Mediating the Tension between Data Sharing and Privacy: The Case of DMA and GDPR

Weigl, Linda ; Barbereau, Tom ; Sedlmeir, Johannes ; Zavolokina, Liudmila

Abstract: The Digital Markets Act (DMA) constitutes a crucial part of the European legislative framework addressing the dominance of 'Big Tech'. It intends to foster fairness and competition in Europe's digital platform economy by imposing obligations on 'gatekeepers' to share end-user-related information with business users. Yet, this may involve the processing of personal data subject to the General Data Protection Regulation (GDPR). The obligation to provide access to personal data in a GDPR-compliant manner poses a regulatory and technical challenge and can serve as a justification for gatekeepers to refrain from data sharing. In this research-in-progress paper, we analyze key tensions between the DMA and the GDPR through the paradox perspective. We argue through a task-technology fit approach how privacy-enhancing technologies-particularly anonymization techniques-and portability could help mediate tensions between data sharing and privacy. Our contribution provides theoretical and practical insights to facilitate legal compliance.

MEDIATING THE TENSION BETWEEN DATA SHARING AND PRIVACY: THE CASE OF DMA AND GDPR

Research-in-Progress Paper

Linda Weigl, University of Luxembourg, Luxembourg City, Luxembourg, linda.weigl@uni.lu

Tom Barbereau, University of Luxembourg, Luxembourg City, Luxembourg, tom.barbereau@uni.lu

Johannes Sedlmeir, University of Luxembourg, Luxembourg City, Luxembourg, johannes.sedlmeir@uni.lu

Liudmila Zavolokina, University of Zurich, Zurich, Switzerland, liudmila.zavolokina@dsi.uzh.ch

Abstract

The Digital Markets Act (DMA) constitutes a crucial part of the European legislative framework addressing the dominance of 'Big Tech'. It intends to foster fairness and competition in Europe's digital platform economy by imposing obligations on 'gatekeepers' to share end-user-related information with business users. Yet, this may involve the processing of personal data subject to the General Data Protection Regulation (GDPR). The obligation to provide access to personal data in a GDPR-compliant manner poses a regulatory and technical challenge and can serve as a justification for gatekeepers to refrain from data sharing. In this research-in-progress paper, we analyze key tensions between the DMA and the GDPR through the paradox perspective. We argue through a task-technology fit approach how privacy-enhancing technologies – particularly anonymization techniques – and portability could help mediate tensions between data sharing and privacy. Our contribution provides theoretical and practical insights to facilitate legal compliance.

Keywords: Data Sharing, Digital Markets Act, Decision Tree, General Data Protection Regulation, Differential Privacy, Privacy-Enhancing Technologies

1 Introduction

Recent geopolitical dynamics in digital markets are steered by the rise of giant technology companies, with revenues exceeding many countries' GDP (Werthner, 2022). With the growth of the information economy triggered by various technological developments, the market dominance and reach of many Chinese and American firms has expanded substantially. For Europe's digital market and its firms, this data asymmetry and potential abuse of market power pose structural problems that threaten contestability and competition (Cabral et al., 2021). In response, the Digital Markets Act (DMA) constitutes a crucial part of the legislative framework addressing the dominance of "Big Tech" in the European Union (EU) (European Commission, 2022). The regulation aims to shift the balance of competition in the European digital market by preventing particularly large companies from abusing their dominance. Google, Amazon, Meta, Apple, Microsoft (aka GAMAM) and others are classified as "gatekeepers" (Article 3 (DMA)) and will be directly affected. Article 6 (DMA) lists several obligations for these gatekeepers to avoid practices that limit contestability, such as the sharing of end users' personal data with business users. At the same

time, the DMA explicitly states that it will remain "without prejudice" to the General Data Protection Regulation (GDPR) (European Commission, 2016). The GDPR aims to protect end-users' privacy by restricting the sharing of personal data by third parties. The DMA, on the other hand, aims at stimulating data sharing among businesses. Thereby, the GDPR frankly targets the protection of end-users, while the DMA targets the protection of business users. Consequently, gatekeepers are inclined to use the argument to protect the privacy of individual end-users as a pretense to prioritize compliance with GDPR over the obligations imposed upon them by the DMA. Thus, it remains unclear whether and how the tension between these two regulations can be resolved in practice (Etteldorf, 2022).

These challenges raise the following research question: *How can technology help satisfy the different regulatory requirements of the GDPR and the DMA and thereby alleviate the tension between data protection and data sharing?* Being aware of the technological requirements needed to address the GDPR, organizations are yet to find means to simultaneously comply with the obligations stipulated in the DMA. Based on requirements derived from legislations and a review of privacy-enhancing technologies (PETs), we contribute a decision tree that helps to resolve the tension between the need for data sharing and upholding data protection requirements. The tentative decision tree we propose as research-in-progress is derived based on task-technology fit (TTF) (Goodhue et al., 1995). It serves as the basis for a paper that refines and evaluates this work in a Design Science Research (DSR) approach with perspectives from gatekeepers, business users, and policy experts (Hevner et al., 2004; Peffers et al., 2007).

