Council of the European Union

Brussels, 25 November 2020
(OR. en)

13351/20
ADD 1

Interinstitutional File:
2020/0340(COD)

TELECOM 227
COMPET 592
MI 521
DATAPROTECT 135
JAI 1021
CODEC 1215
IA 95

COVER NOTE

From:	Secretary-General of the European Commission, signed by Ms Martine DEPREZ, Director
date of receipt:	25 November 2020
To:	Mr Jeppe TRANHOLM-MIKKELSEN, Secretary-General of the Council of the European Union
No. Cion doc.:	SWD(2020) 295 final
Subject:	COMMISSION STAFF WORKING DOCUMENT IMPACT ASSESSMENT REPORT Accompanying the document Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European data governance (Data Governance Act)

Delegations will find attached document SWD(2020) 295 final.

Encl.: SWD(2020) 295 final

EUROPEAN
COMMISSION

Brussels, 25.11.2020
SWD(2020) 295 final

COMMISSION STAFF WORKING DOCUMENT

IMPACT ASSESSMENT REPORT

Accompanying the document

Proposal for a
REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

on European data governance
(Data Governance Act)

{COM(2020) 767 final} - {SEC(2020) 405 final} - {SWD(2020) 296 final}

Table of contents

1. INTRODUCTION: POLITICAL AND LEGAL CONTEXT 1
2. PROBLEM DEFINITION 8
   - 2.1. What is the problem? 8
   - 2.2. What are the problem drivers? 11
   - 2.3. How will the problem evolve? 16
3. WHY SHOULD THE EU ACT? 17
   - 3.1. Legal basis 17
   - 3.2. Subsidiarity: Necessity of EU action 18
   - 3.3. Subsidiarity: Added value of EU action 18
4. OBJECTIVES: WHAT IS TO BE ACHIEVED? 19
   - 4.1. General objective 19
   - 4.2. Specific objectives 20
5. WHAT ARE THE AVAILABLE POLICY OPTIONS? 21
   - 5.1. What is the baseline from which options are assessed? 22
   - 5.2. Description of the policy options 22
   - 5.3. Options discarded at an early stage 31
6. WHAT ARE THE IMPACTS OF THE POLICY OPTIONS? 31
   - 6.1. Economic impact 31
   - 6.2. Social and environmental impact 43
   - 6.3. Impact on SMEs 45
   - 6.4. Member States' and stakeholders' views 46
7. HOW DO THE OPTIONS COMPARE? 48
8. PREFERRED OPTION 51
9. HOW WILL ACTUAL IMPACTS BE MONITORED AND EVALUATED? 54
   - 9.1. Monitoring of the specific objectives 55
   - 9.2. Monitoring of the preferred option 55

ANNEX 1: PROCEDURAL INFORMATION 58
1. LEAD DG, D*E*CIDE PLANNING/CWP REFERENCES 58
2. ORGANISATION AND TIMING 58
3. CONSULTATION OF THE RSB 58
4. EVIDENCE, SOURCES AND QUALITY 64

ANNEX 2: STAKEHOLDER CONSULTATION 66

ANNEX 3: WHO IS AFFECTED AND HOW? 74
1. PRACTICAL IMPLICATIONS OF THE INITIATIVE 74
2. SUMMARY OF COSTS AND BENEFITS 77

ANNEX 4: ANALYTICAL METHODS 81
ANNEX 5: SUBSIDIARITY GRID.....87
ANNEX 6: GOVERNANCE FRAMEWORK IN THIRD COUNTRIES 98

1. INTRODUCTION: POLITICAL AND LEGAL CONTEXT

This Impact Assessment accompanies the proposal for a Regulation of the European Parliament and of the Council[1] on data governance. It is the first of a set of measures announced in the 2020 European Strategy for Data[2]. The instrument aims to stimulate the availability of data for use and to strengthen data governance mechanisms in the EU. It would facilitate the following situations:

- the sharing of data among businesses, against remuneration or because of other benefits they derive from sharing;
- making public sector data available for reuse, in situations where such data is subject to the rights of others[3];
- allowing the reuse of personal data with the help of a 'personal data space', designed to help individuals exercise their rights under the General Data Protection Regulation (GDPR);
- making data reusable for altruistic purposes.

1.1 Technological, economic and societal context

An evolving technological landscape

In our increasingly connected world, more and more data is being generated, originating in factories or on farms, in cars or household appliances. The availability of such data is a critical enabler for data-driven innovation, including the development of more personalised and cheaper products, not least using artificial intelligence (AI) and related technologies.

Europe has missed the first wave of innovation based on data, mainly data collected from individuals over the Web 2.0. But a second wave of innovation is emerging from objects connected to the Internet-of-Things (IoT). It is expected that the volume of data produced annually in the world will grow from the 33 zettabytes in 2018 to 175 zettabytes by 2025[4]. The European Strategy for Data indicates that opportunities arise both from the increasing data volumes that are generated in fields in which the EU has a strong basis (such as manufacturing) and the changing technological landscape for data use will offer opportunities for European companies in the data economy.

The importance of data for the economy

In her 2020 State of the Union address, President von der Leyen stated that 'A real data economy […] would be a powerful engine for innovation and new jobs.' According to a study by the International Data Corporation (IDC) for the European Commission, the data economy was

[1] The final form of the legal act will be determined by the content of the instrument.
[2] COM/2020/66 final.
[3] "Data the use of which is dependent on the rights of others" or "data subject to the rights of others" covers data that might be subject to data protection legislation, intellectual property, or contain trade secrets or other commercially sensitive information.
[4] Reinsel D., Gantz J., and Rydning J., (2018). *The Digitization of the World. From Edge to Core*, International Data Corporation White Paper No. US44413318.

estimated to be worth over EUR 324.86 billion at the end of 2019[5], representing 2.6% of the Gross Domestic Product (GDP) of the EU-27. The slow-down caused by the COVID crisis in 2020 is expected to be followed by a rebound.

Data is the basis for new digital products and services. It is essential for training AI systems. An example is the self-driving car: in addition to the data generated by the car itself, additional third-party data are required for this type of system to operate securely, irrespective of weather conditions, visibility or road-surface quality[6]. Moreover, the use of data drives productivity and resource efficiency gains across all sectors of the economy. Research by the Organization for Economic Cooperation and Development (OECD) suggests that companies that invest in data-driven innovation and data analytics exhibit faster productivity growth than those that do not by approximately 5% to 10%[7].

Data is a critical resource for startups and SMEs, in particular as a business can be set up with very low initial capital. Over 99% of data supplier companies and over 98.8% of data user companies in the EU are SMEs[8]. Some 85% of new jobs created in the data economy over the last years have been created by SMEs[9].

Some 93% of the EU executives surveyed in a recent study by McKinsey believe that better access to data would be important to their organisation (with approximately 40% designating this as very important). More than 50% would be willing to share their data if they either received access to similar data from competitors in return or were paid for the data[10]. It is important to note that the term 'data sharing' does not imply that all data will be available for free for all, but may include situations of data exchanged against reward.

Societal impact of data

A better use of data can lead to improvements in health and well-being, a better environment, strengthened climate action, more efficient public services and safer societies. As demonstrated during the COVID-19 crisis, data is an essential asset for tackling emergencies such as pandemics. More generally, in the health sector, data can help develop better and more personalised treatments. McKinsey estimates that data and digital technologies could lead to savings of approximately EUR 120[11] billion a year in the EU health sector.

In the mobility sector, as well as saving more than 27 million hours of public transport users' time[12], up to EUR 20 billion a year could be saved in labour costs of car drivers thanks, amongst

[5] European Commission (2020a). *Final Study Report of the Updated European Data Market Study,* SMART 2016/0063.
[6] OECD (2019). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris.
[7] OECD (2015). *Data-driven innovation: big data for growth and well-being*, OECD Publishing, Paris.
[8] European Commission (2020a). *Final Study Report of the Updated European Data Market Study,* SMART 2016/0063.
[9] European Commission, *Entrepreneurship and Small and medium-sized enterprises (SMEs)*.
[10] European Commission (2020). *Shaping the digital transformation in Europe*, study prepared by McKinsey.
[11] McKinsey (2020). *Shaping the digital transformation in Europe.*
[12] Huyer E. (2020). *The economic impact of open data: opportunities for value creation in Europe*, European Data Portal Study.

others, to real-time navigation that reduces time stuck in traffic[13]. In turn, this has benefits in terms of tackling climate change, due to reduced CO_2 emissions and air pollution.

Data is also at the core of the open marketplaces that facilitate the collaborative or sharing economy. An example of such marketplaces is *Dawex*[14], which acts as orchestrator between data holders and data users and facilitates the exchange of data between companies and organizations. It is estimated that this saves up to 7% of household budget spending and reduces waste by 20%[15].

1.2 Political context

Already in March 2019, the European Council conclusions stated that '*the EU needs to go further in developing a competitive, secure, inclusive and ethical digital economy with world-class connectivity. Special emphasis should be placed on access to, sharing of and use of data, on data security and on AI, in an environment of trust'*[16].

The European Strategy for Data of 19 February 2020 responded to such political calls to strengthen Europe's position globally by making better use of data-driven innovation. In particular, it calls for the creation of common European data spaces.

In its conclusions of 2 October 2020[17], the European Council welcomed the data strategy. It stressed the need to make high-quality data more readily available and to promote and enable better sharing and pooling of data, as well as interoperability. It also welcomed the creation of common European data spaces in strategic sectors.

The role of common European data spaces

The European Strategy for Data proposes to establish sector- or domain-specific data spaces, as the concrete arrangements in which data sharing and/or data pooling can happen beyond one single Member State. A common European data space will be composed of a secure IT environment for processing of data by an open number of organisations, and a set of rules of legislative, administrative and contractual nature that determine the rights of access to and processing of the data. Data will be made available on a voluntary basis and can be reused against remuneration or for free, depending on the data holder's decision.

The present instrument proposes an ***overarching framework encompassing horizontal measures relevant for all common European data spaces****. The framework will leave room for sector-specific rules, governance mechanisms and standards where relevant. The objective of the initiative is not to create the common European data spaces by law, but to enhance their development by strengthening trust in data sharing and in data intermediaries.*

[13] Idem.
[14] See Dawex *website* for more info.
[15] Rademaekers K. *et al.* (2017). *Environmental potential of the collaborative economy*, European Commission.
[16] Council of the European Union *Conclusions* (22 March 2019).
[17] Council of the European Union *Conclusions* (2 October 2020).

The European Strategy for Data was welcomed by the Member States in the Council Conclusions of 9 June 2020. They called on the European Commission '*to present concrete proposals on data governance and to encourage the development of common European data spaces for strategic sectors of the industry and domains of public interest'*[18].

In his opinion on the Data Strategy, the European Data Protection Supervisor (EDPS) underlined the political relevance of working towards common European data spaces: '*one of the objectives of the Data Strategy should be to prove the viability and sustainability of an alternative data economy model - open, fair and democratic. Unlike the current predominant business model, characterised by unprecedented concentration of data in a handful of powerful players, as well as pervasive tracking, the European data space should serve as an example of transparency, effective accountability and proper balance between the interests of the individual data subjects and the shared interest of the society as a whole'*[19].

As stated by President von der Leyen in her State of the Union speech, 'Europe must now lead the way on digital – or it will have to follow the way of others, who are setting these standards for us.' The EU must seize the opportunity of this pivotal moment and ensure that it is at the forefront of the second wave of innovation based on data. This urgency is confirmed by the COVID-19 crisis, which has demonstrated the importance of data for an effective response to a global health crisis. Effective responses can only be identified if as much evidence (data) is available as possible to test out as many hypotheses as possible.

> This was the essence for example of the Exscalate4COV initiative[20]: In the initiative, an ad hoc consortium of 18 partner organisations tested available molecules with drug-like properties in order to identify new treatments against COVID-19. They have been able to identify several molecules for treatment against the virus that are now being tested in clinical trials. This was only made possible because pharmaceutical companies 'donated' information on these molecules to European research centres. In the absence of established processes for the sharing of such data, it took 3 months to obtain it.

It is also essential that the EU acts quickly because data will play a key role in the economic recovery, not least because of its importance for small and medium-sized enterprises (SMEs) and startups. The Communication 'Europe's moment: Repair and Prepare for the Next Generation'[21] gives a prominent place to measures that accelerate the development of the data economy, including legislative action on data sharing and governance, which is the subject of this impact assessment.

[18] Council of the European Union *Conclusions* (9 June 2020).
[19] European Data Protection Supervisor (2020). *Opinion 03/2020 on the European strategy for data*.
[20] Exscalate4COV *webpage*.
[21] COM(2020) 456 final.

1.3 Legal context

1.3.1 Horizontal legislation

The current initiative covers different types of data intermediaries, handling both personal and non-personal data. Therefore, the interplay with the legislation on personal data is particularly important. With the General Data Protection Regulation[22] and ePrivacy Directive[23], the EU has put in place a solid and trusted legal framework for the protection of personal data and a standard for the world. The legislative framework for the common data spaces would work within the rules of the existing legislation on the protection of personal data. In particular, it would remain the responsibility of each party to identify the suitable legal basis for the processing of personal data within a common European data space.

The proposal will build on the mechanisms present in the existing legislation, in particular the portability right under Article 20 GDPR, that give individuals more control over how their data is used. Article 20 of the GDPR gives data subjects the right to move their data (e.g. their social media data, mobility or health data) to another service, or to allow a third party to access that data. This right has a strong potential for reuse of personal data, as identified, amongst others, in the report on competition policy and the digital era prepared for Commissioner Vestager in 2019[24]. Additionally, this right would give individuals the possibility to make some of their data, such as their mobility or health data, available for the common good, if they wish to do so.

Similarly, the initiative would not amend existing competition law provisions, and would be designed in full compliance with Articles 101 and 102 of the Treaty on the Functioning of the European Union (TFEU), which prohibit anti-competitive agreements and the abuse of dominant market power, respectively.

The initiative would also be in full compliance with the EU's international obligations, notably in the multilateral agreements in the World Trade Organisation and in its regional trade agreements.

The current proposal would complement the Directive on open data and the reuse of public sector information (Open Data Directive)[25]. The Open Data Directive deals with data for which public sector bodies have all the relevant rights. It does not, however, cover public sector data subject to the rights of others (e.g. personal data, data protected by intellectual property rights or trade secrets). Due to these third party rights, such data cannot be made available as open data, i.e. with as little usage restrictions as possible. By facilitating the secure access to such datasets, this proposal encourages the exploitation of data whose reuse is not regulated by the existing Directive. As a consequence, the Implementing Act on High-Value Datasets under the Open Data Directive, which is expected to be adopted in 2021[26], will also be fully complementary with this initiative.

---

[22] OJ L 119, 4.5.2016, p. 1-88.
[23] OJ L 201, 31.7.2002, p. 37-47.
[24] Crémer J., de Montjoye Y.-A. and Schweitzer H. (2019). *Competition policy for the digital era,* Report prepared for Commissioner Vestager.
[25] OJ L 172, 26.6.2019, p. 56–83.
[26] See COM/2020/66 final.

1.3.2 Sectoral legislation

Sector-specific legislation on data access is in place to address identified market failures in fields such as automotive[27], payment service providers[28], smart metering information[29], electricity network data[30], intelligent transport systems[31] and electronic freight transport information[32]. The legal instrument for the common European data spaces would support the use of data made available under such rules without altering them or creating new sectoral obligations.

1.3.3 Relationship with other planned initiatives

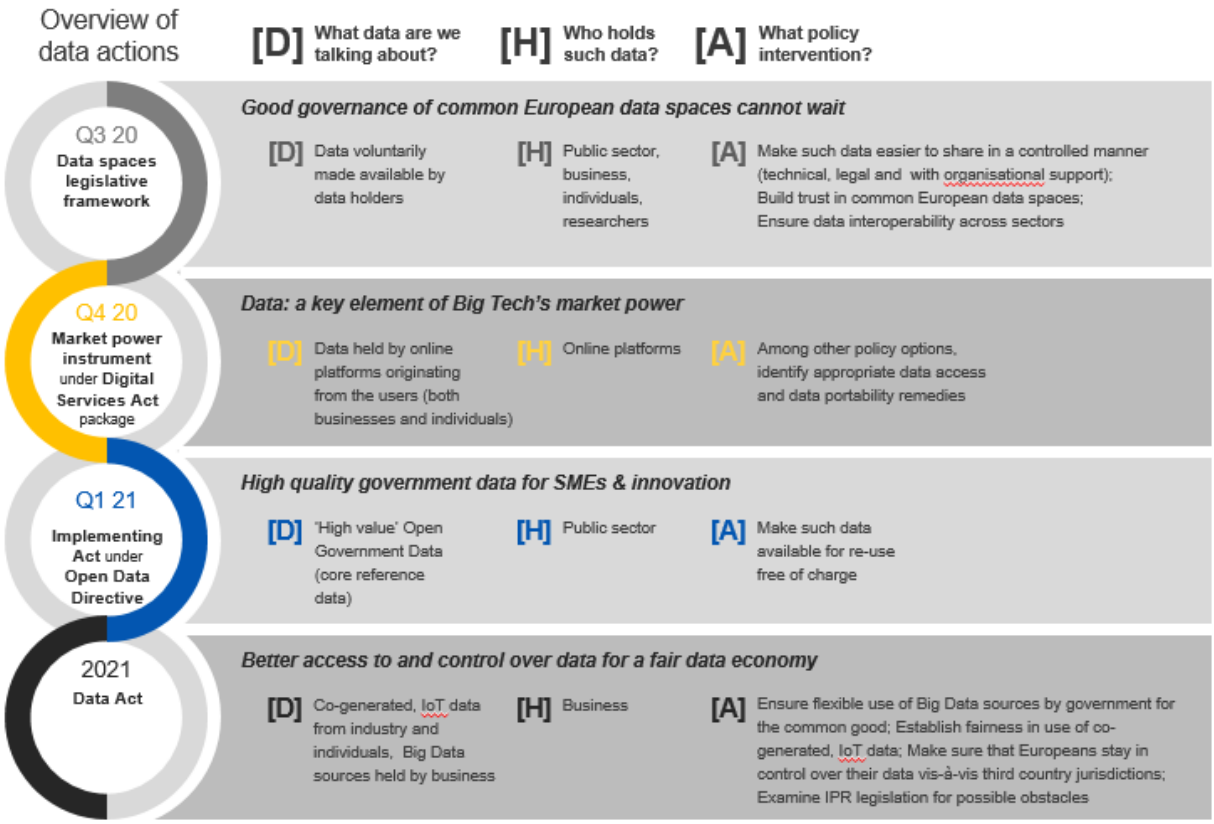
The current legislative initiative has logical and coherent links with the other initiatives announced in the European Strategy for Data. It aims at improving voluntary data sharing within and across common European data spaces. This would be achieved by supporting the emergence of data intermediaries that could organise data spaces as trusted third parties and provide relevant technologies. In addition, it would support the development of technical and legal standards relating to the means of the data exchange which, in turn, will enhance trust in data sharing.

The current initiative is a first step in the two-step approach announced in the European Strategy for Data. The initiative will address the urgent need to facilitate data sharing through an enabling governance framework. In a second step, the Commission will address issues about who controls or 'owns' the data, i.e. the material rights on who can access and use what data under which circumstances. The introduction of such rights will be examined in the context of the **Data Act** (2021)[33]. Diverging interests of the stakeholders and different views on what is fair in this respect make these issues subject to intense debate, which warrants taking more time.

While offering an alternative model to the data handling practices of the Big Tech platforms, the current legislative initiative is also clearly distinct from the **Digital Market Act** (DMA) and the **Digital Services Act** (DSA). The DMA, foreseen for Q4 2020, will combine two elements to ensure the proper functioning of the internal market by promoting effective competition in digital markets: (i) a set of clear-cut prohibitions and obligations to address known unfair practices of online platforms with a gatekeeping role, resulting from, among other factors, their control of large amounts of data; and (ii) a market investigation regime which would allow tackling existing and emerging market failures in digital markets, including in relation to data access and use. The DSA, also foreseen for Q4 2020, intends to clarify the responsibilities and obligations of digital services, and in particular online platforms, based on, amongst other elements, an evaluation of the e-Commerce Directive.

The interplay with the other initiatives announced in the Data Strategy is illustrated in the image below:

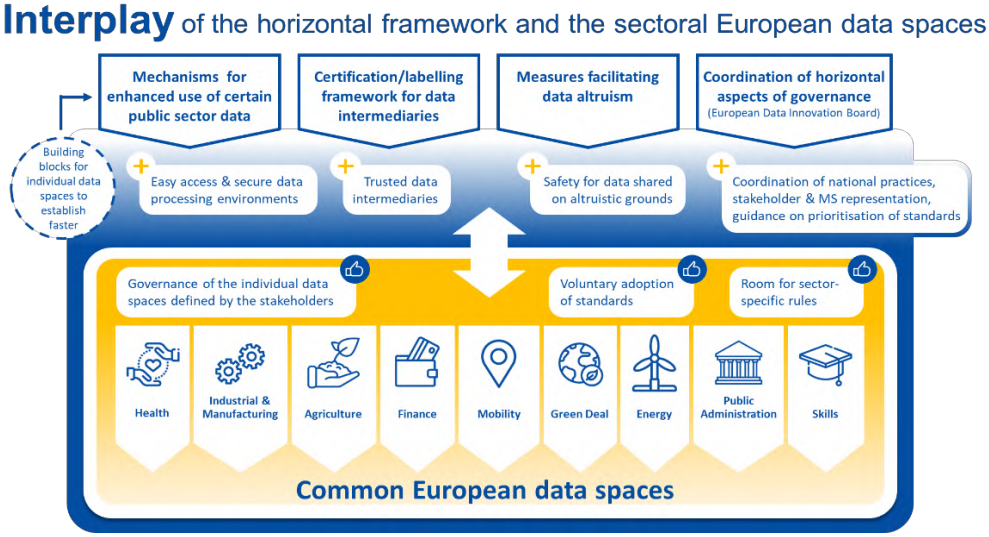
[27] OJ L 188 18.7.2009, p. 1 as amended by OJ L 151, 14.6.2018, p. 1.
[28] OJ L 337, 23.12.2015, p. 35-127.
[29] OJ L 158, 14.6.2019, p. 125-199; OJ L 211, 14.8.2009, p. 94-136.
[30] OJ L 220, 25.8.2017, p. 1-120; OJ L 113, 1.5.2015, p. 13-26.
[31] OJ L 207, 6.8.2010, p. 1-13.
[32] OJ L 249, 31.7.2020, p. 33–48
[33] See COM/2020/66 final.



Overview of envisaged data actions. Source: European Commission

The European Strategy for Data also proposes the creation of sectoral data spaces in areas such as mobility and health, and announces sector-specific initiatives, including legislative action for the specific sectors. The Commission is, for example, working on a review of the current EU type approval legislation for motor vehicles. The initiative aims at ensuring fair and safe access to vehicle data and ultimately to offer better access to more services based on car data. This initiative is envisaged for 2021.

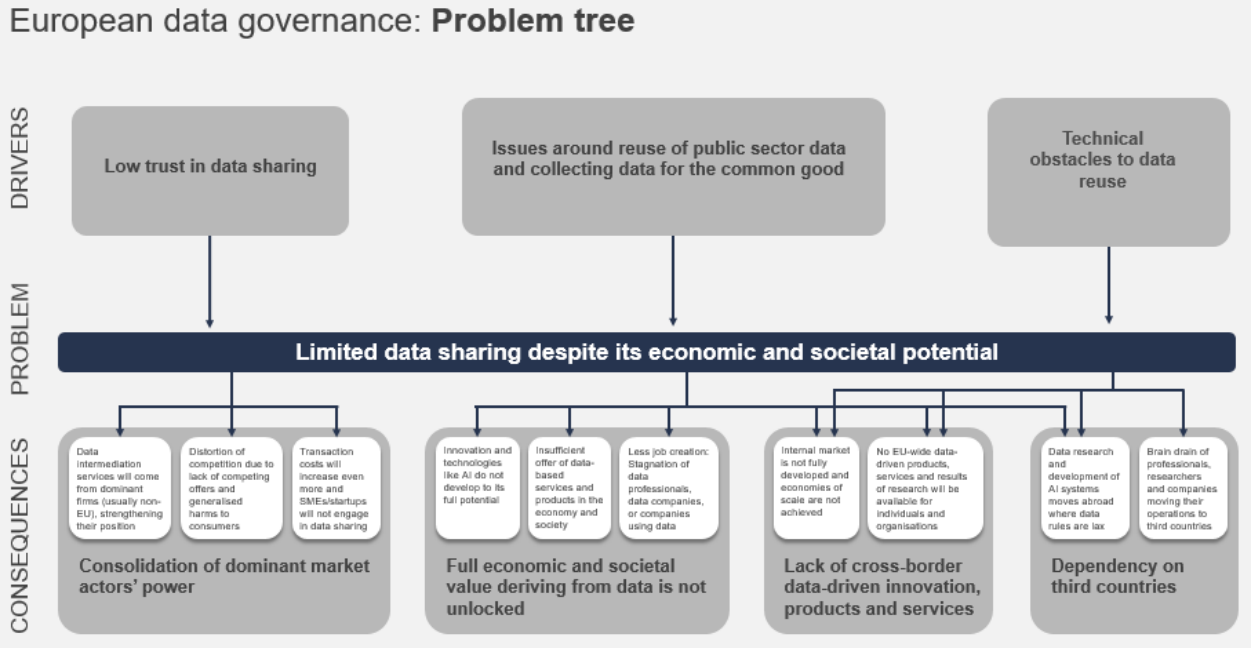
The image below shows the interplay of the horizontal framework with the sectoral initiatives.



Source: European Commission

The development of common European data spaces will be supported financially under the Digital Europe programme and the Connecting Europe Facility[2]. The current legislative initiative and the financing from these programmes will mutually reinforce each other.

2. PROBLEM DEFINITION



Source: European Commission

2.1. What is the problem?

As described in Chapter 1, the economic and societal potential of data use is enormous, in terms of new products and services based on novel technologies, more efficient production, and tools for combatting societal challenges. The problem that this initiative addresses is that this potential is not realised due to limited data sharing in the EU. A number of obstacles (low trust in data sharing, issues related to the reuse of public sector data and data collection for the common good, technical obstacles) stand in the way of data sharing becoming more prevalent. These problem drivers are described in section 2.2.

The importance of data sharing

In order to leverage the value of data in the economy and society, more economic operators and organisations promoting societal interests need to be able to use data. This will include data held by others, as it is not cost-efficient if every company or organisation collects similar or even identical data in parallel to others. Digital data can be copied at virtually no cost, and can be used simultaneously by different actors for an unlimited numbers of times. These characteristics distinguish data from traditional economic resources. In order to harness such potential, more data needs to be shared among operators and organisations, including against monetary and other rewards, to have a sufficient amount of data available for innovation in the market. Given the

fact that the availability of resources feeding data-driven innovation benefits the entirety of the data-driven economy in the EU, the socioeconomic benefits will positively impact all the vertical sectors directly or indirectly linked to the ever-growing EU data economy.

According to the OECD, data access and reuse could generate social and economic benefits worth up to 1.5% of GDP in the case of publicly held data, and between 1% and 2.5% of GDP when also including privately held data[34]. Data access and sharing can increase the value of data to holders (direct impact), but it can also help create 10 to 20 times more value for data users (indirect impact), and 20 to 50 times more value for the wider economy and society (induced impact)[35].

An unfulfilled potential

Difficulties in accessing and using data held by others have been reported repeatedly. According to the OECD, 'individuals, businesses, and governments often face barriers to data access, which may be compounded by reluctance to share'[36].

In the recent public online consultation on the Data Strategy, almost 80% of companies reported problems in data access. When asked about the nature of such difficulties, 72.1% of these companies reported 'technical aspects relating to both data interoperability and transfer mechanisms' and 43.5% the 'impossibility to find data of the relevant quality' (multiple choices possible). Other issues relate to outright denial of data access (65%) or prohibitive prices or other conditions (41.7%). This suggests that technical difficulties are an important barrier to data sharing.

