

Analytical Report n 12



Analytical Report 12: Business-to-Government Data Sharing

This study has been prepared by Capgemini Invent as part of the European Data Portal. The European Data Portal is an initiative of the European Commission, implemented with the support of a consortium led by Capgemini Invent, including Intrasoft International, Fraunhofer Fokus, con.terra, Sogeti, 52North, Time.Lex, the Lisbon Council, and the University of Southampton. The Publications Office of the European Union is responsible for contract management of the European Data Portal.

For more information about this paper, please contact:

European Commission

Directorate General for Communications Networks, Content and Technology
Unit G.1 Data Policy and Innovation
Daniele Rizzi – Policy Officer
Email: daniele.rizzi@ec.europa.eu

European Data Portal

Gianfranco Cecconi, European Data Portal Lead
Email: gianfranco.cecconi@capgemini.com

Written by:

Esther Huyer
Email: esther.huyer@capgemini.com
Gianfranco Cecconi
Email: gianfranco.cecconi@capgemini.com

Last update: 15.07.2020
www: <https://europeandataportal.eu/>
@: info@europeandataportal.eu

DISCLAIMER

By the European Commission, Directorate-General of Communications Networks, Content and Technology. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use, which may be made of the information contained therein.

Luxembourg: Publications Office of the European Union, 2020
© European Union, 2020



OA-BF-20-012-EN-N ISBN: 978-92-78-41997-4 ISSN: 2600-0601 doi: 10.2830/078126



The reuse policy of European Commission documents is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Except otherwise noted, the reuse of this document is authorised under a Creative Commons Attribution 4.0 International (CC-BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

Table of Contents

- 1 Introduction 4
- 2 Benefits 6
 - 2.1 Creating value for the public 6
 - 2.2 Creating value for businesses 8
- 3 Challenges for data providers 10
 - 3.1 Organisational considerations 10
 - 3.2 Technical requirements 12
 - 3.3 Governance and legal challenges 13
- 4 Models and examples 15
 - 4.1 Multi-party data sharing agreement 15
 - 4.2 Data donorship 16
 - 4.3 Data partnerships 17
 - 4.4 Data Intermediaries 18
 - 4.5 Data sharing by regulation 19
- 5 A six-step action plan for B2G data sharing 20
- 6 Summary and the way forward 23

1 Introduction

*As we enjoy great advantages from the inventions of others,
we should be glad of an opportunity to serve others by any
invention of ours, and this we should do freely and generously.*
Benjamin Franklin

Data sharing is a key enabler of growth, employment and competitiveness for Europe, as well as for the Digital Single Market¹ envisaged by the European Union. The non-rivalrous nature of data, combined with technological innovations such as the availability of big data analysis and artificial intelligence applications, enable maximising the value of data. Re-using data can save costs, time and lives.²

The benefits from data re-use are not only reserved to the private sector. In fact, to become more cost-efficient and to provide effective services for citizens, public sector bodies can benefit greatly from data sharing and need to exploit the potential of new data sources.³

Data can be sourced from the private sector, academia, NGOs or the public sector itself. Much of the data generated in the public sector is already made open for re-use, encouraged by the Directive on Public Sector Information⁴.

However, there is also data that cannot be made open because of sensitivity or confidentiality. It is common to find cases of this in the private sector, for all of those data sets that describe the customer in industries such as telecommunications, transport, car manufacturing and services, retail and health. This data can, for example, be generated by:

- the Internet of Things (e.g. the data collected by your car's onboard computer as you drive),
- IT systems supporting logistics and services (e.g. payment card transaction data),
- websites and social media platforms, as they monitor their users' activity (e.g. search engine usage).

Despite the fact that this privately-held data cannot always be made open, it can still provide insight into behavioural patterns of citizens and businesses across social, political, historical or environmental factors. The insights can help public sector organisations understand, evaluate, predict and prepare for certain situations and scenarios, for example:

- Understanding commuting patterns can support urban planning, road safety and traffic management, as well as environmental protection.
- Insight into a population's health conditions, diagnosis and medical treatments can improve public health care and lead to a more effective response to epidemics.

¹ European Commission – Shaping the digital single market. Available at: <https://ec.europa.eu/digital-single-market/en/policies/shaping-digital-single-market>

² European Data Portal: (2017): <https://www.europeandataportal.eu/en/highlights/creating-value-through-open-data>

³ European Commission (2001). Data for Policy: A Study of big data and other innovative data-driven approaches for evidence-informed policy-making, report by Technopolis, Oxford Internet Institute, CEPS, drafted for the European Commission. Available at: <https://ofti.org/wp-content/uploads/2015/05/dataforpolicy.pdf>

⁴ European Commission - European legislation on the re-use of public sector information. Available at : <https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information>

- Market monitoring and payment patterns can help detect fraud and increase consumer protection.

Since much of this data cannot be made open, it can only be shared under special conditions and to a restricted and controlled set of users in the public sector in order to leverage their value. Personal or confidential data increases the complexity of data sharing. Legal, technical, and organisational factors must be considered when setting up a framework for data sharing between businesses and public organisations. Research already identified and reviewed different approaches, which aim at explaining, encouraging, guiding and safeguarding Business to Government (B2G) data sharing.

Research background

To further facilitate the [European Data Economy](#)⁵, the European Commission initiated the [Digital Single Market strategy in 2015](#)⁶. The 2017 Communication "[Building a European Data Economy](#)"⁷ discussed these opportunities, and in the mid-term review of the Digital Single Market strategy, the Commission committed to exploring this issue further.

In addition, the results of the public consultation undertaken in the context of the review of the Directive on the re-use of public sector information showed support for the idea of improving access to private sector data for public authorities for public interest purposes in general ("business-to-government" or "B2G" data sharing).

The European Commission performed a [public online consultation](#)⁸ and held a [stakeholder dialogue](#)⁹ between April and June 2017 including activities such as the [Workshop on access for public bodies to privately-held data of public interest](#)¹⁰ to debate the best practices to share and access private sector data for public interest purposes. Workshop participants voiced the preference for enhanced guidance rather than stricter regulation. The later [Communication: Towards a common European data space](#)¹¹ and a Staff Working Document "[Guidance on private sector data-sharing](#)"¹² (April 2018) took this feedback into account.

These publications provide guidance in the form of principles of B2G data sharing, models, legal and practical considerations and technical means. The documents aim at supporting the supply of private sector data to public bodies under preferential conditions for re-use.