This research-in-progress paper is structured as follows. Section 2 introduces background on Europe's regulatory response to online platforms and the tensions between data sharing and privacy from the paradox perspective. Section 3 presents our TTF approach. In Section 4, we present the regulatory requirements derived from the GDPR and DMA. We then discuss our proposed solution in terms of selected technical tools and our tentative artifact – a decision tree that reflects a suitable subset of PETs – in Section 5. We conclude by discussing our contribution so far and the potential implications of our planned research for both theory and practice in Section 6.

2 Background

2.1 Europe's regulatory response to online platforms

A key strategy of EU policy makers to address the issue of increasingly dominant digital platforms and corresponding anti-competitive behavior in the digital market is the implementation of digital policies (Bradford, 2020; Metakides, 2022). This materializes in a series of regulatory measures, including the GDPR, the Data Governance Act, the Data Act, the new Artificial Intelligence Regulation, the proposed European Digital Identity Framework, as well as the Digital Services Act (DSA) and the DMA (Codagnone et al., 2023). The DMA was introduced by the European Commission in December 2020 and is in force since November 2022. It comprises several provisions that impose obligations on gatekeepers. At the regulation's core, gatekeepers need to provide some end-user-related data to their business users (European Commission, 2022). The legal obligations imposed on these gatekeepers share the objective to weaken companies' exclusive access to large sets of collected data and to eliminate corresponding market distortions inside the platform. In particular, Article 6(10) (DMA) obliges gatekeepers to provide business users with effective, high-quality, continuous, and real-time access to data, including personal data. Fines for non-compliance can amount to up to 10 % of gatekeepers' total global turnover in the preceding financial year, and up to 20 % in the case of recurring infringements (ibid.).

Both the access to and the sharing of end-user-related data may constitute the processing of personal data subject to the GDPR. Article 5 (GDPR) stipulates several data protection principles that both data sharing and receiving entities need to comply with in general and under the DMA specifically. Article 5(1)(b) (GDPR) requires data controllers to process personal information only for "specified, explicit and legitimate purposes" that are compatible with the original purpose of collecting the data.

Under the DMA, this means that both gatekeepers who share the data and business users who receive that data may find it challenging to comply with the purpose specification and limitation principles. Data recipients are additionally required to ensure an appropriate legal basis for their processing activities. If personal data is no longer necessary for the purpose it was originally collected and processed, an individual has, under the GDPR, the right to have their personal data erased. This 'right to be forgotten' is stipulated in Article 17 (GDPR). When gatekeepers are required to share data with a potentially indefinite number of recipients under the DMA, the erasure of data "without undue delay" may be difficult to achieve.
Overall, the ongoing legal construction of digital policies yields a complex and fragmented landscape, "making at times regulatory coherence and consistency hard to be achieved" (Codagnone et al., 2022, p. 9). On the one hand, this could mean that if market entrants, as data recipients, do not have the resources for sophisticated legal analysis, drive smaller and medium-sized enterprises may be driven into non-compliance, and thus fail to foster contestability or fairness as originally envisioned in the DMA. On the other hand, such legislative inconsistencies, including the pretext of preserving user privacy, can be used by gatekeepers to justify insufficient implementation, as seen during the public consultation. This refers in particular to the feedback addressed to the European Commission by Google (European Commission, 2020a) and Meta (European Commission, 2020b) on 30 June 2020, and to the feedback provided by Microsoft (European Commission, 2020c) on 3 May 2021. Despite the surge of digital policy made in the EU over the last decade, little research has focused on using technological tools to approach and facilitate regulatory compliance for concerned actors.

2.2 Tensions between data sharing and privacy

The data-sharing literature agrees that legislation is an important factor for guiding how data are shared and used (Dawes, 1996, 2010; Mayer-Schönberger et al., 2018; Sokol et al., 2021), notably in light of the emergence and increasing influence of platform intermediaries (Parker et al., 2017). Policy frameworks provide an attempt to maintain a commitment to transparency when sharing data, all while complying with data privacy regulation (Dawes, 1996). This understanding is at the core of the DMA (Larouche et al., 2021). Challenges related to conflicting regulations and lack of guidance for data sharing projects are not a new development per se (Dawes et al., 2009; Nelson, 2004; Weber et al., 2008). In the case of the DMA, a crucial tension can be identified between the obligation to provide access to data (DMA) for business users and upholding data protection principles (GDPR) to safeguard the privacy of end-users. Specifically, this tension relates to the obligation to comply with purpose specification and limitation principles and to obtain explicit consent from data subjects when sharing and processing continuous, real-time data. In this case, on the one hand, the purpose as perceived by the data-sharing entity and the data recipient may not be identical. On the other hand, obtaining legally valid consent from each data subject within huge data sets may pose another challenge. It should be mentioned that, though not adopted in the final legislative text of the DMA, the European Parliament suggested an amendment to Article 6(10) (DMA) so that business users would have the "possibility and tools to access and analyze data 'in-situ' without a transfer from the gatekeeper". Such an in-situ mechanism should increase data security by bringing the business users' algorithms to where the data is stored. However, this relatively simple solution is subject to limitations, as queries on collections of personal data may very well have personally identifiable results.
We analyze the key tensions between the DMA and the GDPR through the *paradox perspective*, which has been applied in information systems (IS) research to identify the nature of tensions and corresponding resolutions (Ciriello et al., 2019). Schad et al. (2016, p. 10) define a paradox as the "persistent contradiction between interdependent elements". According to Poole et al. (1989), the paradox perspective consists of (1) a contradiction between two propositions (the tension) and (2) a resolution of the tension. To resolve a paradox, Poole et al. (1989) provide four approaches: (1) Acceptance: keeping the opposing theses separate and their contrasts appreciated; (2) Spatial separation: situating the opposing phenomena at