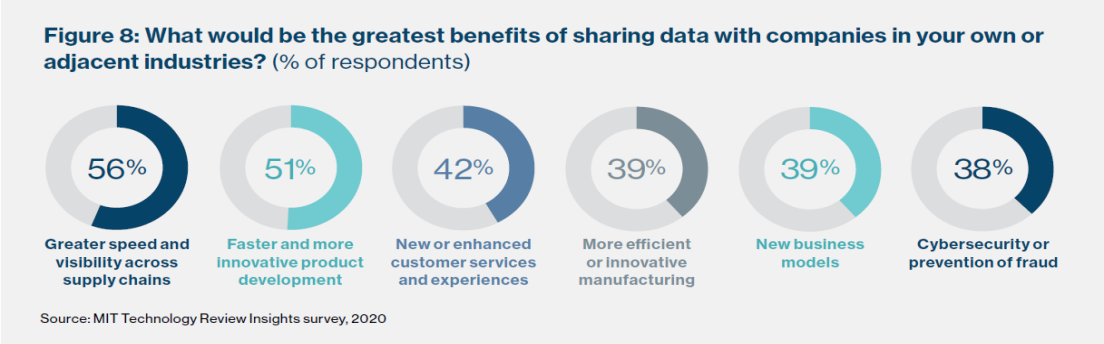
A 2018 report by Deloitte[37] highlights the considerable potential for increasing the level of data sharing, in particular of machine-generated non-personal data, in Europe over the next decade. The report suggests that only between 43% and 58% of the potential of data sharing along a value chain is realised and only 20% and 40% of the potential of sharing between sectors[38]. The report estimates that leveraging this potential would create, in monetary terms, EUR 35 billion of value in agriculture by raising yields; reduce costs from road vehicle damage, maintenance and repairs by EUR 40 billion; and generate efficiencies in resource management and prevent drug counterfeiting in the healthcare sector, saving EUR 14 billion. It could create as much as EUR 1.3 trillion of value in manufacturing by improving productivity by 2027. The untapped potential of data sharing is confirmed by a recent study on ecosystems (focusing on health, construction and automotive and mobility) carried out by McKinsey[39].

---

[34] OECD (2019). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris.
[35] Idem.
[36] Idem.
[37] Deloitte (2018). *Realising the economic potential of machine-generated, non-personal data in the EU*, Report for Vodafone Group.
[38] Horizontal data sharing is defined as sharing between organisations involved in the same commercial or non-commercial point of the value chain, e.g. businesses selling the same product in the same market place. Vertical data sharing is defined as sharing between organisations who have a customer or supplier relationship, directly or indirectly.
[39] McKinsey (2020). *Shaping the digital transformation in Europe.*

Why would companies and individuals share their data?

What are the incentives for companies to share their data? Direct monetisation is currently not the main reason for data sharing. Respondents to a survey conducted by the *MIT Technology Review* indicate that data sharing helps to obtain greater speed and visibility across supply chains, and to support faster and more innovative product development[40]. Incentives for companies to share data include increased access to data of other contributors in exchange for giving access to their own data, analytical results derived from the shared data, the availability of services such as predictive maintenance services or licence fees, as well as reduced time and costs of product marketing.



The incentive for individuals to share data can vary. It can come from the wish to contribute to research on rare diseases or to make local transport more efficient (in the case of data altruism). It can also be driven by possibilities to obtain better advice on their personal situation or more personalised or cheaper services in exchange for the use of the data.

The role of platforms in the data economy

The consumer-oriented data economy has given rise to the development of intermediaries that cover the entire value chain, from data collection (collection through websites, smartphones or connected objects such as thermostats) to storage and processing (cloud infrastructures) and services. This has led to economies of scope and scale in terms of data (i.e. the capacity to not only have large volumes of data at their disposal but also data on a variety of human activities), resulting in huge advantages in rolling out additional data-driven services, including in the field of AI.

As European industry begins to interconnect factories, suppliers and other business partners and clients, to deliver better, more personalised products in a more efficient (and thus cheaper) manner, questions related to the organisation of such data flows arise. According to many bilateral interactions with stakeholders, there is a high level of distrust in integrated tech service providers as platforms for industrial data exchange. Large players like Airbus, Siemens, GE or MAN therefore sometimes opt for creating their own platforms. However, these can be exposed to similar criticism from their business partners (notably SMEs, but also suppliers), who may be in a weaker bargaining position to determine data use by such platforms.

[40] McCauley D. (2020). *The global AI agenda*, MIT Technology Review Insights.

The current initiative thus represents an important first step in creating a new model for the data economy. This has the potential to meet new market demands and allow the EU to become more competitive in the data-driven world economy, while maintaining its data sovereignty and its full compliance with its international obligations in trade agreements. Such a model is necessary as an alternative to the current business model dominated by Big Tech platforms. It would be built on a division of functions and the development of common European data spaces as collaborative ecosystems in which data would be usable by a broader range of organisations (public and private) based on a collective governance of data sharing. These data spaces will constitute the core tissue of an interconnected and competitive data economy in the EU.

2.2. What are the problem drivers?

2.2.1. Low trust in data sharing

Based on the views expressed in the series of stakeholder consultations organised by the European Commission on data sharing, companies do not necessarily trust that, if they share data, the reuser will use it in line with the contractual agreement. For example they fear that their data could be made available to third parties.

As shown in a 2017 Commission consultation, 20% of companies do not engage in data sharing because of this fear[41]. Some companies fear that they might lose their competitive advantage within their market or in prospective markets if they engage in data sharing. The OECD also reports that this is one of the major concerns for both organisations and individuals with regard to data-sharing constellations[42]. Furthermore, in the 2020 consultation, several companies highlighted the difficulty they face when trying to access datasets of other companies, which may be reluctant to share data due to this fear[43]. For example, one insurance company explained that: 'Companies are reluctant to share data since it is the fundamental basis of their competitive advantage. Therefore, it is critical to introduce appropriate safeguards to develop a trusted environment.' Some emerging technologies can track and trace data use within a data ecosystem. This can improve trust, but the use of these technologies in more open ecosystems is not yet widespread[44].

The lack of trust leads to high transaction costs, related to finding a suitable data-sharing partner; negotiating, drafting and monitoring the contract, and; developing interoperability solutions for transferring, transforming and cleaning the data[45]. This has been highlighted by stakeholders (especially SMEs) since 2017[46]. The OECD confirms that these high transaction costs might heavily affect those in a weaker position, notably individuals (consumers) and SMEs[47].

[41] European Commission (2017). *Synopsis report consultation on the 'building a European data economy' initiative*.
[42] Idem.
[43] European Commission (2020b). *Outcome of the online consultation on the European strategy for data*.
[44] E.g. the Connector Architecture of the International Data Spaces Association, part of the Gaia-X initiative.
[45] Deloitte (2018). *Realising the economic potential of machine-generated, non-personal data in the EU*, Report for Vodafone Group.
[46] European Commission (2017). *Synopsis report consultation on the 'building a European data economy' initiative*.
[47] OECD (2019). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris.

Bringing the offer and demand for data together in new market places is a pre-requisite to solving the problem. These new market places are, however, at risk of not being able to scale up sufficiently due to the lack of sufficient trust in them. It is clear that this can happen only upon the condition of a sufficient level of trust in intermediaries for market actors to buy-in[48]. Without this, it is unlikely that they will be able to scale up.

Both existing companies and start-ups propose data marketplaces, platforms or trusts and personal data intermediaries as a means to improve findability of relevant data, lower the costs of transacting in data and propose exploitation of shared resources[49]. These data intermediaries can reduce transaction costs, for example by proposing standardised clauses in data-sharing contracts, providing a platform for data sharing, offering solutions for data interoperability, and helping data holders who may not have the necessary skills to ensure compliance with data protection law[50]. These facilitators generally aim to remain neutral in the data exchange that they accommodate, meaning that they do not accumulate data or monetise on the data exchanged[51].

Personal data intermediaries are a specific category of data intermediaries. They seek to empower individuals to exercise their rights under data protection law and manage their own personal data[52]. Already in 2016, the EDPS highlighted the potential of these solutions[53]. The 2020 online consultation showed that close to 80% of the 201 citizens responding consider that 'it should be made easier for individuals to give access to existing data held about them, in line with the GDPR'. In the same group of citizens, 43% considered that this could be achieved through practical solutions that allow individuals to exercise control, such as mobile and online dashboards or apps[54].

2.2.2. Issues related to the reuse of public sector data and collecting data for the common good

Data sharing is hampered by an absence of appropriate structures and processes, notably to facilitate data altruism and the reuse of publicly held data that is subject to the rights of others.

a) Limited data-handling capacity and reuse culture in the public sector

[48] Idem; Joint Research Centre, *Business-to-business data sharing: An economic and legal analysis,* 2020.
[49] COM/2017/09 final; SWD/2017/02 final; the importance of data sharing platforms or institutions is discussed in the French, UK and German data or AI strategies: Villani, Donner un sens à l'intelligence artificielle, 2018, Hall/Pensenti, Growing the artificial intelligence strategy in the UK, 2017, Report German Datenethikkommission, 2019; and in the following reports: OECD (2019). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*; Open Data Institute, Designing trustworthy data institutions, 2020.
[50] Joint Research Centre, *Business-to-business data sharing: An economic and legal analysis,* 2020.
[51] This will avoid the problem of platformisation that profits from network effects and economies of scope. There is a risk that these neutral entrants may realise that it could be necessary to leave their neutral position and monetise on the data exchanged through their service by offering added value services in other markets. This would lead to problems of data aggregation analysed in consumer platforms discussions. See also JRC (2020); COM/2020/66 final.
[52] European Commission (2016). *An emerging offer of "personal information management services" - Current state of service offers and challenges*; Ctrl+Shift (2014). *Personal Information Management Services: An analysis of an emerging market*;
[53] European Data Protection Supervisor (2016). *EDPS Opinion on Personal Information Management Systems*, Opinion 9/2016.
[54] European Commission (2020b). *Outcome of the online consultation on the European strategy for data*.

The GDPR has increased awareness on personal data protection in companies, data subjects, the public sector and academia alike[55]. While this is a very positive development, the increased awareness is not always matched by a high level of expertise in the public sector on the rules and exceptions. The consultation supporting the review of the Public Sector Information Directive showed that public sector bodies had difficulties in managing risks related to the reuse of data subject to the rights of others, especially personal data[56]. In addition, public sector bodies have signalled[57] that dealing with requests to reuse this specific category of data represents a major issue for them, as they lack the technical and legal capacity to process these requests.

The potential value of this type of data held by the public sector (such as health data or micro-statistics) is often high for machine learning and research. The challenge is to find ways to make it possible to extract knowledge from the data, while fully preserving privacy or other rights that may be attached to the data. Technical mechanisms exist that allow controlled processing of data that is subject to the rights of others ('safe reading rooms'). Some Member States (notably France, Finland and Germany) have established specific bodies underpinned by legislation that offer such technical mechanisms, creating secure and privacy-enhancing conditions for the reuse of such data.

> In 2018, the **Centre d'accès sécurisé aux données**[58] (Centre for secure access to data) was established by the French government and the National School for Statistics and other partners to allow the secure processing of statistical micro-data. Finnish legislation recently established the **data permit authority Findata**[59] with the aim of providing researchers with a one-stop-shop service for receiving a permit to process data from a range of public registers for health and social protection. Similarly, the German government has adopted legislation that will enable research on the basis of medical reimbursement data. In Germany 38 **Forschungsdatenzentren** (secure data research infrastructures) have been set up in order to facilitate access to sensitive data for researchers and more are being established.

In spite of these initiatives, the overall capacity of the public sector across the EU to handle these types of requests remains low, since public sector bodies are often not equipped to make the data available for use in a way that is compliant with data protection rules. At the same time, offers to public sector bodies (cities, hospitals) from large companies to collaborate on projects involving data can lead to situations in which the company gets de facto an exclusive access to the data[60].

[55] COM/2020/264 final.
[56] European Commission (2018a). *Synopsis report of the public consultation on the revision of the Directive on the reuse of public sector information*.
[57] Idem.
[58] Centre d'accès sécurisé aux données *website*.
[59] FINDATA *website*.
[60] European Commission (2018d). *Study to support the review of Directive 2003/98/EC on the re-use of public sector information*, study prepared by Deloitte.

b) Lack of means to manage consent-based sharing of personal data at scale

Individuals are increasingly willing to share their personal data for the common good and research[61]. This is confirmed by a 2017 public consultation that gathered more than 1 400 replies, in which 81% of respondents believed that 'sharing of health data could be beneficial to improve treatment, diagnosis and prevention of diseases across the EU'[62]. In addition, in the 2019 Eurobarometer[63], six in ten respondents indicated that they would be willing to securely share some of their personal information to improve public services.

Pilot initiatives for individuals to give access to [64] their data for altruistic reasons do exist. They remain, however, limited in scale. One example is a citizen-driven model of collaborative governance and management of health data called *Salus Coop*[65]. Other examples are the pilot projects in La Rochelle, Nantes and Lyon. Nantes has used data made available by citizens to develop an energy transition scheme for the city. La Rochelle intends to improve mobility services and public transport through insights gained from such data. Lyon aims to help socially excluded families and to simplify the life of citizens who do not speak French[66]. During the COVID-19 crisis, the German *Robert Koch Institut* developed the *Corona Datenspende-App*, allowing individuals to provide their fitness tracker and smart watch data to help determine patterns of the spread of the virus[67]. Such opportunities were already identified in the Villani Report[68], which recognised the potential of 'civic data sharing', i.e. data contributed by individuals for the benefit of public services or research.

Despite these efforts, researchers, innovators and public sector organisations lack the means to collect personal data at scale, based on the consent of the data subject or following the exercise of their right to data portability provided by the GDPR. This is confirmed by the 2020 public consultation, which showed that almost 70% of participating citizens considered there are not enough mechanisms to give their consent to the processing of their data or they simply do not know about them (18.4%). Therefore, personal data sharing for the common good remains underdeveloped and it is difficult to establish sufficiently large data pools[69].

There are currently no clear rules and processes in place in a large majority of the Member States that address the issue of data altruism. For health data, only Denmark has already put in place a

[61] Recital 33 of the GDPR; Halvorson G. C., Permanente K. and Novelli W. D. (2014). *Data altruism: honouring patients' expectations for continuous learning*, Institute of Medicine Commentary, Washington, DC.
[62] European Commission (2018b). *Synopsis report of the public consultation on Digital transformation of health and care in the context of the Digital Single Market*.
[63] European Commission (2019c). *Special Eurobarometer 503: Attitudes towards the impact of digitalisation on daily lives*.
[64] The notion of 'donation' should be used with care as it may suggest that consent for the processing of data can no longer be withdrawn, which would be counter to Article 7(3) GDPR.
[65] SalusCoop *website*.
[66] CNIL (2017) *La plateforme d'une ville*, CAHIERS IP N.5.
[67] Corona-Datenspende *website*.
[68] Villani, C., (2018). *For a meaningful artificial intelligence: towards a French and European strategy*, Report from a parliamentary mission from 8 September 2017 to 8 March 2018.
[69] BBVA (2019).*The case for a regulation on data sharing by users to increase European competitiveness*, Policy Paper sent to DG CNECT.

data altruism mechanism and Germany plans to roll it out in 2023, while 15 EU Member States are viewing the idea favourably[70].

A key barrier is that there are currently no mechanisms to examine and attest whether the organisations behind such schemes are trustworthy and actually use the data for the proclaimed altruistic purposes[71].

2.2.3 Technical obstacles to data reuse

a) Interoperability problems for data use across sectors

The 2019 workshops on common European data spaces[72] highlighted a series of issues regarding standardisation within the different sectors. For instance, in industrial and agricultural settings, data and service providers have selected architectures, ways to describe the data and data formats for their platforms, which make it difficult to exchange data.

This problem is even stronger at the cross-sectoral level. The value of data is often derived from combining datasets from diverse sources, possibly coming from different sectors. A 2018 study by Deloitte estimated that, depending on the sector, between 24% and 36% of the benefits of data sharing will come from sharing **between** the sectors[73]. Standards are an important tool for this to happen, both from a technical and legal point of view. However, commonly accepted standards are failing to emerge in domains where stakeholders have conflicting interests.

The OECD states that 'one of the most frequently cited barriers to data sharing and reuse is the lack of common standards, or the proliferation of incompatible standards'[74]. This is confirmed by the 2020 public consultation, where 91.5% of the respondents agreed that standardisation is necessary to improve interoperability and ultimately data reuse across sectors. Some 91.1% of respondents agreed that future standardisation activities need to better address the use of data across sectors of the economy or domains of society[75].

b) Limited findability of data that is fit for a given purpose and the related uncertainty about data quality

Companies often struggle to find or obtain the data that they need. In the public online consultation on the Data Strategy, almost 80% of companies reported problems in data access. When asked about the nature of such difficulties, 43.5% of those companies signalled the 'impossibility to find data of the relevant quality' (multiple choices possible).[76] In the course of

[70] EU Health Consortium (2020). *Assessment of the EU Member States' rules on health data in the light of the GDPR*.
[71] European Commission (2020). *Support Study to this Impact Assessment*, SMART 2019/0024, prepared by Deloitte.
[72] European Commission (2019b). *Reports of the workshops on common European data spaces*.
[73] Deloitte (2018). *Realising the economic potential of machine-generated, non-personal data in the EU*, Report for Vodafone Group.
[74] OECD (2019). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*.
[75] European Commission (2020b). *Outcome of the online consultation on the European strategy for data*.
[76] European Commission (2020b). *Outcome of the online consultation on the European strategy for data*.

targeted consultation activities undertaken in 2019[77], stakeholders confirmed that findability of data is one of the main barriers to trading data. For example, stakeholders in the environmental field stated that, while there is no shortage of environmental data, these data are not easily findable, comparable and accessible, that there is a need for harmonised standards and work on data quality, and that these issues could be addressed through the common data spaces[78]. The measurable cost of not having research data compliant with standards developed by the FAIR initiative (aiming to make research data findable, accessible, interoperable and reusable)[79] is estimated at EUR 10.2 billion per year in Europe[80].

Next to findability or discovery of existing data, information about the quality of data is key. It allows the reuser to assess whether certain data is fit for the given purpose. This is especially important for big data analytics and AI, including machine learning, to avoid bias in the results. A low level of data quality can have particularly severe consequences in certain sensitive domains, such as health or critical infrastructures[81]. A 2018 study on data sharing between companies in Europe on behalf of the Commission confirmed the importance of data quality: 73% of the 129 companies surveyed indicated that poor or insufficient data quality hampers data sharing[82]. In the 2020 online consultation, stakeholders also signalled this problem[83]. As an example, the Netherlands Vehicle Authority (RDW) commented that: 'The main condition to ensure the reuse of a dataset is availability in general, quality of data, quality of meta-data, findability, actuality and accuracy.'

2.3. How will the problem evolve?

According to the OECD, with the increasing use of AI and the Internet of Things (IoT) the supply of, and demand for, data will increase even in traditionally less data-intensive fields, and this to a level that very few organisations will be able to meet alone[84]. Therefore, even in the absence of EU action, the use of data and data sharing are expected to grow, but would encounter the following limitations:

1. Consolidation of market actors' power: without measures to overcome the generalised low trust and uncertainty related to data sharing, the high transaction costs (see 2.2.1) are unlikely to change. Big Tech platforms already enjoy a high degree of market power in several digital markets. In the absence of measures, including this initiative, they could enter the data-sharing market and offer services as data intermediaries[85] without substantial competition - an evolution

[77] European Commission (2019b). *Reports of the workshops on common European data spaces*.
[78] Idem.
[79] FORCE11 (2020). *The FAIR data principles*.
[80] PwC (2018), *Cost of not having FAIR research data*, Study prepared for DG RTD.
[81] That applies also for public procurement, which ensures the functioning of many critical public services such as health, education, construction, mobility, security, defence, emergency response, etc.
[82] Everis (2018). *Study on data sharing between companies in Europe*, Study prepared for DG CNECT.
[83] European Commission (2020b). *Outcome of the online consultation on the European strategy for data*.
[84] OECD (2019). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*.
[85] Tombal T. (2020). *Economic Dependence and Data Access*, International Review of Intellectual Property and Competition Law, Vol. 51.

that some industrial players, for example in the banking sector[86], fear. Smaller companies would then be confronted with a business dilemma of not being able to offer new data-based services or products, or of sharing data through such Big Tech platforms even if this means losing some of the value of the data to the platforms. Some stakeholders consider the significant amounts of data held by Big Tech platforms to be the single biggest barrier to entry in the digital economy[87].

2. Full economic and societal value derived from data remains untapped: limited data availability would lead to less innovation in the EU and a slower development of AI, as underlined in the Villani Report[88]. It is likely that the public sector in some Member States would facilitate the use of data that cannot be made available as open data, while others would not, thus creating a growing gap between the Member States. Data altruism would take off through more standalone initiatives, but this would not lead to data pools of the necessary scale and cross-border dimension.

3. Lack of cross-border data-driven innovation, products and services: cross-sector standardisation efforts would be slow or could be dominated by large players who are already working across the different sectors, as indicated in the support study for this Impact Assessment[89]. Without a harmonised set of rules, Member States would continue to legislate in highly diverging ways, which would lead to an even more fragmented landscape. This would make it difficult for companies to develop pan-European products and services, and research results would not be representative for the whole of the EU.

4. Dependency on third countries: data research and the development of AI systems would move abroad[90], where rules are less stringent. The EU could experience a brain drain of professionals, researchers and companies[91] moving their operations to third countries. The ambition to foster data infrastructures that would make Europe more autonomous, as announced in the Data Strategy, is also relevant for limiting the dependency on third countries.

3. WHY SHOULD THE EU ACT?

3.1. Legal basis

This initiative is part of the 2020 European Strategy for Data that aims to reinforce the Single Market for Data. With a growing digitalisation of the economy and society, there is a risk of Member States increasingly legislating data-related issues in an uncoordinated way, which would intensify fragmentation in the internal market. Setting up the governance structures and mechanisms that will create a coordinated approach to using data across sectors and Member States would help stakeholders in the data economy to capitalise on the scale of the internal

[86] Padilla J. (2020). *Big Tech "banks", financial stability and regulation*, Financial Stability Review, Issue 38.
[87] Furman (2019). *Unlocking digital competition - Report of the Digital Competition Expert Panel*.
[88] Villani, C., (2018). *For a meaningful artificial intelligence: towards a French and European strategy*, Report from a parliamentary mission from 8 September 2017 to 8 March 2018.
[89] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.
[90] Bughin J., Seong J. *et al.* (2019). *Tackling Europe's gap in digital and AI*, McKinsey Global Institute, Discussion Paper.
[91] Delcker J. (2018). *Merkel warns of AI brain drain to foreign tech companies*, Politico.

market. This would be in full respect of the provisions on anti-competitive practices and the ban on the abuse of dominant market power, as laid out in Articles 101 and 102 of the Treaty on the Functioning of the European Union (TFEU). Thus, Article 114 TFEU is identified as the relevant legal basis for this initiative.

3.2. Subsidiarity: Necessity of EU action

In the EU, the key sectors of the economy span across borders, with the suppliers, producers and clients established in different Member States. Data flows form an intrinsic part of digital activity and mirror these EU-wide supply chains and collaborations. Any initiative aiming to organize such data flows must address the EU single market in its entirety.

Companies active in the data economy should be able to benefit from the size of the internal market by rolling out EU-wide products and services. Datasets available in individual Member States often do not have the richness and diversity to allow big data pattern detection or machine learning. In addition, data-based products and services developed in one Member State may need to be customised to the preferences of customers in another, and this requires local data. Data needs to be able to flow easily through EU-wide and cross-sector value chains, making it easier to launch a cross-border service or to replicate an existing data-based service from one Member State to another.

As they become increasingly aware of the importance of data sharing, including the reuse of data held by public sector bodies, Member States have started to legislate on different aspects of the data economy[92]. This creates a risk of legislative and administrative fragmentation. France, Germany and Finland, for example, are setting up administrative structures and processes to allow the reuse of publicly held data, the use of which is subject to the respect of rights of others. In Denmark, the Statistical Office has the role of granting permits for research to be carried out using several data sources, including smart metering information[93]. Other Member States have not taken any legislative action in this field. EU action would offer a common vision to these national endeavours and ensure that the barriers and bottlenecks which are common across the entire EU economy can be tackled in a coherent manner across the internal market.

EU-level intervention is ultimately best suited to increase the levels of data reuse across the economy, as it can lay down the elements that ensure comparable access and use conditions in all data spaces. It is unlikely that national intervention would be equally efficient.

3.3. Subsidiarity: Added value of EU action

Considering the importance of economies of scale for the development of data technologies and services, coordinated action at EU level can bring greater value to the European economy and society than action by individual Member States. A Single Market for Data would ensure that data from the public sector, businesses and citizens can be accessed and used in the most

[92] See on this and the European Commission (2020). *Support Study to this Impact Assessment*, SMART 2019/0024, prepared by Deloitte.
[93] Statistics Denmark, *Data for research*.

effective and responsible manner possible, while businesses and citizens keep control of the data they generate and investments made into their collection are respected. Companies would be able to market their products and services in all Member States. Companies and research organisations would advance representative scientific developments and market innovation in the EU as a whole, which is particularly important in situations where EU coordinated action is necessary, like the COVID-19 crisis.

Furthermore, only concerted action by the Member States can ensure that a European model of data sharing, with trusted data intermediaries for B2B data sharing and for personal data spaces, can take off. Mutual recognition of certification/labelling mechanisms and of a trust scheme for data altruism will make it possible to collect and use the data at the necessary scale. A 'light touch' enabling legislation, as proposed in this initiative, will ensure that the Member States move in the same direction and at the same speed.

In the 2020 consultation, 86.4% of the respondents agreed that data governance mechanisms are needed to capture the potential of data, in particular for cross-sector data use[94]. This shows a clear added value and relevance of EU intervention to establish a coordinated approach to data sharing.

4. OBJECTIVES: WHAT IS TO BE ACHIEVED?

4.1. General objective

The general objective of this intervention is to **leverage the potential of data for the economy and society**. This would be brought about by facilitating a higher level of data sharing across the entire EU Digital Single Market.

The economy would be boosted by increased innovation and competitiveness. Such benefits would, for example, materialise in terms of better and personalised products for customers and of important efficiency gains in industry. Society as a whole would benefit from evidence-based policies and from the availability of more data that would help to address societal challenges (e.g. combating climate change, improving healthcare systems, addressing the challenges of ageing societies across the EU).

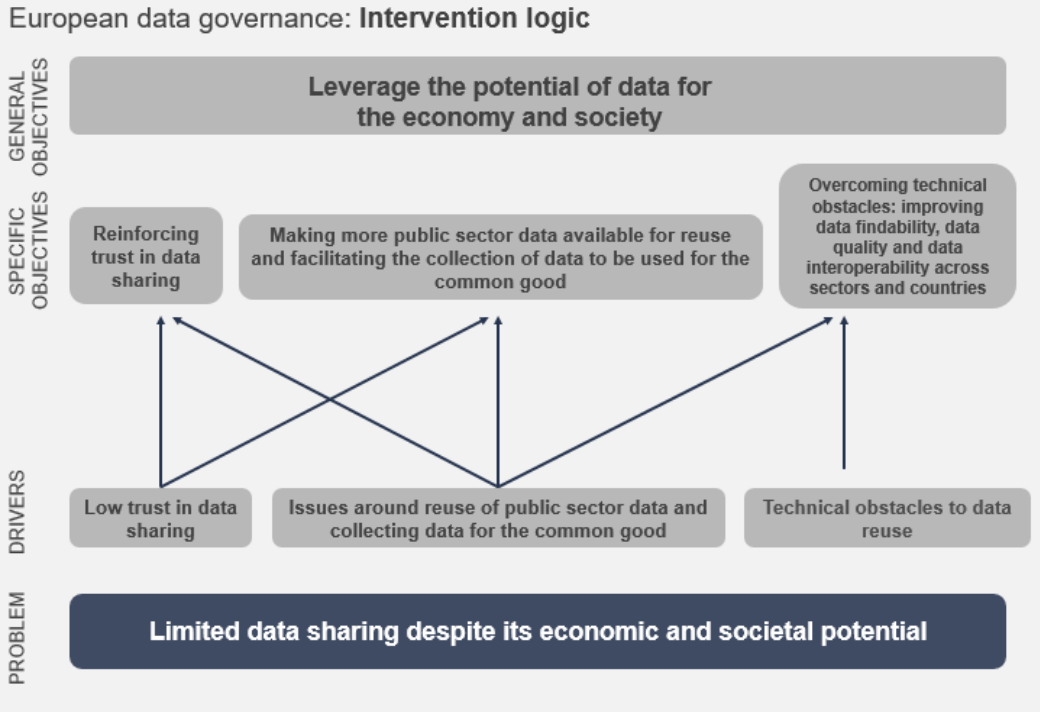
This initiative would lay the foundations to tackle the problem drivers identified in Chapter 2, and ensure that the Member States' actions on data are aligned. It would benefit the different common European dataspaces, by:

- increasing trust in data sharing;
- strengthening mechanisms that increase data availability;
- overcoming technical obstacles to the use of data.

