Data & Policy (2024), 6: e30
doi:10.1017/dap.2024.20

RESEARCH ARTICLE

Private sector trust in data sharing: enablers in the European Union†

Jaime Bernal

Joint Research Centre, European Commission, Varese, Italy
Email: jaimebernal743@gmail.com

Received: 05 November 2023; **Revised:** 15 March 2024; **Accepted:** 18 March 2024

Keywords: data altruism; data asymmetry; data governance; data intermediaries; data sharing

Abstract

Enabling private sector trust stands as a critical policy challenge for the success of the EU Data Governance Act and Data Act in promoting data sharing to address societal challenges. This paper attributes the widespread trust deficit to the unmanageable uncertainty that arises from businesses' limited usage control to protect their interests in the face of unacceptable perceived risks. For example, a firm may hesitate to share its data with others in case it is leaked and falls into the hands of business competitors. To illustrate this impasse, competition, privacy, and reputational risks are introduced, respectively, in the context of three suboptimal approaches to data sharing: data marketplaces, data collaboratives, and data philanthropy. The paper proceeds by analyzing seven trust-enabling mechanisms comprised of technological, legal, and organizational elements to balance trust, risk, and control and assessing their capacity to operate in a fair, equitable, and transparent manner. Finally, the paper examines the regulatory context in the EU and the advantages and limitations of voluntary and mandatory data sharing, concluding that an approach that effectively balances the two should be pursued.

Policy Significance Statement

It is necessary to widen access to valuable data to contribute to the public good and address a persistent data asymmetry in society that disempowers consumers and consolidates the power of a few market players. Given that most data are generated by the private sector, a critical policy challenge is to increase trust in data sharing among firms. Yet, when offered opportunities to engage in data sharing, firms can hesitate, as they might perceive that they lack enough control over their data in the face of unmanageable risks. This paper addresses this question by proposing, analyzing, and assessing the trustworthiness of seven trust-enabling mechanisms.

1. Introduction

Society needs to urgently broaden access to valuable data generated in the private sector to contribute to the public good (Kirkpatrick, 2013; Alemanno, 2018; Susha et al., 2019a). Fortunately, data can be exploited multiple times by different actors "without reducing the amount of data available to anyone else" (Tonetti and Jones, 2020, p. 2819). Data sharing is expected to accelerate scientific research in key

† The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

areas such as medicine and environmental sciences (Tenopir et al., 2011), while also proving beneficial in addressing global challenges such as pandemics and migration (Spyratos et al., 2019; Iacus et al., 2020). From an economic perspective, the aggregation of data from different sources has the potential to generate economies of scope, which exist when the cost of production is lowered through diversification (Teece, 1980). Instead of keeping data separated in silos, two or more parties can share complementary datasets to derive more insights and value while decreasing the total amount spent per dataset (Martens, 2021). In short, data sharing is practically feasible, economically viable, and advantageous in numerous instances.

Most data sharing occurs in a bilateral Business-to-Business (B2B) context, as businesses possess the necessary technological means and a unique position in society to collect, retain, share, and profit from data generated by a large number and range of actors (Spiekermann, 2019). However, despite the abundance and value of datasets, widespread data sharing and trading remain limited (Koutroumpis et al., 2020). The status quo threatens to further widen a data asymmetry in society that disempowers consumers and consolidates the power of a few market players (Alemanno, 2018; Zuboff, 2019). Thus, unleashing the potential of private sector data sharing becomes a critical policy challenge (Koutroumpis et al., 2020).

Recognizing the risks and benefits at stake, the European Commission designed the European strategy for data, which includes the creation of common European data spaces for sector and domain-specific initiatives to spur competition and innovation while addressing societal challenges and improving public services (European Commission, 2020a). Essential to the strategy are the Data Governance Act and Data Act, which have entered the implementation phase and govern two regulatory models for data sharing—voluntary and mandatory.

While opening channels for businesses to share data is an integral tactic to this strategy, a significant barrier arises due to the distrust among firms that sharing data might harm their business interests. To overcome this trust deficit, policymakers can implement "trust-enabling mechanisms," which can be succinctly defined as measures and tools to build business confidence around data sharing.

1.1. Trust as the glue for data sharing

In the official communication of the European strategy for data, lack of trust among businesses is cited as a major cause for the suboptimal levels of data sharing (European Commission, 2020a). The rationale behind the policy interventions of the Data Governance Act and Data Act is rooted in the recognition that without enough legal clarity about who can use the shared data for what objectives, businesses will fear that their data will be used for unagreed purposes by third parties against their business interests. This deficit of trust is present in both B2B and Business-to-Government (B2G) data sharing relations (Klievink et al., 2018; Spiekermann, 2019).

Trust is understood as a "relationship in which an agent (the trustor) decides to depend on another agent's (the trustee) foreseeable behaviour in order to fulfil his expectations" (Taddeo and Floridi, 2011, p. 1). In this context, the trustor is a private firm that holds valuable data but is vulnerable to the behavior of a trustee, who might defy its expectations by exploiting data at its expense. Because data sharing is still a relatively new phenomenon, this lack of trust in data sharing could be attributed to the absence of opportunities to cultivate trust in the past. Yet, building trust will continue to be particularly challenging in scenarios where data sharing creates unmanageable conditions of uncertainty for the trustor.

Under such circumstances, the "perceived risk" by a firm, which refers to the perception of experiencing a loss as a result of an uncertain event, ought to be taken into account (Agahari et al., 2022, p. 1582). Equally important is the notion of "usage control," or a firm's perceived capacity to influence how the data are used by others (Zrenner et al., 2019, p. 478). These two perceptions enable and disable trust, indicating that balancing the two is necessary. Nevertheless, the notion of "trust" transcends interpersonal relations. In this context, this one refers to the subjective beliefs of decision-makers in a company. Therefore, a more comprehensive understanding is required to account for the dynamics of trust production when an external agent, like a regulator, implements trust-enabling mechanisms to instil trust.

While understanding businesses' risk and usage control perceptions toward trust-enabling mechanisms is key, assessing the trustworthiness of these and the regulatory attempts aimed at producing trust

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

through them is also crucial. Thus, the paper considers whether such measures and tools can operate in a fair, equitable, and transparent manner, which opens up the space to examine the institutional dimension of trust production (Bodó, 2021).

1.2. Methodological remarks

Addressing the question of how to overcome the barriers to enable trust in data sharing scenarios involving businesses becomes crucial to contributing to the success of the European strategy for data and similar initiatives. In light of this challenge, the paper poses the following guiding research question: *Can trust-enabling mechanisms to increase data sharing be trusted?*

Section 2 presents the problem statement by illustrating how three types of common risks—competition, privacy, and reputational—generate distrust when businesses engage in three emerging approaches to sharing data: data marketplaces, data collaboratives, and data philanthropy, respectively. The selection of these exemplary cases has been made taking into consideration that, despite being perceived as avenues to unlock the value of private sector data, they have failed to scale to a satisfactory extent (Lev-Aretz, 2019; Spiekermann, 2019; Ruijer, 2021).

Section 3 departs from the premise that, in order to increase trust, it is essential to first enable it by increasing the usage control perceived by firms and mitigating perceived risks. Given the multidimensional complexity of data sharing and the interplay of technological and social factors that shape trust in this practice, this is expected to be achieved by exploiting technological, legal, and organizational trust-enabling mechanisms. Seven diverse types have been selected and analyzed from grey and academic literature due to their potential to increase data sharing. From a critical angle, the capacity of these to operate in a trustworthy manner is assessed.

Finally, Section 4 discusses the regulatory context in the EU, applying the insights drawn from the previous sections to the Data Governance Act and Data Act and assessing the regulatory attempts to build trust around data sharing.

2. Perceived risks

Among the most common risks faced by businesses hindering data sharing are competition, privacy, and reputational risks. To exemplify how these risks can affect a firm's behavior, they are, respectively, introduced in the context of three emerging models for sharing data: data marketplaces, data collaboratives, and data philanthropy. It is worth noting that the list of risks extends well beyond those discussed in this section and that there are important synergies among them—they can empower each other.

2.1. Data marketplaces and competition risks

As multisided platforms for connecting data holders with data users, data marketplaces facilitate the collection and aggregation of data from various sources, allowing it to be processed, refined, and traded (Stahl et al., 2016; Abbas et al., 2021). The well-functioning of data marketplaces is perceived as essential for unlocking the value of business data across sectors (European Commission, 2018). However, their unsatisfactory growth shows that, despite the possibility of making profits by monetizing their data, firms continue to be discouraged due to persistent competition risks (Spiekermann, 2019). Although significantly different across industries, data marketplaces in the automotive industry are relatively more developed (Bergman et al., 2022), offering an example to analyze how competition risks work in practice.

