**A Novel Data Governance Scheme Based on the Behavioral
Economics Theory**

Bohan Hou

PnS Institute

bohanhou@outlook.com

Jan.27, 2021

Contents

Executive Summary

Glossary

Background

1 Introduction

2 The Unsolvable Yet Unavoidable Problems in Data Laws

3 Re-Construct Problems

4 Behavioral Economics Solutions

5 Components of the New Data Governance

6 Potential Scenarios under the New Data Governance

7 Conclusion

8 More Scenarios

References

Executive Summary

The digital economy has become one of the most important sectors in global GDP. Personal data is the new asset class that creates value through the applications of cybertechnologies and Artificial Intelligence. However, there are increasing concerns over the privacy invasions and human rights violations associated with the exploitation of personal data.

Various data laws were made in nations to balance the data fluidity and privacy protections. However, most laws have inherent limitations and underenforcement issues that fail to achieve their aims and protection principles. Utilizing a behavioral economics theoretical framework, this study categorizes the issues and causes to Information Asymmetry, Bounded Rationality, Power Imbalance, and Technical Incapacity.

The study makes a novel contribution by proposing a global data governance scheme to address the limitations of data laws. The scheme adopts a Libertarian Paternalism approach and develops seven principles in the framework design. Elements and components in the scheme include individuals, data controllers, privacy rating frameworks, meta-data and privacy configuration, reports, Automated Consent Management (ACM), Bureaus, and signatures, etc. The components will operate on an interoperable and global data management platform. Visual diagrams are developed to describe the various forms of interactions between components and procedures.

A balance between privacy protection and data fluidity is found through experimental scenarios such as Ordinary Data Request, Sensitive Data Request, Inconsistency Checks, Data Rights Exercise, Monitored Data Transfer, Broadcast and Notice.

Data Governance Scheme

The scenarios analyzed are not exhaustive but serve as the meaningful starting point to inspire more designs and discussions from scholars.

Keyword: Big data, Data Protection, Behavioral Economics, Architecture, Interoperability

Glossary

Automated Consent Management (ACM) – an automated agent that helps individuals manage their personal data and realize most functions of the data governance in conjunction with bureaus.

Bureaus – the digital authorities in the data governance that makes up the power imbalance between individuals and data controllers and enhance enforcement power through participating in different data processing scenarios.

Central Technical Council – an international, neutral, trustworthy group that monitors all encryption-related activities through dispatching Central Technical Experts to data controllers.

Central Technical Experts (CTE) – the specialized personnel who are loyal to the Council and are assigned to different data controllers to monitor encryption-related activities.

Data Controller -- the natural or legal person, public authority, agency, or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data.

Data Management Platform – a global, interoperable digital platform that provides interfaces for individuals, data controllers to manage their data and supports components in the novel data governance design.

Individual (Data Subject) – the identifiable natural person that the personal data relates to.

Privacy Configuration – a set of privacy preferences individuals set in the data management platform, which is used by ACM as a reference to manage consents.

Privacy Policy – a set of documents data controllers use to specify their processing means and purposes in order to obtain individuals' consents for data processing.

Background

The digital age has arrived. With the wide applications of cybertechnologies and data technologies, human society now enters a period that communicates all information through the Internet (Bukht, R. and Heeks, R., 2017). We leave communication records on emails and social media accounts, which are stored in digital forms. Our behaviors leave digital traces through browsing web pages, listening to music on apps. We have our location and biometrics data tracked by our smart devices even when we are disconnected from the Internet. The data will constitute a digital ecosphere that reshapes the social structure, revolutionizing the social, economic, and political activities (Albrecht, J.P., 2016).

The massive amount of data has tremendous values. Humans cannot process a large amount of information and excavate patterns out of a huge volume of data. But more messages and insights are deeply embedded in the thousands of engine sensors, the millions of user clicks, and the streaming media's megabytes of data flow. Big Data applications become our ability to collect and process the information to generate insights and benefit parties, fostering the industry and the digital economy (Sagiroglu, S. and Sinanc, D., 2013).

Data is the central resource and the key factor of production to the data-driven economy (Aaronson, S.A., 2019). The size of the digital economy occupies up to 15.5% of the global GDP, with over 11.5 trillion U. S. dollars in values (United Nations, 2019). In particular, personal data is one important dimension of data that drives economic activities. However, the way to handle personal data raises privacy and security concerns. Users still give out their personal data and sacrifice privacy in exchange for free services provided by digital and tech companies (United Nations, 2019).

In addition, big data analytics generate values for businesses, transforming data into a part of the firm's competitive advantage (Caribou Digital, 2016; Evans and Gawer, 2016). Digital platforms occupy dominant importance in the digital economy, achieving

monopoly positions. Global digital companies such as Google can occupy up to 90 percent of the market share (Khan, L.M., 2018). The network effect and the platforms' huge computational resources contribute to such prosperity. On the one hand, tech monopolies exploit their market positions and engage in anti-competitive behaviors to hamper economic growth (Kenney, M. and Zysman, J., 2016). On the other hand, they integrate seamlessly into all aspects of people's lives and use data to gradually influence and even shape individual behaviors (Graef, I., 2018).

The increasing connectivity and digitization also have implications for national security and sovereignty (Crampton, J.W., 2015). The information environment of citizens is impacted by the data-intensive technologies from foreign parties, which in turn affect citizens' thoughts and opinions (Van Dijck, J., 2014). National secrets, key population statistics, individual security also become accessible with a giant amount of computational resources. There is a global trend towards "data localization", which urges data maintenance and processing to happen within a nation's geographic boundaries (Chander, A. and Lê, U.P., 2014). No international agreement on cross-border data transfers exists. However, data localization will adversely affect Internet users, as well as preventing global collaborations and innovations (Aaronson, S.A., 2019).

The digital economy raises many challenges in policies and legislations. There is a lack of global privacy protection standards. Different countries have different data protection rules, making the international businesses hard to operate and ineffectively protect individual privacy (Mitchell, A.D. and Hepburn, J., 2017).

The future of the digital economy demands a more comprehensive regulatory framework to protect human rights, create market order, and generate more values to benefit the public.

1. Introduction

With the growing importance of the digital economy and increasingly frequent use of data technologies, our lives were transformed, and new challenges were raised. Big Data collects information about individuals through various channels and devices and tracks people in all aspects of their lives, including web browsing activities, media information, financial information, location information, etc. (Allen, A.L., 2016). However, concerns about privacy violations, mass surveillance, manipulations, and algorithm prejudices also become the central agendas of public discussions.