The initiative would underpin a new, 'European' approach for data that would work as an alternative to an integrated platform model, dominated by Big Tech but potentially also by any other player with a high degree of market power. The policy interest behind the advancement of

[94] European Commission (2020b). *Outcome of the online consultation on the European strategy for data.*

such a model is to truly empower individuals to exercise their rights under the GDPR and to give companies more control over the data that they generate and over its value.



Overview: Intervention logic. Source: European Commission

4.2. Specific objectives

4.2.1 Reinforcing trust in data sharing

Trust is a key prerequisite for data sharing. Therefore, the first specific objective of the initiative is to create trust in data sharing as such. Common European dataspaces should become environments in which businesses and individuals can trust that the data they exchange or pool is secure and processed in compliance with applicable legislation as well as with the conditions they set on the use of such data. Where such a data space is organised by a specific intermediary, businesses and individuals should be able to trust those. When businesses prefer not to make recourse to an intermediary, they would benefit from a framework composed of technical standards and their related governance ('data-sharing schema') as an alternative means to create trust.

Increased trust in data sharing and the citizens' and companies' assurance that their 'sensitive data' (personal data, commercially sensitive information) is processed in line with relevant legislation and the limitations they set in contractual obligations, would serve as an incentive to share the data with selected partners. Such assurances should also apply with respect to demands to access data by governmental authorities, including from third countries that do not comply with due process requirements. The considerations that are at the heart of the CJEU judgment in

case C-311/18[95] (Schrems II), examining the impact of broad investigative powers of intelligence authorities and invalidating the Privacy Shield are very relevant in this context.

4.2.2 Making more public sector data available for reuse and facilitating the collection of data to be used for the common good

The second specific objective is to make more data available for reuse for businesses, in particular SMEs and start-ups, public administrations and researchers, by addressing the problems caused by mainly the lack of institutional capacities and expertise in the public sector, as well as by the lack of mechanisms for collecting data for the common good.

Actions should focus on data that could be made available for reuse by others on the basis of the existing legislative framework and where data holders could agree to this. This concerns i) data held by public sector bodies that cannot be made accessible as 'open data'[96] but could be used under legal and technical restrictions/processed in trusted and secure environments, and ii) data that individuals would agree to make available for reuse[97] for research, official statistics or other altruistic or innovative purposes. Targeted measures would address situations in which more technical, legal and organisational support would be necessary to make such data available.

4.2.3 Overcoming technical obstacles: improving data findability, data quality and data interoperability across sectors and countries

Interoperability of data is a precondition for using the data in different contexts. Therefore, the third specific objective is to improve data interoperability to increase data sharing across sectors, Member States and different types of organisations.

First, interoperability and generic standards could allow data to be reused across sectors and Member States more smoothly. Second, businesses, researchers and other actors should be able to easily find the data they need and ascertain that the quality of such data is fit for purpose.

Concerns about data quality affect all sorts of data-sharing situations, including for public interest purposes. This, together with the issues around interoperability, stand in the way of increasing data sharing.

5. WHAT ARE THE AVAILABLE POLICY OPTIONS?

In Chapter 2 the following problem drivers were identified: **lack of trust** in data sharing, **issues related to the reuse of public sector data** and **collecting data for the common good**, and **technical obstacles** to data reuse.

The specific objectives defined in Chapter 4 aim to overcome these issues by: **reinforcing trust** in data sharing, **making more data available** for use in the common European data spaces, and **overcoming technical obstacles**.

In order to achieve the specific objectives, four intervention areas were identified:

[95] ECLI:EU:C:2019:1145.
[96] OJ L 172, 26.6.2019, p. 56-83.
[97] Article 20 of the General Data Protection Regulation.

- mechanisms to **enhance the reuse of public sector data** that cannot be made available as open data: these would ensure that more data from public sector databases becomes available for use in a trusted way, and would increase the findability of such data;
- measures addressing **data intermediaries**, both in situations of B2B and C2B data sharing: this intervention area would increase trust in data sharing;
- measures to facilitate **data altruism**: these would ensure that more data becomes available for the common good, and would increase trust in altruism schemes;
- mechanisms to **coordinate and steer horizontal aspects of data governance**: these would contribute to overcoming technical problems, in particular issues related to interoperability.

The work of the support study for this Impact Assessment was split into four separate streams, corresponding to these four intervention areas. A cost-benefit analysis and a multi-criteria analysis was carried out for each of the areas.

The study gathered evidence through case studies and workshops (on the use of data subject to the rights of others and on possible structures of data governance) as well as market research (on data intermediaries) and legal analyses (in particular on data altruism).

5.1. What is the baseline from which options are assessed?

In the baseline scenario[98], existing horizontal EU practices such as the exchange of good practices between sectors and Member States would continue. An EU study predicts[99] that in the absence of policy and legal frameworks supporting the data economy, the value of the data economy, i.e. the overall impact of the exchange of data on the economy, would still increase from its current 2.6% of EU GDP, but only to 3.9% of EU GDP.

On a more granular level, in the baseline scenario Member States would remain free to take their own approach with regards to the reuse of data held by public bodies and the use of which is subject to the rights of others. As a result, it is uncertain whether such data would become more available for reuse for research and development purposes. More likely, the currently observed difficulties in allowing the reuse of datasets containing personal data[100] would remain unchanged. Likewise, interoperability issues across sectors and Member States would likely persist while data holders would have no incentive to ensure their data is of the highest possible quality and accuracy. Fragmentation as regards access to, and combination of, data of sufficient quality would continue.

5.2. Description of the policy options

A number of possible policy options of different strengths can be considered. In addition to the 'no action' scenario, they vary from putting in place an EU-level coordination mechanism underpinned by non-binding acts to fully-fledged legislation. The legislative intervention in turn can be split into measures of lower and higher intensity, measured against the extent to which

[98] European Commission (2020). *Support Study to this Impact Assessment*, SMART 2019/0024, prepared by Deloitte.
[99] European Commission (2020a). *Final Study Report of the Updated European data market study*, SMART 2016/0063.
[100] European Data Portal (2018). *The PSI directive and GDPR*.

they deviate from the organisational and legal status quo. The range of options can be expressed as follows:

- Policy option 0: No horizontal action at EU level – baseline scenario
- Policy option 1: Coordination at EU level and soft regulatory measures
- Policy option 2: Regulatory intervention with low intensity
- Policy option 3: Regulatory intervention with high intensity

The measures under all policy options would be combined with investments in common European data spaces foreseen under the Digital Europe programme (DEP) and the Connecting Europe Facility 2 (CEF). More specifically, the European Commission plans to invest EUR 2 billion to foster the development of data processing infrastructures, tools, architectures and mechanisms for data sharing and to federate energy-efficient and trustworthy cloud infrastructures and related services. In particular, the Commission will co-finance technical solutions (common standards, profiles and technical specifications) for federating European cloud capacities in order to ensure portability, trust, data protection, security and interoperability. Such funding is independent from this initiative, since it will support the creation of a technical infrastructure for the development of common European data spaces, without directly affecting the legal and governance arrangements for data sharing. On the other hand, the effects and the efficiency of the spending will be augmented by the impact of the legislation, which aims at reinforcing trust in data sharing and increasing the overall amount of data being shared.

The assessment does not take into account the interplay with further legislative measures, in particular the Data Act, which will deal with the issue of fairness in data access and use, and the rights and obligations of persons and organisations on data.

5.2.1 Policy option 0: Baseline scenario - No horizontal action at EU level

In the baseline scenario, no horizontal action is taken at EU level on data governance and interoperability of common European data spaces. However, action may be taken at sectoral or Member State level.

5.2.2 Policy option 1: Coordination at EU level and soft regulatory measures only

EU coordination and soft measures have been used in the area of data sharing over the past decade, with limited impact. Under this scenario, the Commission would adopt a Recommendation or guidelines. An exchange of good practices could be organised between the Member States on an ad hoc basis. Investment actions would support the deployment of data infrastructures that can underpin the common European data spaces in specific sectors.

5.2.3 Policy options 2 and 3: Regulatory intervention with low or high intensity

Policy options 2 and 3 consider regulatory intervention of low and high intensity respectively. Both policy options have similar objectives, but may lead to a different level of impact in terms of costs, benefits and administrative burden.

Both regulatory options are presented together in the following sections in order to clarify the difference between them.

A) **Mechanisms for the enhanced reuse of public sector data**

This intervention would allow for more data to be made available for reuse. It builds on the experience of some Member States in the areas of statistics[101], mobility[102] and health[103], which have spurred further data-driven innovation in their countries[104]. The regulatory intervention would require that Member States ensure that public sector bodies set in place the organisational structures and mechanisms to enhance the use of public sector data, the reuse of which is conditional on the respect of rights of others, and which therefore cannot be made available for reuse under the Open Data Directive. This concerns public sector databases that include information on individuals or company information (e.g. information on financial systems, or on the approval of pharmaceutical drugs). Very few public sector bodies currently have mechanisms in place that allow certain types of data analyses (e.g. data mining) on such data.

The policy options considered are as follows.

- Under the low intensity option, individual public sector bodies allowing the reuse of data subject to the rights of others would need to allow such reuse in line with a set of harmonised conditions. Member States would have to establish one single entry point (one-stop shop) through which reusers can contact public registers holding the data. The one-stop shop would provide advice to reusers. Member States would have to comply with a broad obligation to have capacity and services in place to facilitate further compatible uses of the data. They would be free to decide upon the exact form of these mechanisms, taking into account differences between individual sectors.
  In practice this would mean that a reuser from Member State A could contact the single entry point in country B to see how to get access to a certain type of data. The single entry point would channel the request to the relevant public sector body or bodies. These public sector bodies would receive technical and legal support for making re-use possible, but would remain responsible for the operations. Member States would have to invest in this support system.
- The high intensity option would oblige Member States to create one single data authorisation body as a central decision-making point. It would be competent to decide on all further compatible uses.
  In practice this would mean that reusers would be served by a single organisation that would offer technical solutions for querying the data (e.g. safe digital reading rooms, or mechanisms to bring algorithms to the data) and would, where relevant, issue permits for data re-use. This would require national legislation to be in place underpinning the operations of the organisation, including in terms of compliance with data protection legislation. It could also imply a change in the way in which public registries are managed.

[101] *RatSWD* (German Data Forum).
[102] Aholainen J., *Finnish solutions for opening up fare data*, TRAFICOM; European Data Portal (2017). *Smart mobility in Finland*.
[103] Cuggia M. and Combes S. (2019). *The French health data hub and the German medical informatics initiatives: two national projects to promote data sharing in healthcare*, Yearbook of Medical Informatics.
[104] SITRA *website*.

Both options would include a prohibition for exclusive arrangements for data held by the public sector and set out the conditions under which situations having de facto the effect of granting exclusive access would be lawful. This would avoid discriminatory practices, and in particular the risk that powerful players in the market get exclusive access to the data (e.g. health data)[105]. Neither option would create a right to re-use. They would both build on situations where re-use of the public data is allowed in the Member States for commercial or non-commercial purposes.

In the design of the options, a number of dimensions were considered, such as centralisation of the decision-making on who can use what type of public register data as compared to the responsibility of individual bodies, and how to assist potential reusers (other than by centralising the decision-making).

Under both options, Member States would have to ensure that these mechanisms comply with a number of harmonised, compulsory criteria that would ensure trust in the mechanisms, also across borders. Member States would be required to provide a secure data processing environment to allow innovative processing of data to which access can be granted under conditions controlled by the public sector (a 'safe reading room for data'). Whenever data is being transferred to a reuser, assurances should be in place that ensure compliance with the GDPR and preserve the commercial confidentiality of the data. Both options would benefit from the use of the technical infrastructures and tools developed with the support of the DEP and CEF programmes to ensure interoperability between the solutions offered by Member States.

B) **A certification/labelling framework for data intermediaries**

In order to lower transaction costs for data sharing or pooling within common European data spaces, businesses and/or individuals may want to have recourse to data-sharing intermediaries[106]. In light of distrust in platform business models[107], novel service providers are emerging but with limited brand recognition. Their business model is based on transaction fees or regular subscriptions. It excludes own use of the data the exchange of which they offer to facilitate.

The emergence of data intermediaries (providers of data-sharing services), such as 'data marketplaces', is largely supported by stakeholders. Almost 60% of respondents to the 2020 online consultation considered that they are useful enablers in the data economy[108]. Both in reply

[105] Similar arrangements have been tested and worked well in the context of re-use based on the Open Data Directive. The need for a similar provision for data not covered by the open data Directive was already signalled in the evaluation report of the Directive.
[106] OECD (2019). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris, pp. 36-39; The ODI (2020). *How do we create trustworthy and sustainable data institutions?* Blog piece.
[107] Bradbury D. (2020). *Distrust of Big Tech is contact tracing's biggest hurdle*, InfoSec Blog; Hutchinson A. (2020). *New report shows universal distrust in social media as a news source*, SocialMediaToday; CISION PR Newswire (2019). *Brand reputation at risk from consumers' data distrust*.
[108] European Commission (2020b). *Outcome of the online consultation on the European strategy for data*.

to the online consultation and in a workshop on data intermediaries held in May 2020[109], stakeholders mentioned that this new type of facilitator would increase opportunities for cross-sectoral innovation. They underlined the importance of the neutrality of data intermediaries as a tool for businesses' and individuals' data sovereignty, contrary to what has been seen until today with existing data aggregators[110].

A **certification or labelling framework** would allow novel data intermediaries to increase their visibility as trustworthy organisers/orchestrators of data sharing or pooling. Legislation would define a set of core criteria that should be met by all certified/ labelled intermediaries in order to **demonstrate their neutrality:** absence of conflict of interest, no competition with data users (e.g. no development of own data apps in competition with others, so as to avoid any risks of self-preferencing) and commitment to not discriminate between companies that would like to offer data services (openness obligation). Two types of intermediaries would be covered by this scheme: those addressing business users and those addressing individuals (providers of 'personal data spaces'). Furthermore, they should be able to ensure through technical and organisational measures that the data are transferred in compliance with the stated preferences of the company or individual.

The two policy options considered are as follows.

- Under the lower intensity option, labelling/certification would be voluntary for actors involved in data sharing. The awarded labels would be equally valid in all Member States.
  In practice, the provision of data intermediary services remains an unregulated activity. Data intermediaries could obtain a label or certificate in order to show that their business model is in line with a series of requirements set at the EU level. The label/certificate would not be a requirement for offering data intermediary services in the EU.
- Under the higher intensity option, the certification of providers of data-sharing services would be compulsory. A compulsory scheme would ensure that all data intermediaries operating in the Union would comply with the requirements. Certification would be complementary to existing certification frameworks (e.g. under the GDPR and the Cybersecurity Act).
  In practice, data intermediary services would have to comply with the requirements of the certification scheme before they would be able to start operating and offering data-sharing services. This would make data sharing a regulated activity.

Under both options, the labels/certificates could be awarded by public authorities or by private conformity assessment bodies, based on criteria developed at the European level.

Self-regulation was discarded, as a) there is no natural industry forum for this emerging market and b) it was deemed that the stakeholders involved would not be able to agree upon strict criteria of neutrality. Self-regulation would also potentially lead to the emergence of different

[109] European Commission (2020c). *Report of the workshop on labels for/certification of providers of technical solutions for data exchange*.
[110] Barclays, BBVA, Deutsche Bank AG, HSBC, ING Group and Banco Santander (2019). *Advancing the EU data framework: user data sharing*, Policy Paper sent to DG CNECT.

solutions in different sectors and countries, thus increasing fragmentation. Currently there is no such self-regulatory certification scheme in place.

C) **Measures facilitating data altruism**

In order to address the lack of means to manage consent-based sharing of personal data and the availability of data in general, the regulatory intervention would require Member States to enable data altruism by putting in place the necessary laws and processes that allow companies and individuals to make their data available for the wider common good based on consent. As the 2020 online consultation showed, a large proportion of respondents (87%) consider that there are not sufficient mechanisms in place for altruistic data sharing, while 83.3% see the need for such enabling tools and mechanisms to be able to share their data for the common good[111]. Thus, it is the lack of tools, not of willingness, that hampers data sharing for the common good.

The two options considered are as follows.

- The lower intensity option would require that Member States have in place certification schemes for data altruism mechanisms and/or organisations offering such mechanisms. Certification would be voluntary. Certificates would be issued by private certification bodies, or by a public authority. Certification would be complementary to existing certification frameworks.
  In practice, an organisation engaging in data altruism could apply for certification to show that it is a trusted intermediary (this would be voluntary). The application would be handled by a public authority or a private certification body.
- The higher intensity option would require Member States to have in place an authorisation scheme for data altruism mechanisms and/or organisations offering such mechanisms. Organisations that seek to perform activities facilitating data altruism would have to comply with the requirements and seek authorisation before launching their operations. The authorisation would be compulsory, and would be handled and issued by a designated national authority (which could be an existing body). The authorisation could be general (allowing data activities to start upon notification) or ex ante (approval by the competent authority is a prior requirement for starting the activities). The granted authorisations would be equally valid in all Member States. Member States' authorities would monitor compliance. As such, a mandatory authorisation regime would act as a filter and an entry barrier for entities that wish to start providing such mechanisms. It would create an intervention that is of higher intensity than a voluntary certification scheme.
  In practice, data altruism activities in the EU could only be carried out by organisations that have sought an authorisation from a public authority.

At the core of both options is the wish to ensure that data altruism mechanisms (operated by public sector organisations, NGOs and other private sector organisations) are truly altruistic. The scheme should prevent attempts to describe a data-sharing activity as altruistic, when in fact the

[111] European Commission (2020b). *Outcome of the online consultation on the European strategy for data.*

consent to commercial usages is 'hidden' in long and hard-to-understand consent statements. Furthermore, organisations engaging in data altruism should be able to ensure through technical and organisational measures that the data are used in compliance with the stated preferences of the company or individual.

In line with the rights conferred by the GDPR, these mechanisms for data altruism would be based on consent under Article 7 GDPR and build on the portability right provided by Article 20 GDPR. In line with the GDPR, they should **also provide for individuals to withdraw consent for the processing of their data.**

Both options also foresee the development of a European data altruism portability and consent form similar to those existing in the field of blood, tissue and organ donation, which could be tailored to specific sectors and types of data, to be able to easily give and withdraw consent.

For the higher intensity option, only a public authorisation scheme would qualify. This would be justified given the importance of trust and the nature of acts of data altruism, where people and companies are making their data (sometimes sensitive data) available for the common good (e.g. improving traffic conditions, contributing to health research). For the lower intensity regulatory intervention, both certification by a public authority and by private conformity assessment bodies are possible. For this lower intensity intervention, the Impact Assessment study focused on certification by private conformity assessment bodies, in view of assessing the impact of two clearly contrasting options, representing the key parameters at stake.

As for the labelling/certification of data intermediaries under B), centralising the authorisation in a European body was discarded due to reasons of costs and political feasibility.

D) **Mechanism to coordinate and steer horizontal aspects of governance (European Data Innovation Board)**

A European Data Innovation Board would be created to coordinate efforts in Member States and at European level to support data-driven innovation, to lower transaction costs and prevent further sectoral fragmentation.

It would coordinate national efforts to make more public sector data available, play a role in the certification or labelling of trusted providers of data-sharing services and take a lead in governing standardisation and the prioritisation on standardisation around data sharing, in particular across sectors. The achievements on prioritisation would feed into the Commission's ICT rolling plan for standardisation.

The two options considered are as follows.

- In the lower intensity option, the European Data Innovation Board would be a formal Expert Group, with a secretariat provided by the Commission. It would include among its members representatives of the sectoral data spaces and other interested parties, aiming for a balanced representation of the different sectors (avoiding duplication with existing expert groups).
  It would be responsible for facilitating the exchange of national practices and policies on data altruism and the use of public data that cannot be made available as open data, and advising on the prioritisation of standards for cross-sector data reuse for the Commission's rolling plan

for ICT standardisation, and on the establishment and maintenance of a schema of standards (technical and legal) related to problems common to all data-sharing situations irrespective of a sector[112]. It would also be tasked with facilitating the exchange of national best practices on voluntary labels for trusted providers of data-sharing services. As a Commission expert group, it would include experts from authorities from each Member State. It would not interfere with the roles and powers of the Member State authorities, given that it would only advise the Commission and support it in facilitating the exchange of national practices.
In practice, the European Data Innovation Board would advise and support the Commission on matters related to data sharing, focusing on cross-sector issues.

- In the higher intensity option, the European Data Innovation Board would be a self-standing European body with legal personality, in which representation of sectoral data spaces would be ensured. It would be supported by a secretariat. In addition to the above functions, it would supervise the award process of voluntary labels and, where relevant, authorisations performed by the designated Member State authorities. In this sense, the relationship between the Board and the Member State authorities would be closer, as they would be overseen by this European level body. It would keep a register of the organisations that obtained a label or authorisation. These functions go beyond the power of a Commission expert group and can only be entrusted to an independent body.
In practice, the European Data Innovation Board would advise the Commission, but also carry out activities autonomously, including supervisory functions.

In designing the options, the possibility of setting up an **informal** Commission expert group was also examined. It was found that the political importance of the subject matter warrants the establishment of a formal group or a self-standing independent body. The examples that served as inspiration both for the form and the tasks of the body are either formal expert groups (e.g. European Union Ecolabelling Board) or independent bodies (e.g. European Data Protection Board - EDPB, European Union Agency for Cybersecurity - ENISA).

Bringing the required functions under the remit of the EDPB was also explored. The EDPB deals with the specific issue of personal data protection. Data sharing additionally requires expertise in competition law, and in sector-specific data access and usage regimes, as well as a technical knowledge of technical sharing mechanisms and standards. Also, the composition of the decision-making instance in the EDPB, currently composed of representatives of national data protection authorities, would need to be modified or at least complemented, which would require amending the legislative text of the GDPR. Therefore, this possibility was not discussed further.

[112] Examples of such standards are: description of actors in a data sharing ecosystem, identification of persons, legal entities and connected objects, authentication, permissions on data use (consent in the case of personal data), portability of permissions (consent).

Summary table

Policy options 2 and 3 for regulatory intervention with lower and higher intensity

Intervention area	Regulatory intervention with low intensity	Regulatory intervention with high intensity
Mechanisms for enhanced reuse of public sector data	The reuse of public sector data that is subject to the rights of others would have to comply with basic EU-wide rules (in particular non-exclusivity). Individual public sector bodies allowing this type of reuse would need to be technically equipped to ensure that privacy and confidentiality are fully preserved. Member States should have a single entry point (one-stop shop) in place for persons or organisations that seek to reuse this data. Member States should have capacity and services in place to support public sector bodies for this type of reuse.	The reuse of public sector data that is subject to the rights of others would have to comply with basic EU-wide rules (in particular non-exclusivity). Member States should create a single data authorisation body competent to licence further compatible uses of data contained in any public register that is subject to the rights of others.
Certification/labelling framework for data intermediaries	Voluntary labelling scheme for data intermediaries offering B2B data-sharing services and those offering personal data spaces. A key criterion to obtain the label/certification: the data intermediary cannot use the data as part of its business model.	Compulsory certification scheme for data intermediaries offering B2B data-sharing services and those offering personal data spaces. A key criterion for obtaining the label/certification: the data intermediary cannot use the data as part of its business model.
Measures facilitating data altruism	Obligation on Member States to have legal and administrative arrangements in place to enable data altruism. Voluntary certification scheme for data altruism mechanisms and/or organisations offering such mechanisms. Certification issued by private certification bodies or a public authority.	Obligation on Member States to have legal and administrative arrangements in place to enable data altruism. Compulsory authorisation scheme for data altruism mechanisms and/or organisations offering such mechanisms. Authorisation issued by a public authority.
European Data Innovation Board	The European Data Innovation Board would be a light coordination mechanism at EU level in the form of a formal Expert Group, hosted by the Commission. It would be composed of representatives of the Member States and of representatives for the different domains (health, statistics, etc). It would facilitate the exchange of national practices on the items covered by the legal	The European Data Innovation Board would be an independent European structure with legal personality and supported by a secretariat (e.g. inspired by the structure of EDPB). In addition to the functions under the lower intensity option, it would be tasked with supervisory functions and keeping registers of awarded labels and authorisations.

	instrument, and would address cross-sector standardisation issues.	

Source: European Commission

Relation with sectoral initiatives: The relation between the horizontal framework and the sectoral initiatives is shown in the image at the end of Chapter 1. The horizontal framework will provide the building bricks for individual data spaces so that they can be established faster. The governance of the individual data spaces should, however, reflect the needs of the sector and the set-up of the stakeholder ecosystem and thus be defined by the sector itself and not be prescribed by the horizontal framework.

A specific data space may have its own standards (which can be either EU standards or standards devised by stakeholders). Existing governance frameworks, such as the eHealth network in the area of health, will not be affected. The European Data Innovation Board will include representatives of individual common European data spaces as they emerge. The horizontal framework would leave room for sector-specific *lex specialis* rules.

5.3. Options discarded at an early stage

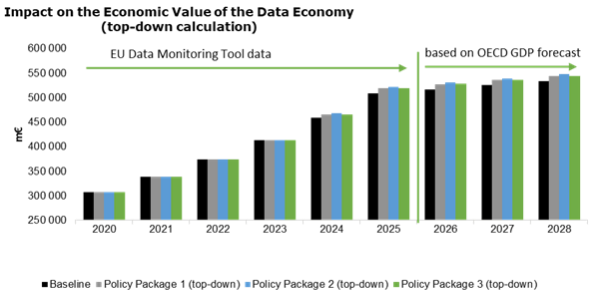
No options were discarded from the outset.

6. WHAT ARE THE IMPACTS OF THE POLICY OPTIONS?

6.1. Economic impact

The Impact Assessment support study considered as the baseline the total economic value of the data economy for the EU-27 in 2020, which is EUR 325 billion (2.6% of GDP). This number takes into account a correction linked to COVID-19's impact on the overall EU economy.

The graphs below illustrate the expected evolution, compared to the baseline scenario, of the direct economic value of data under the lower and higher intensity scenarios, as well as the preferred option of a package of lower and higher intensity interventions (see Chapter 8). The fact that the results of the top-down (based on contribution to GDP) and bottom-up (validation calculation based on cost-benefit analyses) approaches[113] are almost identical confirms the solidity of the methodology.



[113] For more information on these approaches, see: European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.

The graph shows the compared economic impact of the different policy packages for the indicated years. Policy package 1 includes low intensity regulatory intervention in all four areas, policy package 2 contains high intensity regulatory intervention in all areas, while policy package 3 denotes the preferred, mixed option. Source: SMART 2019/0024

In 2028, the value of the data economy would increase from EUR 533.5 billion[114] (3.87% of GDP) under the baseline scenario:

- to between EUR 540.5 and 544.0 billion if the lower intensity regulatory intervention was introduced (from 3.92% to 3.94% of GDP);
- to between EUR 542.7 and 547.3 billion if the higher intensity regulatory intervention was introduced (from 3.93% to 3.97% of GDP).

The impacts are calculated until 2025 on the basis of the value of the data economy as projected by the International Data Corporation (IDC) for the baseline. The IDC forecast projects a growth of the data economy of approx. 8% per year[115]. The IDC forecast for the growth of the EU data economy, however, ends in 2025. In order to calculate impacts beyond 2025, the support study took a conservative approach and calculated the impacts on the basis of the GDP growth rate forecast of the OECD (1.5% per year)[116]. For this reason, the impacts beyond 2025 are based on a much lower per annum growth rate.

At first sight, the gains (of between EUR 7 and EUR 10.5 billion for the lower intensity option and between EUR 9.2 and EUR 13.8 billion for the higher intensity option) seem relatively small compared to the overall size of the data economy that is taken as the baseline for the calculation. It should, however, be borne in mind that they are based on a conservative approach, focusing on the direct impact, and only to a very limited extent on the indirect impact, of the set of measures under consideration. Indeed, the calculations in the support study do not cover the full range of potential impacts of the measures on the economy and society. This also explains the large gap between the impact, as calculated in the impact assessment study, and more general studies that look at the potential of data sharing for the economy and society. Examples are the estimated EUR 1.3 trillion in increased productivity by 2027 in manufacturing through IoT data[117], or savings of approximately EUR 120[118] billion a year in the EU health sector. Ultimately, as estimated by the OECD, the economic value of improved data sharing could amount to up to 2.5% of GDP[119].