⁵ Building a European data economy (2019) <https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy>

⁶ A Digital Single Market for Europe: Commission sets out 16 initiatives to make it happen (2015) http://europa.eu/rapid/press-release_IP-15-4919_en.htm

⁷ Building a European Data Economy (2017) http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=41205

⁸ Summary report of the public consultation on Building a European Data Economy (2017) <https://ec.europa.eu/digital-single-market/en/news/summary-report-public-consultation-building-european-data-economy>

⁹ Stakeholder Dialogue on Building a European data economy (2017) <https://ec.europa.eu/digital-single-market/en/news/stakeholder-dialogue-building-european-data-economy>

¹⁰ Workshop on access for public bodies to privately-held data of public interest (2017) <https://ec.europa.eu/digital-single-market/en/news/workshop-access-public-bodies-privately-held-data-public-interest>, webcast available at <https://webcast.ec.europa.eu/expert-workshop-on-access-for-public-bodies-and-scientists-to-commercially-held-data-of-public-interest>, summary at http://ec.europa.eu/information_society/newsroom/image/document/2017-28/final_-_report_from_reverse_psi_workshop_B7FA94EE-FA15-1929-8BBA2754D0D2FBE9_45916.pdf

¹¹ Communication "Towards a common European data space" (2018) <https://ec.europa.eu/digital-single-market/en/news/communication-towards-common-european-data-space>

¹² Staff Working Document - Guidance on sharing private sector data in the European data economy (2018) <https://ec.europa.eu/digital-single-market/en/news/staff-working-document-guidance-sharing-private-sector-data-european-data-economy>

Other stakeholders like [GovLab](http://www.thegovlab.org/)¹³ and [The Lisbon Council](http://www.lisboncouncil.net/)¹⁴ (member of the European Data Portal consortium) also dedicated their research to analyse different aspects of B2G data sharing and collected good practices. GovLab, for example, formulated eight phases of B2G data sharing and, as well, categorises different models of B2G data sharing in their DataCollaboratives¹⁵ project.

Reviewing the different findings enables us to compare and test the applicability of theoretical models and principles of B2G data sharing to how they address legal, technical and organisational challenges and other matters that may be yet unaddressed. A six-step plan for successful and sustainable B2G data sharing we developed from our research completes the report.

2 Benefits

Benefits of B2G data sharing accrue in a variety of ways. We differentiate a) the benefits for those businesses providing the data, and b) the public sector re-using the data and create benefits for the public. For the former, the value of B2G data sharing can realise in an improved public image - due to a proof of their commitment to corporate social responsibility, new insight from existing or additional data and its analysis and potential revenue from exchanging the data and reusing it. In the latter, advantages come from improved decision making and service offering. Moreover, advantages come from increased efficiency, higher innovation power in research, technological advancements and policy making and the manifold benefits of further developing a European data sharing culture beyond what was already achieved with PSI and open data. The benefits of B2G data sharing for businesses and public organisations are interlinked and affect each other. For example, higher innovation power in the public sector creates new opportunities for businesses and vice versa.

2.1 Creating value for the public

Data-informed decision making and policies

Re-using relevant privately held data increases the public sector's ability to understand, assess and predict different situations and phenomena that affect the citizens. It enables more logical and fact-based decisions, and at a higher pace. It helps react to and prepare for scenarios, like a financial crisis, ecological disaster, urban planning, crime, etc. The impact and reaction to decisions and actions can be monitored and evaluated to learn from it or to intervene. It creates a solid foundation for day by day decision making as well as for strategic regulatory initiatives and new policies.¹⁶

Starting with Rio De Janeiro during the 2016 Olympics, Waze has been sharing with local governments the data its users collect as they drive, under the "Connected Citizens Program"¹⁷. Together with partners such as with Waycare¹⁸ they also set up a platform that is intended to support public bodies with urban planning, law enforcement, traffic operations, traffic engineering, emergency services, etc.

¹³ GovLab: <http://www.thegovlab.org/>

¹⁴ Lisbon Council: <http://www.lisboncouncil.net/>

¹⁵ Data Collaboratives (2018): <http://datacollaboratives.org/canvas>

¹⁶ European Commission (2001). Data for Policy: A Study of big data and other innovative data-driven approaches for evidence-informed policy-making, report by Technopolis, Oxford Internet Institute, CEPS, drafted for the European Commission. Available at: <https://ofti.org/wp-content/uploads/2015/05/dataforpolicy.pdf>

¹⁷ Waze Connected Citizens Program: <https://www.waze.com/ccp>.

¹⁸ Waycare: <http://waycaretech.com/>

Improved public services

The availability of data to the national statistics agencies can increase the quality and scope of their official statistics, reduce the cost of producing them, and make situation-analysis and citizen-behaviour-analysis more insightful. Enabled to better understand, assess and predict behaviour, public services can then be designed or adapted accordingly.

The Dutch Government organisation [Centre for Big Data Statistics](#) combines the large repository of government held data about communities and individuals with data from several private partners, e.g. Google, to effectively observe, measure and describe developments in society and create high-quality visualisations and community data for citizens.

Increased efficiency

European Data Portal's research has already highlighted how greater efficiency can be achieved by successful data sharing even just of *open* data, in both internal administrative processes and the delivery of services to the public. It is forecasted that the accumulated cost savings for public administrations in the EU28+ in 2020 will be of 1.7 bn. EUR.¹⁹ Extending the re-use of data to private sector resources can only expand this potential.

Leadership

Data sharing will make public organisations recognised as leaders in this relatively unexplored space. In turn, this will make them worth the trust of organisations and individuals. Further experience will be developed, and experience will feed the process by which regulation adapts to, and supports data sharing.

Data sharing requires both that experience – and the guidance sourced in that experience - and regulation to ensure success.

Since the models and practices for B2G are so new, experts and role models are needed, and their stories documented. This will create a corpus of knowledge and guidance on top of which further experience and understanding will be developed.

Their digital leadership, in combination with an effective regulatory framework, can increase awareness and trust in B2G and data sharing in general. This would strengthen the opportunities to effectively shape the future of the EU data economy.

A healthy European data sharing culture

One of the ultimate goals of any society striving for human development is the empowerment of all its citizens, through access to and use of information and knowledge: “The freedom to seek, receive and impart information and ideas through any media and regardless of frontiers.”²⁰ and “to share in scientific advancement and its benefits.”²¹ are human rights declared by the United Nations. The Charter of Fundamental Rights in the European Union also highlights that “Everyone has the right to

¹⁹ EC/EDP creating value through Open Data, available at:

https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

²⁰ UN (1948): Article 19 of the Universal Declaration of Human Rights

²¹ UN (1948): Article 27(1) of the Universal Declaration of Human Rights

(...) receive and impart information and ideas without interference by public authority and regardless of frontiers.”²²

The work done by the European Union to develop a data sharing culture started well before B2G could be even conceived, and focused initially on public sector Information and open data, aware that sharing data encourages a culture of openness and accountability, strengthens the European spirit and allows individuals and organisation to engage in a digital world. B2G data sharing adds a whole new dimension to the European data sharing culture. The amount of data that could be used for the public benefit – though not in the open because of confidentiality or privacy considerations – is vast, but its protection is paramount, and its exploitation must take place in the light of ethical considerations and within the limits set by law. Together with its benefits, the development of B2G data sharing implies also developing awareness and responsibility of those constraints and obligations. Privacy, for example, is also a fundamental right in the European Union.²³

2.2 Creating value for businesses

Businesses, whether active at a local, national and international level, also benefit from the indirect benefits created by the public sector re-using private sector data. Below are a few initiatives enabled by B2G data sharing that create value for businesses sharing their data with a public organisation.

