and data interoperability mechanisms. At present, there is limited good practice guidance to support cross-sector data projects like CReDo, to scale.

Lessons emerging from this snapshot:

1. Incentives matter

The incentives for data sharing matter, whether organisations are driven by mandatory regulation, voluntarily for public good, or for commercial reasons. A high-level assessment of initiatives shows that data sharing can be driven by industry and government and can also be regulator led. The water sector provides an excellent example of where data sharing has been led through collaboration across the industry and supported by the regulator.



2. Bespoke rules create uncertainty and add cost

As identified in the Stream initiative, there are currently no common rules in place on identity management.

Melissa Tallack, co-lead of Stream said:

“ We would much rather know that we’re doing it (identity management) in a way that is consistent with other sectors and in the absence of a common approach to identity management, we would have to create our own. ”

In the absence of common frameworks, guidance and standardisation, data sharing schemes and initiatives must develop their own approaches to managing the governance, legal, data interoperability and trust and security challenges. This adds extra cost to any data sharing initiative and means that where public funding is used to support data sharing initiatives it may be used to support activities which are duplicated across multiple initiatives. Because funding grants do not generally factor in resources and budgets for actively sharing outputs which might require coaching others to use a data model or to adapt a data licence, it means that data sharing initiatives do not have the capacity to share their learnings properly to other initiatives.

3. Trust Frameworks need further development

Sharing fast-moving and sensitive data, such as smart meter data, is not straightforward and Trust Frameworks that enable this need to include legal redress if the rules of sharing are broken.

Trust Frameworks which cover multiple sectors need common Trust Framework guidelines. For example, sharing data between energy and water sectors requires a common Trust Framework for energy and water with common rules for sharing data. If there are incompatible terms in sector Trust Frameworks then it is more difficult to enter into data sharing schemes under Trust Frameworks. Therefore, there is an urgent need to develop a national standardised Trust Framework to enable the development of sector based frameworks and cross-sector data sharing schemes, all working to the same terminology and guidance.

4. Ontologies matter

Ontologies can help standardise a data definition and structure to help achieve interoperability. Ontology is a crucial foundation for Data Sharing Infrastructure, providing a common language and framework that enables interoperability and understanding between diverse systems and stakeholders.

Ontologies are key to making Data Sharing Infrastructure effective, secure and resilient for the following reasons:

- Interoperability and standardisation
- Knowledge representation
- Data quality and consistency
- Scalability and flexibility
- Enhanced collaboration and knowledge sharing

In essence, ontologies act as the connective tissue that links Data Sharing Infrastructure together, ensuring that data is understandable, accessible, and valuable for a wide range of applications. Achieving a core ontology is ideal but fraught with real-world challenges such as domain specificity, cultural and regional differences, evolving knowledge and competing interests.

Strategies to create modular ontologies, alignment and mapping, stakeholder engagement, evolution along with ontology governance mechanisms can help to promote greater interoperability and understanding between different systems and stakeholders.



5. Data standards

Sectoral level data standards are required but must be developed through consensus and must be usable. Feedback from stakeholders strongly resisted the creation of multiple new standards. Many industries are starting from a low base in developing data models and would benefit from guidance. Sectors such as energy are recognising the urgent need to review existing standards²². As energy becomes a case study sector with a level of complexity beyond that of the Open Banking example, it is important to find a way to share the learnings from developing a Data Sharing Infrastructure in energy with other sectors at the same time as looking at how to make energy data interoperable with other sectors. The development of the Information Exchange Standard (IES4)²³ presents an opportunity to test out ease of use and adoption.

6. Data institutions and data rights

Where a data sharing scheme is not a legal entity, it can be difficult to enter into legal agreements with other data sharing schemes. Creating data institutions may help to resolve this. Work from the Open Data Institute ([Data institutions | The ODI](#)) and ADA ([Exploring legal mechanisms for data stewardship | Ada Lovelace Institute](#)) could contribute to shared guidance.