This broader potential should be kept in mind when assessing the effects of the measures. The initiative is a necessary first step in the process of creating common European data spaces. It can

[114] For more information on these approaches, see: European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.
[115] European Commission (2020a). *Final Study Report of the Updated European Data Market Study,* SMART 2016/0063.
[116] OECD (2019). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris.
[117] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.
[118] McKinsey (2020). *Shaping the digital transformation in Europe.*
[119] OECD (2019). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*.

make data markets in different sectors function better by creating trust. However, the full range of benefits from the measures rely on actors in the common data spaces seizing the opportunities offered by these building blocks. The measures taken by the initiative would act as a catalyst to increase data sharing across the EU[120]. This would lead to the creation of more efficient services and new products based on data, including AI. This catalyst effect would not only benefit the data economy, but the EU economy and society as a whole.

The Commission has announced that it intends to invest EUR 2 billion in data infrastructures through the DEP and CEF programmes. These investments in the creation of a European data sharing and processing infrastructure will lower the cost for technically implementing the policy options proposed in the current instrument. At the same time the proposed legislation will reinforce the impact of the investment by increasing trust and making more data accessible. The exact effect of the envisaged investments on the different options are hard to establish, and have not been taken into account into the calculations.

The support study to this Impact Assessment shows that SMEs in particular stand to benefit from the initiative: in addition to benefits from higher interoperability, standardisation and simplified access to public sector data, they would benefit from the certification/labelling schemes. The one-off costs of certification/labelling (EUR 20 000-50 000 for a voluntary label, and EUR 35 000-75 000 for a compulsory certification) for data intermediaries would be countered by the high gains in both client base and revenue (25-50% increase)[121].

Member States would incur costs to establish the necessary mechanisms to provide services and to carry out the different tasks, in particular in relation to the measures to facilitate the use of public sector data that cannot be available as 'open data'. However, as outlined below, the direct economic gains alone would outweigh these costs under both scenarios[122]. Besides, Member States would recuperate a large part of the investments through fees for the different services related to reuse, and data holders would benefit from significant cost reductions.

A harmonised horizontal governance framework across the EU would create a level playing field for all the Member States. The initiative would ensure a minimum level of harmonisation across the EU, while leaving a certain leeway for the Member States in terms of how to organise public registries and authorisation mechanisms, building on existing structures. It would create certainty for data users across the EU and for those who want to make data available, and increase trust in data intermediaries and in mechanisms for making more data available for use. Even though some differences between the Member States would remain (e.g. in terms of the supply of public sector data), overall the measures would be an important step towards a more harmonised framework and the creation of a real internal market for data.

The positive impacts of the initiative are expected to be spread across the EU rather than benefiting specific countries. Data-focused start-ups are emerging across the EU, in larger and in

[120] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.
[121] Idem.
[122] Idem.

smaller Member States. For example, in Romania and Bulgaria the share of data companies' total revenues in 2019 as part of the total revenues of all companies was 3.5% and 3.9% respectively, which is similar to France (3.8%)[123]. Highly harmonised conditions for reuse would help reusers from all Member States, regardless of their size or economy.

6.1.1. Baseline scenario

In the absence of EU intervention, the data economy would continue to grow to an estimated EUR 533.5 billion in 2028[124]. Without an alternative European model, there would be a risk of platformisation and the hegemony of the Big Tech companies in this field. There would only be a moderate increase in data use, which would limit the capacity for productivity gains in all sectors, in particular in the traditional sectors that are currently undergoing a major paradigm shift due to data-driven innovation.

Industry-driven initiatives paired with national initiatives to support data sharing in sectors that are key to the particular Member State (e.g. industrial manufacturing in Germany and France, logistics or agriculture in the Netherlands, forestry in Finland) would emerge, but remain limited in terms of impact. Companies would remain wary of data sharing: they would either encounter significant costs in doing it themselves, or face the choice of relying on integrated tech vendors (which would have stronger negotiating power) or on start-ups (with no brand recognition and capacity to become a relevant player in facilitating data sharing).

Individuals may come across initiatives, for example driven by the research communities, asking them to make available data on altruistic grounds, but would not be provided with trusted means to do so. Similarly, researchers would be faced with uncertainty when collecting consent on this basis and would thus be more reticent to make use of this mechanism, resulting in losses in advances in science.

Overall, this would lead to a scenario in which large, integrated tech companies that have already collected large volumes of data would further strengthen their position to decide on data access. They would become centre points of additional ecosystems as they expand into new activities such as health, insurance or finance. Furthermore, they would be able to reinforce their position by acquiring additional data or start-ups that are dependent on them. This would have an impact on the quality of machine-leaning outcomes (e.g. facial recognition algorithms), and result in concerns regarding data quality and bias.

6.1.2. Coordination at EU level and soft regulatory measures

The impact of this policy option depends on the uptake of the Commission's Recommendations or guidelines by Member States. Experience with the two existing soft law measures related to data sharing[125] shows that, due to their non-binding character, they have been taken up with

[123] E.g. in Romania, and Bulgaria the share of data companies' total revenues in 2019 were 3.5% and 3.9% respectively, similarly to Member States such as France (3.9%). Source: *The European Data Market Monitoring Tool*.

[124] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.

[125] SWD(2018) 125 final; OJ L 134, 31.5.2018, p. 12-18.

different intensities and at a different pace by actors in the data economy and Member States. Therefore, this policy option would be unlikely to provide the swift and harmonised action necessary for the EU to become a key player in this emerging market. Soft measures alone cannot be relied upon to prevent the further development of regulatory divergences between Member States. Additionally, coordination at EU level would also be achieved under policy options 2 and 3.

Mechanisms to enhance the reuse of public sector data subject to the rights of others

The impacts of this policy option would rely on the willingness of Member States to set up structures (such as a one-stop shop akin to the Health Data Hub, or a single data authorisation body like Findata), which would also be subject to a set of uniform conditions. According to a workshop organised in the context of the support study, only an estimated nine to 13 Member States would likely implement such recommendations[126]. In addition, the level of ambition of such guidelines or recommendations would likely be inversely proportional to the number of Member States adopting them. Ensuring that similar requirements in relation to the use of such public sector data are available throughout the EU is a matter of legislation and cannot be achieved by soft law.

Certification/labelling framework for data intermediaries

Stakeholders interviewed in the context of the support study generally considered that this policy option would have little added value compared to the baseline scenario. Developing requirements for any label in this emerging market would be difficult as it was deemed that industry would not be able to agree upon strict criteria of neutrality. Moreover, soft measures would not guarantee a fair and representative selection of certification criteria/requirements for the various types of data intermediaries active in the EU. This policy option could lead to the adoption of different labels in the Member States and sectors and, thus, to further fragmentation.

Measures facilitating data altruism

Similarly to the baseline scenario, individuals would not have trusted means to share their data on altruistic grounds. The Commission could host expert exchanges and publish guidance to support individuals and researchers, who would also face uncertainty. Such guidance would, however, not provide sufficient assurances to consumers or researchers for concrete use-cases.

Some interviewed Member States considered that an EU-level coordination mechanism for data altruism mechanisms would reduce their workload by avoiding multiple bilateral discussions, but would not necessarily accelerate the discussions. In addition, they considered that only the Member States that are already actively pursuing data altruism mechanisms would likely participate. Private sector interviewees considered that coordination at EU level could take very long and not result in concrete action. As adoption of the measures would be voluntary, this could widen the data altruism gap between different Member States and companies.

[126] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.

Last, but not least, protecting individuals against data altruism mechanisms that are not truly altruistic is a matter of legislation and cannot be guaranteed by soft-law measures.

European level mechanisms for coordination of standardisation

The costs of an informal group of 10 experts participating in four 3-day meetings per year would amount to around EUR 24 000 per year. In addition, the employers of the participating experts would incur costs. The costs related to the adoption of standards would be borne by industry, as this option does not entail any mandatory standards[127].

However, it was found that the creation of an informal expert group for governance aspects of data sharing would be unlikely to have any effects at all, due to its informal nature. Indeed, according to the study team's estimates, it would increase the number of data users only by 0.1%[128].

Multi-criteria analysis

A multi-criteria analysis (see Chapter 7) was not performed for this policy option because, as indicated above, its effectiveness is limited and dependent on uptake by Member States (and hence it is not quantifiable). Furthermore, during the interviews and workshops conducted with stakeholders in the context of the support study, shortcomings were identified for each intervention area.

6.1.3. Policy option 2: Lower intensity legislation

This policy option entails the softer and less expensive options in all four intervention areas. The measures are expected to yield considerable benefits in the form of enhanced reuse of public sector data, elevated trust in data intermediaries and data altruism, as well as better coordination in the field of standardisation. Compared to the baseline scenario, it would also directly contribute to the growth of the data economy by between EUR 7 and EUR 10.5 billion in 2028[129].

Mechanisms to enhance the reuse of public sector data subject to the rights of others – one-stop shop

Under this option, public sector bodies which grant permissions for the reuse of data subject to the rights of others would need to be technically equipped in a way that ensures that privacy and confidentiality are fully preserved. Member States would have to set up a one-stop shop for reusers and support mechanisms to provide public sector bodies with the necessary legal and technical expertise.

This option would foster trust through transparency between data reusers and data holders, as well as trust among the general public – particularly if the one-stop shop provides expert guidance to citizens on their rights under data protection laws.

[127] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.
[128] Idem.
[129] Idem.

Benefits to public sector bodies, researchers and businesses as reusers include:

- Time and resources saved in identifying the data holder with the desired data and in accessing already interoperable data across sectors;
- Increased fairness in access to such data, i.e. all reusers would have equal access to valuable information on how to acquire permission to re-use that data, which would likely result in an increase in such reuse;
- Access to expert guidance, potentially resulting in time and resources savings related to legal training;
- Access to data of a higher quality (since holders would have an incentive to ensure quality knowing that the data would be reused) and potentially to better tailored data (since data holders would have a clearer views of reusers' needs).

Benefits to individual public sector bodies (as data holders) would include:

- Access to expert guidance, potentially resulting in time and resources savings related to legal training;
- Access to technical guidance on how to allow data reuse, resulting in a decreased risk of data breach and the associated costs;
- Time and resources saved by not providing, and maintaining, a secure data processing environment;
- Access to an increased amount of research resulting from a higher demand for such data – leading to better policymaking.

In order to determine and calculate the costs and benefits of this intervention area, national experiences with such mechanisms (e.g. Findata, Centre d'accès sécurisé aux données, Forschungsdatenzentren) were taken into account.

Mechanisms to enhance the reuse of public sector data subject to the rights of others: Costs and benefits[130]	
Costs	Benefits
One-off investment of €10.6 million for Member States to establish the mechanisms to handle the data and create the one-stop shops.	Income across the Member States for providing services would amount to approximately €41.8 million per year (assuming an average fee of €500 per application).
Annual maintenance costs for Member States of €600 000 per year.	Public sector bodies across the EU would save around €684 million/ year due to the lower cost of data processing and management.
	Cost savings for reusers of €49.2 million /year as a result of easier reuse of data (e.g. easier data discovery).

[130] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.

Voluntary certification/labelling scheme

Data intermediaries would bear the **cost of obtaining and maintaining the certification or label** as well as **implementation costs** to ensure compliance with the requirements. However, as this would be a purely voluntary mechanism, SMEs would not be disproportionately burdened. Data reusers might be impacted by indirect transaction and implementation costs, as the certified intermediaries might increase the user charges to cover the cost of certification.

The benefits mainly include the increased **trust** between actors, leading to further **efficiency gains**, **time savings**, **increase of the client base** and data transactions and therefore **increase in revenues**, allowing data intermediaries to scale up, both regarding gains in client base and revenue. As an indirect benefit, there would be a competition increase for data intermediaries in both the B2B and C2B markets[131].

Certification/labelling would have a **cumulative benefit in terms of company growth**[132]. Increased trust in the market could also lead to an **increase in funding**, as investors would consider it safer to invest in certified companies. Data holders would have the opportunity to monetise more from data sharing while more individuals would be willing to share their personal data through the certified platforms.

Since certification would be voluntary, the positive impacts of this policy option depend on the number of data intermediaries that decide to obtain certification. Given that this is an emerging market, stakeholders indicated that they favour a voluntary scheme, as it would provide an opportunity to see what works and what does not, without disturbing data markets[133].

The contrast between the seemingly low overall impacts and the high benefits for the individual intermediaries is justified by the narrow scope of the specific intervention measures. Even though the individual benefits are high, the total impact on the overall value of the data economy remains low.

Voluntary certification/labelling scheme: Costs and benefits[134]	
Costs	Benefits
One-off cost of €20 000-50 000 for obtaining the label/certification.	Efficiency gains, time savings: 25%-50% business development time acceleration.
€20 000-35 000/ year for renewing the label/certification.	Increase in client base and data transactions leading to increase of revenues (25%-50% expected increase in revenues and client base).
	Up to 25% competition increase for data intermediaries

[131] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.
[132] European Commission (2020c). *Report of the workshop on labels for/certification of providers of technical solutions for data exchange*.
[133] Idem.
[134] European Commission (2020). Support Study to this Impact Assessment, SMART 2019/0024, prepared by Deloitte.

	in both the B2B and C2B markets in a 2-5 years' timeframe.
	In the first year after certification, company revenue is expected to double, the following year to increase by 50%, and the third year to increase by 25%.
	Additional revenues of €48 million per intermediary in the year 2028.

Voluntary certification framework for data altruism services

An obligation on Member States to implement a voluntary certification scheme for **data altruism** mechanisms would create trust, and would result in more data being made available for the common good. As certification could be done also by public sector bodies, costs could be subsidised for certain entities depending on size, for example for an SME (if for-profit approaches are also allowed) or NGO. In addition, NGOs could receive an approximately 10% discount for authorisation. Alternatively, certification could be acquired without any fees, for free.

A certification mechanism would allow a new category of entities in the data ecosystem to flourish (e.g. data charities, data cooperatives for the common good, certification entities and the development of a new non-for-profit business opportunity for existing NGOs). It is expected that by the year 2028 there would be around 1 250 intermediaries facilitating data altruism, with around 5 million citizens and 500 companies participating in such schemes[135]. This policy option would ultimately streamline data altruism and reduce organisational, technical and legal costs in the long run.

As for the altruistic individuals and companies, this would ensure that their data is secure and the mechanism is legally compliant and resilient to cyberattacks, thereby increasing transparency of and trust in data altruism. Internationally, this could offer an opportunity for the EU to be a front-runner in privacy-enhanced and secure data altruism and to set global standards, attracting foreign researchers and innovators to the EU. The benefits of this policy option relate to providing an easy and transparent way to access data from various fields, contributing to research and development as well as improving decision-making. This only includes to a very limited extent the downstream societal benefits, such as faster research, better cures for diseases or improved mobility, due to the lack of available data to quantify this.

Voluntary certification framework for data altruism services: Costs and benefits[136]	
Costs	Benefits
One-off cost for data holders or data altruistic	€22 million for the period 2024-2028, based on the

[135] Idem.
[136] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.

organisations to obtain certification of €20 000-50 000. If the certification is carried out by private sector bodies, costs could be lowered to €3 800-10 500 for SMEs and €3 420-9 450 for NGOs.	predicted revenues of the intermediaries and the total value of data.
€20 000-35 000/ year for renewal.	

European Data Innovation Board as an expert group

Costs related to the **creation of a formal expert group** including all the Member States would stem mainly from organising and participating in the meetings as well as the related activities.

As for the benefits, traditional businesses would benefit from an increased adoption of standards by the standardisation organisations, leading to a reduction in costs for acquiring, integrating and processing data. Estimates from individual case studies show that adoption of standards for data sharing results in increased data-sharing activities. Benefits are calculated on the assumption that through interoperability made possible by standardisation, 800 companies would save 15% of EUR 50 million operational costs over 5 years[137].

European Data Innovation Board as an expert group: Costs and benefits[138]	
Costs	Benefits
€280 000/ year (including travel costs amounting to ca. €50 000-70 000 and operational costs of €180 000-210 000).	€1.2 billion in 2028.

6.1.4. Policy option 3: Higher intensity legislation

As the analysis carried out in the support study shows, opting for the higher intensity regulatory intervention is expected to produce the highest costs, due to the establishment of mechanisms that would generate more expenses. However, it would also potentially create the highest net benefits. Compared to the baseline scenario, it would contribute to the growth of the data economy by between EUR 9.2 and EUR 13.8 billion in 2028[139]. For this option, the legal and political feasibility as well as its efficiency were also thoroughly considered.

A single data authorisation body in each Member State to enhance the reuse of public sector data subject to the rights of others

The costs of a single authorisation body would be higher than in policy option 2, due to the need to create a standalone body, which would perform more activities than under policy option 2. It

[137] Idem.
[138] Idem.
[139] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.

would employ an estimated 25 FTEs once fully running, with each FTE costing approximately EUR 75 000 and requiring 1 to 2 weeks of training[140].

In addition to the benefits described for policy option 2, dealing with one single authorisation body would lead to substantial savings for reusers. One stakeholder estimates that not having to pre-process data from different holders would save them several days of work each time. Not having to submit separate data access applications for a given research project would save about half the overall time spent applying.

As data holders, public sector bodies would gain time and resources as a result of lower costs for data processing and management (EUR 1 253.4 million /year).

A single data authorisation body to enhance the reuse of public sector data: Costs and benefits[141]	
Costs	Benefits
One-off costs for the establishment of data authorisation bodies of approximately €21.2 million.	Costs saving of approximately €167 million/ year for the EU-27.
Annual running costs of approximately €12.2 million.	Additional gains for data holders of €212.7 million /year revenue from application fees (assuming an average fee of €500 per application).

A compulsory certification/labelling framework for data intermediaries

Compared to the lower intensity policy option, costs are expected to be higher. This would make a compulsory certification/labelling framework potentially problematic for smaller data intermediaries.

The benefits of this policy option remain similar to policy option 2, with 25%-50% expected increase in revenues and client base and up to 50% business development time acceleration, which is related to the increased trust between the actors that would result from compulsory certification.

The certification could be done by a public authority or a private conformity assessment body. While setting up an accreditation process for this new field of activity for private conformity assessment bodies would be time-consuming, certification by an existing public authority could be preferable in view of a quick start of the functioning of the scheme.

A compulsory certification/labelling framework for data intermediaries: Costs and benefits[142]	
Costs	Benefits
One-off costs of €35 000-75 000 for obtaining the	€16.7 million per intermediary of additional revenues in

[140] Idem.
[141] Idem.
[142] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.

certification.	the year 2028.
€20 000-50 000/ year for renewal of the certification.	

Compulsory authorisation framework for data altruism services

A **compulsory authorisation framework for data altruism** schemes could ensure generalised trust in data altruism within society. However, the public sector would incur costs for creating a national authorisation scheme as part of the one-stop shops.

All organisations (including those that collect data for their own use as well as those that purely serve as intermediaries) would need to cover costs due to the mandatory authorisation. As authorisation would be done by public sector bodies, costs could be subsidised for certain entities depending on size, for example for an SME (if for-profit approaches are also allowed) or NGO. In addition, NGOs could receive a discount of approximately 10% for authorisation. Alternatively, authorisation could be acquired without any fees, for free.

At the same time, the benefits would be considerably higher than for policy option 2. Thanks to increased trustworthiness, security and awareness of the data altruism schemes, SMEs, NGOs and citizens would be more likely to be willing to share data. It is estimated that by 2028, there would be more than 7 million citizens and more than 700 companies taking part in data altruism, 'donating' their data. The authorisation and the trust it brings (regarding their compliant and trustworthy data handling and processing) would relieve citizens and companies from the burden of verifying the legitimacy of the operations and purposes of the organisations, and it would also bring a considerable benefit for these users in the form of time and effort savings. Under this policy option, benefits include an easy and transparent way to access data from various fields, contributing to research and development as well as improved decision-making[143]. As for policy option 2, the figure in the table below only includes to a limited extent the downstream societal benefits due to the lack of quantifiable data.

Compulsory authorisation framework for data altruism services: Costs and benefits[144]	
Costs	Benefits
€3 800-10 500 for SMEs and €3 420-9 450 for NGOs.	€300 million in the period 2024-2028.
€5 000/ year for the maintenance of the authorisation.	

[143] Idem.
[144] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.

European Data Innovation Board as a self-standing entity

This option would help achieve the objective of increasing data sharing by facilitating the development of relevant cross-industry standards. However, as for policy option 2, its success depends on the ability to ensure the participation of companies in defining and adopting standards.

The cost of setting up and running an independent body are higher than that of a formal expert group. The budget of comparable bodies (such as the European Data Protection Board) amount to EUR 3.5 million per year, which is more than 10 times higher than that of a formal expert group. Other costs are highly variable and are difficult to estimate, such as the costs of documentation and education, guidelines, toolkits, tutorials or webinars.

Under this scenario, the benefits to traditional businesses arise from an increased adoption of standards by the standardisation organisations, and the resulting reduction in costs for acquiring, integrating and processing data. This figure in the table below is calculated on the assumption that through interoperability made possible by standardisation, 900 companies would save 15% of EUR 50 million operational costs over 5 years. This underlines the importance of interoperability. The small difference in benefits between the lower and the higher intensity regulatory intervention is explained by the marginal increase in the number of companies taking up standards: it is estimated that 800 and 900 companies would be concerned respectively for the lower and higher intensity options, as the existence of a more formal and stronger body would only marginally increase the uptake of standards by industry. The form of the mechanism enhancing standardisation would not have a decisive impact on the benefits as long as there is such a mechanism in place with the necessary industry representation.

European Data Innovation Board as a self-standing entity: Costs and benefits[145]	
Costs	Benefits
€3.5 million/ year for setting up and running costs.	€1.35 billion for traditional businesses in the year 2028.

6.2. Social and environmental impact

The study team contracted to carry out the impact assessment support study was unable to quantify the environmental and social benefits of the different policy options due to the lack of available data. However, based on their research and interviews with stakeholders, they provided a qualitative assessment of the likely impact of the different options.

6.2.1. Baseline scenario

The baseline scenario will see a slower realisation of the potential benefits of data. In the absence of coordinated EU action for **the reuse of data subject to rights of others** in the Member States and **data altruism mechanisms**, the societal and environmental benefits would be limited. Thus,

[145] Idem.

the potential value of data altruism in the EU, in particular for scientific research and improved public policy and services, would not be unlocked [146].

6.2.2. Coordination at EU level and soft regulatory measures

As mentioned above, the impact of this option is contingent upon uptake by Member States.

If all Member States decide to set up structures to facilitate the **reuse of publicly held data subject to the rights of others**, environmental and social benefits could be similar to those under policy options 2 and 3. However, this is unlikely to materialise.

The certification of intermediaries via an **industry-driven self-regulatory certification framework** may incentivise individuals to share their personal data. However, as explained above, this option is expected to have little added value as compared to the baseline.

Indirectly, the creation of an **informal Expert Group** could lead to new discoveries for health, environmental efficiency and other new products.

6.2.3. Policy option 2: Lower intensity legislation

The creation of measures to facilitate the **reuse of publicly held data subject to the rights of others** would result in positive social and environmental impacts due to the increased availability and reuse of such data.

The societal benefits of setting up a **voluntary certification/labelling framework** would be twofold: on the one hand, society would benefit as the potential of the European data market would be unlocked through certification, while on the other hand data flows through intermediaries serving societal purposes (i.e. health, research) would increase.

This policy option would lead to positive societal benefits from **data altruism mechanisms**, from personalised medicine and treatment to finding new forms of renewable energy. The availability of data would allow researchers to gather the necessary data at the necessary scale for insights and conclusions to be representative and solid. Data altruism would also help public authorities in taking evidence-based decisions as well as improving the efficiency of their public services thanks to representative insights from individuals. Another important societal benefit is that individuals would have more opportunities to make their data available for the common good, and would be confident that reuse of the data takes place in line with EU data protection legislation. Indeed, at a more general level, the proposed measures will contribute to generate trust in data sharing, and ensure that European companies and citizens are in control of the data they generate.

Beyond the direct impacts of data sharing, these measures would indirectly benefit both society and the environment. The creation of new products and services based on data would lead to, for example, better healthcare and mobility, as well as energy savings. At the same time, more data use would lead to more energy consumption, which underlines the importance of making data

[146] European Commission (2020). *Support Study to this Impact Assessment*, SMART 2019/0024, prepared by Deloitte.

processing and data centres more energy efficient, as indicated in the European Digital Strategy[147].

6.2.4. Policy option 3: Higher intensity legislation

The environmental and societal benefits of a **single data authorisation body** would mirror those presented in policy option 2, as would the establishment of a **compulsory certification/labelling framework**.

This policy option would also create similar societal benefits to policy option 2 for **data altruism**. However, due to the trust in public authorisation schemes, it would increase trust and security in data altruism schemes, which may lead to more individuals making their data available for the common good.

Finally, increased data sharing would also lead to the abovementioned indirect benefits to society and the environment.

6.3. **Impact on SMEs**

This initiative would have an impact on SMEs, both in their capacity as data intermediaries as well as data reusers. In general, more data availability through an increase of trust in data sharing will benefit SMEs proportionally more than large organisations, as it is critical to their survival.

However, as a Commission consultation shows, 40% of SMEs[148] struggle to access the data they need to develop data-driven products and services, because of a lack of financial resources and because they do not have the power to negotiate with data holders. In the absence of EU action, SMEs would continue to suffer from the imbalance between them and large reusers that have the resources to reuse data and adapt to change. As such, SMEs are the main beneficiaries of the proposed instrument.

Many data intermediaries are SMEs. The instrument would give a boost to such SMEs and startups in the data economy. The one-off costs for certification/labelling (EUR 20 000-50 000 for a voluntary label, and EUR 35 000-75 000 for a compulsory certification) for data intermediaries and renewal costs would be countered by the high gains in both client base and revenue (25-50% increase), as well as by a higher possibility to attract investors[149].

Conversely, a more level playing field and reduced legal uncertainty created through EU action would allow SMEs and startups to flourish in the EU data market. Taking into account the associated costs, the majority of stakeholders consulted (including SMEs)[150] agreed on the perceived benefits of such scheme (especially if certification/labelling is voluntary).

---

[147] European Commission (2020). *The European Digital Strategy*.
[148] European Commission (2019a). *SME panel consultation B2B data sharing - Final Report*.
[149] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.
[150] European Commission (2020c). *Report of the workshop on labels for/certification of providers of technical solutions for data exchange*; European Commission (2020). *Support Study to this Impact Assessment*, SMART 2019/0024, prepared by Deloitte.

As regards the **reuse of public data subject to the rights of others**, businesses have to navigate through the same challenges related to finding and accessing datasets as researchers. SMEs, many of which have a business model that is based on the use of public sector data, do not always dispose of the resources and awareness needed to face these challenges, resulting in an unequal access to data that is subject to the rights of others and therefore reduced innovation and business opportunities. This impact is cumulative, since in effect larger companies are in a better position than small ones to innovate and to develop new products and services.

During a workshop organised in the context of the support study, stakeholders indicated that an **industry driven self-regulatory certification framework** could give big industry players a stronger role, which would potentially influence the outcome of the discussions taking place in the stakeholder forum[151]. However, at the same time, the market is not mature enough for a compulsory certification scheme, as it would likely prevent many new businesses from entering the market.

In conclusion, SMEs would benefit both as data intermediaries and as data users. As data intermediaries, they would primarily benefit from the voluntary labelling scheme. As the scheme would be voluntary, it would not pose a general market barrier. Such a voluntary framework is specifically supported by SMEs, as evidenced in the report on SME Panel Consultation[152] (end 2018-early 2019), as well as the workshop on certification/labelling in May 2020[153]. As data users, SMEs would benefit from the easier availability of more data (public, personal and non-personal). The importance of the economies of scale of data sourcing and processing would not diminish – but with easier access, and by facilitating the balance of supply and demand for such data and the value derived from it (e.g. by supporting data marketplaces), smaller companies would have better access to the value of big data.

6.4. Member States' and stakeholders' views

As described in Annex 2, the consultation process sought to collect the views of EU Member States and stakeholders by means of an online consultation and several workshops and meetings.

The analysis contributed to the assessment and the choice of the preferred option on the basis of the four different intervention areas. The consultation actions tried to reach out to various stakeholders from the public and the private sectors and citizens.

The public online consultation was the main consultation action targeting the **citizens**, and in total 201 citizens took part, all from the EU. Most of the views were aligned with those of the other stakeholders in general. They expressed strong support for the overall strategy on data and the development of common European data spaces. Their positive assessment of the initiatives on the reuse of data subject to the rights of others for research and innovation purposes, as well as on data altruism, was very strong, especially for the purpose of health-related research, as well

[151] European Commission (2020). *Support Study to this Impact Assessment, SMART 2019/0024*, prepared by Deloitte.
[152] European Commission (2019a). *SME panel consultation B2B data sharing - Final Report*.
[153] European Commission (2020c). *Report of the workshop on labels for/certification of providers of technical solutions for data exchange*; European Commission (2020, *forthcoming*).

as aspects relating to the city/municipality/region of the individuals (including mobility, environment).