Connected cars produce massive amounts of in-vehicle data that can be used by businesses to optimize performance, ensure safety, enhance automatization, etc. (Siegel et al., 2017). Car manufacturers or original equipment manufacturers (OEMs) have exclusive access to the generated data as it can be directly transferred to their servers. As a result, other firms within the ecosystem are excluded from competing, leading to less consumer choice and innovation (Kerber, 2018). As a result of this market failure due to monopolistic dynamics, a policy intervention was deemed necessary. In response, the EU formulated the Sustainable and Smart Mobility Strategy and proposed a common European mobility data space to

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

broaden access to these data, among others, through the promotion of data marketplaces (European Commission, 2020b).

A major reason contributing to the unsatisfactory growth of data marketplaces in the automotive industry is that OEMs are concerned about losing control over their data, as this one could be exploited against their business interests (Agahari et al., 2022). Retaining access to the insights, their data might generate becomes a safer bet than trusting others with it, even if this entails renouncing profits. Furthermore, even if an OEM sells its data through a data marketplace to a selected third party that perceives it poses no competition risks a priori, there is no absolute guarantee that these data will not be leaked and fall into the hands of a direct business competitor. In other words, firms have strategic concerns about trusting others with their data, as there is a possibility of losing their competitive advantage if commercially sensitive data fall into the hands of business competitors (van den Broek and van Veenstra, 2015; Martens and Zhao, 2021).

2.2. Data collaboratives and privacy risks

Data collaboratives, which are understood as cross-sector partnerships involving private and public actors in the collection, processing and/or exchange of data to address a specific social issue, have been underscored by the expert group on B2G Data Sharing appointed by the European Commission (Susha et al., 2017a, p. 2691; European Commission, 2020c). Exploiting data collaboratives and similar forms of data-driven social collaborations that engage a variety of actors in the sharing of data is a major component of the European strategy for data (European Commission, 2020a). Businesses can be incentivized to participate in these partnerships through both financial and non-financial means (Susha et al., 2017b). Nonetheless, privacy risks are an important barrier to the growth of this novel approach to data sharing.

It makes a big difference for firms to share personal or non-personal data. Due to the presence of privacy regulations, when personal data enter the equation, companies adopt a more stringent and hierarchical control mechanism (even in situations where there are no apparent risks of privacy breaches) (van den Broek and van Veenstra, 2015). Processing and sharing personal data in the EU entail complying with the General Data Protection Regulation (GDPR), which establishes six legal bases for processing data, enforces seven data protection principles, and grants eight rights to data subjects (European Commission, 2016). To avoid going through the complex process of adhering to this regulation and not facing any penalties for non-compliance, a company might refrain from participating in data collaboratives. Moreover, even for non-personal data, re-identification risks continue to persist, given that the combination of data from various sources (e.g., geographical data containing addresses) can lead to such risks (van den Broek and van Veenstra, 2015).

When the sharing of data occurs across multiple entities, the likelihood of security breaches and unauthorized access that exposes the identity of individuals increases. Things get even trickier when a data collaborative involves the sharing of data across borders, as participant firms risk overlooking other countries' privacy laws. Ultimately, privacy risks can result in regulatory risks (i.e., the legal and financial repercussions of failing to adhere to data protection regulations) as well as reputational risks.

2.3. Data philanthropy and reputational risks

Data philanthropy, also referred to as data donorship, is the donation of data by firms for a humanitarian purpose (Kirkpatrick, 2013; UN Global Pulse, 2014; Taddeo, 2017). The purpose of a data philanthropy project is closely linked to that of data collaboratives, namely, contributing to a social cause (Susha et al., 2019b). Yet, a fundamental difference between the two approaches is that data philanthropy exclusively engages private firms in the sharing of data without a profit incentive (Taddeo, 2017). In addition, data collaboratives are more encompassing and imply a greater degree of collaboration among participants than data philanthropy (Lev-Aretz, 2019).

The expert group on B2G data sharing considered data philanthropy to be a form of corporate social responsibility key for making data available during public emergencies (European Commission, 2020c).

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

An example is Facebook's Data for Good initiative, which contributed to the response to the COVID-19 emergency by making mobility datasets available to Italian health researchers (Kang-Xing and McGorman, 2020; Scotti et al., 2022). In this apparent win–win data sharing scenario, researchers could access valuable data and Facebook could expand its market reach and improve its brand's reputation. Nonetheless, despite providing many marketing opportunities for firms, data philanthropy can also produce negative outcomes that damage their reputation (Lev-Aretz, 2019).

Firms are exposed to reputational risks when data sharing results in negative public perceptions of them. When sharing data with others, there is a risk that these data are used for purposes other than those that were agreed upon, including unethical ones. Such mishandling of data could produce unintended consequences that might result in a public backlash. In addition, firms might be hesitant to engage in data philanthropy because they might not want to disclose what type of data they collect and how much of it. Revealing this information could create negative perceptions and open the door for misconceptions. Lastly, a company may fear sharing inaccurate or incomplete data that lead to erroneous conclusions.

3. Trust-enabling mechanisms

The previous three examples show that the success of emerging forms of data sharing with the potential to contribute to tackling societal challenges (as envisioned by the European Commission) is dependent on mitigating associated business risks. Against this backdrop, trust-enabling mechanisms enter the picture due to their capacity to foster trust by mitigating perceived risks and increasing usage control.

Relying exclusively on technological solutions to solve social problems can fail to address the issue and lead to miscalculated outcomes (Morozov, 2014). Building trust should not exclude the more social and cultural aspects in which this form of social capital is intrinsically embedded; hence, technological solutions must be complemented by governance and external accountability mechanisms (Bodó, 2021). This more holistic approach aligns more adequately with the European Commission's emphasis on creating a data sharing culture (European Commission, 2020a). Furthermore, the cost of certain trust-enabling technologies is often established by market dynamics and can entail significant expenses.

Without rigidly classifying them as either technological, legal, or organizational, seven trust-enabling mechanisms (privacy-enhancing technologies, data intermediaries, data exchange platforms, government support, data sharing agreements, regulatory sandboxes, and data stewardship) to increase data sharing are introduced, emphasizing their individual and collective contribution to mitigating perceived risks and increasing usage control, as well as their capacity to operate in a fair, equitable, and transparent manner.

3.1. Privacy-enhancing technologies

Privacy-enhancing technologies **(**PETs) consist of a range of cryptographic as well as non-cryptographic techniques and methodologies to protect data from malicious forces and safeguard privacy (Heurix et al., 2015). In a data sharing scenario, the involvement of a trusted third party may be required to mediate between two or more parties and to possibly act as a certification authority to handle user registration and authentication (Heurix et al., 2015).

PETs have varying degrees of applicability, so there is no one-size-fits-all solution to every data sharing scenario. The suitability of each PET is highly context-dependent, as each has its own specific strengths and limitations. For instance, secure multiparty computation,[1] federated learning,[2] and trusted execution

[1] Secure multiparty computation refers to a group of techniques in which distrusting parties jointly compute a function without revealing any more information about their input than what can be inferred from the output of the computation (Beaver, 1992; Keller et al., 2018). This technique has grown notably in the past years and has been applied in practical scenarios successfully without the need for a trusted-third party, including in the automotive industry (Agahari et al., 2022). However, despite receiving industry recognition, it requires great computational effort to employ secure multiparty computation due to the increased complexity of operating with encrypted data compared to unencrypted data (Lindell, 2021).

[2] Federated learning allows multiple data holders to collaborate in building a machine learning model through multiple independent servers (Kairouz et al., 2021).

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

environment[3] regulate access to the data (Beaver, 1992; Sabt et al., 2015; Kairouz et al., 2021), while homomorphic encryption[4] and zero-knowledge proofs[5] hide the data (Rivest et al., 1978; Goldwasser et al., 1989) and differential privacy[6] and anonymization[7] camouflage it (Samarati and Latanya, 1998; Dwork and Roth, 2013). Moreover, in a data sharing scenario, the use of multiple PETs may be required (even in combination with other technologies like blockchain) (Jia et al., 2022).

PETs can be difficult and costly to implement (Eurich et al., 2010) and may contradict ethical principles such as data minimization (i.e., limiting the collection of data to what is absolutely necessary). Because PETs cannot be flawless, they can produce a false sense of security (Renieris, 2021). Furthermore, it is unclear whether the mere presence of PETs would enable trust or if a sufficient understanding of the technique is necessary by the parties involved in the sharing of data. In addition, another persistent challenge in implementing PETs is securing information while maximizing the utility of the data. Ultimately, as PETs continue to evolve and be employed in combination with other technologies, questions about their efficacy remain, especially in light of the present and emerging instruments employed for re-identification purposes (Ohm, 2010).