In 2016, the notorious Facebook-Cambridge Analytica data scandal culminated the worries (Isaak, J. and Hanna, M.J., 2018). The app “thisisyourdigitallife” collected Facebook user data and developed psychological profiles, generating up to 87 million user profiles. The data was used in the political campaigns to promote political advertising, manipulate voters, and consequently influence the U. S. presidential election. Just like the name of the app, individuals are living their “digital lives” in this era. Our personal information and identity become digitized and accessed by organizations without safeguards. Novel policies and laws are required to guarantee protection and prevent human rights violations.

Data laws are the laws that regulate digital activities, studying privacy, personal information and non-personal information as the subjects, exploring the privacy and security problems incurred in the life cycle of data processing (Bygrave, L.A., 2002). Early in 1969, the UN’s International Conference on Human Rights first introduced the concept of data protection. In the year of 1970, the federal state of Hessen passed the first national data protection law in the world (Tikkinen-Piri, C., Rohunen, A. and Markkula, J., 2018). Following the establishment, more countries publish their data protection acts. Influential data laws include the General Data Protection Regulation (GDPR) enacted by the EU commission starting from 2018. It became the blueprint of the modern data protection framework and was widely discussed by legal professionals and scholars (Goddard, M., 2017). Many countries afterward published regulatory frameworks similar to the GDPR

(Buttarelli, G., 2016).

The Universal Declaration of Human Rights establishes privacy as one of the inalienable human rights to individuals (United Nations. General Assembly, 1949). Businesses and government organizations collect personally identifiable information from individuals to perform analytics, at the same time violating the right to privacy. Using data to conduct user profiling and perform data analytics can also influence individual behaviors and autonomy, raising complex challenges with our freedom of thought and expression (Richards, N.M. and King, J.H., 2014). We need data protection laws to protect individuals' right to privacy and freedom.

At the same time, a solid data protection framework will underpin the thriving digital economy with great data fluidity and accessibility (Zhang, M.L. and Chen, M.S., 2019). Data constitutes a novel form of the factor of production (Mueller, M. and Grindal, K., 2019). Its relevant capitalization system and property rights protection framework will be essential to achieve data commercialization and stabilize the market order. Data law is also important to balance privacy protection and data fluidity.

Scope and Definitions

Countries around the world have different data protection frameworks, but in general the scopes and principles have overlapped. The regulations cover the protection of the processing and free movement of personal data and protect fundamental rights and freedoms of natural persons.

“Personal data” refers to any information relating to an identified or identifiable natural person (European Parliament and Council of European Union., 2016). In different protection frameworks, it is also used interchangeably with “personal information”. In this paper, the two terms will all appear but should be treated as synonyms.

“Data subject” is the identifiable natural person that the personal data relates to. An

identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier, or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. In different protection frameworks, “individual” is also the synonym for the data subject. In this paper, “individual” will be used primarily.

“Processing” means any operation or set of operations which are performed on personal data or sets of personal data, whether or not by automated means, such as collection, recording, transfer, storage, etc. (European Parliament and Council of European Union., 2016).

“Data Controller” means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data. Data Controller might not directly engage in processing activities but appoint a “Data Processor” to processes personal data on behalf of the controller. Data controllers and processors usually have different responsibilities and roles. However, for simplification, we will assume that the data controllers are also the data processors. In this paper, the term “data controller” will cover entities who both determine the purposes and means of the processing and perform the processing activities.

Processing Principles

During data processing, there are several principles which the data controllers shall follow. Though varying in different countries protection acts, eight general principles are developed:

Principle of Consent: Data controllers and processors must acquire their consents before collecting and processing the data subjects’ personal data, aside from a few exceptional circumstances specified in data laws.

Principle of Transparency: Data controllers and processors must process data subjects' personal data transparently and legally, such as making explicit their privacy policies and means for collecting and processing data

Principle of Purpose Limitation: Data controllers and processors must process data subjects' personal data on a lawful, specific, clearly defined purpose, and must not process in a way that's incompatible with the original purpose.

Principle of Accuracy: Data controllers and processors are obliged to assure the accuracy of the data, keep it up-to-date, and implement legitimate measures to correct or remove inaccurate data in time.

Principle of Storage Limitation: Data controllers and processors can't store data subjects' personal data longer than what is necessary for achieving the original aim of data collection.

Principle of Integrity and Confidentiality: Data controllers and processors must process data subjects' personal data in a way that guarantees data security, including adopting measures to prevent accidents such as data breaches and illegal authorization of data usage.

Principle of Individual Participation: Data subjects are provided with various rights to control how their data is used. This includes the right to access their data upon request, the right to rectification and erasure of their data, and the right to object their data from being processed. Data controllers and processors must respect data subjects' rights and respond to their requests in time.

Principle of Accountability: Data controllers and processors must undertake the responsibilities and fulfill obligations listed by the principles above, make public the measures for data subjects to report their misconducts, and demonstrate compliances and compensations when they fail to do so.

Loopholes of Data Laws

Just like any other laws, the data laws also have loopholes and enforcement difficulties due to practical constraints. Some limitations cannot be fixed by simply amending the clauses but need to devise solutions from comprehensive perspectives.

This paper makes its contribution by introducing a novel data governance scheme based on behavioral economics theory to address the data law loopholes and enforcement challenges. The scheme consists of various components and features that will interact with each other to enhance privacy and data protection, and at the same time improving data fluidity and accessibility.

The paper will be structured as follows: Chapter 2 will explore the existing problems in data laws raised by the failures to fulfill processing principles. Chapters 3 and 4 will reconstruct the problems and introduce the behavioral economics theoretical framework for implementing libertarian paternalism interventions, or *Nudge*. Chapter 5 will introduce the governance scheme and describe the components in great detail. Chapter 6 will analyze the interactions between the components by exploring several scenarios related to the data processing to help readers better understand the system's working. While Chapter 7 concludes the study's findings, Chapter 8 calls on scholars to contribute to the scheme with more features; it attempts to inspire readers by discussing two potential systems to be included in the future.

2. The Unsolvable Yet Unavoidable Problems in Data Laws

Failure to Fulfill Data Processing Principles

In theory, obliging the processing to the principles set forth will guarantee the legitimacy and protect the rights of the individuals. Yet in reality, some of the principles have little restrictive power. Legal obligations become only ostensible and lead to urgent loopholes that the laws cannot address.