These findings should be qualified by the fact that individuals would be willing to share their data for research and for the common good if the right privacy-preserving and secure conditions are put in place. This finding is in line with the result of the 2017 public online consultation[154].

Member States were consulted both as policymakers and as data users. Regarding **public authorities of the Member States**, a workshop on a common European data space for the public services took place on 10 September 2019. The discussions showed that stakeholders welcome the funding of actions that would facilitate the participation of the public sector in a common European data space, cater for many flows (G2G, B2G, C2G, as well as G2B, and G2C) and span a spectrum from bulk data transfer, to moving algorithms all the way to the Once-Only-Principle. They also concluded that a data space that would allow a one-stop shop for 'data use permits' across the public sector, could put an end to data duplication and facilitate reuse of data by others, including the public administration itself.

In the online consultation, public authorities appeared as strong supporters of developing governance mechanisms supporting standardisation activities for interoperability purposes (especially application programming interfaces (APIs) and metadata). Even more than the rest of stakeholders, they considered that EU or national government bodies have a role to play in prioritisation and coordination of standardisation, and in the clarification of the legal rules. Finally, they appeared as strong supporters of public authorities making a broader range of data that is subject to the rights of others available for R&I purposes and for the public interest.

Industry organisations, including SMEs and business associations, agreed that the European Union needs an overarching data strategy to enable the digital transformation of the society (99% of business respondents to the 2020 online consultation). However, they are particularly affected by the problem of accessing data, highlighting technical problems (interoperability and transfer mechanisms) or simply denied access.

During the workshops conducted in 2019 on common European data spaces, feedback from representatives of the private sector showed the sectors have different levels of maturity and needs, but that there is in general a need for ensuring fair competition on data markets. In the data economy in general, one can observe big companies keeping control over large quantities of data. In the workshops, companies confirmed common data spaces should allow more data to be shared with all types of European actors (including SMEs) and across sectors, allowing new market dynamics to be created. The idea of a voluntary certification scheme for data intermediaries was supported by SMEs during the course of the workshop on certification/labelling held in May 2020[155].

---

[154] European Commission (2018b). *Synopsis report of the public consultation on Digital transformation of health and care in the context of the Digital Single Market*.

[155] European Commission (2020c). *Report of the workshop on labels for/certification of providers of technical solutions for data exchange.*

Academic and research institutions would benefit directly from the decisions on secondary use of data and data altruism, as they could considerably lower their compliance costs related to using data. Unsurprisingly, this stakeholder category agree with facilitating the reuse of data subject to the rights of others for research and innovation purposes, and support the data altruism concept. Academic and research institutions see potential for the use of such data in areas that are similar to those where citizens think their 'donated' data could be useful: health-related research and for aspects relating to the city/municipality/region of the individuals (including mobility, environment).

The European Data Protection Supervisor (EDPS)

On 16 June 2020 the European Data Protection Supervisor adopted Opinion 3/2020[156] on the European strategy for data. The approach of the EDPS towards the strategy in general is positive, considering that the implementation of the strategy will be an opportunity to set an example for an alternative data economy model (see complete reference in Chapter 1).

The opinion raises several practical issues that should be taken into account when moving forward with a possible legislative framework. For example, the EDPS considers that companies participating in data spaces should be subject to a 'vetting' process and data traceability tools and obligations could facilitate the role of data controllers when personal data are processed with a space. The EDPS also considers that the notion of data altruism and its interplay with the GDPR should be clearly defined (since it depends on the consent of the data subject and the portability right under article 20 GDPR). It underlines that exceptions for research on personal data cannot lead to a broad exemption of the scientific sector from GDPR obligations. GDPR obligations.

Inception Impact Assessment

Stakeholders also provided feedback to the Inception Impact Assessment (published on the Better Regulation Portal between 3 and 31 July 2020). The contributions reflected the replies to the online questionnaire, as well as the papers. The feedback dealt with all aspects and measures foreseen in the initiative. Most of the contributions expressed support to the initiative and contained general comments, underlying the importance of fair, transparent and non-discriminatory access to data, of voluntary data sharing (from private entities but also from individuals) and of standards and interoperability.

7. HOW DO THE OPTIONS COMPARE?

In line with the European Commission's Better Regulation Guidelines[157] and its toolbox[158], most importantly tool 63, the Impact Assessment study carried out a multi-criteria analysis (MCA)[159]

[156] EDPS (2020). *Opinion 03/2020 on the European strategy for data*.
[157] SWD/2017/350.
[158] Idem.
[159] European Commission (2020). *Support Study to this Impact Assessment*, SMART 2019/0024, prepared by Deloitte.

in order to take full account of the complexity of the subject matter and the level of granularity of the analyses carried out.

As mentioned in section 6.1.2, the option of soft measures only was not analysed further as part of the multi-criteria analysis, given that such measures would not provide for a uniform structural enabling framework that is essential to achieving both the general and the specific objectives in a timely manner. Therefore, the table below contains an analysis of the lower and higher intensity regulatory interventions.

	Regulatory intervention with low intensity	**Regulatory intervention with high intensity**
Efficiency	This option presents a favourable ratio of costs and benefits where, in most of the areas, the benefits generated will likely significantly outweigh the costs. It is expected to generate slightly lower direct and indirect economic benefits than the higher intensity option (increase from 3.87% to between 3.92% and 3.94% of GDP by 2028). For the enhanced reuse of public sector data, the lower intensity option would cause a one-off cost of EUR 10.6 million for the establishment of the mechanisms, with an annual maintenance cost of EUR 600 000 per year. This would be contrasted by the EUR 41.8 million as direct benefits, the EUR 684 million per year benefit of cost savings and the benefits to reusers in the amount of EUR 49.2 million/year. A voluntary certification/labelling framework for data intermediaries would cost around EUR 20-50 000, with an annual maintenance cost of EUR 20-35 000. The benefits would materialise in the form of a 25%-50% expected increase in revenues and client base and up to 50% business development time acceleration. Voluntary certification of data altruism schemes would cost approximately EUR 20-50 000, with the recurrent costs of between EUR 20-35 000 for maintaining it. However, benefits would only be around EUR 22 million. In the case of the European Data Innovation Board, the costs for the set-up and operations of a formal expert group are limited (around EUR 280 000 per year) while the benefits (EUR 1.2 billion) are similar to those expected under the higher regulatory intervention.	The option also presents a favourable ratio of costs and benefits. It is expected to generate higher costs than the lower regulatory interventions, but also greater overall direct and indirect economic benefits (increase from 3.87% to between 3.93% and 3.97% of GDP by 2028). Setting up a central data authorisation body to enhance the reuse of public data would produce EUR 21.2 million establishment and EUR 12.2 million maintenance costs, and benefits of EUR 1 253.4 million per year in cost savings and EUR 212.7 million per year in the form of revenues from application fees. This would make this option less efficient. A compulsory certification framework would generate a EUR 35-75 000 one-off cost, with a yearly EUR 20-50 000 as a recurrent cost. With the benefits very similar to the lower intensity regulatory option (25%-50% expected increase in revenues and client base and up to 50% business development time acceleration), this would be a less efficient option. Compared to the lower intensity option, a compulsory authorisation for data altruism schemes, would produce a one-off cost of a range between EUR 3 420 – 10 500 to obtain the authorisation, with an annual recurrent maintenance cost of EUR 5 000. At the same time, the increase in altruistic data sharing would create much higher direct benefits, in the amount of EUR 300 million, making it a more efficient option. For the European Data Innovation Board, the costs of setting up a self-standing body would be much higher (EUR 3.5 million) than a Commission expert group, with marginally higher benefits (EUR 1.35 billion).
Effectiveness	This option could significantly contribute to the general objective of leveraging the	This option is also expected to address the need to leverage the potential of data for the EU

	potential of data for the EU economy and society as well as the three specific objectives of reinforcing trust in common European data spaces, making more data available through technical, legal and organization support as well as overcoming technical obstacles (e.g. interoperability) across sectors. This policy option would also further contribute to setting the foundations of a Single Market for Data and strengthening the EU data economy, since the European data market overall will be significantly boosted through the voluntary certification/ labelling of data intermediaries, which will increase the volume of data flows. Similarly, the new opportunity to access and reuse non-open public data as well as the new data sharing schemas appearing due to the voluntary certification would also contribute to data access and use across the EU. The effectiveness of an expert group to facilitate the development and adoption of standards would be limited, given the key role of industry's willingness to take up such standards.	economy and society as well as the three specific objectives of reinforcing trust in common European data spaces, making more data available through technical, legal and organization support as well as overcoming technical obstacles (e.g. interoperability) across sectors. The creation of a single data authorisation body allows the centralisation of reuse requests that might contribute to further effectiveness. However, concerns have been raised with regard to the ability of a compulsory certification scheme for intermediaries to effectively build common data spaces, as the higher certification costs and the compulsory nature might prevent smaller industry players from getting into the market. On the other hand, it would establish clear rules for how data intermediaries are supposed to act in the European data market. Regarding the compulsory authorisation of data altruism schemes, it is expected to be more effective than the lower intensity voluntary certification, given that the trust generated by it would incentivise more citizens and businesses to altruistically share their data. The effectiveness of a self-standing body to facilitate the development and adoption of standards would be limited as well, given the key role of industry's willingness to take up such standards.
Coherence	This option is in line with the EU data strategy's objective of creating common data spaces as well as other horizontal and sectoral legislation currently in effect. It does not create coherence issues with major EU law. The sharing of "sensitive" data held by the public sector or personal data shared under a data altruism scheme can be done in line with GDPR requirements. This option has the potential to minimise friction with national law compared to the higher intensity intervention with regard to the possible flexibility in the set-up of structures and mechanisms to share "sensitive" public data. However, more flexibility would result in a lower level of harmonization of the horizontal governance.	This option is also in line with the EU Digital strategy of creating common data spaces and the horizontal and sectoral legislation currently in effect. This option does not create coherence issues with major EU law. The sharing of "sensitive" data held by the public sectors or personal data shared under a data altruism scheme can be done in line with GDPR requirements. The far-reaching horizontal measure proposed in the higher intensity regulatory intervention could be difficult to reconcile with some national laws that can limit the reuse of (sensitive) data held by the public sector for strictly non-commercial purposes. However, it would also provide for a higher level of harmonization and a more seamless single market for data.
Legal/political feasibility	This option is both politically and legally feasible. The lower intensity regulatory intervention presents a clear advantage over the higher intensity regulatory intervention as concerns the possible flexibility in the setting up of structures and mechanisms for	This option is legally feasible, although for the public data creating one single data authorisation body could imply considerable legal or organisational challenges. Concerns might be raised by stakeholders regarding the compulsory nature of certification for

	the reuse of public sector data, as well as the voluntariness of the certification/ labelling mechanisms for data intermediaries and data altruism schemes. The setting up of a Commission expert group presents no political obstacle, given that it is established by a Commission decision. Industry representatives also welcome the setting-up of a formal expert group that supports the coordination of standardisation rather than a stronger role from the EU under the high regulatory intervention.	intermediaries, as it would prescribe requirements for market entry. Setting up a self-standing European central body that would lead efforts in the domain of standardisation would also be less politically feasible due to the high costs it would attain.
Proportionality	This option is proportionate and is limited to the stakes at hand. It presents a balanced yet focused policy intervention. Allowing flexibility both regarding the structures and mechanisms for the reuse of public data and the voluntariness of the certification/ labelling of the data intermediaries as well as the data altruism schemes ensures that the proposed measures would not have impacts beyond what the initiative aims to achieve.	This option, although stronger, is still proportionate to the stakes at hand. Neither of the planned measures in the intervention areas would go beyond what is necessary to achieve the objectives of the initiative. Given that the measures proposed under this option would only represent a more centralized or binding variant of the measures under the lower intensity regulatory intervention, the difference in proportionality between the two options is insignificant.

Source: European Commission, based on the support study SMART 2019/0024

	Efficiency	Effectiveness	Coherence	Legal/political feasibility	Proportionality
Regulatory intervention with low intensity	++	++	+	+	+
Regulatory intervention with high intensity	+	++	+/-	+/-	+

Source: European Commission, based on the support study SMART 2019/0024

For efficiency, effectiveness and coherence, the scores are given on the expected magnitude of impact as explained above: ++ being strongly positive, + positive, and – negative. For legal/political feasibility and proportionality, + means that the assessment is positive, and – means that it is negative.

8. PREFERRED OPTION

Based on the evidence presented above, a mixed package of lower and higher intensity regulatory interventions is the preferred option. Although, based on the cost-benefit and multi-criteria analyses, for three of the intervention areas the lower intensity option is more favourable, the higher intensity intervention for data altruism would yield higher economic and societal benefits, while incurring fewer costs.

For the **enhanced reuse of public sector data**, the lower intensity regulatory option would cause a one-off cost of EUR 10.6 million for the establishment of the mechanisms and the single entry point, with an annual maintenance cost of EUR 600 000 per year for Member States. This would

be contrasted by the EUR 41.8 million as direct benefits, the EUR 684 million per year benefit of cost savings and the benefits to reusers in the amount of EUR 49.2 million/year. This option would be more favourable than the higher intensity regulatory intervention, which would produce EUR 21.2 million establishment and EUR 12.2 million maintenance costs for Member States, which would be partially counterbalanced by the EUR 1 253.4 million per year in cost savings and EUR 212.7 million per year in the form of revenues from application fees[160].

For **increasing trust in data intermediaries**, in a voluntary certification/labelling framework would cost around EUR 20 000-50 000 to obtain the certificate/label, with an annual maintenance cost of EUR 20 000-35 000. The benefits would materialise in the form of a 25%-50% expected increase in revenues and client base and up to 50% business development time acceleration. Compared to this, a compulsory certification framework would generate an amount of EUR 35 000-75 000 one-off cost for obtaining the certificate, with a yearly EUR 20 000-50 000 as a recurrent cost for maintaining it. With the benefits very similar to the lower intensity regulatory option (25%-50% expected increase in revenues and client base and up to 50% business development time acceleration), the more favourable option would be option 2[161]. However, this option could be considered as an alternative given its structuring function for the European market for data intermediaries, which would lead to higher trust in these intermediaries.

To obtain a certificate under a voluntary **certification mechanism for data altruism services** would cost approximately EUR 20 000-50 000, with the recurrent costs of between EUR 20 000-35 000 for maintaining it. However, benefits would only be around EUR 22 million. A compulsory authorisation would produce a one-off cost ranging between EUR 3 420 – 10 500 to obtain the authorisation if the public sector decides to apply fees, with an annual recurrent maintenance cost of EUR 5 000. Higher trust is expected to lead to an increase in altruistic data sharing. This would create much higher benefits in the order of EUR 300 million, making it a more favourable option[162].

The creation of the **European Data Innovation Board** in the form of a formal Commission expert group would trigger a yearly cost of EUR 280 000, while yielding around EUR 1 billion in benefits through standardisation. In contrast, the set-up of an independent body would cost EUR 3.5 million, more than ten times the amount for an expert group, while at the same time, benefits would remain around 1.2 billion for the year 2028[163]. Thus, the less costly expert group would achieve the same goals, with more efficiency.

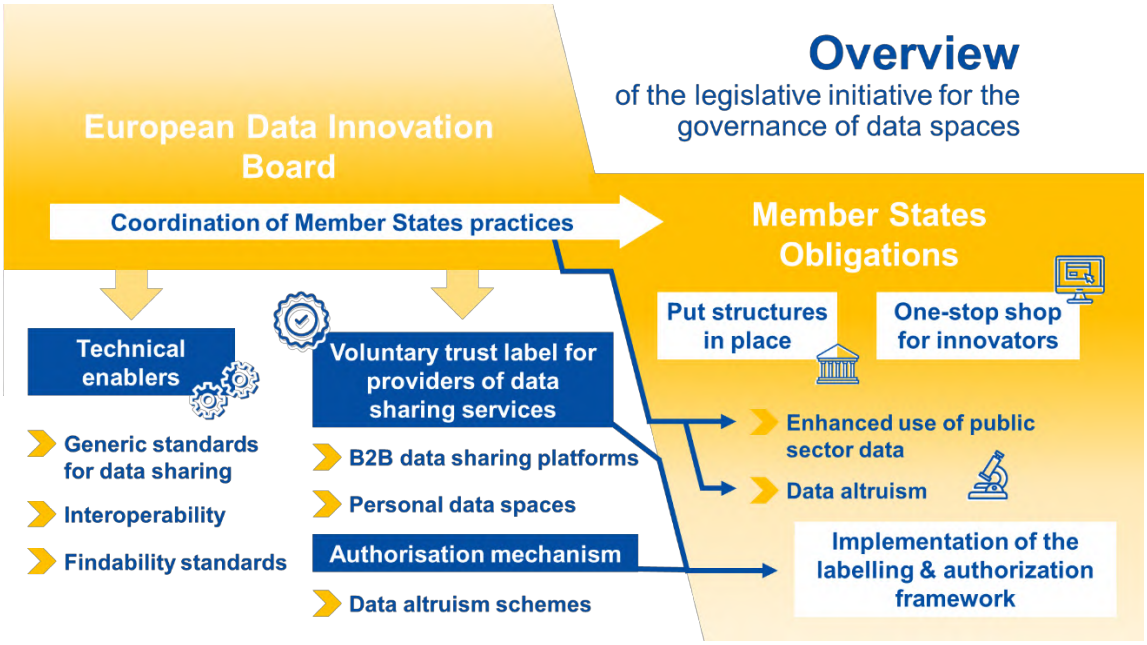
Packaging the lower intensity options together with the higher intensity regulatory option for data-altruism allows for a targeted and proportional intervention, taking into account the different impacts of the individual policy options on the intervention areas, which will lead to a significant

[160] European Commission (2020). *Support Study to this Impact Assessment*, SMART 2019/0024, prepared by Deloitte.
[161] Idem.
[162] European Commission (2020). *Support Study to this Impact Assessment*, SMART 2019/0024, prepared by Deloitte.
[163] Idem.

improvement over the baseline scenario. It is broadly acceptable to stakeholders and can be realistically enacted within a reasonable timeframe, which is critical given the expected value of the initiative in post COVID-19 recovery programmes.

This leads to a preferred option that is based on the following elements (schematically captured in the image at the end of this section):

- **Mechanisms for enhanced reuse of certain public sector data:** a lower intensity regulatory intervention would prescribe for Member States to provide services to facilitate the reuse of publicly held data that is subject to the rights of others in accordance with a set of conditions, without determining the exact institutional and administrative form.
- **Certification/labelling framework for data intermediaries:** the lower intensity regulatory intervention would create a voluntary certification/labelling mechanism, where the designated authorities would handle the application process and award the labels/certificates to the compliant data intermediaries.
- **Measures facilitating data altruism:** the higher intensity regulatory intervention would provide for a compulsory European authorisation scheme as a requirement to offering services facilitating data altruism. The processing and issuing of authorisations would be handled by designated authorities.
- **European Data Innovation Board:** as part of a lower intensity regulatory intervention, it would function as a formal expert group, with a secretariat provided by the Commission. Its functions would include facilitating standardisation and the enhancement of interoperability, and the facilitation of the exchange of national practices.



Source: European Commission

8.1. Estimated impact of the preferred option

The Impact Assessment support study[164] indicates that, while in the baseline scenario the data economy and the economic value of data sharing are expected to grow to an estimated EUR 533.5 billion (3.87% of the GDP) by 2028, this would increase to between EUR 540.7 and EUR 544.4 billion (3.92% to 3.95% of the GDP) under the preferred option.

At the same time, this policy option would make it possible to create an alternative European model for data sharing to the current business model for Big Tech platforms, through the emergence of neutral data intermediaries. This initiative can make the difference for the data economy by creating trust in data sharing as a precondition for the development of common European data spaces, where individuals and companies are in control of the data they generate, and are comfortable with the way in which the data are used in innovative ways.

Indeed, as indicated in section 6.1, the actual impact of this initiative is likely to be far greater than the benefits that can be directly attributed to its different elements. By increasing trust in data sharing, the initiative would function as a catalyst for the data economy. It would facilitate data sharing across the EU, unleashing the power of data-based innovation and supporting the creation of new services and products and more efficiency in industry. It would also contribute to new tools for tackling societal challenges, such as climate change, and to better policymaking.

9. HOW WILL ACTUAL IMPACTS BE MONITORED AND EVALUATED?

Due to the dynamic nature of the data economy, monitoring the evolution of impacts constitutes a key part of the intervention. To ensure that the selected policy measures actually deliver the intended results and to inform possible future revisions, the Commission would set up the monitoring and evaluation process described below.

The European Data Innovation Board would bring together evidence about the situation in the Member States and in the different sectors. It would compile best practice examples based on feedback from Member States on their implementation measures, and the relative strengths and weaknesses of these measures. Member States would be asked to report regularly on the efficiency and impact of the different strands of action in their data market.

This would help the Commission to closely monitor the uptake of the measures in Member States and amongst stakeholders, also in view of compliance. If necessary, the Commission would launch infringement procedures.

Through the Support Centre for Data Sharing, which is planned to be established under the DEP, evidence from stakeholders will be gathered on the market efficiency and effectiveness of measures taken under this initiative to enhance the reuse of public sector data, data altruism and a labelling scheme for data intermediaries.

The monitoring is divided into **two operational parts**: monitoring of the **specific objectives** identified in the section 4.2 and monitoring of the individual components that constitute the

[164] European Commission (2020). *Support Study to this Impact Assessment*, SMART 2019/0024, prepared by Deloitte.

preferred policy option described the Chapter 8. For both parts, the tables below present **operational objectives** corresponding to the identified specific policy objectives/preferred option, **indicators** that would be used to monitor progress as well as **sources of information**.

9.1. Monitoring of the specific objectives

Specific objectives	Operational objectives	Indicators	Sources of information
Reinforcing trust in data sharing	Trust in data sharing increases as clear rules are available for data exchanged or pooled by data holders to be secure and processed in compliance with applicable legislation as well as with the conditions they set on use of such data.	Increase in the level of trust in data sharing reported by data users and suppliers.	Representative survey among stakeholders carried out by the Support Centre for Data Sharing under DEP and an evaluation study to support the review of the instrument within 3 years of its adoption.
Making more data available for reuse within the common European data spaces	More data are made available for reuse on voluntary grounds based on the existing legislation and where data holders agree to this.	Volume of data processed in secure data processing environments, data collected by authorised data altruism mechanisms and data shared among business partners and/or contributed to data pools.	Records of the European Data Innovation Board and the Support Centre for Data Sharing under DEP on reuse of such data reported by the dedicated national authorities. Representative survey among stakeholders carried out by the Support Centre for Data Sharing under DEP and an evaluation study to support the review of the instrument 4 years after its date of application.
Ensuring interoperability across sectors and countries	Interoperability and generic standards contribute to reduction of transaction costs and allow data to be reused across sectors and Member States.	Decrease in the share of stakeholders that have encountered difficulties in using data from other organisations.	Evaluation study to support the review of the instrument 4 years after its date of application.

Source: European Commission (also of the table just below)

9.2. Monitoring of the preferred option

Area	Operational objectives	Indicators	Sources of information
Mechanisms for enhanced reuse of certain public sector data	Reusability of publicly held data subject to the rights of others is ensured by Member States having secure data processing environments in place. Findability of such data is increased as national single entry points for data reusers to contact the public sector are available.	Number of data reuse permits to process data in secure data processing environments issued, processing sessions carried out.	Data on reusability of such data reported by the dedicated national authorities to the European Data Innovation Board, analysed by the Support Centre for Data Sharing under DEP.
Certification/	Novel types of data intermediaries	Number of	Data on the certification/labelling

labelling framework for data intermediaries	are able to scale up at the sufficient speed to provide a viable alternative to the platform model.	organisations awarded a trust label.	framework reported by the dedicated national authorities to the European Data Innovation Board, analysed by the Support Centre for Data Sharing under DEP.
Measures facilitating data altruism	Companies and individuals are able to make their data available securely for the wider common good through trusted data altruism mechanisms.	Volume of data contributed through the authorised data altruism mechanisms.	Data on data altruism reported by the dedicated national authorities to the European Data Innovation Board analysed by the Support Centre for Data Sharing under DEP.
European Data Innovation Board	The European Data Innovation Board ensures effective coordination of the labelling and the authorisation scheme for data intermediaries and data altruism mechanisms; prioritisation of standards for cross-sector data reuse; and maintains the European data-sharing schema to support peer-to-peer data sharing without an intermediary.	Assessment of the support received by the dedicated national authorities. Number of contributions to the Rolling Plan for ICT standardisation. Number of functioning peer-to-peer data sharing schemes in place.	Survey among the dedicated national authorities. Records of the Data Innovation Board on prioritisation of standards and publication of technical guidance for interoperability/peer-to-peer data sharing schemes. Evaluation study to support the review of the instrument 4 years after its application.

Glossary

Term or acronym	*Meaning or definition*
Common European Data Space	An arrangement composed of an IT environment for secure processing of data by an open and unlimited number of organisations, and a set of legislative, administrative and contractual rules that determine the rights of access to and processing of data.
Data altruism	The act of granting access to and sharing of data held by individuals or companies, without seeking direct reward, for the common good.
Data-driven innovation	The use of data and analytics to improve or create new products, services, markets and organisational methods.
Data intermediary	An entity (of either the public or the private sector) that facilitates data sharing, access and use by data holders and data users.
Data portability	Capacity to transfer data to which an individual or entity has a specific relationship from one IT environment (or similar) to another, based on legislative rights (e.g. Article 20 of the GDPR) or contractual agreement.
Data sharing	An act of the data holder, data producer, or data intermediary providing access to a data user for the purpose of joint or individual use of the data, based on voluntary, commercial or non-commercial agreements, or mandatory rules. It should not be understood as making data available for free and to an undefined group of users.
Data the use of which is conditional on respecting the rights of others	Data that might be subject to data protection legislation, intellectual property and main contain trade secrets or other commercially sensitive information.
Internet of Things (IoT)	A network of physical devices, vehicles, home appliances and other items embedded with connectivity software, which enables these objects to connect and exchange data.
Secondary use or reuse	The use by persons or legal entities of documents held by public sector bodies, for commercial or non-commercial purposes other than the initial purpose within the public task for which the documents were produced.

ANNEX 1: PROCEDURAL INFORMATION

1. LEAD DG, DECIDE PLANNING/CWP REFERENCES

The legislative proposal on the governance of common data spaces was prepared under the lead of the Directorate-General Communication Networks, Content and Technology. In the DECIDE Planning of the European Commission, the process is referred to under item PLAN/2020/7446. The Commission Work Programme for 2020 includes a legislative action on data, under the header "10. A European approach to AI".

2. ORGANISATION AND TIMING

An Inter-Service Steering Group (ISSG) assisted DG Communication Networks, Content and Technology in the preparation of the Impact Assessment and legal proposal. It included Commission services of 18 Directorate-Generals, together with the Commission's Legal Service and Secretariat General.

Work for the preparation of this initiative started with the design of the European Strategy on data, adopted in February 2020, which announced measures for a cross-sectoral governance framework for data access and use. Discussions were initiated during the Inter-Service Consultation in view of the strategy (January 2020). Subsequently, the ISSG contributed to the initiative preparation in March 2020 (discussion on the consultation strategy and the Inception Impact Assessment), and in July 2020 (discussion on the draft Impact Assessment).

An Inception Impact Assessment was published on 3 July 2020 and was open to feedback from all stakeholders on the Better Regulation Portal for a period of 4 weeks.

The draft Impact Assessment report and all supporting documents were submitted to the Regulatory Scrutiny Board (RSB) on 20 July, in view of a hearing on 9 September 2020. After a negative opinion, the report got improvements, mainly through the strengthening of the narrative and the clarification of the problem definition and expected impacts. The second opinion delivered by the Board on 5 October 2020 was positive with reservations. The report was further improved on the basis of the comments provided.

An Inter-Service Consultation took place, with all services that are members of the inter-service group on data, and closed on 28 October 2020.