3.2. Data intermediaries

Data intermediaries are an emerging type of actor in the data economy that can be broadly defined as mediators "between those who wish to make their data available and those who seek to leverage that data" (Janssen and Singh, 2022a, p. 2). In certain commercial and non-commercial settings, public and private actors can coordinate the supply and demand of personal and non-personal data more efficiently through a data intermediary that provides the necessary infrastructure to match data holders and users (Richter and Slowinski, 2019).

Data intermediaries can reduce transaction costs (e.g., by guaranteeing interoperability), exploit economies of scale, and capture positive externalities through the application of new technologies and organizational techniques (Martens et al., 2020). They can provide a robust contractual framework for enforcing obligations and facilitating compliance with data protection regulations (von Grafenstein, 2022; Richter, 2023). Regarding risks related to data security and privacy, neutral tools for managing data access and permissions can be supplied by an intermediary, which could also act as a trusted third party to supervise the use of PETs. Furthermore, intermediaries can incorporate in data sharing agreements contractual transactions absent in B2B contexts that mitigate post-contractual risks (Martens et al., 2020, p. 29). For example, *Advaneo*[8] is a data marketplace for businesses to monetize and manage their data in a privacy-preserving way, setting rights and obligations for buyers and sellers. Thanks to its

[3] Trusted-execution environment provides a secure environment that is isolated and protected from intrusions and attacks (Sabt et al., 2015). While the trusted-execution environment is not considered a technique in the same way as the others, it is a broader technology framework that can be used to provide an additional layer of security.

[4] Homomorphic encryption allows data users to conduct operations on encrypted data without having to decrypt it (Rivest et al., 1978; Fontaine and Galand, 2007). This technique is set to become a boon in the healthcare business for its capacity to restrict access to encrypted data to service providers only while maintaining the privacy of patients (Munjal and Bhatia, 2022).

[5] Zero-knowledge proofs are a set of techniques that allow distrusted parties to prove the validity of a statement without disclosing any extra information to each other (e.g., sharing a password without revealing it) (Goldwasser et al., 1989). Zero-knowledge proofs can be particularly costly to implement and apply to large datasets, as an excessive number of complex operations would be necessary to be computed on specialized devices.

[6] Differential privacy is a technique whereby noise is added to the data to prevent the identification of an individual while still making it possible to perform statistical analyses with the data (Dwork and Roth, 2013). Differential privacy has been employed to train machine learning models based on neural networks for its capacity to protect crowdsourced and sensitive information with a minimal amount of privacy loss (Abadi et al., 2016).

[7] Anonymisation is the process of removing identifiable information about an individual. However, reidentification risks are notable when using this technique (Bayardo and Agrawal, 2005). Despite the fact that the k-anonymity model, which states that each data record should be indistinguishable from at least k-1 records, can be used to assess the risk of re-identification (Samarati and Latanya, 1998), the model suffers from serious theoretical and empirical limitations when de-anonymizing diverse and highly dimensional input datasets (Abadi et al., 2016, p. 316).

[8] https://www.advaneo.de/en/#

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

architecture, data providers always retain sovereignty over the raw data, which is only accessible through metadata.

Although the central position of data intermediaries like data marketplaces is advantageous for mitigating power and information asymmetries, a key informational challenge persists due to the intrinsic characteristics of data. The complexity of estimating its value prior to its utilization makes data an asset that can be more difficult to share and trade than tangible goods (Spiekermann, 2019). Buyers might not seek to acquire certain data without a clear vision of the benefits it can provide. Meanwhile, a seller might decide not to assume the abstract risks (e.g., competition, privacy, and reputational risks) involved in disclosing data without a sufficient understanding of its concrete value (von Grafenstein, 2022). As a result, data can be perceived as overvalued for buyers and undervalued for sellers. Therefore, intermediaries cannot avoid that when data are priced according to market dynamics, in which buyers and sellers negotiate its cost, this process will reflect the relative bargaining power and information held by the parties.

While at first, a data intermediary may appear to be neutral in a way that avoids prioritizing the interests of the parties involved, its role, model of governance and, incentive structures are highly contextual, so the strategic aims of the organization will be reflected in its business model (Richter and Slowinski, 2019). In other words, the growth objectives of the entity behind the data intermediary will influence the way it operates, possibly affecting its neutrality. How a data intermediary configures its services, pricing strategies, partnerships, technologies and other aspects of its business is likely to reflect its long-term goals, which can influence its capacity to enable trust.

3.3. Data exchange platforms

Data exchange platforms provide businesses with the software tools to share their data. To enable trust, their architecture ought to be designed in a way that maximizes usage control and minimizes perceived risks, with a data intermediary potentially serving as the organization responsible for overseeing the management of these data sharing systems.

Merely determining who can access specific data may be insufficient for a business that considers it crucial to also define the purposes for which its data must be used or the duration for which the data are made available (Pearson and Casassa-Mont, 2011). To further increase usage control, data exchange platforms could allow a company to monitor data usage in real time. This way, a company will be more likely to trust other agents with its data, as it can reverse or amend its decisions if expectations are not met (Carballa Smichowski et al., 2021, p. 4).

An example of a data exchange platform is the *Snowflake Data Exchange*[9], which allows organizations to share data with business partners. This private platform gives users the possibility to supervise access, audit usage, and implement security controls. Regarding the latter, secure authentication mechanisms and intrusion detection systems can provide a greater sense of control to enable trust. Other platforms like *Dawex*[10] operate with distributed technologies such as blockchain to lessen the need to place trust in others by constraining data users' actions. To ensure privacy and further protect data, PETs can be integrated into these platforms. Such a mix of technological solutions employed to enable trust in data sharing illustrates the societal shift from interpersonal trust relations mediated by humans to trust produced by technological intermediaries.

Even if a platform claims to have robust security measures, the lack of control over the data can be unacceptable for certain businesses, which must ensure business confidentiality and comply with relevant privacy and data protection regulations. Furthermore, purchasing the services of a platform can be costly. Hence, governments can intervene by constructing infrastructure to operate data exchanges and offering them as a public service alternative. *Gaia-X*[11] is the most prominent example in the European context. In addition, notable examples include the *Asynchronous Data Exchange*

[9] https://docs.snowflake.com/en/user-guide/data-exchange
[10] https://www.dawex.com/en/blockchain-data-exchange/
[11] https://gaia-x.eu/what-is-gaia-x/about-gaia-x/

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

(ADEX)[12] in Singapore, the *Amsterdam Data Exchange (AMdEX)*[13], and the *Shanghai Data Exchange*[14] (which saw the trading of data products surpass $1 billion in 2023).

3.4. Government support

Government support for enabling trust in private sector data sharing can be implemented through regulatory frameworks that establish guidelines for sharing data within or across different sectors. This is crucial, as the lack of legal recognition might have contributed to the unsatisfactory growth of the three approaches previously discussed: data marketplaces, data collaboratives, and data philanthropy (Lev-Aretz, 2019; Spiekermann, 2019; Susha et al., 2019b). A regulatory framework can increase accountability, an important element to enable trust (Bodó, 2021).

A more structured regulatory environment with clearer guidelines can pave the way for government support to take the form of incentives, including subsidies, tax breaks, and grants. For example, *Data4Industry-X*[15] is a decentralized data exchange for Industry 4.0 that has been backed by the French government's France 2030 initiative and the Next Generation EU funding program.

Nevertheless, unlocking the value of data is partly the result of constant engagement among stakeholders, trial-and-error experimentation and harmonizing conflicting interests (Günther et al., 2017). Taking shortcuts through government support is susceptible to several challenges. This one can backfire if it is perceived as overly stringent and an impediment to firms' ability to engage in data sharing independently. Moreover, an enforceable legal framework like the Data Governance Act and Data Act can lead to risk aversion among firms that want to avoid complications. Government support can also lead to regulatory capture, where the interests of certain industry players are prioritized over others. Moreover, in highly evolving data-driven industries, support can arrive late due to the difficulties of reaching agreements swiftly. Ultimately, since public institutions can build and undermine trust, supporting data sharing can backfire and erode trust in public institutions (Bodó, 2021).