Principle of Consent

Served as one of the most widely used legal basis for data processing, the principle of consent is central to protect the rights and property of individuals. The right to consent to process their data is the natural extension of individuals' right to informational self-determination (Malgieri, G. and Custers, B., 2018). The principle of consent asserts that once individuals are informed and agreed to the processing, their consents represent the true intentions to allow their data to be processed, so the processing respects the individuals' will and ensures the protection of their autonomy and dignity.

However, in practice this principle does not always effectively protect individuals' rights. Individuals' consents do not necessarily represent their true intention. There are two main sources that weaken the consent's representativeness of individuals' intentions (Schermer., Bart W., Bart C., and Simone, H., 2014). First, the information asymmetry between the individual and the data controller can lead to the individual being less informed of the privacy policies, which are the common forms that data controllers use to acquire individuals' consents (Koops, B.J., 2014). The privacy policies can be too long in length and too many in numbers to overwhelm the individuals, causing "consent desensitization" that refers to people blindly giving out consents (Schermer., Bart W., Bart C., and Simone, H., 2014). In addition, the individuals might fail to sufficiently understand the privacy policies due to the complexity and legalese they involve (Solove, D. J., 2013). The second source is the power imbalance between the individual and the

data controller. It states that the individuals do not have the adequate power to authorize the processing of their data, unlike how it is assumed by the principle of consent. Data controllers' privacy policies are often presented in the form of standardized contracts, so individuals have little to no opportunities to negotiate (Muravyeva, E., Janssen, J., Dirkx, K. and Specht, M., 2018). They either agree to the terms or completely dismiss the processing and the associated services. Further, in cases where the services feature a few or no alternatives, the individuals generally have no room to reject handling out their data (Schermer., Bart W., Bart C., and Simone, H., 2014).

Also, the data controllers themselves are incentivized to weaken the principle of consent's power. The cost for the controllers to acquire a "freely given, specific, informed and unambiguous" consent is often high. In some laws it requires a rigorous examination to demonstrate that the consent represents individuals' free will (Bendiek, A. and Römer, M., 2019). In alternative cases, the law permits other legal grounds for the processing, such as "legitimate interests" as specified by GDPR (European Parliament and Council of European Union., 2016). The data controllers are prone to choose alternative legal grounds and avoid the high cost of acquiring consents. For laws that do not involve a heavy burden of proof or impose a strict consent condition, the data controllers can develop measures to reduce the cost of acquiring consent, such as adopting opt-out designs that process individuals' data by default (Acquisti, A., Adjerid, I., Balebako, R., et al., 2017). Under the opt-out design, only if individuals choose to withdraw their consents, they otherwise agree to the processing by default. Reducing the cost also enables data controllers to incorporate terms in their privacy policies that are essentially against the individuals' will. But due to the information asymmetry and power imbalances, individuals often fail to act according to their real preferences and choose to agree to these terms.

The fall of the principle of consent leads to two significant negative consequences. First, individuals might agree to process sensitive information they do not intend to or fail to

recognize the dangers of the processing when they easily hand out their sensitive information (Willoughby, A., 2017). The disclosure of sensitive information, such as ethnicity, data related to biological characteristics, and financial status, might result in significant disadvantage and discrimination towards the individuals. Second, individuals might unintentionally allow data controllers to share their data with untrusted third parties, whose purpose and means of processing remain lucid and potentially hazardous. This often results in individuals receiving communications that include their sensitive personal details from strange data controllers and put their assets, health, and security at risk (Batista, M., Fernandes, A., Ribeiro, L.P., et. al., 2020).

Principle of Individual Participation

The principle of individual participation grants individuals' rights to their processed data by making requests to data controllers. Specifically, the rights commonly include the right to access their copied information, the right to erasure, the right to correct, and the right to withdraw consent, etc (Acquisti, A., Adjerid, I., Balebako, R., et al., 2017).

Individual participation serves as an integral part of ensuring the order of the digital market, aside from laws and regulations imposed by nations (McDermott, Y., 2017).

Through exercising their rights and reporting the misuse of personal data by individuals, authorities can effectively discover and regulate the illegal data controllers, thereby protecting more individuals and establishing more trust in the economy.

However, on both the individual and the controller sides, barriers to fulfilling this principle exist. Individuals are generally subjected to information asymmetry and power imbalance between them and the controllers. First, they can even be unaware of their rights to exercise, do not know how to exercise the rights, or just feel reluctant to go through the complex process to exercise their rights. Individuals usually need to email their request to the data protection team the controllers appoint to and negotiate (Mittal, S., 2017). They also might not know what and by whom their data is processed. When individuals are subjected to automated processing such as individualized

recommendation, unless specially noted, some might recognize the individualized recommendation as generic recommendations and become unaware of themselves under automated processing (Wachter, S., Mittelstadt, B. and Russell, C., 2017). Even if they exercise their rights and get feedback from the controllers, they sometimes cannot interpret what the feedback entails, especially in the case that involves automated decision-making that is based on particular algorithms. Lastly, if data controllers unfairly reject individuals' requests, individuals have no effective ways to enforce their rights (Altorbaq, A., Blix, F. and Sörman, S., 2017). In fact, data controllers can use a reversible algorithm to process all the stored data by combining different individuals' data and erase all the original data. Individuals cannot exercise their rights over the derived data as data controllers can claim that they now contain others' data which individuals have no rights over, fundamentally disabling them to exercise the rights (Medium, 2017).

For data controllers, there exists incapacities and risks to respond to individuals' requests. Some small data controllers are technically infeasible to identify the requesting individual from their database and perform relevant actions (Guanglan, T., 2020). In the case where the requesting individual does not provide sufficient contact information, the data controller as well cannot know to whom to respond. Even if it is technically feasible for the controller to respond to the request, there are still risks involved (Bozdag, E., 2018). A cybercriminal can request others' information in their names and violate others' property or commit other crimes using their personal details. A terrorist can also exercise her right to erasure to delete all her online traces, making it hard for security officials to track her (Ausloos, J., 2018). This poses a critical national security risk as it makes officials harder to discover the signs of terrorist activities and enact preventive measures. At last, there is a cost for the data controller to organize a special data protection team and communicate with the individuals to deal with their requests. The risk of potentially disclosing their trade secrets also makes them reluctant to respond to the requests (Banterle, F., 2018).

Principle of Purpose Limitation and Data Minimization

Principle of purpose limitation and data minimization make sure that data controllers process individuals' data no more to the agreed purpose and only process the necessary data. However, since data processing is a technical process set up in the backend, it often remains opaque to the individuals (Koning, M.E., 2020). There is unlikely a practical way for them to verify if the two principles are fulfilled.