- **Corporate Social Responsibility (CSR):** As a form of CSR in the data age, collaborating by offering data and data skills – e.g. by offering the time of your data scientists or, in general, your most data-savvy employees - can improve companies’ reputation and enhance community relationships.

Since 2011 non-profits such as DataKind have been mobilising data scientists to volunteer for other non-profits that could not afford their work, or whose maturity is not yet ready to justify the investment of a “data function” in the organisation. They implement a “data science for social impact” model for collaborative philanthropy aimed at accelerating the use of data by empowering non-profit, civic, and public sector organisations with the tools, expertise, and other capabilities they need to help solve the world’s most pressing challenges. Their success is such that DataKind has been recently awarded a USD 20m grant by The Rockefeller Foundation and the Mastercard Impact Fund to support further growth.²⁴

- **Employer branding:** Data sharing can attract and retain talent. An improved image and better community relationships is not only vital for customer attraction and loyalty but also for employer branding. A positive image supports recruiting and retaining employees for the organisation.
- **Insight:** By sharing their corporate data, or the data they know about their customers and suppliers in ways that are respectful of confidentiality and privacy, businesses enable public sector organisation to answer questions they could not have found by working in isolation.

²² EU (2012): Article 11 of the Charter of Fundamental Rights of the European Union.

²³ EU (2012): Article 8 of the Charter of Fundamental Rights of the European Union.

²⁴ DataKind (2019): A Step Change: DataKind Raises \$20M Investment to Support the Data Science for Social Good Ecosystem, <https://www.datakind.org/blog/a-step-change-datakind-raises-20m-investment-to-support-the-data-science-for-social-good-ecosystem>

In many cases, public sector organisations may be the only bodies the public would trust to, and feel comfortable with, setting up and managing such an endeavour. If the same was done by a private organisation, the same trust would be harder to gain, and concerns around the exploitation of the data for commercial purposes would easily arise, and possibly be sufficient to obstruct the initiative.

The struggle of past initiatives in this space illustrates the challenge private and public organisations still face in make data sharing collaboration models work and achieve the trust of the public.

Lots has been written, for example, about the collaboration between the UK National Health Service (NHS) and DeepMind Health, a Google company. This was not a B2G project, but a more “conventional” deal in which a private organisation processes data for a public sector body – patients’ data in this case – for some specified purpose.

Conceived before the EU General Data Protection Regulation became enforceable, the project raised many concerns around the privacy of patients.²⁵ Moreover, the NHS – DeepMind deal could be considered an example of potential data lock-in, restricting opportunities for other organisations to research the data and benefit the public.

At the same time, the British public is reported to be welcoming sharing their very personal and sensitive health data with the health authority²⁶, and early trials suggest significant efficiencies, e.g. in the successful automation of eye disease identification.²⁷

- **Reciprocity:** Public sector organisations own significant physical and digital assets and manage many aspects of the citizens’ lives, directly and indirectly, through services they provide to the public. B2G data sharing can be a means by which the private sector reciprocates for the value they get from these services.

Local government is often responsible for licensing taxi services, controlling the volume of cars circulating in the city, setting up quality standards, controlling and enforcing its rules. Innovation in mobility has challenged traditional models and way of working, often creating tension between public sector, traditional and new operators.²⁸

In an attempt to attenuate this tension, in 2018 major ride-hailing service Uber launched “Movement”²⁹: a web portal providing data from over two billion trips covering, in Europe, cities such as Amsterdam, London and Paris. Anonymised, aggregated data is made available

²⁵ Read for example C. Stokel-Walker “Why Google consuming DeepMind Health is scaring privacy experts”, Wired.co.uk, 2018, <https://www.wired.co.uk/article/google-deepmind-nhs-health-data> ; M. Murgia, “Inside DeepMind as the lines with Google blur”, Financial Times, 2018, <https://www.ft.com/content/c26893d0-e9b0-11e8-a34c-663b3f553b35> and A. Ram, R Waters, “Alphabet AI unit urged to clarify its business model”, Financial Times, 2018, <https://www.ft.com/content/215062da-6fe3-11e8-852d-d8b934ff5ffa> .

²⁶ M. Field, “Over half of Brits happy to share health data with NHS for AI, new report finds”, The Telegraph, 2018, <https://www.telegraph.co.uk/technology/2018/09/04/half-brits-happy-share-health-data-nhs-ai-new-report-finds/> .

²⁷ M. Burgess, “Now DeepMind’s AI can spot eye disease just as well as your doctor”, Wired.co.uk, 2018, <https://www.wired.co.uk/article/deepmind-moorfields-ai-eye-nhs> .

²⁸ E.g. read about New York City in S. Wodinsky, “In major defeat for Uber and Lyft, New York City votes to limit ride-hailing cars”, 2018, The Verge, and London in A. Ram and T. Powley, “New TfL rules put further brake on Uber’s ambitions”, 2018, Financial Times, <https://www.ft.com/content/d4a808a8-1258-11e8-8cb6-b9ccc4c4dbbb> .

²⁹ Uber Movement: <https://movement.uber.com/> .

to the public as open data, while they make themselves available to government experts, transportation professionals, and academic organisations to discuss research opportunities.

Uber's declared ambition is to support urban planning around the world, as access to its data can contribute to inform decisions about how to adapt existing infrastructure and invest in future solutions to make cities more efficient.

Think tanks and common opinion in the public occasionally push the concept so far as to say that public sector organisations should *demand* private organisations to share their data, in a context of reciprocity,³⁰ the reasoning being that data such as geospatial and mapping data, for example, is vital to a country's economy, and the formation of "data monopolies" stifles innovation. The UK government has previously estimated that if such data was shared, it could generate between GBP 7 billion and GBP 14 billion per year of economic and social value for the country.³¹

- **Commercial:** Where a business case for it exists and is solid and sustainable, public sector may justify acquiring commercially the right to access data sourced from private organisations. The formulas by which this can take place are multiple and go beyond simply "purchasing" a data set, e.g. the data provider could be granted preferential conditions or tax discounts.