3.5. Data sharing agreements

An agreement can enable trust by increasing transparency and further guaranteeing accountability among the parties, thereby contributing to firms' confidence that their rights are defined and backed by legal assurances. Such a framework clarifies the roles of the parties involved, the purposes of sharing data, the procedures to be followed and the standards to be met (Information Commissioner's Office, 2021; Sitra, 2022). Regarding special categories of data, agreements are complemented with additional clauses that deal with relevant data protection principles. In the case of ambiguities, the data holder can impose specific restrictions, exceptions, or territorial limitations on the use of data as well (Association of Banks in Singapore, 2019).

The scope of the data sharing agreement can be delineated and any technical terms can be defined in a glossary. Responsibilities should be assigned to the parties and their rights and obligations should be defined. Additional clauses of the agreement can include elements such as confidentiality, intellectual property, liabilities, force majeure, auditing, termination, validity, and dispute resolution. For certain activities, agreements state the specific purposes for which the data will be used and the reasons why that particular data are needed to help ensure that the data will not be misused. Furthermore, a section stating the penalties for non-compliance can be added. Finally, any relevant laws and regulations should be invoked.

Nonetheless, even a well-defined agreement has important drawbacks to consider. Since contracts will always be open to interpretation, there will inevitably be room for disputes. Additionally, the framework

[12] https://www.developer.tech.gov.sg/products/categories/sensor-platforms-and-internet-of-things/asynchronous-data-exchange/overview.html

[13] https://amdex.eu/

[14] https://www.shanghai.gov.cn/nw48081/20230428/c62aaf00a259476b8926368361ec4ab5.html

[15] https://www.data4industry-x.com/

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

cannot be future-proof because of its inability to account for unforeseen events. This situation creates a paradox: The agreement must strike a balance between clarity to enable trust while remaining flexible enough to accommodate the latest developments in the changing technological landscape of data-driven industries. Furthermore, as business strategies change, what was once a mutually beneficial agreement could turn into a situation where the interests of the parties diverge and, in cases where the contract is not adequately enforced, trust will be eroded.

3.6. Regulatory sandboxes

Regulatory sandboxes are a novel regulatory tool in emerging industries to foster innovation while ensuring consumer protection (Allen, 2019). They provide a controlled environment for companies to test new products and services under the supervision of regulators (European Council, 2020). By granting exemptions from certain regulatory requirements, this tool offers an opportunity for firms to engage in less risky data sharing initiatives in order to understand how new data sharing technologies work in practice (Datasphere Initiative, 2022).

For instance, the government of Singapore invited firms interested in working with PETs to participate in a sandbox.[16] In a similar vein, to foster innovation, the government of Japan encouraged firms operating in the country to test Internet of Things (IoT), artificial intelligence (AI), and blockchain technologies.[17] The field of IoT can especially benefit from regulatory sandboxes, as there is an increasing number of data-driven social partnerships between public, private, and non-governmental actors based on the collection and aggregation of data from devices (Susha et al., 2019a).

Although sharing data in a supervised setting within a specified territory for a predetermined period can enable private sector trust, it should be noted that sandboxes can result in a kind of government-granted privilege in favor of participant organizations at the expense of non-participant competitors and newcomers (Poncibò and Zoboli, 2022). A paradox arises when a sandbox achieves its goal of being attractive enough for firms to want to join but weakens those firms that did not join (Poncibò and Zoboli, 2022). Despite lowering the entry barriers in innovation industries for newcomers, they can also weaken competition, creating a division between participants and non-participants. Since sandboxes can be considered a form of government support, it could be the case that, when not implemented properly, trust in both data sharing and public institutions is eroded.

3.7. Data stewardship

Data stewardship is an approach to the governance of data that integrates technical and organizational infrastructures in an organization to responsibly collect, store, use, and share data (Rosenbaum, 2010; Stalla-Bourdillon et al., 2021). The hands-on management and care of data on a day-to-day basis can be taken on by the role of a data steward (GovLab, 2019), an agent responsible for the operational aspects of an organization's data governance, that is, the actual policies, methods, and procedures to manage data (Rosenbaum, 2010; Plotkin, 2014). The expert group on B2G data sharing referred to data stewards as "champions of data sharing" (European Commission, 2020c, p. 37).

Lessons drawn from the field of medical data donation show that accountability and transparency are key normative trust-enablers to increase data sharing (Vayena et al., 2018). To ensure accountability, a data steward can be in charge of a system to handle complaints, investigate incidents, and perform other functions for managing conflicts within or across organizations. Meanwhile, stewards can promote transparency by providing clear information about the data, including how it was collected, what it represents, and how it can be used. Furthermore, by adhering to the findability, accessibility, interoperability, and reusability (FAIR) principles for data management (Wilkinson et al., 2016), data stewards can build trust and encourage firms hesitant to share data to engage in this practice (Ada Lovelace Institute, 2021).

[16] https://www.imda.gov.sg/-/media/Imda/Files/Programme/PET-Sandbox/PET-Sandbox-CFP.pdf
[17] https://www.jetro.go.jp/ext_images/en/invest/incentive_programs/pdf/Detailed_overview.pdf

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

The implementation of data stewardship in an organization can be complex and costly. It might require organizational restructuring and extensive training. Aside from a company's own policies, stewards need to be knowledgeable about the latest data privacy and security regulations. As a new occupation, the role of data stewards can be unclear and confused with other ones present in an organization. Ultimately, it should be avoided that the responsibilities of a data steward are not ever-expanding in a way that they result in a single point of failure.

4. Regulatory context

The expert group on B2G data sharing emphasized the importance of promoting voluntary collaboration to the greatest extent possible, aligning the interests of involved parties for the common good (European Commission, 2020c). However, voluntary collaboration is not always adequate, as there are instances where the commercial interests of a firm cannot be prioritized over the general welfare of society, which forces the government to interfere and oblige organizations to share data (Mercille, 2021).

Despite the need to move beyond voluntary data sharing initiatives, aggressively advocating for mandatory B2G data sharing can be regarded as too interventionist in the EU, where industry self-regulation and market processes tend to set the standards for sharing data (Martens and Zhao, 2021, p. 8). Part of the European tradition of regulating the economy is a bottom-up, industry-driven approach, which, despite being slower than other approaches in producing results due to conflicting interests among stakeholders, is relatively more successful in protecting private property, privacy, and other individual rights (Martens and Zhao, 2021; Roberts et al., 2021a). Hence, in the EU, accessing private sector data must be justified on the basis of a well-defined public objective; in other words, mandatory B2G is often regarded by public authorities as a "last resort" among the policy options to increase access to private sector data (Richter, 2021, p. 538).

This contrasts, for example, with the Chinese approach, which "seeks to maximize the social value of data" through channels that would be politically unfeasible in the EU, which place national interest and collective welfare above private interests and individual welfare (Martens and Zhao, 2021, p. 3). Yet, observations made by Chinese researchers suggest that top–down mandates are not enough to boost data sharing among the scientific community and that more guidelines based on principles as well as incentive mechanisms are needed (Li et al., 2021). Thus, there is a need to find an adequate balance between the voluntary and mandatory models (Vigorito, 2022).

Shkabatur (2019, p. 354) proposes that to get access to private sector data, policymakers can adopt an "open, collaborative, and incentives-based stance." Richter (2021) also emphasizes that governments can make use of soft law approaches that are less interventionist and more based on incentives. These incentives can better contribute to building a data sharing culture where more stakeholders are aware of the benefits of sharing data and are willing to engage in collaboration (European Commission, 2020c). In the long run, it is expected that the framework and procedures established to govern the sharing of data between private firms and public authorities will also contribute to the development of B2B data relations (European Commission, 2020c).

Broadening access to private sector data through voluntary and mandatory approaches shares similarities with other regulatory challenges where there is a need to find a middle ground between prioritizing public or private interests. The example of intellectual property rights highlights this tension. Like with data, their exclusive control limits public access to valuable information, reflecting the anticommons dilemma: "socially suboptimal information availability because of excessive privatisation" (Koutroumpis et al., 2020, p. 657). The future of private sector data sharing might be characterized by the pursuit of a middle ground that accommodates the interests of private and public actors. Ideally for the European Commission, this one would foster a data sharing culture.

4.1. Data governance act

As the first legislative act of the European strategy for data, the Data Governance Act recognized that "action at Union level is necessary to increase trust" (European Commission, 2022a, p. 4). The Data

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

Governance Act sets a harmonized regulatory framework for facilitating voluntary data sharing, promoting the rise of data intermediaries and advancing data altruism, a principled approach to data sharing. The regulation entered into force in June 2022 and became applicable in September 2023.

4.1.1. Data intermediaries

Chapter III of the Data Governance Act introduces new harmonizing requirements for providers of data sharing services (data intermediaries) with the intention of ensuring "the trustworthy exchange of data" (European Commission, 2022a, p. 13). These new rules translate into a soft law approach to increasing data sharing by supporting the emergence and uptake of these promising actors in the data economy.