3. CONSULTATION OF THE RSB

The Impact Assessment report was reviewed by the Regulatory Scrutiny Board on 9 September 2020. Based on the Board's recommendations[165], the Impact Assessment has been revised in accordance with the following points:

[165] *url to be added when created*

Comments of the RSB	*How and where comments have been addressed*
(B) Summary of findings	
(1) The report does not explain the problem clearly enough and why the EU should promote a new model for data sharing.	Chapter 2 has been substantially reworked to better explain the problem, the problem drivers and their interrelation. The key element of **trust** has been made much more prominent as a separate problem driver in section 2.2.1, and has been disentangled from the more technical issues. The explanation on why there is a need for a new European model for data sharing has been reinforced by an analysis of the role of Big Tech platforms in this area and of the lack of trust in data-sharing solutions they may provide (sections 2.1 and 2.3). The report now also explains better why a model of neutral data intermediaries is preferable to the current model of Big Tech platforms (section 2.1) in terms of fostering trust.
(2) The report does not elaborate in sufficient detail the design and composition of the options and how they would work in practice.	The description of the policy options has been further detailed (section 5.2). This section now also explains the reasoning behind the design of the options and, where relevant, the composition of the options is explicit. The description of how each option would work in practice for the different intervention areas has been fine-tuned (section 5.2). This is also reflected in Chapter 6. An analysis of the soft law option has been included, also covering each of the different intervention areas (section 6.1.2).
(3) The scale of the quantified direct impacts is not in line with the impacts presented in the text.	The report now also describes the indirect impact that the initiative could have as a catalyst of seamless cross-border cross-sector data sharing, which would result in wider economic and societal benefits (section 6.1). These benefits are expected to be substantially higher than the direct impact on the data economy.
(4) The analysis is not sufficiently granular to underpin the choice of the preferred option.	Chapters 6 and 7 now go into greater analytical depth on all intervention areas, thus better underpinning the choice of the preferred mixed package in Chapter 8.
(C) What to improve	
(1) The report should better describe the current situation on data sharing in Europe. It should explain why it does not examine	The description of the current situation on data sharing has been improved by adding information and reorganising the 'Problem definition' in

the creation of data markets. It should analyse drawbacks and risks stemming from the current role of data intermediaries. It needs to provide more evidence on the insufficiency of the existing arrangements, for example regarding findability, quality and neutrality of data. The report should inform on the current tendencies of concentration of data supply by intermediaries. It should expand on the problems arising from access to data being concentrated outside the EU. The report should elaborate on the problems that emerging European data sharing initiatives are facing and their internal market dimension. The report should detail the governance problems of data intermediation.	Chapter 2, in particular the problem drivers. Section 2.1 of the report now describes the risks related to the current role of data intermediaries under a dedicated sub-heading entitled 'The role of platforms in the data economy'. It signals the risk of generalising the business model of Big Tech platforms from outside Europe that concentrate large volumes of data to the area of data sharing (sections 2.1 and 2.3). This issue is interlinked with the low level of trust in data sharing, which now appears as a main problem driver in section 2.1, and which is disentangled from the more technical issues (interoperability, findability). The same section also explains in more detail the importance of the neutrality of data intermediaries as a means to increase trust. The section dedicated to the necessity for EU action (section 3.2) better explains how, in order to roll out EU-wide products and services based on data, businesses should be able to benefit from the size of the internal market.
(2) The report should be clear on the objective of the intervention. It could explain that the initiative might help to mitigate the Covid-19 and climate crises. However, the resolution of these crises does not form an integral part of the intervention logic and should therefore not be the general objective. In addition, the report should make evident that the initiative is not about 'free data for all'. The objectives should also better consider the importance of access to data for competitiveness.	In section 4.1, the general objective of the initiative has been reformulated, linking a higher level of data sharing (which is not a goal in itself) to realising the enormous potential of the use of data for the EU's economy and society. The report now explicitly states upfront under the subheading 'The importance of data for the economy' (section 1.1.) that data sharing does not imply that all data will be available for free reuse by all. This is further exemplified in the report, for example in the box on common European data spaces (section 1.2) and by highlighting the incentives for companies and individuals to share data in a separate sub-heading (section 2.1). The report emphasises the essential role of access to and use of data for competitiveness, including innovation in areas such as artificial intelligence, more efficiency across industry, and data as a critical resource for SMEs and start-ups (sections 1.1 and 2.1).
(3) The report should explain the interaction between the investments in common European data spaces by the Digital Europe programme and the Connecting Europe	The report now describes in more detail in section 5.2 how the Commission will invest through the Digital Europe Programme and the Connecting Europe Facility2 in the development of data

Facility, and this initiative. It should include their effects in the baseline and the analysis of options.	processing infrastructures, tools, architectures and mechanisms for data sharing. It also describes the interplay of these investments with the current initiative: the impacts of spending on common European data spaces will depend on the efficiency of the measures under the current initiative.
(4) The report should better explain the composition and completeness of the options. It should justify why it discards all soft regulatory measures upfront. It should elaborate the reasons for the combinations of measures under the 'low' and 'high' intensity options, and explore further if the set of options is complete. The report also needs to explain clearly how each option would work in practice. In particular, it should describe in more detail the role and functioning of the different supervising and coordinating bodies that are under consideration. It should also clarify to what extent the initiative would rely on altruism, and whether this poses concerns regarding supply and scarcity of data. It should explain how control interests of primary data suppliers would be protected. It should consider the possible role of the public sector as a data intermediary with the digitalisation of public administrations.	It has been clarified that none of the policy options were discarded upfront. An analysis of the soft law option has been included, also covering each of the different intervention areas (section 6.1). As part of the soft law/coordination measures, a system of industry-driven certification/labelling of data intermediaries has been considered (section 6.1.2). Section 5.2 of the report (description of the policy options) has been reinforced with additional explanations on all policy options. The low and high intensity options for each measure are now described in more detail, as well as the reasoning behind the design of various combinations. The report also explains more clearly in section 5.2 how the different options would work in practice, including the role of the supporting and supervising bodies at the national level, as well as the European Data Innovation Board (status, role, composition, who is responsible for the secretariat). Data altruism specifically addresses data availability for the common good. The need to protect the interests of the data suppliers in the context of data altruism is now more explicit in sections 5.2 and 6.2, and is a key element for the retained option for this intervention area. The notion that organisations engaging in data altruism should ensure that the data is used in compliance with the stated preferences of the company or individual giving the data has been added (section 5.2.3.C). The roles of the public and private sectors in relation to data intermediary functions have been further calibrated (section 5.2).
(5) The report should explain why the calculated economic benefits of the options are marginal compared with the expected evolution of the data sector. If necessary, it	Section 6.1 of the report now explains in further detail the methodology for calculating the economic impacts of the policy options, concentrating on direct impacts, and to a limited

could rely more on qualitative arguments. The analysis should look into effects on SMEs and costs for Member States. The report should better justify the benefits of creating the European Innovation Board.	extent on the indirect impacts. At the same time, section 6.1 (echoed in Chapter 8) now indicates why the overall benefits of the initiative are expected to be significantly higher than the direct impact: the measures would act as a catalyst to increase data sharing across the EU, which would benefit not only the data economy, but the EU economy and society as a whole. Chapter 6 of the report now better explains the expected consequences of each policy option on the economy, as well as on society and the environment. The section on the impacts on SMEs (section 6.3) has been further enriched. The notion that the initiative is likely to benefit companies from across the EU and not only from some Member States has been added (section 6.1). A reference to national initiatives has been added in relation to the calculation of the costs (and benefits) for Member States (section 6.1.3). The potential benefits of a European Data Innovation Board, in particular in terms of standardisation, are further elaborated in section 6.2 of the report.
(6) The report therefore, needs to present a more granular analysis of the impacts of the different intervention areas to better justify the choice of the preferred option.	Chapter 6 of the report now investigates in more detail the economic, social and environmental impacts of each policy option, including soft law measures. In addition, the multi-criteria analysis in Chapter 7 has been enriched, thus providing a stronger foundation for the chosen package in Chapter 8.

The Regulatory Scrutiny Board delivered a second opinion that was positive, provided that the following recommendations were taken into account in the report.

Comments of the RSB	*How and where comments have been addressed*
(B) Summary of findings and (C) What to improve	
Options (B1) Options are not sufficiently clear on how they would work in practice. The justification for the composition of the options is not always convincing. (C1) The report should further clarify the content of the options. It should explain how the self-regulation option would differ from current practices (which are part of the	Chapter 5 of the report (description of policy options) now specifies, for each intervention area, how the policy options 2 and 3 would work in practice. Examples are provided to clarify what is expected and by whom. All of the comments mentioned in (C1) have been addressed in section 5.2.3 (description of regulatory intervention with low or high intensity).

baseline). For the options on reuse of public data, it should justify why other possible dimensions of the options were considered, but not further analysed. It should better explain how the high intensity option would work in practice. […] Regarding the European Data Innovation Board, the report could further specify its foreseen functioning under the options, including its role and powers vis-à-vis Member State authorities.	It also better explains how the high intensity option would work in practice. The role of the European Data Innovation Board vis-à-vis national authorities has been clarified under each policy option.
Impacts (B2) The analysis lacks depth regarding impacts on SMEs, Member States and the internal market. (C2) The report should deepen the analysis of SME specific impacts and costs for Member States. It should analyse the possible impact on the internal market of different implementation approaches across Member States. It should explain better why the expected benefits in the impact assessment are much smaller than in the referenced research studies. (C4) The report needs to present a more granular overview of the impacts of the different intervention areas in tabular form.	Chapter 6 of the report provides more detailed information on the expected impacts on SMEs in the EU (impact of the policy options), as well as on the expected costs for Member States. The explanation on the difference between the expected benefits in the Impact Assessment and the referenced studies is strengthened in section 6.1. In the same section, the impact on the internal market of diverging approaches between Member States is addressed. To address comment (C4), eight tables have been inserted in Chapter 6. These tables summarise the economic impact (costs and benefits) of the different policy options for each intervention area.
Funding Programmes (C3) The report should better integrate the expected effects of the Digital Europe programme and the Connecting Europe Facility in the analysis of options.	Chapter 5 describes in more detail how the different options, including the baseline scenario, would benefit from the technical infrastructures and tools developed with the support of the DEP and CEF programmes. This concerns in particular actions to ensure interoperability and the use of common standards across Member States. The description of the expected economic impacts (section 6.1) now includes information on how investments in the creation of a European data sharing and processing infrastructure would help lower the costs related to the technical implementation of the measures proposed in this initiative.
Options on data altruism (B3) The report does not convincingly argue the choice of the preferred option for data	The reasoning for selecting the preferred option for data altruism has been strengthened in

altruism. (C1) For the options on data altruism, the report should better justify why the low intensity option foresees voluntary private certification and the high intensity option compulsory public authorisation. It should consider including a voluntary public certification option as an alternative. (C4) It should better justify its choice for the high intensity option for data altruism, especially as it does not analyse a voluntary public certification option (see above).	Chapters 6 and 8. In particular, section 6.1.4 (on the high intensity option) highlights the benefits for citizens, companies and data users of this option. Section 5.2.3 of the responds to comment (C1) on the justification of the low and high intensity options.
Monitoring and evaluation (C5) The report should examine in more depth how it intends to organise future monitoring and evaluation on an ongoing basis. Given that it is experimenting with new, untried approaches, waiting five years for their evaluation seems a rather static approach. It should clarify how increased trust in data sharing will be measured and monitored. It should describe how the effectiveness of these new approaches will be assessed in a timely manner.	Chapter 9 of the report now describes in more depth how the impact of the initiative will be monitored and evaluated on a regular basis. It explains that the European Data Innovation Board will collect experiences from the Member States and assess the effectiveness of their practices. The Support Centre for Data Sharing, which is planned to be established under the DEP, will ensure a similar role with stakeholders.

4. EVIDENCE, SOURCES AND QUALITY

Evidence-collection process

Extensive work was carried out during the previous Commission's mandate to identify the problems that are currently preventing Europe from realising the full economic and societal potential of data-driven innovation, in particular by ensuring greater access to and use of data. This work resulted in earlier Commission policy documents[166], the consultation of stakeholders and extensive exploratory study work[167]. The analyses have identified technical barriers (interoperability, safety and security requirements), legal obstacles (uncertainty about access and

[166] COM/2017/9; COM/2018/232 .

[167] Everis (2018). *Study on data sharing between companies in Europe*, Study prepared for DG CNECT.
European Commission (2018c). *Study on emerging issues of data ownership, interoperability, (re-)usability and access to data, and liability*, study prepared by Deloitte.
European Commission (2017). *Synopsis report consultation on the 'building a European data economy' initiative*.
European Commission (2019a). *SME panel consultation B2B data sharing - Final Report*.
European Commission (2018d). *Study to support the review of Directive 2003/98/EC on the re-use of public sector information*, study prepared by Deloitte.

use rights in relation with the data, the costs of compliance with existing legal obligations as well as costs of licensing), organisational challenges, the difficulty to control downstream use, the fear of misappropriation and the limited availability of skilled labour.

Through an Impact Assessment support study (SMART 2019/0024), additional evidence was gathered in terms of the specific costs and benefits of the concrete elements of the instrument. These costs, benefits and burden reduction/simplification potential were identified and quantified for each of the measures proposed in the initiative (secondary use of sensitive data held by the public sector, data altruism, governance aspects of data sharing, and certification options for data intermediaries). The contractors analysed each of these elements, combining desk research with surveys, interviews and focus groups with representatives of businesses.

The findings of the EC-funded European Data Market study measuring the size and trends of the EU data economy (SMART 2016/0063) also fed into the preparation of the initiative. Based on alternative development paths driven by different macroeconomic and framework conditions, the study monitors several indicators. This included the number of data workers, data companies and their revenues, data user companies and their spending for data technologies, the market of digital products and services, the data economy and its impacts on the European economy, and medium-term forecast scenarios of all the indicators, based on alternative market trajectories.

Stakeholders' consultation process

Recent stakeholder consultation processes provided input: the 2017 public consultation on building a European data economy, the 2018 public consultation on the revision of the Directive on the reuse of public sector information, and the 2018 SME panel consultation on the B2B data sharing principles and guidance.

In addition, a series of 10 workshops on common European data spaces took place in 2019 and an additional one in May 2020. The stakeholders generally supported the creation of common European data spaces, and considered that they should provide for the clarification and harmonisation of data governance models and practices, as well as for the necessary infrastructures for the sharing of good quality and interoperable data.

Together with the adoption of the European Strategy on data, an online public consultation, targeting all stakeholders, was launched on 19 February 2020. It ran until 31 May 2020. The consultation explicitly indicated it was launched in view of preparing the current initiative, and addressed the items covered in the initiative with relevant sections and questions. The feedback on the Inception Impact Assessment also targeted all types of stakeholders, as did the Eurobarometer on the impact of digitisation.

ANNEX 2: STAKEHOLDER CONSULTATION

INTRODUCTION

The stakeholders' consultation process aimed at understanding how stakeholders consider that data governance mechanisms and structures can best maximise the social and economic benefits of data usage in the EU. This provided valuable input for the preparation of the proposal for a Regulation on the governance of common European data spaces.

The proposal also builds on past consultation actions, such as the 2017 public consultation on building a European data economy, the 2018 public consultation on the revision of the Directive on the reuse of public sector information, and the 2018 SME panel consultation on the B2B data sharing principles and guidance.

The consultation actions conducted between July 2019 and June 2020 covered general and horizontal issues, such as the design and main features of the common European data spaces, relevant data governance mechanisms, as well as more specific horizontal questions (standardisation, secondary use of data, data altruism and data intermediaries). Some consultation actions covered the specificities of sectors, and the conditions to facilitate such spaces in domains of public interest. Therefore, the consultation process targeted all types of stakeholders (Member States and public authorities, academic and research institutions, industry stakeholders/businesses including SMEs, as well as individuals) across the EU, and across sectors. Different types of stakeholders are interested in the initiative for different reasons:

- Member States are interested in the initiative from a policy perspective, given the potential of data for the economy and society. Their public authorities are directly concerned by the proposed measures on unlocking and reusing more public sector data. EU level coordination would facilitate the work of public sector bodies in streamlining how data can be used and who can access it.
- Academic and research institutions as well as researchers will directly benefit from the measures on secondary use of data and data altruism, which will considerably lower their compliance costs related to using data.
- Industry stakeholders/ businesses, including SMEs, in the different sectors (e.g. agriculture, finance/banking, energy, transport, sustainability/environment, public services, smart manufacturing and data market places) will in particular benefit from the opportunities provided by easier cross-sectoral data use. They will also be providers of data and as such need to be aware of the rules and limitations on data sharing.
- Individuals will be empowered to allow use of data related to them for the public good (e.g. people with rare or chronic diseases allowing for use of data in order to help cure or improve treatment of those diseases) ('data altruism') and more in general by the opportunities to get more control over their data, e.g. through personal data spaces.

The consultation actions foreseen in the Consultation Strategy, discussed in an inter-service group in March 2020, were carried out. However, due to the COVID-19 outbreak, some actions were modified (i.e. workshops turned into webinars).

Online consultation

A public online consultation was published on the day of adoption of the European strategy for data[168] (19 February 2020) and closed on 31 May 2020. The consultation explicitly indicated it was launched in view of preparing the current initiative, and addressed the items covered in the initiative with relevant sections and questions. It targeted all types of stakeholders. In addition to issues related to the governance of common European data spaces, it gathered input on the EU-wide list of high-value datasets that the Commission will draw up under the Open Data Directive, and explored issues related to cloud computing. Furthermore, it contained some generic questions on the European data strategy

In total, 806 contributions were received, of which 219 were on behalf of a company, 119 from a business association, 201 from EU citizens, 98 on behalf of academic / research institutions, and 57 from public authorities. Consumers' voices were represented by 7 respondents, and 54 respondents were non-governmental organisations (including 2 environmental organisations). Amongst the 219 companies / business organisations, 43.4% were SMEs. Overall, 92.2% of the replies came from the EU-27. Very few respondents indicated whether their organisation had a local, regional, national or international scope.

230 position papers were submitted, either attached to questionnaire answers (210) or as stand-alone contributions (20). The papers provided different views on the topics covered by the online questionnaire, in particular in relation to the governance of common data spaces. They provided opinions on the key principles for those spaces, and expressed a high level of support for the prioritisation of standards as well as the data altruism concept. They also indicated the need for safeguards in developing measures related to data intermediaries.

Inception Impact Assessment

As foreseen by the Better Regulation guidelines, an Inception Impact Assessment was published on the Better Regulation portal on 3 July 2020, and was open for feedback for 4 weeks. It also targeted all types of stakeholders. The Commission received 107 contributions on the Better Regulation Portal[169], mainly from businesses (35%) and associations representing businesses (29%). Other types of stakeholders participated, although in a smaller proportion: non-governmental organisations (11%), academic/research institutions (6%), consumer organisations (3%), EU citizens (2%), trade unions (2%) and others (9%). Some of these stakeholders had also contributed to the public online consultation.

Expectedly, the contributions reflected the replies to the online questionnaire, as well as the papers. Most of the contributions expressed support to the initiative and contained general comments, underlying the importance of fair, transparent and non-discriminatory access to data, of voluntary data sharing (from private entities but also from individuals) and of standards and interoperability. The feedback dealt with all aspects and measures foreseen in the initiative.

[168] COM/2020/66 final.
[169] https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12491-Legislative-framework-for-the-governance-of-common-European-data-spaces

Stakeholders highlighted some concerns and strong needs they have as regards access to and re-usability of data. They also indicated the need for guidance to accompany any legislation.

Other consultation actions

- *Series of workshops on "common European data spaces"*

In order to explore with the relevant experts the framework conditions for creating common European data spaces in the identified sectors, a series of 10 workshops was organised between July and November 2019.

Gathering in total more than 300 stakeholders, mainly from the private and the public sectors, the workshops covered different sectors (agriculture, health, finance/banking, energy, transport, sustainability/environment, public services, smart manufacturing) as well as more cross-cutting aspects (data ethics, data market places). The different DGs concerned were involved in these workshops. A report is available online.

- *The latest Eurobarometer on the impact of digitisation*

This general survey on the daily lives of Europeans includes questions on people's control on and sharing of personal information. The report, published on 5 March 2020, provides information on the willingness of European citizens to share their personal information and under which conditions.

- *Workshop on labels for or certification of providers of technical solutions for data exchange*

Around 100 participants from businesses (including SMEs), European institutions and academia attended this webinar, on 12 May 2020. Its aim was to examine whether a labelling or certification scheme could boost the business uptake of data intermediaries by enhancing trust in the data ecosystem. A report is available online.

- *BDVA Survey*

The survey (May-July 2020) designed by the Big Data Value Association (BDVA) aimed to capture the current state of data-sharing practices by businesses, research institutions, governmental or non-governmental organisations. The objective was to understand the predominance of data sharing and exchange activities, the value that such practices bring to organisations and the difficulties faced by stakeholders, as well as to gather insights into what needs to be done to increase participation in data sharing, in view of the ever increasing need for greater access to data.

- *The Opinion of the European Data Supervisor on the European strategy for data*

On 16 June 2020, the European Data Protection Supervisor adopted Opinion 3/2020 on the European strategy for data. The approach of the EDPS towards the strategy in general is positive, considering that the implementation of the strategy will be an opportunity to set an example for an alternative data economy model.

- *Position of the Member States*

The European Strategy for Data was welcomed by the Member States in the Council Conclusions of 9 June 2020, specifically calling the European Commission *"to present concrete proposals on data governance and to encourage the development of common European data spaces for strategic sectors of the industry and domains of public interest"*[170]. On 9 July 2020, the Digital Single Market Strategic Group, composed of Member States representatives, was also presented initial ideas for the legislative framework on the common European data spaces and expressed a strong support to the improvement of data governance at EU level.

RESULTS OF THE CONSULTATION PROCESS

- ***On the challenges around data sharing***

The 2017 consultation process on 'Building a European Data Economy' investigated the magnitude of data-sharing limitations and the relevance of measures envisaged by the Commission to foster a thriving EU data economy (synopsis report available online). Through the online questionnaire, meetings and workshops, virtually all stakeholders confirmed that more data should be made available for reuse in B2B contexts. Most stakeholders also shared the view that the European Commission should be cautious when taking any measures to make more data available for reuse, stressing that the main issue is how to maximise and organise access to and reuse of data, rather than questions about data access rights.

On the basis of the business-to-business data-sharing principles and guidance that the Commission issued in the April 2018 data package, further consultation actions, including an online consultation (October 2018 to January 2019), provided the views of 979 SMEs (report available online). Some 39% of responding SMEs encountered difficulties in accessing data from other companies.

The 2020 online consultation confirmed this statement, with almost 80% of the 512 respondents to the question indicating that they have encountered difficulties in using data from other companies. These difficulties relate to technical aspects (data interoperability and transfer mechanisms), denied data access, and prohibitive prices or other conditions considered unfair or prohibitive. Some companies also fear that they might lose their competitive advantage within their market or in prospective markets if they engage in data sharing. In this online consultation, some companies highlighted the reluctance of other companies to share data due to this fear, as well as more technical problems such as data quality and granularity. At the same time, the position papers received showed that many stakeholders consider that data sharing should remain voluntary.

The 2019 workshops on the common European data spaces revealed that companies often struggle to find or obtain the data that they need, including from different markets. The findability issue was also raised during the workshop on labels for/certification of providers of technical solutions for data exchanges in May 2020.

[170] Council of the European Union *Conclusions* (9 June 2020).

- ***On the need for common European data spaces and data governance mechanisms***

Throughout the various 2019 and 2020 consultation actions, stakeholders strongly supported **common European data spaces as a concept**. During the 2019 workshops, they stated that common European data spaces should help to establish data governance models leading to more standardised approaches for data sharing, and should provide the necessary infrastructures, including pan-European sustainable cloud federations, for the sharing of good quality and interoperable data.

Stakeholders considered that such data spaces could become the key instance for clarifying data control rights and rules on data access and use. This particularly relates to areas where control rights are an important concern because of sensitive data at stake (e.g. health), or because of existing competition between the different actors within these sectors (e.g. agriculture, transport, energy). In the workshops and in the position papers several stakeholders expressed the concern that Big Tech platforms could move into their sectors, and would get an undue advantage based on the use of data.

The results of the online consultation, conducted from February to May 2020, confirmed those trends. From the 772 respondents to the question, 90% considered that data governance mechanisms are needed to capture the enormous potential of data, in particular for cross-sector data use, and 86% supported the development of common European data spaces in strategic industrial sectors and domains of public interest. In the papers received, stakeholders described in more detail the key principles that they consider should underpin the data spaces: open to all/ non-discriminatory, voluntary, preserving 'sovereignty' of the data provider, agile, decentralised, based on trust and transparency, ethical framework, human-centric, industry-led and inclusive, accountable. They indicated that any legislation should be accompanied by clear guidance, and that an EU level coordination body should be established to maximise the benefits of data.

From the 554 respondents to the question, 91% of stakeholders considered **standardisation** to be necessary, in view of improving interoperability and ultimately data reuse across sectors. Only a very small share (1,6%) of all respondents considered that EU or national government bodies should take no role in standardisation. Public funding was considered necessary to open standards and for testing, and EU and national bodies are expected to take an active role in the prioritisation and coordination of standardisation needs, as well as the creation and updating of standards.

The papers received provided additional input: whereas many stakeholders consider that interoperability (both legal and technical) is a key challenge for EU businesses, there are concerns that the implementation costs will unfairly impact SMEs and may ultimately fall on consumers. They explained that standards should be market-led and global, building on existing standards (e.g. ISO), and that the role of the EU is to coordinate the prioritisation of standards and ensure that they are not imposed by big market players.

The 2019 workshops also revealed that interoperability and data quality issues could be addressed through the common data spaces. For instance, in the workshop on labels for or certification of providers of technical solutions for data exchange, it was highlighted that an

obligation for interoperability with other providers of data-sharing services would be difficult to certify in the absence of standards for interoperability. Participants in various workshops also stated that there is a need for a structured prioritisation of standards on data, especially in view of increasing the opportunities for cross-sectoral reuse.

- ***On the secondary use of public sector data***

On enhancing the secondary use of public sector data that is subject to rights of others (personal data, trade secrets and other commercially confidential data), the Commission has interacted with national organisations that have set up technical mechanisms allowing controlled processing of such data, for instance in the fields of statistics (Germany), mobility (Finland) and health (France). This provided a better understanding of how privacy-preserving technologies can help to allow the extraction of certain insights from the data under controlled conditions while preserving information privacy.

In the online consultation question about making a broader range of 'sensitive' public sector data available for R&I purposes for the public interest, more than three quarters of respondents to the question considered that public authorities should do more, especially mentioning the anonymisation of specific data for concrete use-cases, and the clarification of the legal rules. Unsurprisingly, the vast majority (87%) of respondents from academic and research organisations agreed on the need to facilitate the reuse of sensitive data for research and innovation purposes. In open questions, stakeholders also suggested that public authorities should support the adoption of private portals for the reuse of data, enabling third party trust and quality services. Sensitive data needs robust governance and can benefit from third parties as gatekeepers. In particular regarding health data, research ethics committees or ethics review boards should be involved.

Papers received confirmed that stakeholders consider that public authorities should do more to make data that is subject to the rights of others available for reuse for R&I purposes, but this should be strictly regulated. Decisions to allow reuse should be based on the public interest (which needs to be defined) and use-case specific risk assessment. Anonymization is important, but it could prevent the data from being reusable. Transparency is essential (how the data will be shared, processed, etc.). Data made available should follow a minimization principle (i.e. defined temporal scope and sensitive data should only be made available when needed).

The European Data Protection Supervisor underlined in his Opinion that the contours of scientific research versus commercial innovation are not clear-cut. Therefore, exceptions for research on personal data cannot lead to a broad exemption of the scientific sector from GDPR obligations.

- ***On data altruism***

In a workshop organised on 24 May 2019, experts discussed issues related to data donation in healthcare. A number of experts questioned the use of the term 'data donation', as it could suggest an irreversible process and could presuppose 'data ownership' (as a result, this initiative now uses the term 'data altruism'). However, according to the GDPR, consent by data subjects to processing of personal data pertaining to them can be withdrawn at any time, including when

such processing is based on altruistic motivations of the data subject. Experts also suggested that better support structures and services are needed for data altruism to become widely accepted and used, including interoperability and standards. More success stories and good practices (e.g. models and incentives for data altruism) are also needed to improve understanding of the governance and support requirements.

In the online consultation, a large proportion of respondents considered that law and technology should enable citizens to make available their data for the public interest, without any direct reward. Citizens, particularly, would be willing to make such data available, especially for health-related research and for aspects relating to the locality they live in (e.g. mobility, environment). More than 60% of all respondents considered that there are no sufficient tools and mechanisms to 'donate' their data. On the mechanisms to support 'data altruism', respondents favoured a European approach to obtaining consent, in compliance with the GDPR as well as the establishment of technical infrastructures such as personal data intermediaries (see below) and information campaigns. In open questions, stakeholders suggested more supporting mechanisms (model contractual clauses or data sharing agreements; mechanisms (e.g. blockchain) to ensure data chain of custody in order to capture where and how the data was used; information and transparency ensured by the public sector with regard to the protection of "contributed" data, e.g. the use of technical safeguards such as pseudonymisation).