different locations; (3) Temporal separation: separate the paradox temporally in the same location; and (4) Synthesis: finding a new perspective that eliminates the opposition between the two phenomena.
For the case of upholding data privacy upon sharing, the first method of acceptance (1) is not recommended because the sharing of data potentially involves personal data, which is regulated under GDPR. It would be simplistic to assume that the two contrasting premises of providing access to data while keeping it confidential at the same time can be accepted. Arguably, the challenge to comply with both the GDPR and the DMA is too consequential for end-users, and the fines too costly for businesses to be simply solved by acknowledgment. By contrast, spatial separation (2) could provide a means to resolve tensions between data sharing and data privacy by splitting the former between the gatekeeper and the business user: (a) have the user download their data, and (b) forward it to the business user. As we discuss in Section 5, this approach based on data portability has some merits and aligns with the GDPR's consent requirement. Note that when accessed once, data can be duplicated or reproduced at negligible costs, and restrictions to exclusive use can hardly be enforced. Poole et al. (1989)'s third method, temporal separation (3), seems problematic. It infringes the real-time requirement as imposed by DMA. Moreover, privacy breaches often cannot be reversed. While previously accessible information can be made inaccessible, existing copies' deletion cannot be enforced practically, and consequences are often irreversible. The fourth method of synthesizing the different challenges by adopting a new perspective (4) appears to be the most promising procedure. The new perspective we propose to supplement the in-situ approach and to resolve the remaining tensions incorporates privacy-enhancing technologies.

3 Research Design

To answer our research question, we adopt the TTF lens. The concept was introduced as "the degree to which a technology assists an individual in performing his or her portfolio of tasks" (Goodhue et al., 1995). TTF has been frequently used in IS research to study the use of technologies in specific contexts. The theory proposes that a technology's use depends on the fit to the task to be performed. By matching the technology's characteristics to the needs of the task, researchers can systematically assess the alignment between technology and task. For this study, we have chosen TTF to map regulatory requirements on data sharing (DMA) and privacy (GDPR) to technical tools. We consider this perspective appropriate since it helps us systematically assess the alignment between the technical tools (such as PETs) and regulatory requirements. It also helps identify potential gaps that need to be addressed by additional technologies. We adhere to the recommendations of Zigurs et al. (1998) to conceptualize the 'tasks' and 'fit'.
Our research process consists of two steps. First, we identify regulatory requirements and create a tentative decision tree on the basis of regulatory documents and literature on PETs. For the research-in-progress paper, we conduct a preliminary keyword search in IS databases that includes "PETs" as well as the specific PETs as surveyed in the systematic literature review by Garrido et al., 2022, such as "multi-party computation" and "differential privacy". We aim to extend our literature review to a systematic literature review based on keywords that we will extract from this preliminary keyword search (as recommended by Kitchenham, 2004) in our future research. The decision tree can be described as "a model of a [...] problem in the form of interpretable and actionable rules" (Osei-Bryson et al., 2011).
Second, as will be done in the future, we seek to iteratively evaluate and revise this decision tree in the context of use cases in the e-commerce sector. According to an analytical paper listed in the impact assessment report for the DMA, a significant share of e-commerce sellers operating on larger platforms, such as Amazon, is dissatisfied with the lack of access to data, including customer data as well as financial, listing and advertising data that they consider essential to maintaining their businesses (Gineikytė et al., 2020). Naturally, the DMA will also have an impact on a number of other sectors, such as the accommodation and hospitality industry, or regarding apps and software development. However, we assume the e-commerce sector to be one of the most non-competitive cases because its business users, unlike app developers, do typically not have the capacities to run data and market analytics and have

insights on the performance of their products. This arguably makes e-commerce platforms more powerful. Moreover, business actors in the hospitality industry appear to have comparatively lower data needs (Gineikytė et al., 2020). Further, the e-commerce sector is particularly interesting, as it covers the frequent dynamics from a B2C perspective (e.g., e-commerce sellers interacting with customers on a larger e-commerce platform) and B2B perspective (e.g., e-commerce sellers interacting with the e-commerce platform on which their business can operate). After the completion of both steps, the study will serve as the groundwork for a DSR paper that substantiates our findings (Hevner et al., 2004; Peffers et al., 2007).