This leaves three major negative consequences. First, data controllers can include unnecessary services to collect individuals' data (Forgó, N., Hännold, S. and Schütze, B., 2017). In some cases, the data might be sensitive information whose disclosure will lead to significant risks. Second, Individuals have no way to negotiate about the unwanted services from data controllers (Zeyu, L., 2018). For instance, the data controller might provide a standardized contract that collects an individual's data for providing some services. But if the individual is uninterested in one of the services provided, she does not have a way to opt-out of that particular service and rejects the corresponding processing. She can either choose to reject all the processing and receive no service at all, or to accept all the processing and receive the services. Third, individuals usually cannot determine if the data controller is over-collecting their data to provide services, and they are subjected to power imbalances and unable to stop overcollection (Mantelero, A., 2017).

Principle of Accuracy

The principle of accuracy assures the data being processed accurately reflects the status of the individual so that the processing does not disadvantage her by the inaccuracy. Inaccurate data might cause unfair treatment that is based on data analytics to impact individual's access to public services and receive unequal legal judgment (Clifford, D. and Ausloos, J., 2018). This principle allows data controllers and individuals to update the processed data, but in reality the updates might not be timely or accurate.

Individuals can have little incentive to update their data in cases they do not see the direct benefits (Acquisti, A., Taylor, C. and Wagman, L., 2016). They might feel reluctant to perform the consuming update process or do not know how to. Data controllers have the cost of updating individuals' data. In cases where data accuracy does not affect the quality of their services, they might tend not to do so.

Principle of Transparency

The principle of transparency enhances the informedness of individuals over the data processing so they can exercise their rights to stop the processing against their will. However, if there is a data black market that trades individual data without their consents, individuals do not have an effective measure to discover the illegal transaction, unless when they receive unsolicited communications from the controllers whom they never authorize to and when the illegal controllers declare that they have individuals' data (Wachter, S., Mittelstadt, B. and Russell, C., 2017). Furthermore, even if individuals discover that their data are being processed by unauthorized parties, they do not know the source of provision and can't effectively terminate the illegal transaction.

In terms of the legal data transfer between different controllers, an opaque data chain might still exist. For instance, if a data controller transfers data to another controller and accordingly informs individuals of the recipient controller and the details of the processing (e. g. the purpose, means of the processing, etc), the recipient controller is obliged to process according to the specified processing conditions and must ask for individuals' consent once they modify the conditions. But the recipient controller might experience troubles identifying the individuals or lacking contact details to inform (Politou, E., Alepis, E. and Patsakis, C., 2018). No effective measures that enforce their communication exist as well. This implies that the recipient controller can easily modify how individuals' data is processed without spending a high cost asking for individuals' consent.

Principle of Integrity and Confidentiality

The principle of Integrity and Confidentiality obliges the processing to be under appropriate security measures to protect against unlawful and unauthorized processing. Security measures can take in the form of information encryption. One encryption measure often adopted is anonymization, which modifies the original data in a way that makes identifying a particular individual incapable. Once this technique is applied, controllers can process the anonymized data for other purposes without asking for individuals' consent, since the data, in theory, can no more link to them.

But in practice, as more data accumulates, recombining a large volume of anonymized data makes it possible to re-identify the individuals (Brasher, E.A., 2018). This property invalidates the effectiveness of anonymization to protect the individuals. Unauthorized parties can access the data by reversing the anonymization process. Recipient data controllers and processors in data transfer can also reverse the anonymized data without the original controller being aware of it (Brasher, E.A., 2018).

In more general terms, ineffective encryptions can also pose a national security risk. For instance, an athlete social network app called Strava can record users' fitness activities, but ineffective encryption makes it possible for analysts to conclude user behavioral patterns and identify military or intelligence agency personnel. In turn, the location of U.S. secret military bases might be unveiled (Wired, 2018).

As applications involving biometric data, data that relates to human characteristics, become more prevalent, the ineffectiveness of encryptions also puts individuals at risk. Biometric data as a category of sensitive information shall be stored in anonymized forms and put under appropriate security measures. However, unauthorized third parties can access the data or unauthorized data transfer can occur by reversing the anonymization process and identify individuals. Common risks of biometrics data

disclosure range from bodily harm, stigmatization, political manipulation to loss of liberty and security (Sepúlveda Carmona, M.M., 2018).

Principle of Storage Limitation

The principle of storage limitation specifies that the stored data must be erased after the specified storage period passes. However, individuals do not have ways to verify if the data is effectively erased, or it can involve technical difficulty for controllers to completely erase the data (Bozdag, E., 2018). The problem of ineffective erasure is not limited to the principle of storage limitation. When individuals request their data to be erased by exercising their rights, controllers can claim that their data is erased as individuals cannot verify it. Likewise, when a controller transfers data to another controller and obliges her to erase the data after the specified processing time, the recipient controller can also claim that the data is erased while not actually doing so. This implies that the controller can process the data according to their will as others suppose it is erased already (Shah, A., Banakar, V., Shastri, S. et al., 2019).

Data Monopoly

Due to the commercial value of data, it becomes an increasingly important competitive advantage for firms. Controllers of data can use the data to develop big data products and generate insights, in turn developing more competitive products to obtain revenues. Major data controllers in a market can build up the cost to entry and establish a data monopoly by refusing to share their data with new firms (United Nations, 2019). The monopoly position is further strengthened by users' tendency to use the same platform or related services, a "winner-takes-all" network effect (Stallkamp, M. and Schotter, A.P., 2019). This will discourage data fluidity and hamper innovations and research in the market, as well as unfair pricing, degrade product quality, etc (Carrier, M.A., 2020). But as data controllers involve costs to collect and process the data, they have property rights over the processed data. We don't yet have definite standards to determine the scope of data sharing.

3. Re-Construct Problems

When introducing the unsolvable yet unavoidable problems by laws, we can spot that common patterns exist. Further analyzing the problems, we find that practical limitations and decision-making biases constitute the essence of the unsolvable problems.

Patterns in Problems

Fundamentally, all aforementioned problems are caused by four practical limitations: information asymmetry, bounded rationality, power imbalances, and technical incapacity.