The papers received confirmed the high level of support for putting individuals in control of their own data: solutions are needed to reconcile privacy rights with the use of data for the common good. Trust is key, and a clear legal basis defining how individuals can make data available for altruistic purposes in full compliance with the GDPR would be welcome. Transparency is also critical: the individual should know what their data is being used for. Individuals should not be 'nudged' into sharing more data than they normally would by labelling such sharing 'data altruism'. Individuals should remain free not to 'donate'. The term 'altruism' was sometimes considered misleading. Lack of representativeness could be an issue. Finally, several stakeholders considered that personal data should not be monetised.

In its Opinion, the EDPS made the comment that the GDPR already provides principles and rules on consent, hence giving the possibility for the 'data altruism' concept. Therefore, the initiative should clearly define the scope, including various purposes.

- ***On data intermediaries***

In the online consultation, almost 60% of respondents to this section considered that emerging novel intermediaries, such as 'data marketplaces', are useful enablers to the data economy, while almost 22% don't know or remain neutral to the question. In open questions and papers received, stakeholders confirmed such intermediaries play an important role in providing fluidity of the data economy, but said a strict accountability framework is needed. A data intermediary could verify the connected datasets of a particular individual, which would increase the reliability and therefore relevance of the data for the recipient. However, an intermediary also adds additional contracts and costs.

Some stakeholders indicated that a uniform and binding definition of data trustee systems should be created and corresponding specifications for certification processes should be developed. There is also the view that the EU should support the development of data platforms and marketplaces allowing private and public sectors alike to sell, trade and access quality datasets.

A workshop organised on "labels for or certification of providers of technical solutions for data exchange" in May 2020 showed strong interest in the topic with almost 100 participants. Participants were mostly companies or initiatives active in the field of data intermediation or sharing both in B2B situations and supporting individuals (personal information management services). They stressed the importance of trust in data sharing and explored mechanisms for creation of such trust, namely neutral data intermediation services but also trust frameworks or data sharing 'schemas' that would lay down relevant technical and legal rules to be respected in data sharing situations. This could ensure 'data sovereignty' by businesses in data-sharing situations and empowerment of individuals with respect to use of their data.

CONCLUSION: CONSIDERATION OF STAKEHOLDERS' FEEDBACK IN THE IMPACT ASSESSMENT

The consultation of stakeholders on the general issues of data sharing (obstacles across borders and sectors, and possible solutions at EU level for enhancing data sharing) has been an ongoing process from 2017 onwards. The concept of common European data spaces has been explored for the preparation of the European strategy for data, notably with workshops of horizontal and sectoral nature held in 2019 and 2020. These results were completed with input from all stakeholders on more technical questions such as governance mechanisms or standardisation, as well as input on specific types of action as data altruism or the enhanced use of certain public sector data.

All consultation actions revealed a strong support for the development of common European data spaces, and the human-centric approach to data sharing in general, as presented in the European Strategy on data.

ANNEX 3: WHO IS AFFECTED AND HOW?

1. PRACTICAL IMPLICATIONS OF THE INITIATIVE

The planned legislative framework will have a range of practical implications for different groups of stakeholders from the entire data value chain: data holders (public bodies), data reusers (businesses and the research community), data intermediaries (public bodies, patients association, health insurance schemes, and research organisations) and data (co-)producers (individuals and other public sector authorities).

The initiative will benefit **public sector bodies** (e.g. health institutions, transport authorities, statistical offices) in a number of ways. EU level coordination will facilitate their work by clarifying how data can be used and who can access it, as well as by facilitating the sharing of legal and technical expertise. With more data available for reuse, the public sector will be able to deliver more efficient public services and make more informed decisions, leading to better policies. This will help enhance public-service delivery and facilitate the identification of emerging governmental and societal needs. It can help improve forecasting and the reliability of infrastructures (such as in transportation and utilities).

In the context of the secondary use of data subject to rights of others, two broad categories of data holders and public data intermediaries can be differentiated: statistical offices and health- and social-related data holders/ intermediaries. In terms of data holders: as regards statistical offices (and other public authorities responsible for the development, production and dissemination of statistics), the European Statistical System keeps an up-to-date list that currently contains 286 entities, of which 27 are related to health (and therefore excluded from this count to avoid double-counting). As a result, the number of data holders in the EU27 when it comes to statistical microdata can be estimated to be around 260. As regards health- and social-related data, there are roughly 530 data holders in the health and social domains. In total, therefore, there are around 800 impacted data holders. In terms of public data intermediaries, there are around 110 public data intermediaries in total[171] (except those with a federal structure).

Academic and research institutions will benefit from the increased availability for reuse of public sector data, the reuse of which can be essential for research purposes for the common good, (e.g. health, location, or social media data) for research and innovative purposes through the proposed measures on the secondary use of data and data altruism. The possibility to reuse new datasets can help review and replicate scientific results, and foster new instruments and methods of data-intensive exploration and scientific experimentation. Academic and research institutions will benefit from the availability of support structures in Member States. In particular, this initiative will help researchers reuse publicly held data subject to conflicting rights under secure and privacy-enhanced environments in a similar way across the EU. This will contribute to the scientific developments and innovation in the EU as a whole, particularly important in situations where EU coordinated action is necessary, like the COVID-19 crisis.

[171] It is unlikely that a given Member State would have more than one public data intermediary for the same domain, since the reason behind their existence is to streamline procedures.

An estimated number of reusers of statistical microdata can be derived from Eurostat's list of recognised research entities, which lists a total of 666 recognised research entities in the EU27. The total number of data reusers for health and social data overlaps with the research entities recognised by Eurostat: 48 of these conduct research in *inter alia* social sciences, while 22 conduct research in a health-related domain. However, it also includes a large number of private companies – that number is estimated to be 147 000 companies[172]. Thus, there are roughly 150 000 data reusers impacted in total.

Based on the input collected from the interviews organised through the IA support study, providing high-quality metadata and documentation for scientific datasets requires 5 to 10% of the total project budget, which represents a substantial expenditure. Other sources estimate that the production of metadata and the contextual descriptions of datasets could span an estimated 2 to 3 weeks from an average of a 2-year research grant application (OpenAire 2019).

Industry stakeholders and/ businesses (including **SMEs** and start-ups) across the economy (e.g. agriculture, finance/banking, energy, transport, sustainability/environment, public services, smart manufacturing and data market places) will in particular benefit the measure taken in this initiative to facilitate cross-sectoral data use at the scale of the EU. They will also be providers of data and as such will be affected by the rules and limitations on data sharing. They will provide information and insights on the type of data governance mechanisms (organisational, technical, legal) needed to capture the potential of data in particular for cross-sector data use, in different data-sharing configurations. Through enhancing interoperability at the technical level and making available generic enabling standards, the initiative will lower transaction costs of data sharing and facilitate cross-sector data sharing. The benefits of standardisation translate into lower technical adaptation costs for a larger range of companies as well as public authorities, lower barriers to enter markets or to develop entirely novel products or services. Such benefits should in particular benefit SMEs that normally cannot influence standardisation prioritisation. SMEs would also substantially benefit from wider availability of public sector data as they typically cannot create large data pools themselves.

Enhanced access to data for reuse will create new business opportunities for smaller and larger firms. Better access to open government data, for instance, will allow entrepreneurs to develop innovative commercial and social goods and services. An example is RowdMap, an analytics company that uses open data to help healthcare plans, physician groups and hospital systems identify, quantify, and reduce low-value care. In July 2017, the company was acquired for USD 70 million by Cotiviti, a provider of analytics-driven payment accuracy solutions.

[172] This estimation was reached using:

a) the number of people employed in the healthcare industry (800,000 in 2012 in the EU). See the *website* of the European Commission on Healthcare Industries.
b) the number of active businesses in the EU (27,5 million in 2017), and
c) the number of employed persons in the EU (150 million persons in 2017). See Eurostat, *Business demography statistics*. These figures were used to reach an average number of employees per active business (150,000,000/27,500,00=5,45); from which the number of healthcare businesses was derived (800,000/5,45=146788.99) and rounded-up.

In the era of the Internet of Things, data collection by sensors will provide **consumers** with innovative smart products and services that will increasingly replace traditional products. Also, the data collected in this industry will be of particular utility to private actors in very different business sectors and to public entities. Hence, data collected by smart products will become an important input both for other businesses and for the public authorities.

In addition, the initiative supports the emerging offer of **data intermediaries** that can make the data economy more fluid, i.e. entities that enable any kind of data holder (persons, business, public sector bodies, academic or not-for profit organisations) to share their data of all types with other organisations, and which may provide additional value-added services. Enhanced data access and sharing will enable many business opportunities for data intermediaries, mobile apps and personal information management systems.

The planned initiative will lower transactions costs in data sharing by supporting the offer of data intermediaries. Companies providing data-sharing services may face an additional burden in terms of certification or labelling, but only if such certification or labelling would become compulsory. Such burden would need to be balanced against the advantages such certification or labelling would provide to them in terms of increased business resulting from more market participants trusting them.

An estimation of the total number of data intermediaries active in the European market could include 100-150 organisations, while the number of data users or data holders affected could entail any European company or individual wishing to buy or sell data through the intermediaries. The companies present big differences in the scale of client base. In particular, Siemens' Mindsphere counted more than 6 100 customers in March 2020; the client base of the personal data operator Peercraft includes approximately 100.000 uses, while Dawex' client base include more than 10 000 organisations. Finally, the example of the data trust UK Biobank holds data from about 0.5m people and it includes the number of 946 researchers using its data in its annual accounts of 2018. This would therefore give a ratio of roughly 50 000:1:1 000 (data holders: data intermediary: data reusers).

The initiative will bring enormous benefits to **individuals**, for example through improved mobility, more personalised medicine, reduced energy consumption and more effective responses to tackle pressing societal challenges, such as climate change and recovering from today's health pandemic. It will be easier for individuals to allow the use of their data for the public good (e.g. people with rare or chronic diseases allowing for use of data in order to help cure or improve treatment of those diseases) ('data altruism'). The framework will also empower individuals interested in reusing the data for their own benefit (e.g. for personalised dashboards, services, etc.). Other benefits to the European economy would include: lower switching costs for users when changing providers; lower entry-barriers for firms in digital markets; increased personalisation of goods and services; and increased innovation driven by valuable user-level insights. In addition, access to a greater variety of data to train models and test results could contribute to the ethical and effective use of AI.

The proposed framework envisages structures, mechanisms, technical guidance and standards so that individuals can exercise their rights in a simple and not overly burdensome way and

organisations, including research ones, can create value for society while respecting the privacy of individuals.

The wider benefits to **society** are manifested in cost reduction, quality improvement and greater choice for consumers. Benefits such as reduced healthcare costs, improved levels of care and reduced environmental degradation that are derived from more intelligent and efficient systems accrue to society as a whole, not just particular sectors or groups of consumers.

2. SUMMARY OF COSTS AND BENEFITS

The figures cited in the tables below illustrate the costs under the preferred option in relation to its specific elements for different types of stakeholders. They are based on the quantitative model developed as part of the Support study by Deloitte.

The overall methodology used by the study to estimate the baseline scenario, as well as the impacts of the policy options, are provided in Annex 4.

Overview of Benefits (total for all provisions) – Preferred Option		
Description	***Amount***	***Comments***
Direct benefits		
Effect on Gross Domestic Product (GDP)	EUR 10.9 billion in 2028 (0.079 % of GDP in 2028).	
Costs Savings and efficiency gains - Easier discovery and reuse of data (due to the creation of mechanisms, including a one-stop shop)	EUR 49.2 million/year	Benefits for data reusers for the EU-27, assuming a saving of 20 hours of work per application.
Costs Savings and efficiency gains - Lower cost of data processing and management (due to the creation of mechanisms, including a one-stop shop)	EUR 684 million/year	Benefits for data holders for the EU-27, assuming that 20% of data holders relinquish their dedicated data processing environment and 30% of the data pre-processing work is passed on to the one-stop shops.
Costs Savings and efficiency gains linked to the set-up of a European Data Innovation Board in charge of enhanced governance of standardisation	EUR 5,335.6 million	Efficiency for participating companies assuming 800 companies and 50M EUR turnover based on IDS examples
Business development linked to data	25%-50% business	

intermediary certification/labelling	development time acceleration for data intermediaries	
Easy and transparent way to access data of various fields, contributing to research and development as well as improved decision-making	EUR 300 million	Improved policy making for government as for example data altruism has proved to be valuable during the COVID-19 pandemic. Other examples are smart city initiatives and environmental data for the public good. These would then be improve public services and goods.
Indirect benefits		
Contribution to societal goals through improved policy- and decision-making	*Not quantifiable due to lack of data*	Especially data altruism could enhance societal goals such as achieving environmental goals, building smart cities of the future and help eradicate pandemics (as is currently the case with COVID-19).
R&I and competition advancement for data intermediaries in the B2B market	*Between 1%-25% competition increase in data intermediaries B2B market, in a 2-5 years' timeframe,* *and between 1%-25% competition increase in data intermediaries B2B market, in a beyond 5 years' timeframe.*	
R&I and competition advancement for data intermediaries in the C2B market	*Between 1%-25% competition increase in data intermediaries C2B market within a one-year timeframe after obtaining the certification/label in 2-5 years' timeframe;* *and between 25%-50%*	

	competition increase in data intermediaries C2B market, beyond 5 years' timeframe	

II. Overview of costs – Preferred option							
		Data holders		*Data intermediaries*		*Data (re)users*	
		One-off	*Recurrent*	*One-off*	*Recurrent*	*One-off*	*Recurrent*
Measures facilitating secondary use of sensitive data held by the public sector (low intensity)	*Concerned parties*	*Public sector bodies*		*Mechanisms (incl. one-stop-shop)*		*Researchers and businesses*	
	Direct costs[173]	-	*EUR 7.6 million/year*	*EUR 286.4 million*	*EUR 16.5 million/year*	-	*EUR 41.8 million/year*
	Indirect costs	-	-	-	-	-	-
Certification/labelling framework for data intermediaries(low intensity)	*Concerned parties*	*Businesses, citizens, academia, researchers*		*Certified/labelled data intermediaries*		*Businesses*	
	Direct costs	-	-	*EUR 20 000-50 000*	*EUR 20 000-35 000/year*	-	-
	Indirect costs	-	-	-	*Approximately 25% decreased market competition in B2B market within the 1st year after obtaining certification*	-	*Non-quantifiable costs due to lack of data*
An EU-wide data 'altruism'	*Concerned parties*	*Citizens, businesses, public sector authorities*		*Public sector authorities, research orgs, businesses*		*Public sector bodies, researchers*	

[173]These numbers show the aggregate amount for the entire EU27, including the costs for all Member States.

scheme (high intensity)	*Direct costs*	*Giving consent to make data available*	*Giving consent to make data available (could be recurrent if it is revoked)* *Non-quantifiable costs due to lack of data*	*Becoming authorized (if applicable)* *EUR 3 800-10 500 depending on the size of the organization* *Establish scheme/authorization process and national oversight body (for public authorities)* *Non-quantifiable, however every EU-27 state has a data authority (or equivalent) that could implement this.*	*Maintain data altruism authorisation EUR 5 000*		-
	Indirect costs	-	-	-	-	-	-
European structure for governance aspects of data sharing (low intensity)	*Concerned parties*	*Businesses*		*Public and private organisations*		*Other businesses and researchers*	
	Direct costs	-	-	-	*EUR 280.000/year for running the group*	-	-
	Indirect costs	-	-	-	-	-	-

ANNEX 4: ANALYTICAL METHODS

1. Overall methodology of the study

For each of the sub-tasks, the Study to support an Impact Assessment on enhancing the use of data in Europe (VIGIE 2020-0694) was carried out in three Phases (inception, data collection, and analysis). With regard to the collection of data, the following key methodological and analysis tool were implemented[174]:

- Desk research;
- Interviews with stakeholders;
- Case studies;
- Workshops with key stakeholders;
- Analysis of the public consultation[175] launched by the European Commission
- Targeted questionnaires to legal experts.

An overview of these data collection tools is provided below.

Tool	Details
Desk research	Desk research was a continuous exercise throughout the study and informed the stakeholder mapping, the preparation of the interview guidelines, drafting of case studies, as well as the draft reporting of findings. It provided information on the state of play and context for each subtask. It was based on academic publications, databases and data marketplaces (e.g. Gartner, Forrester Research, Economist Intelligence Unit).
Interviews	Semi-structured interviews were conducted to collect first-hand material from key stakeholders, both on the state of play of the topic concerned and the impact of the different policy options. Interviews were particularly useful to discuss the costs and benefits of the different options. Interviews were conducted with the following types of stakeholders: • Data holders • Data (re)users • Data intermediaries
Workshops	Two workshops were organised to enable an in-depth discussions with key stakeholders on certain topics: • Measures facilitating secondary use of sensitive data held by the public sector • Establishment a European structure for governance aspects of data sharing

[174] The data collected through the implementation of the above tools will be analysed through the application of the following analytical methods and processes: Legal analysis; Triangulation; Analysis of costs and benefits; and Multi-Criteria Analysis.

[175] https://ec.europa.eu/digital-single-market/en/news/online-consultation-european-strategy-data

Tool	Details
Case studies	Case studies (i.e. in-depth and detailed investigations) were carried out to demonstrate what is going on in certain domains, what works, what does not work and whether 'types' of approaches can be discerned. As such, they were particularly useful for defining the baseline scenarios for the different sub-tasks and developing hypotheses on the impact of the different policy options.
Public consultation analysis	A public consultation on the European strategy for data was carried out from 19 February 2020 to 31 May 2020. The questions included in the public consultation were taken into account when the interview guides were prepared in order to avoid duplication. The results of the public consultation were used to support the analyses during the study.

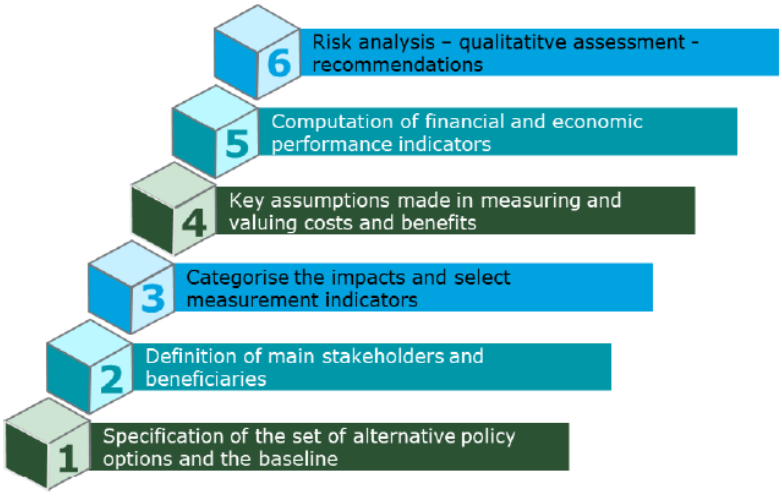
2. Data analysis activities

A market analysis was carried out for sub-task 1.4 ('Data intermediaries') to better understand the business environment and data based value chains as well as to identify the key players and key positions on the market.

A legal analysis was carried out for all sub-tasks, with a more in-depth assessment for sub-task 1.2 ('Establishing an EU wide "Data Altruism" scheme').

The cost-benefit analysis was elaborated individually for each of the sub-tasks. The evaluation process considered the costs and benefits for the different (main) stakeholders associated with each task. The stakeholders were divided into the following categories: data holders, data co-producers, data reusers, and data intermediaries. Impacts on society, environment, economy and fundamental rights are also taken into account.

The key steps in the CBA are outlined in the figure below.



Source: Deloitte

It is in general possible to calculate the project economic performance measured by the following indicators[176]:

— **Economic Net Present Value (ENPV):** The ENPV is defined as the difference between the discounted total socio-economic benefits and the discounted total costs. The ENPV is comparable with the Net Present Value in financial analysis, but it also takes into account the broader socio-economic effects. A positive (economic) net present value indicates that the projected benefits/earnings generated by a project or investment (in present euros) exceeds the anticipated costs (also in present euros). Generally, an investment with a positive ENPV/NPV will be a profitable one and one with a negative ENPV/NPV will result in a net loss. This concept is the basis for the Net Present Value Rule, which dictates that the only investments that should be made are those with positive NPV values.

— **Economic Rate of Return (ERR):** The ERR is defined as the rate that produces a zero value for the ENPV; it is comparable with the ROI (Return on investment) respectively the IRR (Internal rate of Return) in financial analysis. It is another metric commonly used as an ENPV/NPV alternative. Calculations of ERR/IRR rely on the same formula as ENPV/NPV does, except with slight adjustments. ERR/IRR calculations assume a neutral ENPV/NPV (a value of zero) and one instead solves for the discount rate. The discount rate of an investment when ENPV/NPV is zero is the investment's ERR/IRR, essentially representing the projected rate of growth for that investment. Because ERR/IRR is necessarily annual – it refers to projected returns on a yearly basis – it allows for the simplified comparison of a wide variety of types and lengths of investments.

— **Benefit/Cost-ratio (B/C-ratio):** The Benefit-Cost ratio is defined as the ratio between the sum of the discounted economic benefits and the sum of the discounted costs. By putting together the outcomes of the several factors analysed and calculated, it is possible to compute and interpret these three pillars of economic analysis. The different expressions are defined as follows.

Calculation of Economic Net Present Value (ENPV):

$$ENPV = \sum_{t=0}^{n} (N_t - K_t)(1+i)^{-t}$$

- N_t = *Social benefits in year t*
- K_t = *Social costs in year t*
- i = *Social discount rate (SDR)*

Calculation of Economic Rate of Return (ERR):

$$ENPV = \sum_{t=0}^{n} (N_t - K_t)(1+ERR)^{-t} = 0$$

Calculation of Economic Benefit-Cost-Ratio (EBCR):

$$EBCR = \frac{ENPV = \sum_{t=0}^{n} (N_t)(1+i)^{-t}}{ENPV = \sum_{t=0}^{n} (K_t)(1+i)^{-t}}$$

[176] Source: European Commission, 2014: Guide to Cost-Benefit Analysis of Investment Projects – Economic appraisal tool for Cohesion Policy 2014 – 2020.

The economic performance indicators were calculated for each task as well as for each stakeholder, to the extent possible. To do so, assumptions were defined considering the limited availability of quantitative data.

Any CBA is based on a number of assumptions (statistical input as well as certain estimations made by the various stakeholders) that could be critical to the outcome of the analysis. As part of the risk and sensitivity analysis, the critical assumptions were identified and their effects on the outcome determined. Various sensitivity/scenario and risk analyses were performed to analyse the robustness and sensitivity of the results with regard to critical variables.

Impacts that could not be monetized were evaluated in a qualitative manner.

Quality standards for impact modelling

Specific data on costs and benefits is often scarce, inconclusive, and patchy. Any CBA is based on a number of assumptions (statistical input as well as certain estimations made by the various stakeholders) that could be critical to the outcome of the analysis, e.g. qualitative information to fill existing gaps. Oftentimes, these assumptions are based on expert judgment. This means that the data used in the underlying formulas is based on the best data available, challenged and refined (where necessary) by the experts of the consortium for this assignment.

Therefore, in practice, the assumptions used for the CBA are subject to an internal, in-depth peer review process. As part of this process, different assumptions are introduced in the model to compare the different outcomes. Thus, the critical assumptions are identified and their effects on the outcome are determined. This means the risk and sensitivity analysis indicates variances of economic effects as a result of changes of operational figures. Various sensitivity/scenario and risk analyses were performed to analyse the robustness and sensitivity of the results with regard to critical variables.

Figure – Abstract for subtask 1.1

Cost-Benefit Analysis													
Total	M€ (constant prices)	PO	MS	Stakeholder	Category	Subcategory	NPV @ 3%	Total	2023	2024	2025	2026	2027	2028
Total	Data authorisation body	PO3	Total	Data authorisation body	Costs	CAPEX	(556,0)	(572,7)	(572,7)	-	-	-	-	-
		PO3	Total	Data authorisation body	Costs	OPEX	(1 466,0)	(1 648,6)	-	(329,7)	(329,7)	(329,7)	(329,7)	(329,7)
		PO3	Total	Data authorisation body	Benefits	REVENUES	945,8	1 063,6	-	212,7	212,7	212,7	212,7	212,7
	Data holders	PO3	Total	Data holders	Costs	OPEX	-	-	-	-	-	-	-	-
		PO3	Total	Data holders	Benefits	OPEX savings	5 573,2	6 267,2	-	1 253,4	1 253,4	1 253,4	1 253,4	1 253,4
	Data re-users	PO3	Total	Data re-users	Costs	OPEX	(945,8)	(1 063,6)	-	(212,7)	(212,7)	(212,7)	(212,7)	(212,7)
		PO3	Total	Data re-users	Benefits	OPEX savings	742,8	835,3	-	167,1	167,1	167,1	167,1	167,1
	Costs total	PO3	Total	Total	Costs	Costs total	**(2 967,8)**	**(3 284,8)**	**(572,7)**	**(542,4)**	**(542,4)**	**(542,4)**	**(542,4)**	**(542,4)**
	Benefits total	PO3	Total	Total	Benefits	Benefits total	**7 261,9**	**8 166,2**	-	**1 633,2**	**1 633,2**	**1 633,2**	**1 633,2**	**1 633,2**
	Net Cashflow NPV	PO3	Total	Net Cashflow NPV	NPV	NPV	**4 294,1**	**4 881,4**	**(572,7)**	**1 090,8**	**1 090,8**	**1 090,8**	**1 090,8**	**1 090,8**
	Benefit/Cost-ratio	PO3	Total	Benefit/Cost-ratio	BCR	BCR	**2,4**							
Total	Data authorisation body	PO2	Total	Data authorisation body	Costs	CAPEX	(278,0)	(286,3)	(286,3)	-	-	-	-	-
		PO2	Total	Data authorisation body	Costs	OPEX	(73,2)	(82,4)	-	(16,5)	(16,5)	(16,5)	(16,5)	(16,5)
		PO2	Total	Data authorisation body	Benefits	REVENUES	185,7	208,8	-	41,8	41,8	41,8	41,8	41,8
	Data holders	PO2	Total	Data holders	Costs	OPEX	(33,8)	(38,0)	-	(7,6)	(7,6)	(7,6)	(7,6)	(7,6)
		PO2	Total	Data holders	Benefits	OPEX savings	3 041,4	3 420,1	-	684,0	684,0	684,0	684,0	684,0
	Data re-users	PO2	Total	Data re-users	Costs	OPEX	(185,7)	(208,8)	-	(41,8)	(41,8)	(41,8)	(41,8)	(41,8)
		PO2	Total	Data re-users	Benefits	OPEX savings	219,1	246,4	-	49,3	49,3	49,3	49,3	49,3
	Costs total	PO2	Total	Total	Costs total	Costs total	**(351,2)**	**(368,7)**	**(286,3)**	**(16,5)**	**(16,5)**	**(16,5)**	**(16,5)**	**(16,5)**
	Benefits total	PO2	Total	Total	Benefits total	Benefits total	**3 226,7**	**3 628,6**	-	**725,7**	**725,7**	**725,7**	**725,7**	**725,7**
	Net Cashflow NPV	PO2	Total	Net Cashflow NPV	NPV	NPV	**2 875,5**	**3 259,9**	**(286,3)**	**709,2**	**709,2**	**709,2**	**709,2**	**709,2**
	Benefit/Cost-ratio	PO2	Total	Benefit/Cost-ratio	BCR	BCR	**9,2**							

The extent to which an effective sensitivity analysis can be conducted is closely linked to the quality of the CBA. Each of abovementioned calculations was carried out within a Microsoft Excel model that was built specifically for this assignment. Deloitte's Excel models generally follow the ***FAST* standard**[177], consisting of practical, structured design rules for financial modelling.

- ***Flexible:*** Model design and modelling techniques must allow models to be both flexible in the immediate term and adaptable in the longer term. Models must allow users to run scenarios and sensitivities and make modifications over an extended period as new

[177] See: http://www.fast-standard.org/wp-content/uploads/2016/06/FAST-Standard-02b-June-2016.pdf

information becomes available - even by different modellers. A flexible model is not an all-singing, all-dancing template model with an option switch for every eventuality. Flexibility is born of simplicity.