4 Identification of Regulatory Requirements

To define our 'task', we identified key regulatory requirements based on the original legislative texts of the DMA and the GDPR as published in the Official Journal of the European Union (European Commission, 2016, 2022). We also studied the accompanying documents of the DMA. These documents were issued by the corresponding European institutions and authorities and are publicly accessible. They include the opinions of the European Data Protection Supervisor, discussions by the Council of the EU, as well as European Parliament amendments and the Commission's Impact Assessment. The regulations themselves in particular, detail the various obligations between the parties and stipulate important data sharing and data privacy principles. The DMA refers to the obligation to comply with the GDPR (therein denoted as Regulation (EU) 2016/679) 22 times, excluding footnotes. Compliance with the GDPR is, therefore, a general premise. At the same time, it is necessary to understand that both the DMA and the GDPR have different legal purposes and different regulatory histories. Although the DMA does not strictly fall under the scope of competition law (Podszun et al., 2021), it is known as a 'competition tool' with the objective to regulate the market. The GDPR, on the other hand, does not aim at regulating the market, but at the protection of individuals. By adopting such protective approach for individuals' privacy, its main function is to further the fundamental right of data protection. When confronted with competition tool and the data protection law, it is important to keep in mind that neither should systematically have priority over the other. Both regulations stipulate requirements that businesses must comply with in order to ensure effective personal data protection as a fundamental right, while at the same time facilitating a competitive market and economic growth in the digital economy as a key function in society. We derived some key regulatory requirements from the DMA (R1 and R3), which are especially relevant for business users, one requirement set out by both regulations (R2) and two regulatory requirements from the GDPR (R4 and R5), which are of particular importance to protecting the privacy of end-users.

R1: Real-time access. Article 6(9) (DMA) specifies that gatekeepers must provide end users, or authorized third parties, with effective and free data portability in accordance with the GDPR, "including by the provision of continuous and real-time access". Article 6(10) (DMA) stipulates that gatekeepers ought to ensure business users and authorized third parties "continuous and real-time access and use of [...] data".

R2: Accuracy. Article 6(10) (DMA) also obliges gatekeepers to provide business users with "effective [and] high-quality [...] access to [...] data, including personal data, that is provided for or generated in the context of the use of the relevant core platform services [...]." Additionally, Article 5 (GDPR) emphasizes that data should be "accurate and, where necessary, kept up to date" In conclusion, the shared data must be accurate, reliable, and reflect the information required by the business user.

R3: Flexibility. In addition to the context-specificity of data sharing processes, the DMA mandates that both "aggregated and non-aggregated data must be accessible to business users" (Article 6(10) (DMA)). In Article 12 (DMA), the Commission further strenghtens its ability to adopt delegated acts that can extend a legal obligation "in relation to certain types of data." From this can be inferred that a certain degree of flexibility regarding the type of data in question – including different algorithms evaluated on data in the in-situ approach – must be supported by gatekeepers to ensure compliance with the DMA.

R4: Purpose limitation. Article 5 (GDPR) encompasses data protection principles that sharing and receiving entities need to comply with. Article 5(1)(b) (GDPR) requires data controllers to process

information for "specified, explicit and legitimate purposes" compatible with the original purpose of data collection. Data recipients are required to ensure an appropriate legal basis for their processing activities. *R5: Consent*. Under the GDPR, user consent constitutes one of the six legal bases for processing personal data. Recital 40 (GDPR) demands that "in order for processing to be lawful, personal data should be processed on the basis of the consent of the data subject concerned or some other legitimate basis". The requirements for the validity of legal consent are specified further in Article 7 and Recital 32 (GDPR).

5 Proposed Solution

It is apparent that the real-time sharing of gatekeepers' raw data about end users is difficult to align with GDPR. The in-situ mechanism could counteract this challenge by eliminating the need to download and evaluate the data on business users' side. Instead, it brings business users' algorithms to the gatekeeper, who then runs the algorithm on end users' data and reports the result to the business user. However, there is no guarantee that the result of the evaluation – which we call "query" for simplicity – is not personally identifiable (e.g., when extracting a single row corresponding to an individual end user from a large dataset) or otherwise sensitive. Hence, a holistic solution yet remains to be identified.

Previous research has come up with many different technical solutions to facilitate data sharing while addressing requirements of protecting sensitive information, termed PETs (Sonehara et al., 2011). Following Garrido et al. (2022), PETs can be split into techniques for anonymization – statistical disclosure mechanisms that break the link between individuals and data points or provide individuals plausible deniability – and secure computation, which includes secure hardware and cryptographic techniques that allow running an algorithm on data in some kind of a black box. As intended obfuscation of sensitive data often demands additional authenticity or integrity guarantees to verify the results, authenticity-enhancing technologies like digital signatures, notarization (e.g., via publishing hashes of data on a blockchain), or zero-knowledge proofs are an important building block of many of these approaches in practice (Garrido et al., 2022; Schellinger et al., 2022). The use of both these types of PETs in data markets and beyond is an active and innovative research field (e.g., Agahari et al., 2021; Garrido et al., 2023). However, there are relatively few investigations on how to integrate them into existing IS to solve organizations' challenges concerning the handling of personal information. One of the few examples was published by Zöll et al. (2021), who investigated organizational adoption barriers of PETs. These difficulties in using PETs and the heterogeneity of solutions suggest that selecting the right technology is crucial for success.