The European Commission defined data intermediaries as "technical enablers" to harness the potential of data (European Commission, 2018, p. 11). Yet aside from the EU, the governments in the United Kingdom and Singapore are directing efforts toward creating the right conditions for data intermediaries to thrive and exchange data in innovative ways (Personal Data Protection Commission, 2020; Centre for Data Ethics and Innovation, 2021). Because legal recognition might not be enough for data intermediaries to address the massive data asymmetry in society that benefits a few hyperscalers, public institutions are supporting their growth with financial incentives such as subsidies, grants, and tax breaks. Nevertheless, it should be kept in mind that over-focusing on data intermediaries and creating excessively high expectations can undermine the development of alternative instruments with greater potential to increase data sharing (Carovano and Finck, 2023).

As mentioned in Section 3.2, while an intermediary can provide the necessary technical and organizational infrastructure to enable trust, an important question to address is whether it will remain neutral and independent (Richter, 2023). In a hypothetical future where data intermediaries have scaled, it should be avoided that there is a concentration of power within a limited number of intermediaries or that these become caught up in political agendas.

4.1.2. Data altruism organizations

Chapter IV of the Data Governance Act introduces data altruism as "the voluntary sharing of data… without seeking or receiving a reward… for objectives of general interest" (European Commission, 2022a, p. 20). The Data Governance Act does not provide a definition of "general interest" but offers a series of examples where the welfare of society is at stake in the areas of health care, sustainability, mobility, and scientific research. Registered data altruism organizations (DAOs) are expected to act as trustworthy data intermediaries in charge of managing donated data. Any entity that aims to register as one must operate on a non-profit basis and simultaneously meet special transparency and technical requirements. An example is *DATALOG,*[18] a platform for citizens to visualize and better manage their energy expenses.

Chapter IV also introduces a data altruism consent form to harmonize the collection of consent from data subjects. A similar procedure would be beneficial in the context of private sector data. More guidance on how to grant, modify, or withdraw access to data can be provided by the rulebook introduced in Article 22 of the Data Governance Act, which will be established through the adoption of delegated acts and will specify requirements to protect the rights and interests of data subjects and data holders (European Commission, 2022a).

However, like with data intermediaries, the neutrality and independence of a DAO will be crucial for enabling trust. In theory, DAOs operate without seeking a reward other than contributing to a social cause and are allowed to share data with third parties for altruistic purposes only, yet there is the possibility that data are used with other intentions. In this regard, it should be taken into account that DAOs could have ties with or even be directly or indirectly funded by big players in the data economy seeking to get access to donated data. For firms, this creates reputational risks that can make them more reluctant to donate data to a DAO.

[18] https://datalog.es/

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

A possible adverse result is that, first, data altruism fails to scale because of a lack of incentives among firms to donate data. Aside from the marketing opportunities that this practice can offer, firms would be more incentivized to participate if they could learn something about their own business or the industry in which they operate. For example, in the field of environmental sustainability, a DAO could allow donors to compare indicators and derive analytical insights in order to optimize their supply chains. Second, another detrimental scenario is that firms do not engage in data altruism due to a lack of trust toward DAOs. In this regard, designing a logo to be exclusively used by a certified DAO was a positive step to building trust (European Commission, 2023).

4.2. Data Act

The Data Act is the second legislative proposal of the European strategy for data and introduces mandatory rules for accessing and using industrial data (European Commission, 2022b). The Regulation aims to make data more accessible by giving access rights, addressing unfair contractual terms that arise from vendor lock-in effects and anticompetitive practices, establishing rules for public bodies to access private sector data, and allowing customers to switch between service providers. EU policymakers came to an agreement in June 2023 that will make the Data Act applicable in 2025.

4.2.1. Mandatory B2B data sharing

The Data Act establishes requirements to share data generated by connected devices across all sectors. It introduces rights for consumers and businesses to access and share data generated by their devices with third parties, including data intermediaries. Enacting these rights would oblige data holders to make data under their control available to other parties.

During the final negotiations of the Data Act, major European and US-based data-rich firms complained that the original draft lacked enough safeguards to protect their competitive interests from potential misuses of their data by third parties (Yun Chee, 2023). The concept of "trade secret holder" was included in the final text of the legislation to let data holders protect the confidentiality of data and decline sharing it if this was likely to result in serious economic damage. As a result, the Data Act gives trade secret holders the possibility to demand that data receivers guarantee the confidentiality of data, for example, by agreeing on contractual terms, confidentiality agreements, access protocols, standards, and codes of conduct. In the event of a disagreement or failure to implement appropriate measures to safeguard the data, the data holder would be entitled to cancel the sharing of the data.

Adding these safeguards to the final text of the Data Act underscores the significance of competition risks and the need to address them. Future research could monitor how frequently the trade secret holder exemption is employed to prevent data sharing and how often its use is justified—it is not uncommon for data-rich firms to invoke trade secrecy laws to justify their exclusive control over the data they hold (Cohen, 2019).

4.2.2. Mandatory B2G data sharing

Chapter V of the Data Act details under what exceptional circumstances it would be mandatory for data holders to make their data available to public bodies. This includes responding to a public emergency or fulfilling "a specific task in the public interest that has been explicitly provided and defined by national law" when none of the following three alternatives for obtaining access were viable: (1) requesting it voluntarily; (2) purchasing it on the market; or (3) relying on existing obligations (European Commission, 2022b, p. 48).

The issue of including or excluding personal data from the scope of Chapter V was a matter of debate during the trilogue negotiations of the Data Act. While the European Commission and the European Council pushed for including personal data in the scope, the European Parliament proposed excluding personal data altogether (Bertuzzi, 2023a). A major EU trade association that represents the interests of several data-rich firms raised concerns about including personal data in the scope, highlighting the risks

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

associated with data leakages and misuses (DIGITALEUROPE, 2023). Indeed, such risks extend beyond B2B contexts and can discourage firms from sharing personal data with public authorities due to privacy and reputational risks. In the end, personal data were only included for responding to public emergencies (Bertuzzi, 2023b).

Lessons from the B2G data sharing initiative between the European Commission and European mobile network operators to predict and contain the spread of COVID-19 show that given the sensitivity of the issue and the type of data involved, it was crucial to guarantee that the reputation of the firms involved was not harmed due to potential misunderstandings about the use of the data (Vespe et al., 2021).

Competition risks transcend B2B relations. A public body may use a firm's data to independently develop public services or to aid a business competitor (Klievink et al., 2018). Furthermore, public bodies themselves can also operate successful commercial services (Carballa Smichowski, 2018; Martens and Duch-Brown, 2020).

5. Conclusion

The upcoming 2024 European Commission will be occupied with implementing the policies that compose the European strategy for data. In order for the Data Governance Act and Data Act to achieve their goal of increasing private sector participation in data sharing, EU policymakers can deploy a targeted set of tactics that leverage the trust-enabling mechanisms discussed in this paper in addition to others (Farrell et al., 2023). Government support (e.g., funding, tax reductions, regulatory sandboxes, etc.) for the growth of trustworthy data intermediaries and DAOs that host data exchange platforms, steward data, guarantee privacy, and security through the use of PETs and provide a data sharing agreement will be essential but not sufficient to satisfactorily boost participation. Increasing awareness about the benefits of this practice and creating incentive structures remain two other major challenges that require closer examination by academia. Future research could take an empirical angle through surveys and interviews to understand firms' perceived risks and incentives toward different modalities of data sharing.

Furthermore, as more data policies enter into force, providing instructional materials for companies will also be key, especially for those companies with more limited resources, such as startups and small and medium-sized enterprises (SMEs), who might require assistance in understanding the requirements, scope, and interplay of regulations, as well as the processes and technologies involved in complying with them.

Considering that the majority of data are produced by businesses, enabling the creation of trust in the private sector can have a positive effect in other realms where data can be shared (European Commission, 2020c). Furthermore, it is essential to encourage the voluntary participation of not only firms but also individuals, non-governmental organizations, and public institutions.

In the short term, mandating extensive access to private sector data appears politically unfeasible. However, as the EU's overarching goal of strengthening its technological sovereignty, competitiveness, and resilience intensifies, in the longer term, mandatory data sharing could become a more attractive option. It would be positive if this struggle to open new channels for data to flow also contributed to tackling societal challenges and reducing the pervasive data asymmetry in society.