Complexity around data ownership can get in the way of sharing if the data is shared across a supply chain and commercially sensitive. Organisations need to learn more about data rights.



Risks identified through the lack of a joined up Data Sharing Infrastructure

1. Benefits of cross-sector data sharing are not achieved as less cross-sector data sharing happens due to costs and barriers and we fail to address cross-sector challenges.
2. Poor data sharing hinders decision making and slows the response of policy interventions at times of crisis.
3. Data siloes cause fragmented decision-making.
4. Inconsistent information in different systems cause unrepeatable business processes.
5. Detriment to public services because essential information cannot be made available.
6. Loss of citizen freedoms as it is impractical to exercise their preferences about how their data is used and the services they wish to engage with.
7. Loss of trading opportunities as the UK is unable to provide access to its data as part of agreements with other nations, resulting in exclusion from trading blocs.
8. Duplication of competing standards entailing a lack of traction and development of multiple platforms.
9. Large software vendors will continue to have vendor dominance stifling innovation of SMEs being able to play in that marketplace.
10. Increased cost from inventing and reinventing governance, legal, data interoperability, security and trust solutions individually meaning the risks of getting it wrong are higher, and incompatibility leading to lost opportunities and delay.
11. Multiple copies of data stored in multiple places.
12. Sustainability risk from storing multiple copies of data.
13. A myriad of security and privacy breaches from multiple bespoke approaches to security and trust.
14. The inability to audit the sharing and use of data.
15. Different consent mechanisms using different terminology asking for the same data.
16. Foreclosure of markets as data sharing becomes prohibitively expensive.
17. Fragmentation of legislative measures.
18. Insufficient funding for scaling up pilot projects.

Key lesson learnt on coordination

The table shows that even a snapshot of data sharing initiatives includes a diversity and variety of approaches. Data sharing initiatives typically have their own versions of governance, legal agreements, data interoperability, security and trust which may not be scalable or replicable to other initiatives or sectors leading to a fragmented landscape and siloed decision making.

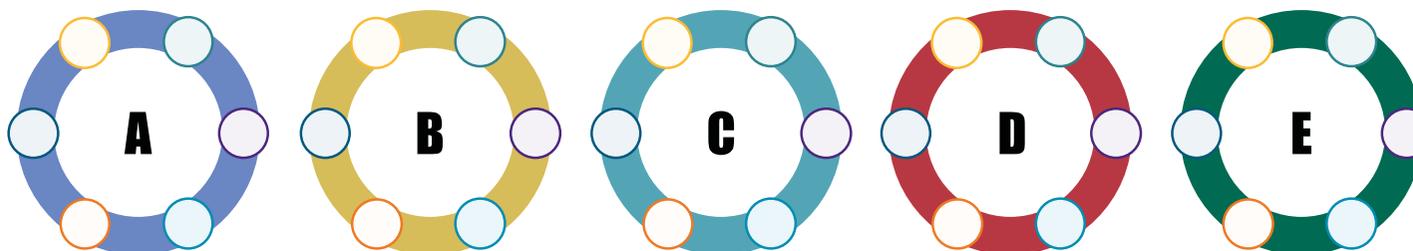
The diagram suggests that there will be more bespoke approaches to data sharing as we do more data sharing. Without guidance, these will not be consistent or coordinated. And each time, we are spending money on common factors like governance arrangements, data licences, data interoperability and security and trust measures when we could be learning from previous and existing initiatives.

As the number of data sharing use cases and initiatives grows in the UK, more and more money will be spent doing the same thing on governance arrangements, data licences, data interoperability and security and trust measures, vastly increasing the cost of data sharing.