Generally, anonymization techniques are characterized by clustering data and adding noise to data or query results to conceal specific information (Garrido et al., 2022). The trade-off is often a decrease in accuracy and/or utility, which tends to become less pronounced when the number of data points is large. The most popular example is differential privacy (DP) as introduced by Dwork (2008). Using DP, gatekeepers would add noise to business users' query results on an 'in-situ' computation. DP has so far been applied in IS to protect sensitive event logs (Mannhardt et al., 2019) or share multiple records of data associated with a user (Kartal et al., 2018). Further anonymization tools have been analyzed to address challenges in deploying PETs in the context of wearables and smart cars (Bondel et al., 2020).

Secure computation, on the other hand, refers to techniques that seek to provide no more information than required to the data processor, i.e., hiding the data on which an algorithm is executed. Corresponding tools comprise multi-party computation for algorithms that involve sensitive data for many entities (Goldwasser, 1997). Within IS, research on secure computing techniques is still nascent (Agahari et al., 2021, 2022a,b). Yet, multi-party computations (MPCs) are not required in a setting where the gatekeeper has all the information that is potentially used for computations. However, when considering the in-situ approach discussed previously, business users may demand some guarantees on the correctness of the computation's result. One of the most established techniques to provide this assurance without leaking the underlying data are zero-knowledge proofs (ZKPs). They have recently also gained some attention in IS research in the context of blockchain solutions (Mattke et al., 2019; Zhang et al., 2021). We found no IS research

applying fully homomorphic encryption (FHE) and trusted execution environment (TEE). FHE allows to perform computations on encrypted data, so the data processor does not learn about data that they are working on. Yet, FHE seems to be more appropriate in a setting where the data provider wants to have some modifications to their data without leaking any information about their data, such as image processing. Finally, TEEs could allow to send encrypted data from gatekeepers to business users and allow them to do only very specific, previously fixed computations. Despite better performance than the previously described secure computing technologies, TEEs are difficult to set up, lead to vendor lock-in, and could be compromised with sophisticated attacks (Garrido et al., 2022).

Henceforth, we assume that the in-situ approach, where the data owner (gatekeeper) does not directly share data but instead shares the results of specific, potentially business user-defined algorithms that they locally run on their data, is indeed the most promising. Yet, specific attention is necessary to ensure that the results of these computations are not personally identifiable anymore. For example, micro-data can sometimes be retrieved from aggregate statistics surprisingly well (Dick et al., 2023). Even applying sophisticated anonymization techniques can fail when there are potential threats of de-anonymization, e.g., linking attacks through using additional data sets that were not anticipated at the time of anonymization – as has happened when combining video streaming usage data from different platforms (Narayanan et al., 2008) or combining mobility data with mobile network usage data (Kondor et al., 2020). In contrast, differential privacy provides privacy guarantees even under linking attacks as long as the composed data is independent, and a privacy budget if the data is correlated (Dwork et al., 2014). However, the tradeoff between privacy, utility, and complexity (Garrido et al., 2023) needs to be considered.

Based on this review of PETs, we consider only simple anonymization (through deleting strongly identifying information) and DP out of the six PETs we analyzed. Yet, our tentative decision tree also accounts for cases where none of these anonymization techniques is suitable, namely, where the privacy guarantees of simple anonymization do not suffice but where accuracy tradeoffs in DP are not acceptable either. In this case, presumably the only way to satisfy requirement 3 (purpose limitation) is getting end users' consent to the specific use of their sensitive information. We suggest levering the portability of end users' data that GDPR and DMA both demand to achieve this: When business users require information, they could directly approach end-users (e.g., during their purchase) and ask them to forward parts of their data directly to the service provider, i.e., the business user. Through this explicit consent and the limitation of the exchanged data, GDPR-related requirements can be addressed. For some standardized and machine-readable data, such as digital attestations of identity attributes, the European digital wallet initiative pushed in the context of the eIDAS 2.0 revision may already provide a good starting point for a technical solution that implements portability (Sedlmeir et al., 2022).

Our tentative decision tree (Figure 1) establishes a fit between the task and the discssed technologies. It structures the decision-making process for gatekeepers in the form of a tree with decision nodes, where a

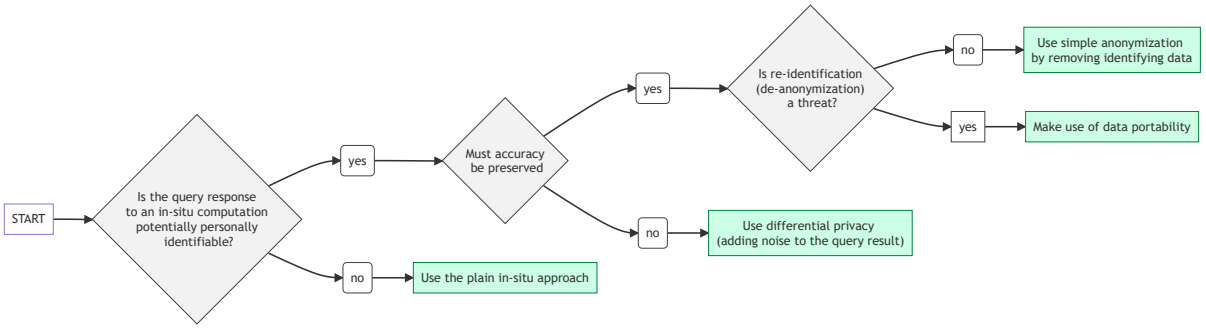

Figure 1. Tentative decision tree.