As a highly strategic resource, data are increasingly the object of competition between sovereign states eager to control its flow (Chander and Sun, 2023). As a result of this struggle, multiple data access regimes continue to emerge and compete with each other (Martens and Zhao, 2021), making distinct sovereignty claims and increasingly geopolitizing data sharing (Amoore, 2018). In the public discourse, the Chinese, European, and American approaches receive special attention (Bradford, 2023).

The latest digital policies pushed by the EU under the digital sovereignty agenda aim to assert greater control over critical infrastructure and reduce the region's heavy reliance on external actors (Von der Leyen, 2019; Roberts et al., 2021b). Yet, in this technological era, the U.S.A has unparalleled leverage over the infrastructure of the European digital economy, not only facilitated by the presence of American companies in key industries such as AI, e-commerce, search engines, social media, and cloud computing, but also by undersea cables, data centers, and communication networks (Farrell and Newman, 2023).

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

Furthermore, China continues deploying digital infrastructure and gaining ground in high-tech not just in Europe but across the world, which could further complicate the pursuit of digital sovereignty.

While data sharing leads to better intelligence, innovation, economic growth, and other key elements for states to remain competitive at a global stage, as explained in the introduction, it is also necessary for tackling the grand societal challenges of the 21st century, which are fundamentally global. To effectively address these challenges, society may need to envision frameworks for facilitating the flow of data between businesses and governments across international borders.

Acknowledgments. The author, Jaime Bernal, initiated work on this article during a scientific traineeship at the Digital Economy Unit of the Joint Research Centre in 2022, and wishes to express gratitude for the guidance provided.

Author contribution. Conceptualization: J.B.; Data curation: J.B.; Funding acquisition: J.B.; Investigation: J.B.; Methodology: J.B.; Project administration: J.B.

Funding statement. This work received no specific grant from any funding agency, commercial or not-for-profit sectors.

Competing interest. The author declares none.

References

Abadi M, **Chu A**, **Goodfellow I**, **McMahan HB**, **Mironov I**, **Talwar K and Zhang L** (2016) Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pp. 308–318. https://doi.org/10.1145/2976749.2978318.

Abbas AE, **Agahari W**, **van de Ven M**, **Zuiderwijk A and de Reuver M** (2021) Business data sharing through data marketplaces: A systematic literature review. *Journal of Theoretical and Applied Electronic Commerce Research 16*(7), 3321–3339. https://doi.org/10.3390/jtaer16070180.

Ada Lovelace Institute (2021) Exploring legal mechanisms for data stewardship. Available at https://www.adalovelaceinstitute.org/report/legal-mechanisms-data-stewardship/ (accessed 20 August 2023).

Agahari W, **Ofe H and de Reuver M** (2022) It is not (only) about privacy: How multi-party computation redefines control, trust, and risk in data sharing. *Electronic Markets 32*(3), 1577–1602. https://doi.org/10.1007/s12525-022-00572-w.

Alemanno A (2018) Big data for good: Unlocking privately-held data to the benefit of the many. *European Journal of Risk Regulation 9*(2), 183–191. https://doi.org/10.1017/err.2018.34.

Allen HJ (2019) Regulatory sandboxes. *The George Washington Law Review 87*, 579. Available at https://digitalcommons.wcl.american.edu/facsch_lawrev/709 (accessed 10 August 2023).

Amoore L (2018) Cloud geographies: Computing, data, sovereignty. *Progress in Human Geography 42*(1), 4–24.

Association of Banks in Singapore (2019) Data sharing handbook for banks and non-bank data ecosystem partners. Available at https://abs.org.sg/docs/library/data-sharing-handbook-for-banks-and-non-bank-data-ecosystem-partners.pdf.

Bayardo RJ and Agrawal R (2005) Data privacy through optimal k-anonymization. In *21st International Conference on Data Engineering (ICDE'05)*. Tokyo, Japan: IEEE, pp. 217–228).

Beaver D (1992) Efficient multiparty protocols using circuit randomization. In *Advances in Cryptology—CRYPTO'91: Proceedings*, Vol. *11*. Berlin Heidelberg: Springer, pp. 420–432.

Bergman R, **Abbas AE**, **Jung S**, **Werker C and de Reuver M** (2022) Business model archetypes for data marketplaces in the automotive industry. *Electronic Markets 32*(2), 747–765. https://doi.org/10.1007/s12525-022-00547-x.

Bertuzzi L (2023a) Data Act: Trade secret safeguards shall be exception not rule, Commission says. Euractiv. Available at https://www.euractiv.com/section/data-privacy/news/data-act-trade-secret-safeguards-shall-be-exception-not-rule-commission-says/ (accessed 12 September 2023).

Bertuzzi L (2023b) EU policymakers to nail down agreement on new data-sharing law. Euractiv. Available at https://www.euractiv.com/section/data-privacy/news/eu-policymakers-to-nail-down-agreement-on-new-data-sharing-law/ (accessed 12 September 2023).

Bodó B (2021) Mediated trust: A theoretical framework to address the trustworthiness of technological trust mediators. *New Media and Society 23*(9), 2668–2690. https://doi.org/10.1177/1461444820939922.

Bradford A (2023) *Digital Empires: The Global Battle to Regulate Technology*. New York: Oxford University Press.

Carballa Smichowski B (2018) The value of data: An analysis of closed-urban-data-based and open-data-based business models. HAL. Available at https://hal.science/hal-01736484/document (accessed 10 August 2023).

Carballa Smichowski B, **Duch Brown N and Martens B** (2021) To pool or to pull back? An economic analysis of health data pooling. JRC Working Papers on Digital Economy. Joint Research Centre.

Carovano G and Finck M (2023) Regulating data intermediaries: The impact of the data governance act on the EU's data economy. *Computer Law & Security Review 50*, 105830. https://doi.org/10.1016/j.clsr.2023.105830.

Centre for Data Ethics and Innovation (2021) Unlocking the value of data: Exploring the role of data intermediaries. Centre for Centre for Data Ethics and Innovation. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1004925/Data_intermediaries_-_accessible_version.pdf (accessed 20 September 2023).

Chander A and Sun H (eds.) (2023) *Data Sovereignty: From the Digital Silk Road to the Return of the State*. New York: Oxford University Press.

Cohen JE (2019) *Between Truth and Power*. New York: Oxford University Press.

Datasphere Initiative (2022) Sandboxes for data: creating spaces for agile solutions across borders. Datasphere Initiative. Available at https://www.thedatasphere.org/wp-content/uploads/2022/05/Sandboxes-for-data-2022-Datasphere-Initiative.pdf (accessed 25 August 2023).

DIGITALEUROPE (2023) Access to data by public bodies is a double-edged tool to use with caution and restraint. Available at https://www.digitaleurope.org/news/joint-statement-access-to-data-by-public-bodies-is-a-double-edge-tool-to-use-with-caution-and-restraint/ (accessed 12 September 2023).

Dwork C and Roth A (2013) The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science 9*(3–4), 211–407. https://doi.org/10.1561/0400000042.

Eurich M, **Oertel N and Boutellier R** (2010) The impact of perceived privacy risks on organizations' willingness to share item-level event data across the supply chain. *Electronic Commerce Research 10*(3–4), 423–440. https://doi.org/10.1007/s10660-010-9062-0.

European Commission (2016) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

European Commission (2018) Commission staff working document guidance on sharing private sector data in the European data economy accompanying the document communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions "Towards a common European data space" SWD/2018/125 final.

European Commission (2020a) Communication from the commission to the European parliament, the Council, the European Economic and Social Committee and the Committee of the Regions "A European strategy for data," Brussels, 19 February 2020, COM(2020) 66. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066 (accessed 2 September 2023).

European Commission (2020b) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Sustainable and Smart Mobility Strategy – Putting European transport on track for the future COM/2020/789 final, 9 December. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020DC0789&from=EN (accessed 1 September 2023).

European Commission (2020c) Towards a European strategy on business-to-government data sharing for the public interest: final report prepared by the High Level Expert Group on Business to Government Data Sharing, Publications Office. https://data.europa.eu/doi/10.2759/406717 (accessed 1 September 2023).

European Commission (2022a) Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act).

European Commission (2022b) European Commission Brussels, 23.2.2022 COM(2022) 68 final 2022/0047 (COD) Proposal for a regulation of the European parliament and of the council on harmonised rules on fair access to and use of data (Data Act).

European Commission (2023) Data Governance Act: common logos to easily identify trusted EU data intermediaries and data altruism organizations to re-use data. Available at https://digital-strategy.ec.europa.eu/en/news/data-governance-act-common-logos-easily-identify-trusted-eu-data-intermediaries-and-data-altruism (accessed 19 September 2023).