It has also to be considered that there is an entire market for B2B data sharing, too, that is outside of the scope of this document. It is a significant potential revenue stream, and the same changes required of an organisation to enable B2G data sharing would also enable them to access B2B data sharing markets.

3 Challenges for data providers

Challenges for B2G data sharing occur on both sides, the data provider (businesses and other private organisations) and the data user (a public body). This section focuses on the organisational, technical and legal challenges to data re-use from the *businesses' perspective*. An evaluation of the same challenges but in the public sector, and solutions to overcome these, can be found in the European Data portal's analytical report: "PSI re-use in the public sector".³²

3.1 Organisational considerations

Costs and benefits

Businesses must weigh the benefits against costs and risks of a given data sharing initiative.

Costs can occur for investments in new data infrastructure to enable data processing, including setting up any channel necessary to share with third parties. A significant cost factor is also perceived to be the cost of resources needed to gather, validate and prepare the data before is shared. Scattered data

³⁰ Open Data Institute, "The UK's geospatial data infrastructure: challenges and opportunities", 2018, <https://theodi.org/article/geospatial-data-infrastructure-report/>.

³¹ UK Cabinet Office, "An Initial Analysis of the Potential Geospatial Economic Opportunity", 2018 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/733864/Initial_Analysis_of_the_Potential_Geospatial_Economic_Opportunity.pdf.

³² European Data Portal, "Analytical Report 11: Re-use of PSI in the public sector", 2019, https://www.europeandataportal.eu/sites/default/files/analytical_report_11_psi_re-use_in_the_public_sector.pdf.

and ineffective organisational structures are holding up sharing data. Research³³ shows that 76% of the companies interviewed have not completely integrated their data sources across the organisation and only 35% have robust processes for data capture, curation, validation and retention. 54% do not have joined project teams where business and IT executives work together on data initiatives. 47% have either scattered pockets of resources or follow a decentralised model for analytics and initiatives.

A barrier in formulating such a business case is that the costs are more tangible and straightforward to calculate but the **value of the benefits** of data sharing can be difficult to put into numbers. A majority of companies does not have well-defined criteria to measure the success of their data initiatives.³⁴

Another barrier is the evaluation of **risks** (e.g. data breach, misuse, non-compliance) and how to measure the opportunity risks of not sharing data. Typical formulas including the likelihood and the severity of implications are challenging to define since there are barely any reference cases.

As a final note, it is important to consider that many of these cost components are likely related just to the process by which organisations need to adapt from an old-fashioned, inward-looking data infrastructure and processes to modern ones, suitable to support data sharing. A modern business can be designed from the ground up with sharing, interoperability and data governance principles in mind, reducing many of the cost components described above to a minimum.

Customer reaction

Beyond the legal and financial risks, companies voice concern around the ethical implications of data sharing for their customers, when that might include customer data. This is an issue even when all the necessary precautions are taken to protect confidentiality and privacy, and when the company puts all the necessary effort in communicating the scope of the data sharing initiative transparently and thoroughly

Whether based on fact or just speculation, the reaction of clients and customers might be negative and damage reputation and perceived corporate social responsibility, as a violation of trust. This is especially necessary to consider if data is shared for a commercial gain. The general lack of data literacy in the public cannot but make the problem worse, as even perfectly legit data sharing initiatives may be misunderstood and considered malicious.³⁵

Uncertainty around potential reputational risks can prevent companies from even start evaluating a business case for data sharing, because the support of the businesses' leadership is likely to fail to materialise.

Capability and culture

³³ Capgemini Insights & Data Research (2015). Cracking the Conundrum: How Successful Companies Make Big Data Operational. Available at: https://www.capgemini-consulting.com/sites/default/files/resource/pdf/big_data_pov_03-02-15.pdf

³⁴ Capgemini Insights & Data Research (2015). Cracking the Conundrum: How Successful Companies Make Big Data Operational. Available at: https://www.capgemini-consulting.com/sites/default/files/resource/pdf/big_data_pov_03-02-15.pdf

³⁵ B. Dykes, "Why Companies Must Close The Data Literacy Divide", 2017, Forbes, <https://www.forbes.com/sites/brentdykes/2017/03/09/why-companies-must-close-the-data-literacy-divide/#6a7eb488369d>.

The absence of **leadership support** and of data sharing initiatives, in general, can also be traced back to a lack of data **literacy, skills, awareness** and a local or corporate culture that does not support data sharing or mistrusts data collaboration.

People are at the heart of the process and empowering people to understand, act and learn from data sharing makes **cultural change** and training necessary. More organisations are developing **training** on business intelligence techniques as well as analytics, agile working methods as well as an evolution of the Human Resources policy to enable to attract, train and retain analysts and data scientists. In the field of data, e-skills need to pursue their growth to ensure all employees become data savvy enough to understand the underlying principles of data and data sharing, the basic logic followed by algorithms as well as the benefits of the data and algorithm economy.³⁶ However, change management is equally essential to unleash the transformative benefits of data and data sharing and feeding the learnings back into the organisation to drive change.

3.2 Technical requirements

A few technical processes are specific to data sharing. The B2G context may add a degree of complication that we try to capture in this section.

Selection and preparation

Data providers face several barriers with regard to basic processes such as gathering and selecting the data to be shared in the first place and then preparing it for sharing.

Shared data is mostly obtained as extracts from an organisation's operational or production data. Hence, data providers must make sure that they do not share data they do not have the right to share (e.g. intellectual property constraints), or that is confidential, e.g. it may reveal components of their competitive advantage, or expose trade secrets of their business model.

When **selected**, and before the actual publication, to ensure its value the data must be **cleaned** up so that it can be processed by users. It needs to be checked for aspects like **accuracy, consistency and completeness**. Common issues that are addressed in the clean-up process include, for example, **data formats** that are inconsistent and not compliant with the standard, **redundant entries and duplicates**, the use of mixed numeric scales, spelling errors etc. Data often needs to be enriched by adding both **documentation** and **metadata** to increase discoverability by and usability for data users.

Anonymisation, pseudo anonymisation and aggregation

Data anonymisation, pseudo anonymisation and aggregation are a few of the most common techniques by which information is sanitised to protect the privacy of the people described in the data. These enable data that would otherwise be considered personal and confidential to be processed without the extra care and legal constraints required by the original datasets. The principles of data protection defined in the EU General Data Protection Regulation (GDPR), for example, do not apply to anonymised information, as long as it is "rendered anonymous in such a manner that the data subject is not or no longer identifiable."³⁷

³⁶ European Commission, Re-Using Open Data, European Data Portal, A study on companies transforming Open Data into economic and societal value (2017): https://www.europeandataportal.eu/sites/default/files/re-using_open_data.pdf

³⁷ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), recital 26, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679>.