decision must be made on a specific regulatory requirement, and leaf nodes, which show the opportunity to include a PET or leverage portability. Our decision tree acknowledges that secure computing techniques may be less suitable for facilitating the real-time sharing of end users' personal information because of the challenges associated with setting them up in a scalable way on both gatekeepers' and in particular business users' systems: Flexibility aims to allow for a broad range of potential evaluations of data. If the result of a computation on this data is not personally identifiable (e.g., because it represents a differentially private query), we can satisfy purpose limitation trivially as the GDPR does not apply. If, in contrast, a query would be personally identifiable, our decision tree proposes to make use of data portability and have the data transmitted to the service provider directly through the affected person, who can then explicitly consent to this data use. Simple anonymization techniques seem preferable when the risk of de-anonymization attacks through linking other datasets seems low because they are simple to use and do not decrease the data's accuracy. When de-anonymization is a risk, differential privacy can be used for mathematical anonymity guarantees, yet at the cost of accuracy, particularly for datasets that involve fewer end users. If this accuracy tradeoff is not acceptable, the portability of data that needs to be implemented in any case as demanded by GDPR and DMA can be leveraged: Business users can ask end users directly for the data that they need, and end users can retrieve it from the gatekeeper for this purpose.

6 Discussion and Future Work

Our paper aims to contribute to research on (1) data sharing and privacy as an envisioned tension, resulting from the DMA, and (2) technical approaches as a potential solution. As for data sharing and privacy, we illustrate how the newly introduced DMA causes regulatory tension and increases compliance issues for market players and markets entrants in particular. Our proposed decision tree incorporates mechanisms of tension resolutions in the form of synthesis (anonymization techniques) and spatial separation (leveraging data portability also for real-time data sharing). Thus, our decision tree aims to serve as a decision support tool for the affected parties. The proposed solution indicates that data sharing and data protection can go hand in hand, and that gatekeepers should not use data privacy protection of users as an excuse not to share data with business users. As for technical approaches, we review modern PETs and propose a useful combination of selected tools to address policy requirements. Furthermore, utilizing the paradox perspective (Ciriello et al., 2019; Schad et al., 2016) with ways to resolve tensions might provide impetus of how PETs can be leveraged by organizations to comply with new regulations, e.g., through temporal and spatial separation or synthesis (Poole et al., 1989). However, our research is limited by the need for further empirical evidence to showcase exactly to which specific business situation the decision tree can be applied, and to what extent such application can be generalized to different industries. This is something that the further development of this research will consider.

We also believe that the end users' perspective on non-anonymized data should be further explored. This becomes relevant in both the consent-based approach (when a user consents to the gatekeeper processing or sharing their data) and through granted portability (a user downloads their data from the gatekeeper, a business can approach the user and ask for parts of this data). In the first case, it may be difficult for the user to assess how much and which data will be shared. In the second case, how the interaction between a user and a business can be established remains unclear. The approach may also trigger interesting ways of data monetization. In this interaction, it may be challenging for service providers to know how to contact the user without immediately exposing the user's identity. Here, digital wallets could again be a remedy for ensuring the end users' herd privacy (Schlatt et al., 2022). These questions should be further explored.

We also aim to contribute to practice by providing policy recommendations for the EU to avoid too general, vague, and fragmenting regulations that, without technical guidance, do not align well with each other and might even harm European players in the single market more than multinational companies. In the next step, we aim to evaluate our decision tree in a series of expert interviews with policy or legal experts, gatekeepers and their potential business users, to assess the feasibility and usefulness of our proposal.

References

Agahari, W., R. Dolci, and G. de Reuver (2021). "Business model implications of privacy-preserving technologies in data marketplaces: The case of multi-party computation." In: *Proceedings of the 29th European Conference on Information Systems*. AIS.

Agahari, W., H. Ofe, and M. de Reuver (2022a). "It is not (only) about privacy: How multi-party computation redefines control, trust, and risk in data sharing." *Electronic Markets* 32, 1577–1602.

Agahari, W. and M. de Reuver (2022b). "Rethinking consumers' data sharing decisions with the emergence of multi-party computation: An experimental design for evaluation." In: *Proceedings of the 30th European Conference on Information Systems*. AIS.

Bondel, G., G. M. Garrido, K. Baumer, and F. Matthes (2020). "Towards a privacy-enhancing tool based on de-identification methods." In: *Proceedings of the Pacific Asia Conference on Information Systems*. AIS.