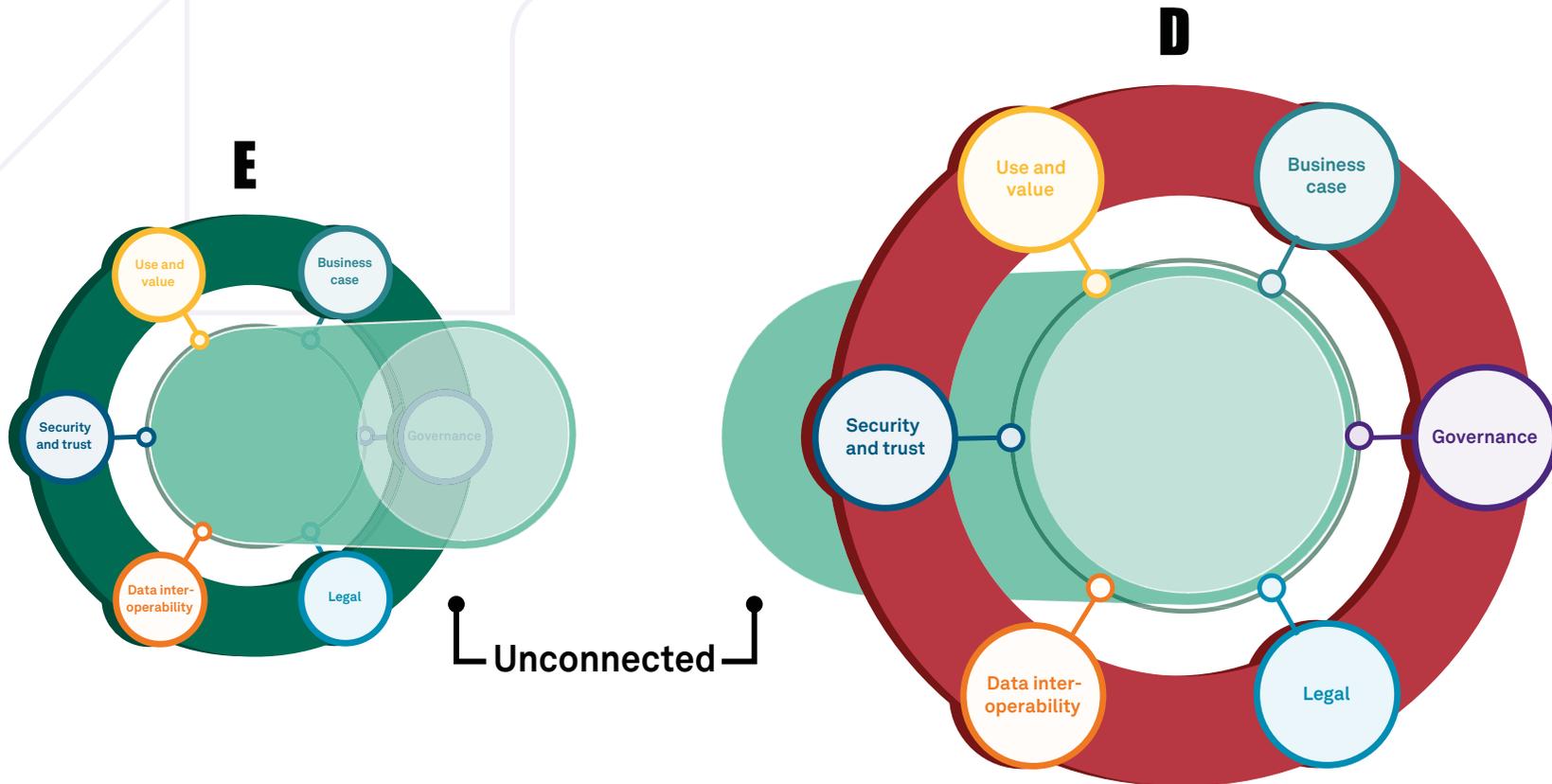
Within the Data Sharing Working Group and wider discussions amongst the practitioner community, individuals and organisations have articulated both desire and ambition to connect up initiatives where possible but acknowledge that support is required to do this.

Practitioners do not advocate for a single solution but a distributed and federated approach where some degree of standardisation across the board and under sector domains will aid progress.