Given a suitable legal basis and ownership of the necessary intellectual property rights, data processed as described may also be published as open data, hence even further multiplying its benefits. This is common practice, for example, of many national statistics authorities, e.g. with census information³⁸. Opportunities of doing the same in the private sector are less common but feasible and sought after. In conversations we have had, for example, with financial institutions, we recorded interest in persuading governments to publish in the open anonymised versions of data that they are required to provide for compliance with applicable regulations.

Technically, **anonymisation** is the process of removing personally identifiable information from datasets so that the people whom the data describes remain anonymous. Similarly, **pseudonymisation** replaces personal identifiable information with artificial identifiers, or pseudonyms; for example, a name might be replaced with a number, but the same number will show up in every instance of that name. This enables the data provider to re-identify the individual if useful. Finally, **aggregation** changes the granularity of the dataset, replacing the data describing a set of individuals with aggregated statistical value that is relevant to the objectives of data processing, e.g. the average age of a group of people, rather than each individual's age. Census data, for example, is usually both anonymised and aggregated, e.g. to protect

Unfortunately, the proper application of these techniques is not trivial, though often underestimated. The incorrect application may enable, for example, the “re-identification” of a few or all the people described in the original dataset.

Access

The model by which users **access** data is also a decision that organisations must take in all sharing scenarios including B2G. Data can be shared in multiple ways, most commonly via download or Application Programming Interfaces (APIs). Each method has its pros and cons, e.g. downloading is suitable to datasets that are static or change rarely and enables the publisher to benefit from savings on the consumption of resources such as bandwidth and computing power. Conversely, the use of APIs is almost inevitable for live data, or to enable the user to access only those parts of large datasets that they really need.

Differently than with open data, sharing data with a controlled audience comes with risks regarding **information security** (infosec). The consistency, integrity and availability of data can be compromised. Infosec must be implemented to provide authorisation, authentication, access control and auditing functionality end to end. Networks and information systems in general need to be protected. Without proper infosec, data providers will be hesitant about sharing their data, and the whole ecosystem risks missing out on the benefits. Security must be consistent across the systems of all participants to a B2G sharing initiative, as issues could be caused by a vulnerability anywhere, in any of the systems and networks involved. This also once again highlights the need for the people involved to be suitably skilled, as humans are often the weakest link in ensuring the security of systems.

3.3 Governance and legal challenges

Data governance

³⁸ E.g. find the UK 2011 census data at <https://www.nomisweb.co.uk/census/2011> .

To safeguard and manage data sharing in the long run, a structured **data governance** set-up is needed in which data stewards and officers are guardians of data generation, preparation, sharing and governing. When the actual sharing of the data is initiated, data can change in structure and content or the related rights can change. Over the course of time, the sharing framework must be adapted and shared data needs to be updated.

Legal

B2G data sharing is still a rather unexplored territory from a legal perspective. There is an absence of a specific legal framework and a lack of reference cases, nor documentation about best practices or standard scenarios. Although different legal regulations, like Intellectual property rights, data protection rights, competitions laws, etc. all cover in different ways certain legal aspects of B2G data sharing, in most cases the support of a legal expert is necessary when setting up a data sharing initiative.

An interesting aspect of uncertainty is **liability**: determining who – if anyone – can be held responsible when bad quality data is shared and contributes to a wrong decision by a public body, possibly damaging citizens or causing financial loss. The risk of being considered accountable, despite having shared data in good faith, can be a strong deterrent.

Finally, another relevant aspect is the possible **inconsistency of relevant legal frameworks internationally**, e.g. across the EU Member States. This naturally creates barriers for B2G data sharing between countries. In addition, uncertainty about how the regulation of such a new space will evolve in the future creates further friction to sharing.

Intellectual property

Another limitation to data sharing, or at least an element of uncertainty, can be intellectual property. Companies may not have the necessary certainty of having the right to share the data they process every day. Company data is often the combination of very heterogeneous sources, from historical archives whose origin is no more known, to datasets that are the results of working with suppliers and contractors, whose relationships may not have been regulated by contracts that were explicit as of the intellectual property of the data produced.

Competition law

Interestingly, during our research we found anecdotal evidence of businesses being concerned that data sharing - whether B2G or B2B – could be potentially interpreted by regulators as anti-competitive behaviour.

Competition law concerns arise when, for example, the nature of information exchanged between current or potential competitors enables them to predict each other's behaviour and adjust their own accordingly, e.g. to fix prices or allocate customers or markets. Occasional exchange of information between businesses already happens every day. However, repeated and substantial sharing of data could be considered in breach of competition law and create exposure to the risk of fines or even criminal sanctions for individuals.

In the case of B2G data sharing, the public sector bodies and the businesses involved may be concerned, for example, that their collaboration could be considered unfair to competitors. For example, this would happen if a privileged relationship was created between the parties, enabling

perhaps one business to access confidential information regarding the public body that its competitors would not have. The upcoming new PSI and Open Data Directive³⁹ addresses some components of this.

Personal data

The importance of respecting the rights of individuals described in the data has been highlighted already many times in this document. In the EU, the **General Data Protection Regulation (GDPR)**⁴⁰, is likely the single most important piece of regulation relevant to this matter. Together with other regulation relevant to privacy protection, the principles captured in GDPR create an environment of “digital trust” that is a key condition for any data sharing model, hence enabling the ideal of the Digital Single Market promoted by the European Commission.

GDPR became enforceable across the EU on 25 May 2018, and includes principles such as purpose limitation, transparency, data minimisation, and security (including integrity and confidentiality). Purpose limitation implies that personal data is not used in any way that’s incompatible with the originally envisaged purposes at the time of collection. Transparency requires that data subjects must be informed appropriately about how their data is processed. Data minimisation means that the personal data must be limited to data which is adequate, relevant and necessary in relation to the purposes for which they are processed.

4 Models and examples

Different actors in research and politics have dedicated their efforts to categorise B2G data sharing models. In general, it can be said that there is not one approach that suits all situations and even the different models must be adapted based on each individual situation. Differentiating different models still makes sense in order to understand different answers to different needs, especially in the state of designing and assessing data sharing frameworks.

As a result of reviewing research findings, we have identified the B2G data sharing models defined below. However, these models are not set in stone, they much more occur in combination or in hybrid forms.

- **Multi-party data sharing agreement**
- **Data donorship**
- **Data partnerships**
- **Data intermediaries**
- **Data sharing by regulation**

In this report, we exclude the open data model, because, as mentioned before, by making data open, many barriers to sharing and re-use are removed.