Bradford, A. (2020). *The Brussels effect: How the European Union rules the world*. Oxford University Press.

Cabral, L., J. Haucap, G. Parker, G. Petropoulos, T. Valletti, and M. v. Alstyne (2021). *The EU Digital Markets Act: A Report from a Panel of Economic Experts*. Tech. rep. LU: Publications Office of the European Union. URL: https://data.europa.eu/doi/10.2760/139337 (visited on 03/27/2023).

Ciriello, R. F., A. Richter, and G. Schwabe (2019). "The paradoxical effects of digital artefacts on innovation practices." *European Journal of Information Systems* 28 (2), 149–172.

Codagnone, C., G. Liva, and T. Rodriguez de las Heras Ballell (2022). *Identification and Assessment of Existing and Draft EU Legislation in the Digital Field*. Study for the Special Committee on Artificial Intelligence in a Digital Age (AIDA). Luxembourg: Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament. URL: https://www.europarl.europa.eu/RegData/etudes/STUD/2022/703345/IPOL_STU(2022)703345_EN.pdf (visited on 03/27/2023).

Codagnone, C. and L. Weigl (2023). "Leading the charge on digital regulation: The more, the better, or policy bubble?" *Digital Society* 2 (4).

Dawes, S. S. (1996). "Interagency information sharing: Expected benefits, manageable risks." *Journal of Policy Analysis and Management* 15 (3), 377–394.

Dawes, S. S. (2010). "Stewardship and usefulness: Policy principles for information-based transparency." *Government Information Quarterly* 27 (4), 377–383.

Dawes, S. S., A. M. Cresswell, and T. A. Pardo (2009). "From "need to know" to "need to share": Tangled problems, information boundaries, and the building of public sector knowledge networks." *Public Administration Review* 69 (3), 392–402.

Dick, T., C. Dwork, M. Kearns, T. Liu, A. Roth, G. Vietri, and Z. S. Wu (2023). "Confidence-ranked reconstruction of census microdata from published statistics." *Proceedings of the National Academy of Sciences* 120 (8).

Dwork, C. (2008). "Differential privacy: A survey of results." In: *International Conference on Theory and Applications of Models of Computation*. Springer.

Dwork, C., A. Roth, et al. (2014). "The algorithmic foundations of differential privacy." *Foundations and Trends in Theoretical Computer Science* 9 (3–4), 211–407.

Etteldorf, C. (2022). "DMA – Digital Markets Act or Data Markets Act?" *European Data Protection Law Review* 8 (2), 255–261.

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 (2022). *Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act)*.

European Commission (2020a). *Feedback from: Google on Digital Services Act package – ex ante regulatory instrument of very large online platforms acting as gatekeepers.* URL: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12418-Digital-Services-Act-package-ex-ante-regulatory-instrument-of-very-large-online-platforms-acting-as-gatekeepers/F535552_en (visited on 03/30/2020).

European Commission (2020b). *Feedback from: Meta on Digital Services Act package – ex ante regulatory instrument of very large online platforms acting as gatekeepers.* URL: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12418-Digital-Services-Act-package-ex-ante-regulatory-instrument-of-very-large-online-platforms-acting-as-gatekeepers/F535672_en (visited on 03/30/2020).

European Commission (2020c). *Feedback from: Microsoft on Digital Services Act package – ex ante regulatory instrument of very large online platforms acting as gatekeepers.* URL: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12418-Digital-Services-Act-package-ex-ante-regulatory-instrument-of-very-large-online-platforms-acting-as-gatekeepers/F2256709_en (visited on 03/30/2020).

Garrido, G. M., X. Liu, F. Matthes, and D. Song (2023). "Lessons learned: Surveying the practicality of differential privacy in the industry." In: *Proceedings of the 23rd Privacy Enhancing Technologies Symposium*.

Garrido, G. M., J. Sedlmeir, Ö. Uludağ, I. S. Alaoui, A. Luckow, and F. Matthes (2022). "Revealing the landscape of privacy-enhancing technologies in the context of data markets for the IoT: A systematic literature review." *Journal of Network and Computer Applications* 207, 103465.

Gineikytė, V., E. Barcevičius, and L. Matulevič (2020). *Platform data access and secondary data sources*. Tech. rep. PPMI. URL: https://platformobservatory.eu/app/uploads/2020/09/Analytical-paper-1-Platform-data-access-and-secondary-data-sources_final.pdf (visited on 03/27/2023).

Goldwasser, S. (1997). "Multi party computations: Past and present." In: *Proceedings of the 16th Annual ACM Symposium on Principles of Distributed Computing*.

Goodhue, D. L. and R. L. Thompson (1995). "Task-technology fit and individual performance." *MIS Quarterly* 19 (2), 213–236.

Hevner, A. R., S. T. March, J. Park, and S. Ram (2004). "Design science in information systems research." *MIS Quarterly* 28 (1), 75–105.

Kartal, H. and X.-B. Li (2018). "Protecting privacy when releasing query results with multiple records per person." In: *Proceedings of the 39th International Conference on Information Systems*. AIS.