Information Asymmetry

Information asymmetry generally refers to the differential access to information of economic agents in a relevant transaction (Acquisti, A., Adjerid, I., Balebako, R., et al., 2017). In the context of data protection, it is the incapacity for both the individuals and controllers to follow the exact flow of data and understand the data processing system from a holistic perspective. Failures to achieve many principles can be partially explained by information asymmetry. For instance, individuals lack information on what rights over their data they can exercise and to whom they should request, which leads to the fall of the principle of individual participation. Individuals are also blind to how controllers process their data and what other controllers participate in the opaque data chain, which leads to the fall of the principle of purpose limitation, data minimization, and transparency. Information asymmetry also affects data controllers. For instance, they might be incapable or hesitant to respond to an individual's request to exercise her rights, due to insufficient contact information of the individual or a lack of authentication to verify her identity (Soh, S.Y., 2019). Controllers might also be unable to inform individuals of data transfer or other new processing actions performed on their data. These problems contribute to the fall of the principle of consent, transparency, individual participation, etc.

Bounded Rationality

Bounded rationality states that individuals have limited mental capabilities to make decisions, which often fail to reach an optimal outcome. Traditional economics views individuals as rational agents who always maximize their utilities regardless of the complexity that their decisions involved (Acquisti, A., Adjerid, I., Balebako, R., et al., 2017). However, in reality, individuals commonly have limited computational resources to understand all information pertaining to a decision and fail to make logical deductions based on relevant factors. In the context of data protection, bounded rationality is most exemplified by the consent desensitization phenomenon with individuals either fail to sufficiently understand privacy policies or feel unwilling to spend abundant time reading (Solove, D. J., 2013). It also explains the fact that individuals might be reluctant to exercise their rights over data through a complex request process. This weakness is further exploited by data controllers through inducing individuals to hand out other sensitive data and consents to third parties to process their data. Bounded rationality is deeply intertwined with information asymmetry. Individuals have inadequate information and are not fully capable of processing information related to their data. They altogether lead to the fall of principles of consent, individual participation, accuracy, transparency, etc.

Power Imbalance

Power imbalance characterizes individuals' inability to enforce controllers when their processing actions deviate from individuals' will and describes the incapacity to reject using their services. Individuals' voices are too insignificant. Even when they exercise their rights and complain about the abuse, data controllers can still ignore them without being effectively enforced (Biancotti, C. and Ciocca, P., 2019). No judge ensures the exercise of rights. For tech giants and platform providers, their services have dominant positions in most individuals' lives. In addition, individuals' acquaintances will still use these services. Rejecting services not only brings inconvenience, but also detachment from individuals' social network (Sagiroglu, S. and Sinanc, D., 2013). The increasingly

intimate connections created by the Internet are socially pressuring individuals. This power imbalance leads to the fall of principles of consent, individual participation, and the problem of data monopoly, etc.

Technical Incapacity

Technical incapacity states that there are practical technical difficulties laws fail to consider, which prevent obligations from achieving effective data protection. For instance, some data controllers are technically infeasible to identify individuals who wish to exercise their rights, thus being incapable of responding to their requests (Bozdag, E., 2018). This incapacity is also exemplified by the problem of ineffective erasure and ineffective anonymization. Individuals are technically incapable of verifying if their data is completely erased. Data controllers also cannot guarantee that the encryption they applied on data won't be deciphered. These problems undermine the principle of integrity and confidentiality, as well as the principle of storage limitation and data minimization, posing negative consequences including national security and disclosure of sensitive data.

4. Behavioral Economics Solutions

Innovative solutions from a libertarian paternalism perspective can be devised to address the law enforcement issues created by practical limitations.

Nudge, or libertarian paternalistic intervention, refers to interventions that attempt to influence the decision-making of individuals to improve well-being (Thaler, R. H., and Cass R. S., 2009). In behavioral economics theory, nudge interventions typically are made by reframing the choices available to individuals when they are in a decision-making context. In the broad sense, a nudge takes account of behavioral and decision-making biases of individuals to improve social welfare (Hansen, P.G., 2016). Unlike paternalistic interventions that force individuals to make decisions believed to be beneficial to them,

libertarian paternalistic interventions only try to increase the likelihood of certain decisions without actually limiting the freedom of choice.

In the behavioral economics theory, the nudge can recognize human limitations and even utilize them to increase social welfare while respecting human autonomy in decision-making (Thaler, R. H., and Cass R. S., 2009). Traditional economics views assert that humans as rational agents could always maximize their utility by their own decisions. Therefore, strictly libertarian approaches are often recommended, which are assumed to be completely nonintrusive of individuals' decisions. Yet in practice, when humans make decisions, they are subjected to various practical constraints and their own mental limitations, such as information asymmetry and bounded rationality. By libertarian paternalistic interventions, these actual limitations will be taken into accounts and sometimes converted as a power to help individuals make better decisions.

A nudge intervention scheme will be effective to address the enforcement problems of legal obligations in the field of data protection. We have analyzed that the unsolvable problems by law are actually created by the practical limitations of individuals and data controllers. Under the nudge intervention scheme, by devising solutions that are targeted to the behavioral and practical limitations, individuals can make better decisions to maximize utility. For instance, we can first identify that information asymmetry sometimes prevents individuals from exercising their rights and becoming benefited. Therefore, we can make up for the information asymmetry by properly educating the individuals and assist them to exercise their rights and bring more welfare. Also, libertarian paternalistic interventions do not serve to regulate people with force but encourage people to make their decisions in a more beneficial way. It might be a more acceptable solution for both the individuals and the data controllers than directly imposing restrictions on their activities.

A New Form of Data Governance based on Nudge

A new form of data governance based on libertarian paternalism should be established to solve the intertwined problems in the field of data protection. First, seven principles for designing the framework can be determined. The components and features of the framework are then designed based on the principles.

Principles of the Framework Design

1) Human-centric

The ultimate aim of this governance framework and the use of big data must focus on benefiting humans. Since a person is linked to a set of data, this principle asserts that individuals' control over data should be holistic. All data centering around an individual are relevant and should not be overlooked and discretized. The solutions take account of the practical imperfections of humans, but also respect and embrace them as an integral part of humanity.

2) Privacy-preserving and Democratic

This framework aims to enhance the protection of individual privacy and autonomy. As an inalienable right to individuals, the right to privacy contributes to our human autonomy, property protection, and happiness. The use of technology should not violate our fundamental human rights. Though a form of governance, it is never meant to be paternalistic, but should be libertarian paternalistic that respects individuals' voices and helps them maximize their welfare. The data controllers and individuals are never forced to join this scheme and they can choose to opt-out at any time. The function of the governance scheme is always to *encourage* individuals to make more privacy-preserving decisions and *encourage* data controllers to operate in ways that better benefit the economy and themselves.