European Council (2020) Council Conclusions on Regulatory sandboxes and experimentation clauses as tools for an innovation-friendly, future-proof and resilient regulatory framework that masters disruptive challenges in the digital age.

Farrell E, **Minghini M**, **Kotsev A**, **Soler Garrido J**, **Tapsall B**, **Micheli M**, **Posada Sanchez J**, **Signorelli S**, **Tartaro A**, **Bernal Cereceda J**, **Vespe M**, **Di Leo M**, **Carballa Smichowski B**, **Smith R**, **Schade S**, **Pogorzelska K**, **Gabrielli L and de Marchi D** (2023) European Data Spaces - Scientific Insights into Data Sharing and Utilisation at Scale. JRC Research Reports. Joint Research Centre (Seville site).

Farrell H and Newman AL (2023) *Underground Empire: How America Weaponized the World Economy*. New York: Henry Holt and Co.

Fontaine C and Galand F (2007) A survey of homomorphic encryption for nonspecialists. *EURASIP Journal on Information Security*, *2007*, 1–10. https://doi.org/10.1155/2007/13801

Goldwasser S, **Micali S and Rackoff C** (1989) The knowledge complexity of interactive proof-systems. *SIAM Journal on Computing 18*(1), 186–208. https://doi.org/10.1137/0218012.

GovLab (2019) Wanted: Data stewards - Recruiting and developing people to manage data responsibly. The GovLab. Available at https://thegovlab.org/static/files/publications/wanted-data-stewards.pdf (accessed 6 August 2023).

Günther WA, **Rezazade Mehrizi MH**, **Huysman M and Feldberg F** (2017) Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems 26*(3), 191–209.

Heurix J, **Zimmermann P**, **Neubauer T and Fenz S** (2015) A taxonomy for privacy enhancing technologies. *Computers & Security 53*, 1–17. https://doi.org/10.1016/j.cose.2015.05.002.

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

Iacus SM, **Santamaria C**, **Sermi F**, **Spyratos S and Vespe M** (2020) Human mobility and COVID-19 initial dynamics. *Nonlinear Dynamics 101*(1–2), 1901–1919. https://doi.org/10.1007/s11071-020-05854-6.

Information Commissioner's Office (2021) Data sharing code of practice. Information Commissioner's Office. Available at https://ico.org.uk/media/for-organisations/guide-to-data-protection/ico-codes-of-practice/data-sharinga-code-of-practice-1-0.pdf (accessed 28 August 2023).

Janssen H and Singh J (2022a) Data intermediary. *Internet Policy Review 11*(1), 1–9. https://doi.org/10.14763/2022.1.1644.

Jia B, **Zhang X**, **Liu J**, **Zhang Y**, **Huang K and Liang Y** (2022) Blockchain-enabled federated learning data protection aggregation scheme with differential privacy and homomorphic encryption in IIoT. *IEEE Transactions on Industrial Informatics 18*(6), 4049–4058. https://doi.org/10.1109/TII.2021.3085960.

Kairouz P, **McMahan HB**, **Avent B**, **Bellet A**, **Bennis M**, **Bhagoji AN**, **Bonawitz K**, **Charles Z**, **Cormode G**, **Cummings R**, **D'Oliveira RGL**, **Eichner H**, **El Rouayheb S**, **Evans D**, **Gardner J**, **Garrett Z**, **Gascón A**, **Ghazi B**, **Gibbons PB**, **Gruteser M**, **Harchaoui Z**, **He C**, **He L**, **Huo Z**, **Hutchinson B**, **Hsu J**, **Jaggi M**, **Javidi T**, **Joshi G**, **Khodak M**, **Konecní J**, **Korolova A**, **Koushanfar F**, **Koyejo S**, **Lepoint T**, **Liu Y**, **Mittal P**, **Mohri M**, **Nock R**, **Özgür A**, **Pagh R**, **Qi H**, **Ramage D**, **Raskar R**, **Raykova M**, **Song D**, **Song W**, **Stich SU**, **Sun Z**, **Suresh AT**, **Tramèr F**, **Vepakomma P**, **Wang J**, **Xiong L**, **Xu Z**, **Yang Q**, **Yu FX**, **Yu H, and Zhao S** (2021) Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning 14*(1–2), 1–210. https://doi.org/10.1561/220000008.

Kang-Xing J and McGorman L (2020) Data for Good: New tools to help health researchers track and combat COVID-19. Meta. Available at https://about.fb.com/news/2020/04/data-for-good/ (accessed 10 August 2023).

Keller M, **Pastro V and Rotaru D** (2018) Overdrive: Making SPDZ great again. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Cham: Springer International Publishing, pp. 158–189.

Kerber W (2018) Data governance in connected cars: The problem of access to in-vehicle data. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law 9*, 310.

Kirkpatrick R (2013) Big data for development. *Big Data 1*(1), 3–4. https://doi.org/10.1089/big.2012.1502.

Klievink B, **Van Der Voort H and Veeneman W** (2018) Creating value through data collaboratives. *Information Polity 23*(4), 379–397.

Koutroumpis P, **Leiponen A and Thomas LDW** (2020) Markets for data. *Industrial and Corporate Change 29*(3), 645–660. https://doi.org/10.1093/icc/dtaa002

Lev-Aretz Y (2019) Data philanthropy. *Hastings Law Journal 70*(6), 1491–1545. Available at https://repository.uchastings.edu/hastings_law_journal/vol70/iss6/3 (accessed 19 August 2023).

Li X, **Cheng G**, **Wang L**, **Wang J**, **Ran Y**, **Che T**, … **Zhao G** (2021) Boosting geoscience data sharing in China. *Nature Geoscience 14*(8), 541–542.

Lindell Y (2021) Secure multiparty computation. *Communications of the ACM 64*(1), 86–96. https://doi.org/10.1145/3387108.

Martens B (2021) An economic perspective on data and platform market power. Proceedings of the JRC Digital Economy Working Paper 2020–09. Joint Research Centre.

Martens B, **de Streel A**, **Graef I**, **Tombal T and Duch-Brown N** (2020) Business-to-business data sharing: An economic and legal analysis. JRC Digital Economy Working Paper 2020–05. Joint Research Centre.

Martens B and Duch-Brown N (2020) The economics of business-to-government data sharing. JRC Digital Economy Working Paper No. 2020–04.

Martens B and Zhao B (2021) Data access and regime competition: A case study of car data sharing in China. *Big Data & Society 8*(2), 1–11. https://doi.org/10.1177/20539517211046374.

Mercille J (2021) Inclusive smart cities: Beyond voluntary corporate data sharing. *Sustainability 13*(15), 8135–8148. https://doi.org/10.3390/su13158135.

Morozov E (2014) *To Save Everything, Click Here: The Folly of Technological Solutionism*. New York: PublicAffairs.

Munjal K and Bhatia R (2022) A systematic review of homomorphic encryption and its contributions in the healthcare industry. *Complex & Intelligent Systems*, 3759–3786. https://doi.org/10.1007/s40747-022-00756-z

Ohm P (2010) Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA Law Review 57*, 1701.

Pearson S and Casassa-Mont M (2011) Sticky policies: An approach for managing privacy across multiple parties. *Computer 44*(9), 60–68. https://doi.org/10.1109/MC.2011.225.

Personal Data Protection Commission (2020) Guide to managing data intermediaries. Personal Data Protection Commission. Available at https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Other-Guides/Guide-to-Managing-Data-Intermediaries–2020.pdf (accessed 29 August 2023).

Plotkin D (2014) Data stewardship and data governance. In Plotkin D (ed.), *Data Stewardship*. London: Elsevier, pp. 1–21. https://doi.org/10.1016/B978-0-12-410389-4.00001-5.

Poncibò C and Zoboli L (2022) The methodology of regulatory sandboxes in the EU: A preliminary assessment from a competition law perspective. EU Law Working Papers No. 61, Stanford-Vienna Transatlantic Technology Law Forum.

Renieris E (2021) Privacy-enhancing technologies: Not always our friends. Ada Lovelace Institute. Available at https://www.adalovelaceinstitute.org/blog/privacy-enhancing-technologies-not-always-our-friends/ (accessed 20 August 2023).

Richter H (2021) The law and policy of government access to private sector data ('B2G data sharing'). In *Data Access, Consumer Interests and Public Welfare*. Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG, 529–572. https://doi.org/10.5771/9783748924999-529.

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

Richter H (2023) Looking at the data governance act and beyond: How to better integrate data intermediaries in the market order for data sharing. *GRUR International 72*(5) 458–470. https://doi.org/10.1093/grurint/ikad014.