Kondor, D., B. Hashemian, Y.-A. de Montjoye, and C. Ratti (2020). "Towards matching user mobility traces in large-scale datasets." *IEEE Transactions on Big Data* 6 (4), 714–726.

Larouche, P. and A. de Streel (2021). "The European digital markets act: A revolution grounded on traditions." *Journal of European Competition Law & Practice* 12 (7), 542–560.

Mannhardt, F., A. Koschmider, N. Baracaldo, M. Weidlich, and J. Michael (2019). "Privacy-preserving process mining: Differential privacy for event logs." *Business & Information Systems Engineering* 61 (5), 595–614.

Mattke, J., A. Hund, C. Maier, and T. Weitzel (2019). "How an enterprise blockchain application in the US pharmaceuticals supply chain is saving lives." *MIS Quarterly Executive* 18 (4), 245–261.

Mayer-Schönberger, V. and T. Ramge (2018). *Reinventing capitalism in the age of big data*. Hachette UK.

Metakides, G. (2022). "A crucial decade for European digital sovereignty." In: *Perspectives on Digital Humanism*. Springer, pp. 219–225.

Narayanan, A. and V. Shmatikov (2008). "Robust de-anonymization of large sparse datasets." In: *IEEE Symposium on Security and Privacy*, pp. 111–125.

Nelson, L. (2004). "Privacy and technology: Reconsidering a crucial public policy debate in the post-September 11 era." *Public Administration Review* 64 (3), 259–269.

Osei-Bryson, K.-M. and O. Ngwenyama (2011). "Using decision tree modelling to support Peircian abduction in IS research: a systematic approach for generating and evaluating hypotheses for systematic theory development." *Information Systems Journal* 21 (5), 407–440.

Parker, G., M. Van Alstyne, and X. Jiang (2017). "Platform Ecosystems." *MIS Quarterly* 41 (1), 255–266.

Peffers, K., T. Tuunanen, M. A. Rothenberger, and S. Chatterjee (2007). "A design science research methodology for information systems research." *Journal of Management Information Systems* 24 (3), 45–77.

Podszun, R. and P. Bongartz (2021). "The digital markets act: Moving from competition law to regulation for large gatekeepers." *Journal of European Consumer and Market Law* 10 (2).

Poole, M. S. and A. H. Van de Ven (1989). "Using paradox to build management and organization theories." *Academy of Management Review* 14 (4), 562–578.

Schad, J., M. W. Lewis, S. Raisch, and W. K. Smith (2016). "Paradox research in management science: Looking back to move forward." *Academy of Management Annals* 10 (1), 5–64.

Schellinger, B., F. Völter, J. Sedlmeir, and N. Urbach (2022). "Yes, I Do: Marrying blockchain applications with GDPR." In: *Proceedings of the 55th Hawaii International Conference on System Sciences*, pp. 4631–4640.

Schlatt, V., J. Sedlmeir, S. Feulner, and N. Urbach (2022). "Designing a framework for digital KYC processes built on blockchain-based self-sovereign identity." *Information & Management* 59 (7), 103553.

Sedlmeir, J., T. Barbereau, J. Huber, L. Weigl, and T. Roth (2022). "Transition pathways towards design principles of self-sovereign identity." In: *Proceedings of the 43rd International Conference on Information Systems*. AIS.

Sokol, D. D. and M. Van Alstyne (2021). "The rising risk of platform regulation." *MIT Sloan Management Review* 62 (2), 6A–10A.

Sonehara, N., I. Echizen, and S. Wohlgemuth (2011). "Isolation in cloud computing and privacy-enhancing technologies." *Business & Information Systems Engineering* 3 (3), 155–162.

Weber, E. P. and A. M. Khademian (2008). "Wicked problems, knowledge challenges, and collaborative capacity builders in network settings." *Public Administration Review* 68 (2), 334–349.

Werthner, H. (2022). "Geopolitics, digital sovereignty... What's in a word?" In: *Perspectives on Digital Humanism*. Ed. by C. Ghezzi. Springer, pp. 241–248.

Zhang, W., C.-P. Wei, Q. Jiang, C.-H. Peng, and J. L. Zhao (2021). "Beyond the block: A novel blockchain-based technical model for long-term care insurance." *Journal of Management Information Systems* 38 (2), 374–400.

Zigurs, I. and B. K. Buckland (1998). "A theory of task/technology fit and group support systems effectiveness." *MIS Quarterly* 22 (3), 313–334.

Zöll, A., C. M. Olt, and P. Buxmann (2021). "Privacy-sensitive business models: Barriers of organizational adoption of privacy-enhancing technologies." In: *Proceedings of the 29th European Conference on Information Systems*. AIS.

Acknowledgements

This research was funded in part by the Luxembourg National Research Fund (FNR) through the PABLO project (grant reference 16326754) and by PayPal, grant reference "P17/IS/13342933/PayPal-FNR/Chair in DFS/Gilbert Fridgen" (PEARL) as well as through the University of Zurich and the Digital Society Initiative under the DIZH postdoc fellowship of Liudmila Zavolokina.