Data Sharing Initiatives



Data Sharing Initiatives

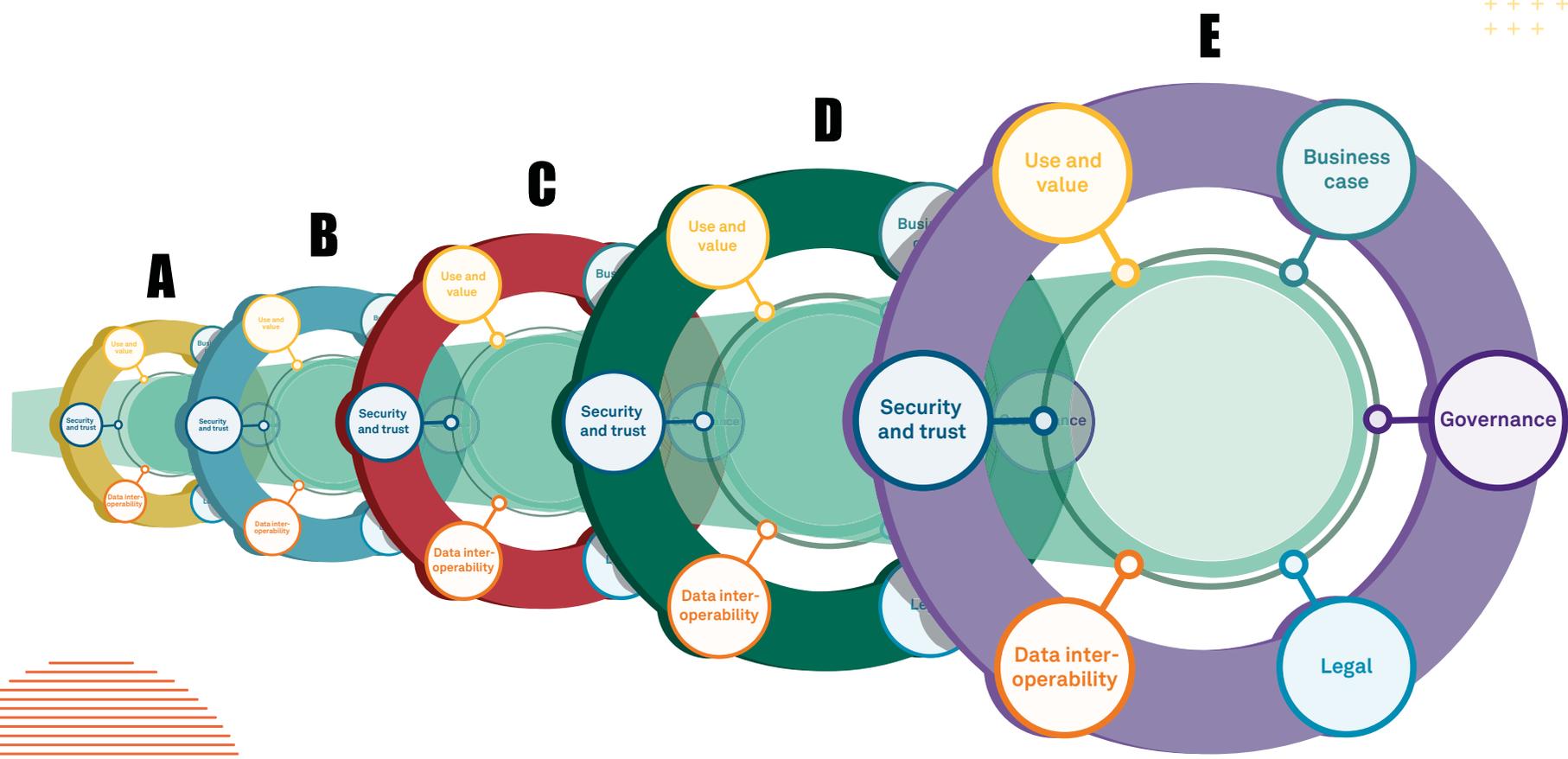


The use case and business case will naturally be particular to each organisation and similarly, the way they use the data and derive value from it will also be particular to that organisation. This landscape snapshot has identified that there would be significant benefit from having a common and connected approach to:

1. Governance
2. Legal
3. Data interoperability
4. Security and Trust

Frameworks, guidance and standardisation supporting these areas that can be used across sectors and organisations as part of Data Sharing Infrastructure would allow organisations to get on with the job of secure data sharing, minimising cost and effort.