Richter H and Slowinski PR (2019) The data sharing economy: On the emergence of new intermediaries. *IIC-International Review of Intellectual Property and Competition Law 50*(1), 4–29. https://doi.org/10.1007/s40319-018-00777-7.

Rivest RL, **Adleman L and Dertouzos ML** (1978) On data banks and privacy homomorphisms. *Foundations of Secure Computation 4*(11), 169–180.

Roberts H, **Cowls J**, **Casolari F**, **Morley J**, **Taddeo M and Floridi L** (2021b) Safeguarding European values with digital sovereignty: An analysis of statements and policies. *Internet Policy Review 10*(3), 1–26. https://doi.org/10.14763/2021.3.1575.

Roberts H, **Cowls J**, **Morley J**, **Taddeo M**, **Wang V and Floridi L** (2021a) The Chinese approach to artificial intelligence: An analysis of policy, ethics, and regulation. *AI & SOCIETY 36*(1), 59–77. https://doi.org/10.1007/s00146-020-00992-2.

Rosenbaum S (2010) Data governance and stewardship: Designing data stewardship entities and advancing data access: Data governance and stewardship. *Health Services Research 45*(5p2), 1442–1455. https://doi.org/10.1111/j.1475-6773.2010.01140.x.

Ruijer E (2021) Designing and implementing data collaboratives: A governance perspective. *Government Information Quarterly 38*(4), 101612. https://doi.org/10.1016/j.giq.2021.101612.

Sabt M, **Achemlal M and Bouabdallah A** (2015) Trusted execution environment: What it is, and what it is not. In *2015 IEEE Trustcom/BigDataSE/ISPA*, pp. 57–64. https://doi.org/10.1109/Trustcom.2015.357.

Samarati P and Latanya S (1998) Protecting privacy when disclosing information: K-anonymity and its enforcement through generalization and suppression. Technical Report SRI-CSL-98-04.

Scotti F, **Pierri F**, **Bonaccorsi G and Flori A** (2022) Responsiveness of open innovation to COVID-19 pandemic: The case of data for good. *PLoS One 17*(4), e0267100. https://doi.org/10.1371/journal.pone.0267100.

Shkabatur J (2019) The global commons of data. *Stanford Technology Law Review 22*, 1–46. Available at https://ssrn.com/abstract=3263466 (accessed 29 August 2023).

Siegel JE, **Erb DC and Sarma SE** (2017) A survey of the connected vehicle landscape–architectures, enabling technologies, applications, and development areas. *IEEE Transactions on Intelligent Transportation Systems 19*(8), 1–16.

Sitra (2022) Rulebook for a fair data economy. Sitra. Available at https://www.sitra.fi/en/publications/rulebook-for-a-fair-data-economy/ (accessed 6 September 2023).

Spiekermann M (2019) Data marketplaces: Trends and monetisation of data goods. *Intereconomics 54*(4), 208–216. https://doi.org/10.1007/s10272-019-0826-z.

Spyratos S, **Vespe M**, **Natale F**, **Weber I**, **Zagheni E and Rango M** (2019) Quantifying international human mobility patterns using Facebook network data. *PLoS One 14*(10), e0224134.

Stahl F, **Schomm F**, **Vossen G and Vomfell L** (2016) A classification framework for data marketplaces. *Vietnam Journal of Computer Science 3*(3), 137–143. Berlin: Springer Science and Business Media LLC. https://doi.org/10.1007/s40595-016-0064-2.

Stalla-Bourdillon S, **Carmichael L and Wintour A** (2021) Fostering trustworthy data sharing: Establishing data foundations in practice. *Data & Policy 3*(e4), 1–8. https://doi.org/10.1017/dap.2020.24.

Susha I, **Grönlund Å and Van Tulder R** (2019b) Data driven social partnerships: Exploring an emergent trend in search of research challenges and questions. *Government Information Quarterly 36*(1), 112–128. https://doi.org/10.1016/j.giq.2018.11.002.

Susha I, **Janssen M and Verhulst S** (2017a) Data collaboratives as "bazaars"? A review of coordination problems and mechanisms to match demand for data with supply. *Transforming Government: People, Process and Policy 11*(1), 157–172. https://doi.org/10.1108/TG-01-2017-0007.

Susha I, **Janssen M and Verhulst S** (2017b) Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development. In *Proceedings of the 50th Hawaii International Conference on System Science Paper. 50th Hawaii International Conference on System Sciences (HICSS 2017), Waikoloa, USA, January 4–7, 2017*, IEEE, pp. 2691–2700.

Susha I, **Rukanova B**, **Ramon Gil-Garcia J**, **Tan Y-H and Hernandez MG** (2019a) Identifying mechanisms for achieving voluntary data sharing in cross-sector partnerships for public good. In *Proceedings of the 20th Annual International Conference on Digital Government Research*, pp. 227–236. https://doi.org/10.1145/3325112.3325265.

Taddeo M (2017) Data philanthropy and individual rights. *Minds and Machines 27*(1), 1–5. https://doi.org/10.1007/s11023-017-9429-2.

Taddeo M and Floridi L (2011) The case for e-trust. *Ethics and Information Technology 13*(1), 1–3. https://doi.org/10.1007/s10676-010-9263-1.

Teece DJ (1980) Economies of scope and the scope of the enterprise. *Journal of Economic Behavior & Organization 1*(3), 223–247. https://doi.org/10.1016/0167-2681(80)90002-5.

Tenopir C, **Allard S**, **Douglass K**, **Aydinoglu AU**, **Wu L**, **Read E**, … **Frame M** (2011) Data sharing by scientists: Practices and perceptions. *PLoS One 6*, e21101. http://doi.org/10.1371/journal.pone.0021101.

Tonetti C and Jones CI (2020) Nonrivalry and the economics of data. *American Economic Review 110*(9), 2819–2858.

UN Global Pulse (2014) Mapping the next frontier of open data: Corporate data sharing. Available at https://www.unglobalpulse.org/2014/09/mapping-the-next-frontier-of-open-data-corporate-data-sharing/ (accessed 24 August 2023).

van den Broek T and van Veenstra AF (2015) Modes of governance in inter-organizational data collaborations. Twenty-Third European Conference on Information Systems (ECIS), Münster, Germany.

Vayena E, **Haeusermann T**, **Adjekum A and Blasimme A** (2018) Digital health: Meeting the ethical and policy challenges. *Swiss Medical Weekly 148*(34), 1–9. https://doi.org/10.4414/smw.2018.14571.

https://doi.org/10.1017/dap.2024.20 Published online by Cambridge University Press

Vespe M, **Iacus S**, **Santamaria C**, **Sermi F and Spyratos S** (2021) On the use of data from multiple mobile network operators in Europe to fight COVID-19. *Data & Policy 3*(e8), 1–10. https://doi.org/10.1017/dap.2021.9.

Vigorito A (2022) Government access to privately-held data: business-to-government data sharing: Voluntary and mandatory models. *European Journal of Comparative Law and Governance* (1), 1–22. https://doi.org/10.1163/22134514-bja10030.

Von der Leyen U (2019) A union that strives for more. My agenda for Europe. Political guidelines for the next European Commission 2019–2024. Brussels: European Commission. Available at https://www.europarl.europa.eu/resources/library/media/20190716RES57231/20190716RES57231.pdf (accessed 2 August 2023).

von Grafenstein M (2022) Reconciling conflicting interests in Data through Data Governance. An analytical framework (and a brief discussion of the Data Governance Act Draft, the Data Act Draft, the AI Regulation Draft, as well as the GDPR). https://doi.org/10.5281/ZENODO.6457735 (accessed 4 August 2023).

Wilkinson MD, **Dumontier M**, **Aalbersberg IJ**, **Appleton G**, **Axton M**, **Baak A**, … **Mons B** (2016) The FAIR guiding principles for scientific data management and stewardship. *Scientific Data 3*(1), 1–9.

Yun Chee F (2023) EU draft Data Act puts trade secrets at risk, Siemens, SAP say. Reuters. Available at https://www.reuters.com/technology/siemens-sap-say-eu-draft-data-act-puts-trade-secrets-risk-2023-05-07/ (accessed 1 October 2023).

Zrenner J, **Möller FO**, **Jung C**, **Eitel A and Otto B** (2019) Usage control architecture options for data sovereignty in business ecosystems. *Journal of Enterprise Information Management 32*(3), 477–495. https://doi.org/10.1108/JEIM-03-2018-0058.

Zuboff S (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs.

Cite this article: Bernal J (2024). Private sector trust in data sharing: enablers in the European Union. *Data & Policy*, 6: e30. doi:10.1017/dap.2024.20