- ***Appropriate:*** Models must reflect key business assumptions directly and faithfully without being overbuilt or cluttered with unnecessary detail. The modeller must not lose sight of what a model is: a good representation of reality, not reality itself. Spurious precision is distracting, verging on dangerous, particularly when it is unbalanced. For example, over-specifying tax assumptions may lead to an expectation that all elements of the model are equally certain and, for example, lead to a false impression, if the revenue forecast is essentially guesswork.
- ***Structured:*** Rigorous consistency in model layout and organisation is essential to retain a model's logical integrity over time, particularly as a model's author may change. A consistent approach to structuring workbooks, worksheets and formulas saves time when building, learning, or maintaining the model.
- ***Transparent:*** Models must rely on simple, clear formulas that can be understood by other modellers and non-modellers alike. Confidence in a financial model's integrity can only be assured with clarity of logic structure and layout. Many recommendations that enhance transparency also increase the flexibility of the model to be adapted over time and make it more easily reviewed.

Multi-criteria analysis

In line with the EC's Better Regulation Guidelines, a Multi-Criteria Analysis (MCA) was carried out, in parallel to the Cost-Benefit Analysis, to identify the preferred policy option for each sub-task.

The MCA is a largely qualitative analysis of the policy options, based on ratings and rankings with quantitative data supporting the assessment. For this reason, MCAs accompany Cost Benefit Analyses and Economic Modelling but do not replace them. As part of the Multi-Criteria-Analysis, the most significant impacts were assessed as a comparison to the baseline scenario:

- Economic impacts;
- Societal impacts; and
- Environmental impacts.

The impacts on Fundamental Rights was used as exclusion criterion.

The following criteria were taken into to assess these impacts:

- Effectiveness, i.e. the extent to which different options would achieve the objectives;
- Efficiency, i.e. comparing the benefits of the options versus the costs (incl. additional and reduced compliance costs);
- Coherence with the overarching objectives of EU policies;
- Legal and political feasibility;
- Compliance of the options with the proportionality principle.

The proportionality principle was used as an exclusion criteria.

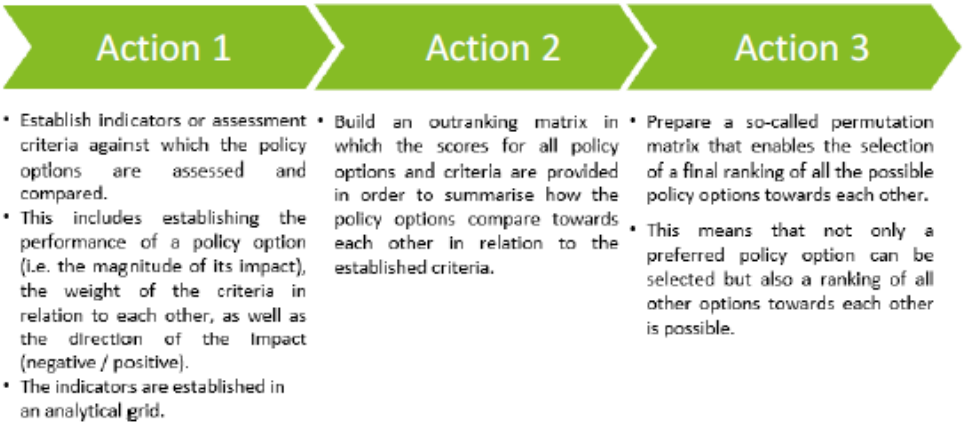
The sources of information were also defined, i.e. existing data (i.e. secondary data from other studies or databases), new data (i.e. primary data) derived from interviews, as well as the workshops.

The same assessment criteria were used for all policy options. Using the same criteria ensures comparability across the policy options, which is imperative for the comparison of the options.

When carrying out the assessments, the expected timing of the impacts (one-off, short term, long term) was taken into account, considering changes in the baseline scenario for the specific time-frame considered.

While the impacts were assessed from the point of view of society as a whole, impacts on different groups of society (e.g. data holders, data intermediaries, data reusers) were differentiated.

The picture bellows summarises the key steps leading to a full Multi-Criteria-Analysis.



Source: Deloitte

ANNEX 5: SUBSIDIARITY GRID

1. Can the Union act? What is the legal basis and competence of the Unions' intended action?
1.1 Which article(s) of the Treaty are used to support the legislative proposal or policy initiative?
This initiative follows from the 2020 European Strategy for data, which aims to create a Single Market for Data. With a growing digitalisation of the economy and society, there is a risk that Member States increasingly regulate data-related issues in an uncoordinated way; this will intensify fragmentation in the internal market. Therefore, this legislative proposal is based on Article 114 of the Treaty on the Functioning of the European Union (TFEU). This article provides for the EU to adopt measures for the approximation of the provisions laid down by law, regulation or administrative action in Member States which have as their object the establishment and functioning of the internal market in the EU.
1.2 Is the Union competence represented by this Treaty article exclusive, shared or supporting in nature?
Digital policies are a shared competence between the EU and its Member States. Articles 4(2) and (3) of the TFEU specify that, in the area of the internal market and technological development, the EU can carry out specific activities, without prejudice to the Member States' freedom to act in the same areas.
2. Subsidiarity Principle: Why should the EU act?
2.1 Does the proposal fulfil the procedural requirements of Protocol No. 2[178]: - *Has there been a wide consultation before proposing the act?*
A consultation process supported the preparation of this proposal and its accompanying Impact Assessment. An online public consultation was launched on 19 February 2020, targeting all types of stakeholders. It ran until 31 May 2020 and collected 806 replies. This consultation was specifically prepared in order to provide input to this initiative. The Inception Impact Assessment was also open to feedback from all types of stakeholders, as was the Eurobarometer on the impact of digitisation. Furthermore, the initiative builds on recent consultation actions, including the 2017 public consultation on building a European data economy, the 2018 public consultation

[178] Protocol (No 2) - On the Application of the Principles of Subsidiarity and Proportionality.

on the revision of the Directive on the reuse of public sector information, and the 2018 SME panel consultation on the B2B data-sharing principles and guidance.

In addition, a series of 10 workshops on common European data spaces took place in 2019 and an additional one in May 2020, in view of exploring with the relevant experts the framework conditions for creating common European data spaces in the identified sectors. Gathering in total more than 300 stakeholders, mainly from the private and the public sectors, the workshops covered different sectors (agriculture, health, finance/banking, energy, transport, sustainability/environment, public services, smart manufacturing) as well as more cross-cutting aspects (data ethics, data market places). The concerned DGs were involved in these workshops.

- *Is there a detailed statement with qualitative and, where possible, quantitative indicators allowing an appraisal of whether the action can best be achieved at Union level?*

The Explanatory Memorandum of the proposal, as well as the Impact Assessment (Chapter 3 – 'Why should the EU act?'), contain dedicated sections on subsidiarity and added value, as explained in section 2.2 below.

2.2 Does the explanatory memorandum (and any impact assessment) accompanying the Commission's proposal contain an adequate justification regarding the conformity with the principle of subsidiarity?

The Impact Assessment accompanying the proposal features a dedicated section on the conformity of the proposed initiative with the principle of subsidiarity (Chapter 3). This is also reflected in the Explanatory Memorandum.

The problem analysis revealed that, despite the growing digitisation of society and the economy, public and private actors in the data economy continue to struggle to access the data they need from other sectors or Member States to develop and roll out data-based products and services. Furthermore, some Member States are already legislating certain aspects or sectors of the data economy while others are not. This increasing legal fragmentation may lead to inconsistent regulatory action across the EU and even potential conflicts of law with the EU *acquis*. In addition to legislative intervention, some Member States are supporting industry-driven approaches to data governance (examples in section 3.2 of the Impact Assessment). This can lead to divergences between sectors and Member States, as they have different priorities. These unresolved problems plead for action at EU level, as was called for by the Member States in their Council conclusions of 9 June 2020[179].

2.3 Based on the answers to the questions below, can the objectives of the proposed

[179] Council of the European Union *Conclusions* (9 June 2020).

action be achieved sufficiently by the Member States acting alone (necessity for EU action)?
The assessment of the barriers related to data sharing in the Impact Assessment has led to the conclusion that the objectives cannot be achieved sufficiently by Member States acting alone for a number of reasons. Big data and AI need large datasets that also contain data on rarer situations (so-called 'long tail'), which are hard to find in individual Member States alone. Also, the development of pan-European data services and products requires data from more than one Member State. Finally, a market for novel data intermediaries can only develop at the scale of more than one Member State. Member States and their public authorities have supported this initiative at the political level (Council conclusions of 9 June 2020), and through the different consultation actions.
(a) Are there significant/appreciable transnational/cross-border aspects to the problems being tackled? Have these been quantified?
Hurdles to data sharing are encountered with regard to sharing between economic operators (public or private), between sectors, but also between Member States. Such obstacles have been investigated extensively since 2017, especially through the consultation that led to the adoption of the Regulation on the free flow of data (addressing the specific issue of data localisation restrictions). The study supporting the Impact Assessment for this proposal, together with the consultation process, showed that individual Member States are pioneering approaches to data governance and related standardisation and starting legislating on enhanced use of 'sensitive' public sector data, with a risk of regulatory fragmentation between Member States and sectors. These differentiated approaches increase the transaction costs when developing new data-related products and services across the EU.
(b) Would national action or the absence of the EU level action conflict with core objectives of the Treaty[180] or significantly damage the interests of other Member States?
Member States action naturally is departing from industrial interests present in a Member State. For example, Franco-German Gaia-X departed from interests of the manufacturing industry, given a strong presence of original equipment manufacturers (OEMs) in Germany and France. Similarly, the Dutch iShare scheme evolved around the logistics industry that has a strong presence in that Member State. Data-sharing initiatives in agriculture are most developed in France (cf. API-Agro platform). These examples illustrate that a European approach is necessary to ensure that there are no – even

[180] European Union: *The EU in brief*.

unintentional – advantages for industrial players in any single Member State, but the approach should serve the interests of businesses in all Member States from the outset. Disadvantages can arise, e.g. if industrial players in one Member State are predominantly suppliers whereas the OEMs are typically present in another Member State (a concern voiced by Danish industry). Disadvantages could also materialize in rules of participation on the governance of common European data spaces, standardisation or other rules-setting on data sharing in such data spaces, which makes cross-sector data use more difficult. Also, data-sharing initiatives are ultimately necessary in all industries. Pace-setting initiatives in some Member States may not focus on the needs of other industrial sectors that are more strongly present in other Member States. Pioneering of data-sharing initiatives, governance mechanisms and technical platforms or architectures through Member States alone risks therefore not to sufficiently factor in industrial interests in all countries. Finally, mechanisms that enhance the use of 'sensitive' public sector data in individual Member States should neither explicitly nor implicitly favour data users of that specific Member State.

(c) To what extent do Member States have the ability or possibility to enact appropriate measures?

The internal market is an area in which Member States and the EU have shared competences. In the absence of European action on data sharing, individual Member States have taken action, in the form of legislation or by supporting industry-driven initiatives. As described above, this entails risks of – even unintentionally – favouring their own industrial players. Actions at national level cannot create access to the pan-European databases that are necessary for big data analyses or machine learning.

On the other hand, while the EU has a competence to regulate under Article 114(1) TFEU where there is an actual or potential obstacle to any of the fundamental freedoms, this proposal does not prohibit Member States' ability to enact further appropriate measures in the data economy. In particular, the proposal will leave considerable margin to Member States on the 'how' of the implementation of the rules, notably how to provide enhanced access to public sector information which is subject to rights of others.

(d) How does the problem and its causes (e.g. negative externalities, spill-over effects) vary across the national, regional and local levels of the EU?

There is no significant variation in the magnitude of the problem of insufficient data sharing and its underlying causes at national, regional or local level. However, evidence shows that data reusers in larger Member States benefit from larger (and therefore often more representative) datasets. EU-wide exchanges would allow more actors to use a large range of datasets for big data purposes (e.g. for research purposes). Also, certain Member States (including France, Germany, the Netherlands, Denmark and Finland) have

developed national strategies or supported industry-driven initiatives on data sharing that will benefit players in these countries, but would not necessarily benefit other industries with strong presence in other Member States. This could lead to further disparities and unequal distribution of benefits of digitisation within the EU.
(e) Is the problem widespread across the EU or limited to a few Member States?
The problem of insufficient data sharing is widespread across the EU, however it affects Member States with different levels of intensity. This is partly linked to a Member State's level of digitisation and the state of its national data economy. Member States are increasingly aware of the growing value of data for their economic and societal development, including for post COVID-19 recovery programmes. They are launching initiatives of varied intensity, including legislation, that aim to resolve different aspects of this problem. There are risks of doing this in potentially diverging ways, but also that economic disparities within the EU will deepen further as some Member States advance faster than others.
(f) Are Member States overstretched in achieving the objectives of the planned measure?
Member States will need to target funding at developing the mechanisms and structures proposed in this instrument. As Member States will remain in control of the extent of their novel services to enhance the better use of certain public sector information, no Member State should be stretched. In particular, the instrument does not prescribe the use of a particular technology or institutional form of the structures/bodies that need to be in place. Additionally, EU funding will be available to Member States to help with the implementation of this and other EU measures. Overall, the economic and societal benefits obtained through this intervention would be significantly higher than the costs, as explained in Chapter 6. As well as economic and societal welfare gains, businesses (in particular SMEs), researchers and citizens stand to benefit. Acting at EU level would achieve greater impact in a more effective and efficient manner.
(g) How do the views/preferred courses of action of national, regional and local authorities differ across the EU?
Member States have been unanimous in asking for action to improve data sharing at the European level. The March 2019 European Council conclusions state that: "*the EU needs to go further in developing a competitive, secure, inclusive and ethical digital economy*

with world-class connectivity. Special emphasis should be placed on access to, sharing of and use of data, on data security and on AI, in an environment of trust"[181].

In the Council conclusions on 'Shaping Europe's digital future' of 9 June 2020, Member States also unanimously called on the Commission to present concrete proposals on data governance and to encourage the development of common European data spaces for strategic sectors of industry and domains of public interest[182]. On 9 July 2020, the Digital Single Market Strategic Group, composed of Member States' representatives, discussed initial ideas for the legislative framework on the common European data spaces, based on the Inception Impact Assessment, and expressed strong support for the improvement of data governance at EU level. Views and actions differ in terms of industries that Member States focus on when supporting industry-driven data sharing and governance initiatives.

The public sector at national, regional and local level was consulted through different actions during the preparation of the initiative (workshops, online consultation). Some Member States (notably France, Finland and Germany) have shared their experience in establishing specific bodies that offer technical mechanisms to create secure and privacy-enhancing conditions for the reuse of data subject to the rights of others. This has proved to drive forward a sharing and reuse culture. In the online consultation, public authorities were clearly in favour of developing governance mechanisms to support standardisation activities for interoperability purposes. They considered that EU or national government bodies have a role to play in the prioritisation and coordination of standards. They also considered that they have a role to play in the clarification of the legal rules. Finally, they agreed that public authorities should make a broader range of sensitive data available for R&I purposes and for the public interest.

2.4 Based on the answer to the questions below, can the objectives of the proposed action be better achieved at Union level by reason of scale or effects of that action (EU added value)?

Considering the importance of economies of scale for the development of data technologies and services, coordinated action at European level will bring higher value to the European economy and society than action by individual Member States. Chapter 6 of the Impact Assessment shows the potential direct and indirect benefits of the initiative, such as greater access to data, costs savings, efficiency gains and economic and societal benefits.

While allowing the Member States to take further action in this area, the proposed instrument will ensure that future national and sectoral legislation flows from a number of horizontal principles that make data sharing fit for cross-sectoral and EU-wide exchanges. It will ensure that all industry players benefit, irrespective of their situation in

[181] Council of the European Union *Conclusions* (22 March 2019).
[182] Council of the European Union *Conclusions* (9 June 2020).

industrial supply chains. It will also ensure that all industrial sectors benefit, taking into account the respective strengths of all Member States. Furthermore, the proposed instrument brings clarity on what can be done with public sector data and by whom.
(a) Are there clear benefits from EU level action?
The instrument would lead to demonstrable benefits compared to the lack of EU action or to other policy options, as explained in Chapter 6 of the Impact Assessment. It ensures that data sharing can happen across sectors and between Member States. These benefits are achievable only at the EU level due to the scale of the internal market and the economies of scale that harmonised initiatives bring to this.
(b) Are there economies of scale? Can the objectives be met more efficiently at EU level (larger benefits per unit cost)? Will the functioning of the internal market be improved?
The measures will improve the functioning of the Single Market for Data and consequently of the Single Market as a whole. Products and services in the future will benefit at various stages (strategic investment decisions on novel products and services, product design, product processing, product surveillance and design feedback loops) from big data analysis, use of sensor (IoT) data and machine learning. For certain products or services, access to large volumes of data are important. These are hard to obtain in small or medium-sized Member States alone. Also, it should become easier for firms to adapt products or services developed in one Member State to the market of another Member State.
(c) What are the benefits in replacing different national policies and rules with a more homogenous policy approach?
A more harmonised way of data sharing across Member States and sectors would present benefits for industrial sectors and players across the value chain, irrespective of their relative strength or presence in individual Member States. Also, only a concerted European approach can present an alternative to the current business models around data dominated by Big Tech platforms and cloud computing hyperscalers. The initiative will reduce fragmentation in the legal and policy governance frameworks for data sharing (including the absence of them), which currently stands in the way of creating the common European data spaces and a data economy that is transparent, effective and accountable. In the different consultation actions conducted to prepare the initiative, stakeholders strongly supported EU measures to create a harmonised and clear set of rules on data sharing, as presented in the Annex 2 of the Impact Assessment.
(d) Do the benefits of EU-level action outweigh the loss of competence of the

Member States and the local and regional authorities (beyond the costs and benefits of acting at national, regional and local levels)?
The proposed initiative does not lead to a loss of competence in the Member States. It will 'Europeanise' the support to industry-driven initiatives such as Gaia-X and ensure that benefits of such initiatives accrue throughout the EU. The proposed initiative intends to clarify different rules and practices across the EU, which currently make it difficult for companies to develop pan-European data-based services and products. It allows Member States and sectors to complement and reinforce the data economy with initiatives that respond to their specificities.
(e) Will there be improved legal clarity for those having to implement the legislation?
Yes, the proposed instrument will ensure that future national and sectoral legislation flows from a number of horizontal principles that make data sharing fit for cross-sectoral and EU-wide exchanges. Furthermore, it will bring legal clarity on what can be done with certain data and by whom. Such clarity will be brought on three levels: First, researchers and innovators seeking to use public sector data that can only be used under strict conditions will be able to obtain guidance from public authorities on the usability and conditions of data use. Similarly, researchers and innovators that seek to use data voluntarily made available by individuals or companies for the public good will have access to schemes that give legal certainty on the rights to use such data. Finally, for B2B and C2B transactions, novel intermediaries will play a functional role in ensuring compliance with relevant laws, in particular data protection and competition law. Regarding data altruism, an authorisation regime with mutual recognition among the Member States would ensure clarity through harmonisation of the requirements necessary to provide these kind of services. Mutual recognition mechanisms will ensure that European data intermediaries can operate across EU countries.
3. Proportionality: How the EU should act
3.1 Does the explanatory memorandum (and any impact assessment) accompanying the Commission's proposal contain an adequate justification regarding the proportionality of the proposal and a statement allowing appraisal of the compliance of the proposal with the principle of proportionality?
The Explanatory Memorandum of the proposal, as well as the Impact Assessment (Chapter 3 – 'Why should the EU act?'), contain dedicated sections on subsidiarity and added value, which also address the proportionality question. The sections of the Impact

Assessment that deal with the impacts of the policy options (Chapter 5) and the way the different actors are affected by the initiative (Annex 3) also demonstrate that the initiative is in line with the proportionality principle.

The initiative is proportionate to the objectives sought. The Member States are the main stakeholders that will be required by the legislation to put in place the necessary processes and structures or bodies to address the problems. However, the initiative will leave a significant amount of flexibility for implementation at national and sector-specific levels, including through the European data spaces. Also, for B2B data sharing, the initiative essentially fosters the emergence of novel intermediaries.

3.2 Based on the answers to the questions below and information available from any impact assessment, the explanatory memorandum or other sources, is the proposed action an appropriate way to achieve the intended objectives?

The proposed legislation will induce financial and administrative costs to be borne essentially by national authorities. However, the exploration of different options and their expected costs and benefits led to a balanced design of the instrument. It will leave enough flexibility for national authorities to decide on the level of financial investment and possibilities to recover such costs through administrative charges or to take additional measures, while offering overall coordination at EU level (e.g. through a European structure for coordinating the governance aspects of data sharing).

(a) Is the initiative limited to those aspects that Member States cannot achieve satisfactorily on their own, and where the Union can do better?

The proposed initiative only focuses on areas where there is a demonstrable advantage in acting at EU level, and on problems identified and described in its Impact Assessment. It pursues objectives that cannot be sufficiently achieved by the Member States acting alone, as described in section 2.3 above, due to the scale, speed and level of transnational coordination needed.

(b) Is the form of Union action (choice of instrument) justified, as simple as possible, and coherent with the satisfactory achievement of, and ensuring compliance with the objectives pursued (e.g. choice between regulation, (framework) directive, recommendation, or alternative regulatory methods such as co-legislation, etc.)?

In line with the Better Regulation Guidelines, a multi-criteria analysis carried out for the accompanying Impact Assessment has explored several options for each of the measures foreseen. A comparative assessment of the merits of each option also included its efficiency, the effectiveness, the coherence, the legal/political feasibility and the proportionality (Chapter 7). The option of coordination at EU level and soft measures

only was discarded based on the cost-benefit analysis, as it would not significantly change the situation as compared to the baseline scenario. Existing soft law measures in the field have shown that, although useful in providing certain clarity and giving a general direction, they have been taken up with different intensities by actors in the data economy and Member States.

The choice of a regulation as the form of the legal instrument is justified by the predominance of elements that should not leave margins to implementation such as the authorisation of data altruism mechanism, labelling of novel data intermediaries and the setup of coordination structures at European level. The direct applicability of the Regulation would avoid an implementation period for the Member States, so that the establishment of the common European data spaces could start very soon, in line with the EU Recovery Plan. At the same time, the provisions of the Regulation are not overly prescriptive and leave room for different levels of Member State action for elements that do not undermine the objectives of the initiative, in particular the organisation of the mechanisms supporting the reuse of 'sensitive' public sector data.

(c) Does the Union action leave as much scope for national decision as possible while achieving satisfactorily the objectives set? (e.g. is it possible to limit the European action to minimum standards or use a less stringent policy instrument or approach?)

The proposed legislation will require that national authorities put in place mechanisms to support the reuse of 'sensitive' public sector data, in order to address the problems presented in the Impact Assessment. However, the initiative will leave Member States some flexibility for implementation of such mechanisms.

(d) Does the initiative create financial or administrative cost for the Union, national governments, regional or local authorities, economic operators or citizens? Are these costs commensurate with the objective to be achieved?

The proposed legislation will create financial and administrative costs, mainly for the Union (creating a European structure for the governance aspects of data sharing) and national governments (putting the structures and processes in place for the reuse of certain public sector data, the authorisation of data altruism schemes and the certification of data intermediaries). However, the legislation will allow for those costs to be recovered directly through administrative charges from the beneficiaries. Discretion on this matter is left to Member States with ceilings set in order to avoid that charges become prohibitive.

The initiative will also benefit public sector bodies and allow them to deliver a better public service around public databases, as they will benefit from technological and legal expertise – also through EU-level coordination. With more private sector data available

for reuse, the public sector will also make more use of such data. This will lead to more evidence-based policy- and decision-making and, ultimately, to better and more efficient public services.

Economic operators will be the main beneficiaries of the initiative. The reuse of public sector data subject to rights of others will be subject to administrative charges. However, it is expected that operators will only reuse such data (and pay the charges) if on balance they expect a positive economic outcome. Similarly, they are expected to incur costs in relation to the authorisation of data altruism schemes, however the benefits are expected to outweigh such costs. Similarly, the certification costs for novel data intermediaries will translate into service fees for business users. It is expected that also here, businesses will only make recourse to such intermediaries if the reduction in friction and transaction costs outweigh those costs. Finally, through enhancing interoperability at the technical level and making available generic enabling standards, the initiative will lower transaction costs of data sharing and facilitate EU-wide and cross-sector data sharing. The prioritisation of standards will in particular benefit SMEs.

(e) While respecting the Union law, have special circumstances applying in individual Member States been taken into account?

In general, there are no such special circumstances. The issues in relation to data sharing are very similar across all Member States. The initiative will leave flexibility to Member States when implementing the legislation, also in order to allow recently established national initiatives on enhanced reuse of public sector data to continue to exist in its present form. It will also build on existing initiatives to support industry-led approaches to data sharing such as Franco-German Gaia-X, Dutch iShare or Finnish Sitra/IHAN.

ANNEX 6: GOVERNANCE FRAMEWORK IN THIRD COUNTRIES

Over the last few years, several countries around the world have put in place mechanisms to boost their data economies by enhancing trust in data sharing. Some of the elements foreseen in this initiative take inspiration from experiences in other countries in the field of data intermediaries and certification schemes for data intermediaries.

Japan: certification schemes for information banks

The Japanese government has tried to increase trust towards Japanese 'information banks' - systems for securely accessing personal data with the data subject's consent[183] - by releasing guidelines on the functions of certification schemes.

Certification is based on government guidelines[184] compiled in June 2018 for information banking services to use personal information while protecting the privacy of individuals. Certification remains voluntary. As part of the certification process, an internal auditing body checks what is done with the personal data. Under the system, information banks allow client firms to tap into their databases only after obtaining consent from data-supplying individuals. In addition, the individuals can select the types of data to be used and grant specific firms access to their data.

In July 2019, the Information Technology Federation of Japan certified FeliCa Pocket Marketing Inc. and Sumitomo Mitsui Trust Bank as information banks[185].

Endorsement by the federation is not compulsory to commercialise personal data, but it does enhance the credibility of companies as safe data providers.

The Aeon Co. unit and the trust bank have started operations using personal data that they hold. FeliCa, which offers reward points and e-money unique to municipalities trying to reinvigorate their local economies, plans to provide its individual customers' data to local retailers and small firms to help them set up business strategies. Sumitomo Mitsui Trust aims to capitalise on its database in the healthcare field.

Republic of Korea: data vouchers as intermediaries

The Korea Data Agency[186] started the 'data voucher' programme to promote data-based innovative businesses and the surrounding data ecosystem. Data vouchers can be used to purchase and process a dataset by small companies that have difficulties in using data, and promoting the data and AI industries by expanding transactions of purchased and processed data.

[183] D. A. Consortium (2018). *Pilot testing begins on an "information bank," a new system for storing personal data*, News Release.
[184] Ministry of Economy, Trade and Industry (2018). *Release of the Guidelines of Certification Schemes Concerning Functions of Information Trust ver. 1.0.*
[185] Jiji (2019). *Japan grants certification for first time to 'information banks'*, The Japan Times.
[186] *Korea Data Agency.*

They play an intermediary role between the companies that hold the data and the companies that need the data. They are part of the Korean Government's (Ministry of Science and ICT) initiative to support SMEs in accessing data. The programme is popular, with on average 1 000 data purchases and 640 data processing activities supported annually. Some 2 795 companies applied for the voucher programme last year[187].

On 11 May 2020, the South Korean Financial Services Commission (FSC) launched a financial data exchange programme to facilitate data transactions between buyers and sellers from financial institutions, fintech firms, retailers and telcos. It will serve as a platform to match data providers and recipients as needed. The FSC will also consider introducing a data voucher programme to help set appropriate data prices and support data purchases[188].

Australia: public sector designation of trusted data-sharing platform

This example shows that governments can act as or create a trusted third party for data-sharing relationships. In 2017, the Australian government initiated the 'Data Integration Partnership for Australia' (DIPA) as an investment to maximise the use and value of the government's data assets[189].

DIPA is a whole-of-government collaboration including more than 20 Commonwealth agencies. It improves technical data infrastructure and data integration capabilities across the Australian public service. The agencies make available important data assets such as in the health, education and social welfare sectors, allowing policymakers to gain insights that were not possible before. Sectoral hubs of expertise, independent entities that are funded by the Commonwealth and denominated Accredited Integrating Authorities, enable the integration of those longitudinal data assets. Individual privacy and the security of sensitive data are preserved, as DIPA only provides access to controlled, de-identified, and confidentialised data for policy analysis and research purposes.

[187] Information provided by the EEAS Delegation to Seoul.
[188] Pulse (2020). *Korea to launch financial data exchange in March*.
[189] Australian Government, Department of the Prime Minister and Cabinet (2017). *Data Integration Partnership for Australia*.