Data Sharing Infrastructure



Recommendations

A common and connected approach to governance, legal, data interoperability and trust and security would lay the foundations for multiple data sharing use cases and initiatives to emerge at lower cost across the public and private sectors. A key element of a common and connected approach must be flexibility and the ability to adapt the approach to the needs of the sector.

The missed opportunity of common frameworks, guidance and standardisation creates risk for all data sharing initiatives. It will be possible to go further faster with streamlined Data Sharing Infrastructure guidance.

At the detailed implementation level, different solutions may work in different organisations and sectors due to cultural differences, but a common and connected framework to guide this will minimise the cost of sharing data and maximise the potential benefit. For example, a more integrated Data Sharing Infrastructure programme could help inform government-based collective purchasing agreements for commercial data.

The UK Government has a huge opportunity to provide guidance, but it cannot do this without input from the community of practitioners. Inaction will result in being left behind the progress of the EU Data Spaces work.

The EU are ahead of the UK because it is recognised in Europe that funding and coordination are required to advance Data Sharing Infrastructure.

In developing a common approach, it is important to consider the appropriate balance of incentives and enforcement, and the roles of regulator, industry and facilitator.



Key considerations

- A regulator led approach could enable faster progress in standardisation and adoption across a sector through enforcement.
- A regulator led approach can also risk slow progress if regulator is not well informed and industry objects
- An industry led approach could make progress faster driven by commercial incentives but is more likely to be siloed, as is the status quo.
- An industry led approach can find it difficult to gain consensus across multiple interests and there is always the risk of capture by larger firms.
- A facilitator led approach has the potential to draw in wide engagement and expertise across industry and can coordinate experiments (through sandboxing) to understand optimal choices.
- A facilitator led approach may suffer from a lack of enforcement capability.

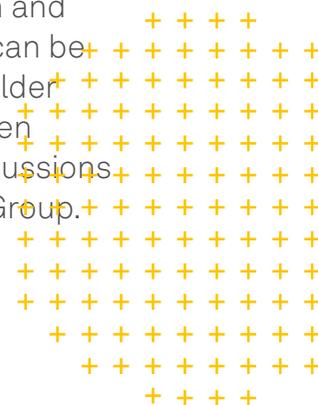
There is no one Department or Regulator in the UK who currently has the job to develop a common and connected approach to governance, legal, data interoperability, and trust and security within a cross-sector Data Sharing Infrastructure.

As Open Banking and the emergence of the energy Data Sharing Infrastructure suggests, enforcement is necessary where collaboration is not automatic. Where enforcement is required, so is funding.

The remit to enforce a common and connected yet flexible approach to governance, legal, data interoperability, trust and security across sectors must be accompanied by an appropriate flow of funding.

This is an opportune moment to put the responsibilities for cross-sector Data Sharing Infrastructure governance under the roof of one regulator or cross sector digitalisation body who would be tasked to work with the regulators of all sectors in promoting safe and secure data sharing.