3) Open

This framework aims to help data laws promote data fluidity and healthy competition in the market, as well as leading to more innovations and social progress. An open and fluid

environment should be established to tackle problems of data inequality and monopoly, giving entering businesses more opportunities and improving social productivity. The new governance should also be interoperable across multiple platforms, achieving a higher degree of fluidity to benefit people.

4) Automated

An automated form of data governance will address many difficult problems that undermine legal enforcement due to the limited mental resources of humans. Individuals have more effective management of their data and more actively exercise their rights. Empowering data governance by automation will also make the flow of data more traceable and efficient, leaving more room for systematic improvement from a technical perspective.

5) Informative

The new form of data governance should be informative to educate individuals with adequate information about their data. It is also reader-friendly and helps individuals process and understand the technical and overwhelming information related to their data. An informative data governance scheme can also inform individuals when they make less desirable decisions and prevent or mitigate negative consequences. This will ensure they make decisions that are closer to their true preferences and maximize social welfare.

6) Trustworthy

This form of data governance should be led by neutral global institutions and not controlled by governments from any particular country, in case of inflicting national security issues. All parties joining this form of governance abide by the same code of ethics and conform to the mutual trust framework. The governing body should make the judgment on a fair basis, not particularly favoring individuals or data controllers, but trying to make up the power imbalances of different parties.

7) Accountability

This form of governance should be transparent and traceable, ensuring the responsibility of different parties to prevent or mitigate risks. An accountable structure strengthens the authoritative power of the governance and enhances it to fully exert its function.

5. Components of the New Data Governance

In this new form of data governance, there are basic components that realize all the functions and enhance data protection as well as achieving a greater data fluidity. The governance scheme and elements it incorporates are established based on the libertarian paternalism principles. Enrollment into the data governance is not mandatory, but on a free basis. It will not impose rigorous regulations on individuals but always encourage them to make more privacy-preserving decisions. The data controllers are also not subject to paternalistic interventions. They can always choose to not follow the policies, but this governance will create a trust and transparent environment in which the controllers' decisions will be open to individuals. Those who decide to not enter this more privacy-preserving environment will be doubted by individuals on their trustworthiness.

Data Management Platform

All the components in the governance scheme will operate on a global, electronic-based data management platform to ensure interoperability and foster transparency. This platform should be run by a neutral, international organization to define standards, code the framework, and structure the components in machine-readable forms. The automated consent management, bureaus, and the Central Technical Council will affiliate to the same international organization that operates this platform. Evaluative frameworks such as privacy rating and formats of reporting will also be globally defined by the platform. The platform and the international organization should have personnel with a balanced representation of backgrounds and possess no complex political motivation.

Individuals and data controllers can choose to join this data management platform. Individuals will need to fill in their personal information and be aware that this platform will help manage their personal information from now on, and the ACMs are their respective agents. Data controllers will register their business information and other relevant details.

The platform will be reader-friendly to individuals that are not tech-savvy. Plain languages and terms to explain processing activities and information technologies will be provided to better educate people and help them make decisions more reflective of their true wills.

Participating data controllers will conduct all of their requests, transfer, data trading through the framework and interfaces of the management platform. Every party is placed under the same trust framework and utilizes an interoperable set of tools, APIs, packages to perform functions and communicate. The data controllers should follow all kinds of regulations issued by the bureaus and be responsive to the policy updates.

The enrollment into this platform is supposed to be opt-out. Individuals and data controllers should join this platform by default but can choose to quit whenever they want on a free basis. However, we should aim for a global standard and promote our efforts. This is a holistic undertaking that requires all governments and international organizations, business unions to push forward.

Individual, Data Controller, Privacy Ratings, Data Traders

Under the new scheme, individuals are also viewed as entities who have a set of personal data linking to them. However, individuals will not respond to any single request from controllers to processing their data or directly negotiate with them to exercise their rights. Individuals will hand out consents and exert their rights over the data in a systematic way under the interventions of Automated Consent Management (ACM) and

bureaus. Data controllers are also rated by privacy ratings that are determined according to their processing activities. Privacy ratings below a certain level will inform individuals of the potential harms associated with the processing activities.

Data traders are market entities that distribute, source, exchange, and commercialize data as a product to data controllers to promote data fluidity and data accessibility. They serve to help generate a greater value out of the existing data and break data isolations. Data traders have all kinds of data available to data controllers and relevant pricing mechanisms. Data controllers can purchase the data they need to expand their dataset and create better products for economic gains.

However, unprotected data transfer will incur significant risks about privacy violations. Individuals' information belongs to themselves and should not be handed to unauthorized data controllers. Bureaus will closely inspect the data traders' framework to protect data and ensure transparency and accountability during transfer.

Personal Data, Meta-data, Privacy Configuration, Reports

An individual has her set of personal data as well as the meta-data and privacy configuration attached to it. The meta-data is the data that describes the properties of the data, such as the time and method for generation, data type, etc (Maguire, S., Friedberg, J., Nguyen, M.H.C., et al., 2015). Meta-data are essentially the ID card of every piece of data. Each time, a new kind of processing activity is done, the meta-data should be overwritten. They are the information that helps bureaus, ACMs, and individuals to keep track of the information flow and ensure accountability and transparency. Under this new scheme, the meta-data might not be simply XML languages, but also involving more advanced mechanisms that help realize functions such as auto-deletion. Privacy configurations are the user preferences that individuals systematically authorize before processing activities. Individuals will make explicit the privacy preferences in the Automated Consent Management, such as the category of data that is processable to the

particular category of data controllers. Most consents are handed out by ACM in an efficient and tractable way. Reports can vary in categories and perform different functions. For example, a report can be a processing request, an update reminder, or a notice about inconsistency detection. In general, reports are the information flow between individuals, ACM, bureaus, and data controllers to send messages and updates.

Automated Consent Management (ACM)

ACM is an automated agent that helps individuals manage their personal data and realize most functions of the data governance in conjunction with bureaus. The roles of ACM include managing individuals' privacy preferences, handing out automated consents to controllers according to individuals' privacy configuration, sending warnings, notices, inconsistency alerts, reports to individuals to help keep track of the data usage, active detection of inconsistency to discover unauthorized data controllers, providing evidence and authentication to exercise individuals' rights.