4.1 Multi-party data sharing agreement

³⁹ See http://europa.eu/rapid/press-release_IP-19-525_en.htm .

⁴⁰ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), recital 26, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679> .

Description	This is the obvious traditional model. Multiple parties, including one or more public bodies, make an explicit agreement regulated by a contract to share a specified set of data, most often triggered by the need of only a few of the participants (the “requesters”, that is the public bodies in a B2G scenario) to access data owned by others (the “suppliers”).
Key characteristics	<ul style="list-style-type: none"> • one-time sharing, time- and purpose-bound • one-directional sharing (from the suppliers to the requesters)
Benefits	<ul style="list-style-type: none"> • the requesters can specify exactly what they need and have direct access to the data for processing • by being a relatively traditional sharing model, it is less subject to public scrutiny or misunderstanding than the more innovative alternatives described in this section
Challenges	<ul style="list-style-type: none"> • complexity of setup, not too dissimilar from a procurement process, with the additional complexity of having to regulate the provision of digital rather than physical assets and services • no incentive for suppliers but for financial gain, which may not be sufficient to build a business case • there is no certainty for the requesters that the data being shared will address their needs, in terms of scope, quality etc.; however, an assessment over samples of the data can mitigate the risk
Example	Public administrations purchase regularly from private sector consultancies the right to access their geospatial data and intelligence services to be used, for example, in urban planning.

4.2 Data donorship

Description	Donorships can be considered as form of corporate social responsibility. Data providers share data with a public body pro bono to serve the greater good.
Key characteristics	<ul style="list-style-type: none"> • one-time sharing • one directional sharing
Benefits	<ul style="list-style-type: none"> • simple set-up of data sharing framework • easier to define organisational and contractual agreements, if any are necessary
Challenges	<ul style="list-style-type: none"> • little incentive for companies, beyond reputation, employer branding, philanthropical reasons, etc. • initiatives need to be communicated properly to achieve public recognition, otherwise impact on the organisation’s brand and image can be limited.

	<ul style="list-style-type: none"> depending on the maturity of the public sector body receiving the data, the donorship context may make the contractual obligations on the data provider looser; e.g. the donor may not be transparent about what data is donated, and it may be selected to give a biased perspective of the phenomenon being described. more generally, a non-commercial set up may misleadingly loosen the participants' attention to the legal implications of sharing, that are not any less, e.g. with regard to privacy
Example	<ul style="list-style-type: none"> Twitter provides UN Global Pulse with access to their data tools to support efforts to achieve the UN Sustainable Development Goals, specified in 2015. Tweets can contain real-time information on many issues including the cost of food, availability of jobs, access to health care, quality of education, and reports of natural disasters. Though the tweets are available in the public domain, Twitter's own tools to access and analyse content are normally available to third parties only commercially. The partnership allows the development and humanitarian agencies of the UN to turn these social conversations into actionable information to aid communities around the globe⁴¹. Supermarket chains in Denmark and the Netherlands shared their scanner data at no cost with the national statistical agencies for economic development monitoring and to produce the consumer price index (CPI).^{42 43} In 2018 ride-hailing service Uber launched "Movement"⁴⁴: a web portal providing anonymised data from over two billion trips covering, in Europe, cities such as Amsterdam, London and Paris. Uber's declared ambition is to support urban planning around the world, as access to its data can contribute to inform decisions about how to adapt existing infrastructure and invest in future solutions to make cities more efficient.

4.3 Data partnerships

Description	Partnerships are characterised by the pursuit of mutual interests and are often found in academia and research-intensive sectors like the pharmaceutical industry and transport.
Key characteristics	<ul style="list-style-type: none"> mutually beneficial mid- to long-term
Benefits	<ul style="list-style-type: none"> individual value proposition for both parties

⁴¹ See <https://www.un.org/sustainabledevelopment/blog/2016/09/twitter-and-un-global-pulse-announce-data-partnership/>.

⁴² UNstat (2010): <https://unstats.un.org/unsd/EconStatKB/KnowledgebaseArticle10379.aspx>

⁴³ Statistics Denmark (2013):

[http://www.ottawagroup.org/Ottawa/ottawagroup.nsf/4a256353001af3ed4b2562bb00121564/8bdac0e73d96c891ca257bb00002fdb4/\\$FILE/Nina%20Gustafsson%20Drawing%20a%20Sample%20from%20Scanner%20Data%20to%20use%20in%20the%20Danish%20CPI%20-%20Ottawa%20Paper%20may%202013.pdf](http://www.ottawagroup.org/Ottawa/ottawagroup.nsf/4a256353001af3ed4b2562bb00121564/8bdac0e73d96c891ca257bb00002fdb4/$FILE/Nina%20Gustafsson%20Drawing%20a%20Sample%20from%20Scanner%20Data%20to%20use%20in%20the%20Danish%20CPI%20-%20Ottawa%20Paper%20may%202013.pdf)

⁴⁴ Uber Movement: <https://movement.uber.com/>.

	<ul style="list-style-type: none"> • collaborative insight generation
Challenges	<ul style="list-style-type: none"> • not less complex set up than multi-party agreements, however, by switching the focus on the mutual benefits, the commercial component of the contract may be easier to agree upon • as with donorship, a non-commercial set up may misleadingly loosen the participants' attention to the legal implications of sharing, that are not any less, e.g. with regard to privacy⁴⁵
Example	Propeller Health, a private company shared their data with the Centre for Disease Control and Prevention (CDC) to support to identify trends for environmental triggers of asthma. In return, they benefited from access to data from the CDC. This enables Propeller Health to help physicians develop personalised treatment plans and spot prevention opportunities. ⁴⁶

4.4 Data Intermediaries

Description	To address a lack of trust to share data, a third party can be tasked to process the shared data to achieve a specified objective – e.g. produce insight about a phenomenon – without any of the participating parties having access to its full set. The intermediary could be an individual acting as a trustee, or organisations offering, for example, a data sharing platform.
Key characteristics	<ul style="list-style-type: none"> • data is shared with a trusted intermediary, and not between the participating parties • multiple layers of intermediaries are possible, e.g. the first layer may have a guarantee role because of their reputation – e.g. an advocacy organisation – while the second layer could provide the actual data processing services⁴⁷ • especially suitable to ensure complex confidentiality / privacy requirements
Benefits	<ul style="list-style-type: none"> • assuming that the intermediary is more trustworthy than the other parties involved, the use of an intermediary makes it possible to protect confidentiality and privacy from both a process and a technical perspective, in ways that are not otherwise possible by assembling the data in the hands of one of the partners • once the intermediary model is understood by the public, relying on a reputable intermediary can be an effective way to reassure the public of the legitimacy of the data sharing arrangement