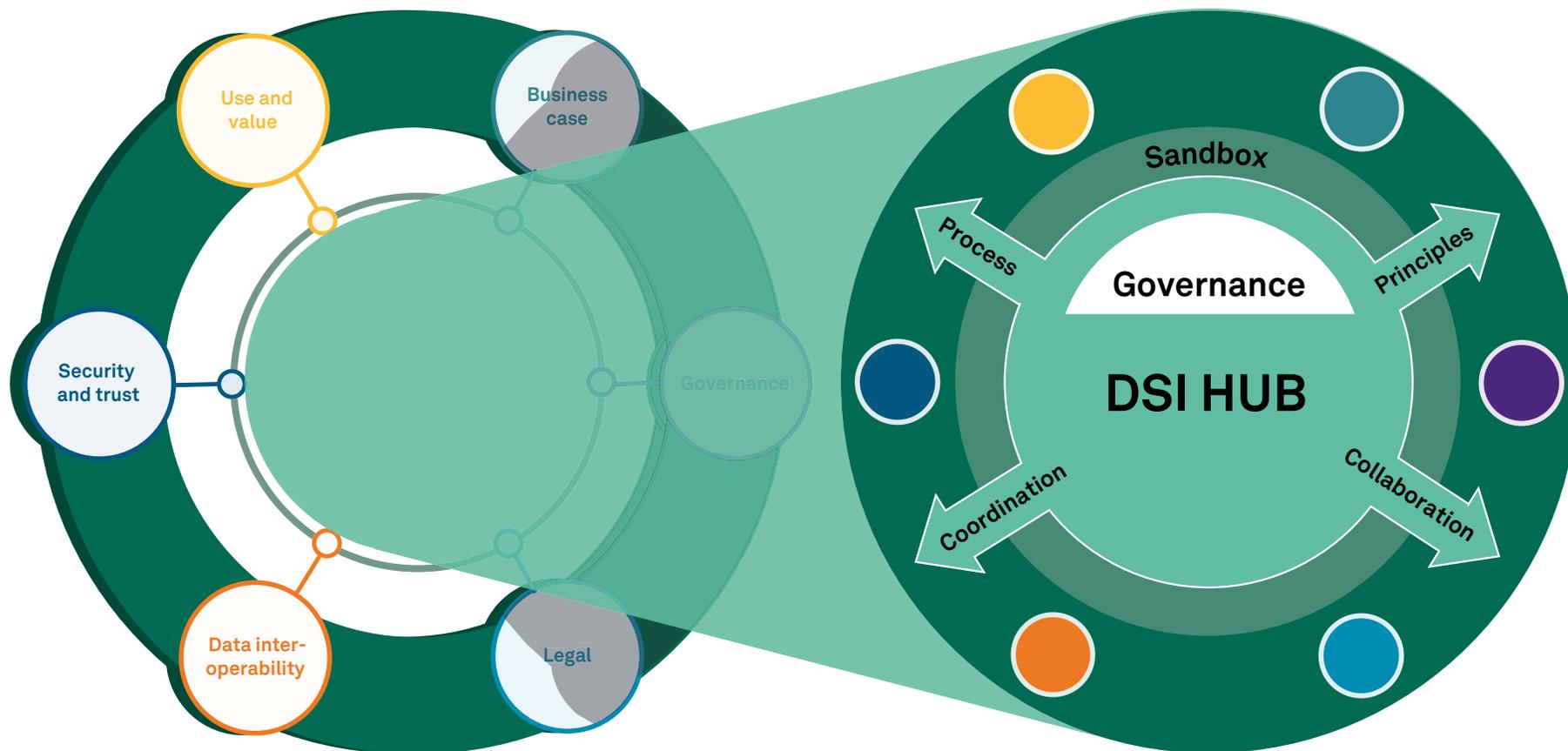
There are a wide range of opinions on how a common and connected approach to Data Sharing Infrastructure can be led. A crucial part of this is convening across stakeholder communities, getting input and having a valid and open debate. A Data Sharing Hub could facilitate such discussions building upon the work of the Data Sharing Working Group. Such a Hub needs funding to function effectively.





A Data Sharing Hub can:

- Feed in learning and best practice from the wider community building on all existing knowledge and good practice. This can enable the development of the rules around how we share and use our data to be collaboratively driven.
- Draw input from a wide resource pool across stakeholder communities and help to set direction for a common approach to governance, legal, data interoperability, and trust and security with expert input from on the ground practitioners and innovators.
- Lead coordination and collaboration across stakeholders and sectors in order to align principles and processes.
- Support ideally one cross-sector digitalisation body or regulator, or multiple regulators, to agree standards across the Hub which are then enforced at the regulatory level.