ACM is the key component in realizing this framework conducted on behavioral economics principles. ACM will speed up the process of handling consents by automated functions. However, it achieves an efficiency boost on the basis of individual autonomy and privacy protection. The ACM will manage consents to abide by the conditions in individuals' privacy configurations. Individuals can adjust the privacy configurations at any time according to their free will, and they sometimes can choose to turn off ACM's functions and even opt-out of the scheme. No paternalistic regulations are imposed. The ACM will only *encourage* individuals to make more privacy-preserving and beneficial decisions.

Bureaus and Signatures

Bureaus are the authorities in the data governance that makes up the power imbalance between individuals and data controllers, balancing various groups of interests, making necessary decisions pertaining to crucial issues such as national security and sensitive

information disclosure, empowering the whole system, and concretizing enforcement powers. There are different kinds of bureaus in different functional scenarios, such as the “data police bureau” for sensitive information protection, “data judge” in exercising individuals’ data rights. They will carry out roles including authenticating data subjects’ identities, checking the consistencies of requests, and interacting with the ACM to communicate.

In reality, the personnel that runs the data bureaus should come from neutral international organizations such as the UN. The data bureaus should not affiliate to a particular organization or possess other motivations, as they have chances of accessing all individuals’ data under the governance.

Signatures are the stamps issued by Bureaus to express their approval over processing actions. Signatures can vary in categories. For instance, for the Data Police Bureau that monitors the sensitive data processing, a signature will be issued to the data controller’s legitimate request, after checking with the ACM. Signatures are specific to processing actions, instead of the entities that perform the processing. Checking signature is a way for the data management platform to verify any processing’s legitimacy and protect individuals’ interests.

Central Technical Council

Due to the fundamental loopholes in anonymization and other encryptions methods, there are always methods to decipher. Therefore, all the encryption-related processing should be run or monitored by individuals from a central, international, trustworthy technical group. The Central Technical Council will train Central Technical Experts (CTE) and ensure that they follow the same set of confidentiality principles. The technical experts must be trusted and always act according to ethical and legal principles.

Each CTE might be dispatched to data controller entities. They are the people responsible for encryption and decryption over data. Professionals who never receive training and approvals from the Central Technical Council are not allowed to engage in integral encryption-related activities. This makes sure that all personnel performing encryptions are trusted or under the supervision of CTEs. The CTEs will also supervise data controllers' other processing actions and make sure they do not exploit their huge computational resources to conduct illegal decryption. Current legislation commonly asks data controllers to appoint data protection officers and designate relevant data security department. The CTE will also coordinate with the department to provide data protection guidance, accessing the controller's processing activities to ensure legitimacy.

6. Potential Scenarios Under the Novel Governance Scheme

Holistic Landscape

Under the governance scheme, different components will interact with each other to perform various functions to enhance data governance. A holistic overview of the potential scenarios in the governance scheme is presented as follows:

Diagram 1 Holistic Scheme

Six scenarios are covered under the governance scheme: Ordinary Data Request, Sensitive Data Request, Inconsistency Check, Data Rights Exercise, Monitored Data Transfer, Notice and Broadcast. Five bureaus will participate in the scenario: Data Police Bureau, Data Judge, Data Exchange Bureau, Central Technical Council, and Publicity Bureau.

Ordinary Data Request

Personal data can be categorized as sensitive personal information and non-sensitive information, which we refer to as “ordinary data” in this case. Processing ordinary data usually does not incur substantial harm to individuals. On the basis of privacy protection and individual autonomy, we should simplify processing ordinary data to increase data fluidity.

Diagram 2 Ordinary Data Request

For the ordinary data request scenario, in step 1, the data controller will make a request to access one’s personal data. The request is directly sent to the ACM in a machine-readable and interoperable form. This request will detail the scope, context, and purpose of processing, the type of information requested, as well as the data controller’s basic business information. The ACM will also ask for the data controller’s privacy rating. In step 2, the ACM will evaluate the request, privacy rating by comparing it with the individual’s pre-configured privacy preference and specific meta-data. For instance, the individual can pre-configure “no permission for data controllers who have three or fewer stars to process financial information”. The ACM can compare if the data controller has met the condition and return a decision in step 3. If the processing is permitted, the personal data and the binding meta-data will be provided by the ACM. The meta-data will specify the time, context, and purpose of the information processing and prohibit

processing means other than the specified ones. The processing record will be stored in the ACM.

Through this scheme, individuals do not need to spend extra time reading the long and incomprehensible privacy policies from data controllers. The ACM will be the housekeeper that manages all human-centric data based on an individual's privacy configuration and preferences. This saves more time and energy for individuals and makes personal information more accessible to data controllers.

An important problem lies in how individuals should pre-configure privacy preferences. Individuals are not experts who know the exact scopes of each information category if defined improperly. For example, individuals might not fully understand what “physical information” entails. They also can be imprecise about the standards to judge the trustworthiness of data controllers. Individuals can underestimate or overestimate the risk of information processing, inducing privacy violation or hampering data fluidity on the other end. Even worse, individuals are subjected to status-quo bias, which in behavioral economics terms refers to the tendency to remain in the status-quo (Thaler, R. H., and Cass R. S., 2009). In our context, once individuals make an undesirable privacy configuration, they are less likely to adjust it, making it less likely to recover the loss. Though we have privacy ratings to enhance decision-making, more help should be provided.

More explanatory information is recommended when individuals pre-configure the preferences. Additionally, due to the status-quo bias, we can provide default privacy configuration or recommendation based on the big data and individual surveys. Before configuring the preferences, the data management platform will conduct surveys to ask individuals questions like how sensitive they care about their personal data, what categories of information they feel most comfortable sharing with, the scope of services they want. The nature of the surveys is like the Risk Tolerance Questionnaire used by

banks to help individuals understand the nature of investment style and risks, objectives of their investment (Yook, K.C. and Everett, R., 2003). Based on the survey results, a privacy configuration recommendation will be presented to the user.

This default privacy configuration can be automatically and dynamically updated by the results of the big data analysis. The data management platform will collect individuals' responses to the default configuration, browsing behaviors, and user trends of configurations. The privacy rating framework will also be timely adjusted to better reflect the trustworthiness of data controllers. From comprehensive evaluations, different privacy configurations will be recommended to individuals over time. Individuals will review the recommendation and decide to adopt the changes or not.

Sensitive Data Request

Processing sensitive information requires a stricter level of intervention to protect individual privacy and security. A Data Police Bureau will participate in the process to help monitor and track the requesting data controller. The data controller will not directly request data from an individual's ACM, but indirectly through the mediation of the Data Police Bureau.