⁴⁵ See for example the case of Dropbox sharing usage data with Northwestern University to support their research on the habits of scientists using the file-sharing service. Described in Wired (2018): <https://www.wired.com/story/dropbox-sharing-data-study-ethics/>

⁴⁶ Propeller health (2018): <https://www.propellerhealth.com/>

⁴⁷ Web Foundation (2015): <https://webfoundation.org/2015/08/open-data-intermediaries-their-crucial-role/>

Challenges	<ul style="list-style-type: none"> • complex set up • qualification of the intermediary • scoping: how can we effectively instruct the intermediary to find the insight we need? Can the data be processed by an intermediary that does not have the subject matter expertise of the partners it provides the service to?
Example	<ul style="list-style-type: none"> • Artstor⁴⁸ is a nonprofit organisation that runs a digital library that includes millions of copyrighted high-quality images that are offered by the respective right holders to support art education and research • NPC Data Labs⁴⁹ enables businesses, non-profits and public sector organisations to better understand their performance in providing services to individuals through their respective channels, without the need of creating thorough unified profiles of personal data that could compromise their privacy. It works by enabling participating organisations to independently share the data describing the individuals – e.g. past offenders with a criminal record in the “Justice Data Lab” project — and the activities that involve them, returning back only aggregated insight. In the “Justice Data Lab” project, for example, the participating parties could see, which activities have resulted in a lower reoffending rate and which not. • In the attempt to address the weaknesses of a model that relies on an intermediary, the OPAL (for "Open Algorithms") is a non-profit project developed by a group of partners around the MIT Media Lab, Imperial College London, Orange, the World Economic Forum and Data-Pop Alliance.⁵⁰ In their model, “the code is sent to the data”, avoiding the need for an intermediary. OPAL combines privacy-preserving technology and a participatory governance system with an ethical oversight body and capacity building activities. It started in 2017 with pilots in Colombia and Senegal in partnership with their governments and local telecom operators.

4.5 Data sharing by regulation

Description	<p>Data sharing can also be obligatory. In sectors like health, pharmaceutical, chemical manufacturing and finance, the government can regulate which data must be shared with whom and in which way.</p> <p>In Europe, beyond sector-specific legislation, the European Commission looks at allowing public sector bodies to gain access to data held by private entities when such data is necessary to accomplish public goals and for public interest reasons (e.g. computing statistical indicators, operation of multimodal transport services, access to vehicle-generated data necessary for traffic) or to put in place and carry out emergency response actions (e.g. associated with imminent risk to public health or security).⁵¹</p>
-------------	---

⁴⁸ Artstor: <http://www.artstor.org/>

⁴⁹ NPC Data Lab: <https://www.thinknpc.org/examples-of-our-work/initiatives-weve-worked-on/data-labs/>

⁵⁰ See <https://www.opalproject.org/>.

⁵¹ Midterm review of the DSM (2017): <https://ec.europa.eu/digital-single-market/en/content/mid-term-review-digital-single-market-dsm-good-moment-take-stock>

Key characteristics	Data sharing is required by regulation, clearly defined and safeguarded
Benefits	<ul style="list-style-type: none"> • data sharing is clearly defined and guided within the regarding legislation
Challenges	<ul style="list-style-type: none"> • not feasible for all data sharing scenarios
Examples	<ul style="list-style-type: none"> • The Markets in Financial Instruments Directive 2 (MIFID2): it requires the disclosure of data on trading activity to the public, and of transaction data to regulators and supervisors • The Payment Services Directive 2 (PSD2): it requires payment service providers, among the other things, to implement API's that give authorised third parties access to their back-end data • The Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) regulation: it requires companies to register chemical substances with European Chemical Agency (ECHA)

5 A six-step action plan for B2G data sharing

Successful and sustainable B2G data sharing must be legally compliant, technically feasible, socially acceptable, financially and commercially viable and has to mitigate risk effectively.⁵²

Our research shows that the legal and technical aspects are relatively easy to overcome, once a data sharing cooperation is agreed upon. Then, it is merely a matter of getting the right resources in place and carrying the costs of any necessary extra effort.

However, projects in B2G data sharing often do not reach a state of maturity. A lack of trust, not only in the sharing partner but in the concept itself of data sharing hinders ambitions since the early stage. Data sharing still has a widespread negative image and businesses are reluctant to share their data because the benefits for their customers are difficult to communicate. This counts for sharing both personal and non-personal data. The negative image comes from a track record of incidents – whether due to human error or negligence, violations of privacy, actors not being transparent about their data practices, lack of compliance with regulation, or the misalignment between regulation and the ethical standards of the public.

That means that compliance with ethical standards and legal regulations in a publicly visible and convincing manner has to be ensured. We have identified six steps that appear to be vital to the success of a B2G initiative.

- 1. Team up with renowned and authoritative third parties**
- 2. Involve the customers**

⁵² Adapted based on Nuria Oliver, director of data science research at Vodafone and chief data scientist at the Data-Pop Alliance. Data-Pop Alliance is a global coalition of big data and development created by the Harvard Humanitarian Initiative, MIT Media Lab and the London-based Overseas Development Institute.

3. **Set up a data sharing framework**
4. **Use a code of conduct**
5. **Specify contractual agreements**
6. **Use new technologies**

1. **Team up:** Rely on the support of renowned, authoritative third parties and existing data sharing collaborations and partnerships that benefit of a good reputation, like the Big Data Value Public-Private Partnership⁵³ by the Big Data Value Association (BDVA) and the European Commission. This partnership of actors from European industry (large players and SMEs), researchers and academia cooperate in data-related research and innovation, enhance community building around data and set the grounds for a thriving data-driven economy in Europe.⁵⁴ A renowned partnership can provide guidance on how to set up a data sharing framework, it can increase the trust of the stakeholders involved directly or indirectly in the data sharing framework and it can support a positive perception of the public.

2. **Involve the customers:** In many cases, companies have the opportunity to share personal data about their customers. Even in those cases where companies do have the right to share, and even when the technology is correctly applied to protect their privacy, the lack of involvement of the customer and of proper communication can have unexpected consequences. For example, the so-called “civic data sharing” movement focuses on enabling the sharing of personal data by actively involving the people. Individuals are encouraged to find out about the data sharing initiatives and to authorise the companies processing their data to share it – even where they legally had already the right to do so. In some cases, by being involved this way, they even benefit from additional insight and services. This addresses the ethical and legal concerns of data sharing and tangibly communicates the benefits.