Please contribute to the evolution of this report and the table on data initiatives through the form:

[Data Sharing Landscape Review](#)

Appendix 1

List of data sharing initiatives (note this is a limited and by no means exhaustive list, there are many thousands of data sharing initiatives underway in the UK). Add further initiatives here: [Data Sharing Landscape Review](#)

Apollo Protocol	StreetManager	UKRI Data Hubs and Compute
DINI/DAFNI	National Travel Survey	Private sector providers, e.g. Telematics (TomTom, INRIX), OEMs (e.g. Mercedes, Ford), ticketing (e.g. Visa) etc
CRoDo	Urban Observatories	OCPI will create open data for Public EV chargers
Energy Data Sharing Infrastructure	TfL open data	D-TROs will provide open data for traffic regulation orders
NDTP	National Highways Open Data	Government Better Off Calculator
NUAR	FOIs can ask for Local Authority data	DWP AW social tariff LITE
Open Banking	ADEPT LL Data Hubs	Social Housing data sharing + Watercare
Open Energy	Future Transport Zones Data Hubs	Council data sharing initiatives (multiple)
Stream	Consumer Data Research Centre	Earth Observation Data Hub
London Datastore	Interdigital Data Platform	
Rail Data Marketplace	NDTP: Integration Architecture	
NAPTAN	Find transport data	
BODS	Convex for vehicles	

Endnotes

1. [National Data Strategy – GOV.UK](#)
2. [Invest 2035: the UK’s modern industrial strategy – GOV.UK](#)
3. [Enabling a national cyber-physical infrastructure \(CPI\) to catalyse innovation – GOV.UK](#)
4. [National Cyber-Physical Infrastructure ecosystem – Digital Twin Hub](#)
5. [Data Spaces Support Centre | European Digital Innovation Hubs Network](#)
6. [Digital Spine Feasibility Study | Arup, Energy Systems Catapult, University of Bath](#)
7. [Apollo Protocol](#)
8. [10061361 | ENA Innovation Portal](#)
9. [National Digital Twin Programme \(NDTP\) - GOV.UK](#)
10. [GitHub - dstl/IES4: The Information Exchange Standard \(IES\) is a standard for information exchange developed within UK Government](#)
11. [National Underground Asset Register \(NUAR\) - GOV.UK](#)
12. [National Underground Asset Register \(NUAR\) - Economic Case Summary - GOV.UK](#)
13. [Home – Open Banking](#)
14. [Open energy datasets](#)
15. [Decision on frameworks for future systems and network regulation | Ofgem](#)
16. [Governance of the Data Sharing Infrastructure | Ofgem](#)
17. [Government response to the energy system ‘digital spine’ feasibility study](#)
18. [Open Energy – Find, access and use Open and Shared energy data](#)
19. [Consultation on Data Best Practice guidance and Digitalisation Strategy and Action Plan guidance | Ofgem](#)
20. [Stream – Portal](#)
21. [Climate Resilience Demonstrator - Digital Twin Hub](#)
22. [Government response to the energy system ‘digital spine’ feasibility study](#)
23. [GitHub - dstl/IES4: The Information Exchange Standard \(IES\) is a standard for information exchange developed within UK](#)



GEMINI

Visit the Digital Twin Hub website

www.digitaltwinhub.co.uk/working-groups/

Follow us on LinkedIn

in [Digital Twin Hub](#)

Email us

 dthub@cp.catapult.org.uk

©Digital Twin Hub

This work is licensed under the Creative Commons Attribution 4.0 Unported License:

<http://creativecommons.org/licenses/by/4.0>

This report draws on inputs from the Digital Twin (DT) Hub's Data Sharing Working Group, the DT Hub Governance and Trust Working Group, DT Hub Board and the Catapult Network.

 **GEMINI**
CC-BY-SA 4.0