3. **Use a code of conduct** that defines rules, values and principles in a data sharing collaboration. The code creates a space where ethical concerns can be included and addressed. The European Commission has defined certain key principles (see the box on the right) for B2G data sharing, to provide guidance for a fair data environment⁵⁵. These principles cover the challenges of B2G data sharing and are a solid base for a code of conduct and convey credibility to the public. In addition, the nine principles for digital development by the UN⁵⁶ can be used as reference, too.

- Proportionality in the use of private sector data
- Purpose limitation
- ‘Do no harm’
- Conditions for data re-use:
- Mitigate limitations of private sector data
- Transparency and societal participation

⁵³ Big Data Value Public-Private Partnership: <http://www.bdva.eu/PPP> .

⁵⁴ DSM Big data value public private partnership, available at: <https://ec.europa.eu/digital-single-market/en/big-data-value-public-private-partnership>

⁵⁵ European Commission – Guidance on sharing private sector data in the European data economy. Available at <https://ec.europa.eu/digital-single-market/en/news/staff-working-document-guidance-sharing-private-sector-data-european-data-economy>

⁵⁶ UN Global Pulse (2015): <https://digitalprinciples.org/principles/>

4. Set up a data sharing framework. A data sharing framework addresses technical aspects, embeds legal regulations and ethical standards and anticipates personal concerns in order to maintain the trust of consumers and citizens and increase the value of data sharing. A framework must support the alignment between all actors and IT systems. Figure 1 suggest aspects to be addressed by a data sharing framework.

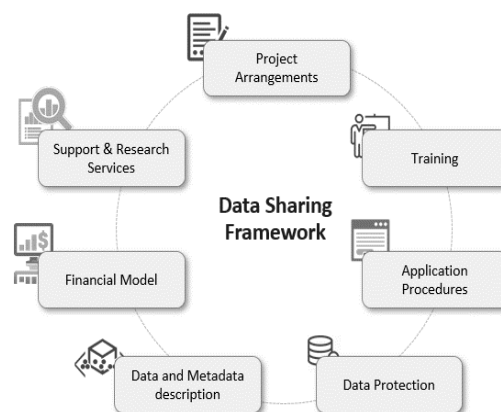


Figure 1 Data Sharing Framework

The European Interoperability Framework revised and updated in March 2017 under the ISA² Programme adopts a broad data sharing concept, defining it as “the ability of disparate and diverse organisations to interact towards mutually beneficial and agreed common goals, involving the sharing of information and knowledge between organisations, through the business processes they support, by means of the exchange of data between their respective ICT systems”. The framework defines interoperability principles in legal, organisational, semantic and technical interoperability to be integrated into a data sharing framework.

5. Specify contractual agreements. Due to the absence of a specific data sharing legal framework, a contract between data provider and data user must specify the agreements and ensure compliance with existing legislation as well as address any ethical concerns. A pre-requisite to achieve this is the recognition of clear responsibilities and accountability. The European Commission published pointers on what a contractual agreement should consider.⁵⁷ Below is a summary highlighting the most important points.

- Specify that after the purpose has been achieved or the limit of the duration has been reached, the data transmitted is to be erased.
- The usage of the same data for a different purpose should be subject to a new or amended collaboration agreement.
- The parties should define the conditions at operational level for the transfer of data: format of the data and metadata, quality, granularity and duration of access and mode of access.
- The parties should determine the compensation.
- The parties should agree on common guiding principles, agree on a code of conduct and appoint an independent auditor to oversee the data use.
- Public bodies put in place the necessary safeguards preventing the misuse of the accessed data for other purposes than the ones defined in the contract.
- The contract should include rules on liability (including its limitation or waiver) for supply of erroneous data.
- The contract should establish the applicable law and the dispute settlement mechanisms.

⁵⁷ Commission Staff Working Document Guidance on sharing private sector data in the European data economy Accompanying the document Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions "Towards a common European data space" (2018): <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1539766272141&uri=CELEX%3A52018SC0125>

6. **Use new technologies** and the novel opportunities they offer that can support data sharing, for example:

Synthetic data: Synthetic data is information that's artificially manufactured rather than generated by real-world events. It is a useful alternative to use real data – and avoid the complication of sharing it – for some applications, like testing the scalability of algorithms and the performance of new software. Synthetic data is produced by observing the properties of the original data and reproducing its characteristics.

Blockchain: Blockchain has the potential to offer solutions for control over information and access, as well as security and privacy of data, in situations where a centralised authority cannot be trusted to store or process the data for all participants. By its decentralised nature, it has the potential to make processes more democratic, transparent and efficient. Blockchain is regarded as a neutral technology with the potential to redefine the 'rules of the game' in terms of data validating, sharing and access. Given the theoretical incorruptible nature of the ledger, blockchain is considered a perfect 'storage place' to ensure the consistency of the non-sensitive elements of a data sharing initiative, such as identities, agreements, property rights or credentials.

6 Summary and the way forward

The value of data sharing between businesses and public bodies creates value for all parties involved. For the public sector, advantages occur in improved decision making and service offering, increased efficiency and the manifold benefits of a European data sharing culture. For businesses, the value of B2G data sharing realises in an improved image (including corporate social responsibility), insight and revenue.

Different aspects can make the set-up and the success of data sharing challenging. The predicted **benefits** must outweigh the **costs and risks**. Transactions and their effects have to be **ethically** unquestionable from the perspective of any shareholders and supported by leadership and the corporate culture.

When agreed upon a data sharing collaboration, the data needs to be **prepared and safely accessed**, and its processes continuously **governed**.

Data rights must be complied with, and responsibilities defined. In the absence of a clear specific legal directive, different rights and legislation have to guide a contractual agreement between all involved parties.

Based on the chosen **model of data sharing** these challenges can be more or less complex. The presented models of data sharing, however, are always a simplification of the real-life situation, that will always include additional elements of complication. In most cases, expert advice is necessary to build the actual model that works for your situation.

We have suggested a **six-step action plan** to support data sharing initiatives, that can help guide collaboration and to set up a solid data sharing framework. Involving customers and shareholders increases transparency and trust. A code of conduct and contractual agreements define ethical standards, values and rules. New technologies can offer additional value and can safeguard data sharing.

Due to its **novelty and complexity**, it can be expected that **new actors** and **new service models** will emerge, and **new regulations** will guide, regulate and support B2G data sharing. We are at the very beginning of an inevitable shift toward achieving the **fifth freedom**: the free movement of data in the European Digital Market⁵⁸.

1 ⁵⁸ European Parliament, “Free flow of non-personal data: Parliament approves EU’s fifth freedom”, 2018, <http://www.europarl.europa.eu/news/en/press-room/20180926IPR14403/free-flow-of-non-personal-data-parliament-approves-eu-s-fifth-freedom> .