Diagram 3 Sensitive Data Request

Step 1, the data controller will initiate a sensitive information processing request to the police bureau, specifying the nature and scope of the processing. The police bureau will keep a record of the request and check if it is legitimate and necessary, following the principle of data minimization. The data controller's background history will also be examined to ensure its trustworthiness. If it is approved, we move on to step 2. The bureau will contact the ACM to request individual information. In step 3, the ACM will evaluate the privacy configurations and return the personal data and meta-data if the processing is permitted. Individuals can also in their privacy configuration choose to receive extra special reports. Each time a sensitive data request is initiated, the individual

can ask the ACM to send a notification alert. This individual could double-check if the sensitive information requested is indeed allowable to be processed, and also inspect the background reports on the data controller to determine its trustworthiness. The ACM should also document and store records of sensitive information processing for the future. In step 4, the Data Police Bureau will return the sensitive information to the data controller along with the Bureau's approval signature and binding meta-data that prohibits processing means other than the agreed ones. This approval signature will serve as the controller's legitimacy certificate. If the signature is found to be missing, the data police bureau will enact measures to punish and restrict controllers from further processing sensitive information to mitigate risks.

Through this scheme, the individual avoids absentmindedly giving out sensitive data and loses track of the information flow. Data controllers who unnecessarily process sensitive information will be spotted and regulated timely. Accessing sensitive information will be a semi-automatic process that involves minor human participation to verify its trustworthiness. Overall, it makes data controllers easier to access data while imposing more rigorous protection standards.

Inconsistency Checks

Though we have bureaus and ACM to assure the processing is according to our privacy configuration, there always can be outliers. Technical incapacity in many ways might undermine the scheme's implementation. For instance, we have meta-data to specify the data's processing conditions, such as to which data controller this data becomes accessible, for what categories of information are processable, etc. But if the data controller decides to ignore the conditions, the meta-data does not have actual restrictive power.

One way to solve this is through a trusted processing container (Maguire, S., Friedberg, J., Nguyen, M.H.C., et al., 2015). All data controllers that enroll in the data management platform should use the same set of processing containers. This container can fulfill processing functions and impose special obligations from the management platform. Ideally, when the controller uses the container to process information, the container can compare if the processed data's meta-data is consistent with the processing means. If it is not, the container should have mechanisms that make the processing incompatible. For instance, when calling the data management platform's library to process data, the processing functions should have codes to compare the loaded data's meta-data with the processing parameters and determine if they are consistent. After processing, this container will also overwrite meta-data to document the processing history. Additionally, the container should be able to perform functions like auto-deletion when the storage time reaches its limit.

Other than the processing container, we have alternative routes to detect inconsistencies. The ACM should proactively reach out to check for inconsistencies. For instance, the ACM can perform web crawling on the data management platform to search for disagreement with the agreed processing conditions (Jensen, C., Sarkar, C., and Potts, C., 2007). Suppose a data controller agrees to delete an individual's stored personal information. The ACM will document that the individual's information should not exist

in this data controller's database. However, if, during web crawling, the ACM ever finds that the data reappears, an inconsistency is detected. It can also check if the data controllers have legal signatures issued by bureaus. If the certification is missing, the ACM can report to relevant bureaus and exercise the individual's right to withdraw processing consent.

Diagram 4 Inconsistency Check

In step 1, the ACM performs web crawling on the data management platform. In step 2, it detects that a data controller's processing means are inconsistent with the records stored in ACM. In step 3, the ACM sends a report to the individual, reminding her to exercise the right to withdraw consent or take other actions.

In addition, the data management platform can regulate the data products to detect inconsistencies. Any data chain consists of its input (personal information from the individual), process (the processing means of the data controller), and output (the derived data obtained by the data controller). We cannot easily detect inconsistencies from the input and the process stage. However, the data controller might send emails to individuals for personal recommendations based on the derived data. The data management platform can ask data controllers to include a line in every email, displaying notices like "I have not authorized this data controller to process my data, why do I keep receiving emails from them" with a hyperlink attached. It will make individuals reflect if

these data controllers have violated their privacy. Using the hyperlink, the individual can easily exercise their rights and protect data. The notice must be in bold and visually significant to readers. By the libertarian paternalism principles, data controllers are free to choose to not include the notice in their products. But individuals will be informed that such data controllers are likely to engage in unauthorized activities. The data controllers will face difficulties developing trust with individuals.

Data Rights Exercise

After detecting an inconsistency that violates the agreed privacy policies, individuals can file a complaint to exercise their data rights. Under the novel governance scheme, individuals will perform this process with the aid of ACMs and the Data Judge.

Diagram 5 Data Rights Exercise

In step 1, the ACM detects an inconsistency and sends a report to the individual to remind her to exercise the right. If the individual decides to complain, in step 2 the ACM

will make a request to the data judge, providing the reasons and evidence for inconsistencies to exercise the rights. The ACM will also present authentication to prove the individual's identity. In step 3, the judge will ask the data controller to provide its processing details and relevant signature to prove legitimacy. The judge will evaluate and compare the records and return a judgement to the ACM in step 4. If the right is exercised, the data judge will supervise the enforcement to make sure the data controller fulfills the individual's requests.

This scheme benefits the data controller and individuals from several aspects. First, the data judge can authenticate if the individual is truly herself exercising the right and determine if the right is legitimate to be exercised. This prevents cybercriminals and terrorists to exploit the data right and threatens security (Sepúlveda Carmona, M.M., 2018). In addition, the data judge will help execute the enforcement and guarantee that individuals' complaints are effectively responded to. For the data controllers that are technically incapable to fulfill individuals' requests, the data judge can also provide technical assistance.

The mechanism will effectively discover unlawful data controllers, regulate the market order, and make the environment and the data chain more transparent.

Monitored Data Transfer

Data transfer is an important channel to promote data accessibility and data fluidity.

There are several scenarios in which data transfer can occur. For instance, a data controller might transfer the collected data to another controller to process the data.

Data controllers might also request or purchase individual information from other data controllers with a Data Trader. When the entity processing the data changes, individuals must be re-informed of the change and acquire their consent. However, this will normally become a complex process as it involves communicating with all concerned

individuals. With the help of the data management platform, this can be done more efficiently.

Diagram 6 Monitored Data Transfer #1

In step 1, data controller #1 sends a request stating the intention to transfer data to data controller #2. The ACM will analyze the scope, context, and nature of the request as well as the kinds of processing activities controller #2 will engage in. The ACM will compare if it meets the conditions and privacy configuration set forth by the individual. If it is consistent, we skip step 3 and jump to step 4, automatically allowing the transfer to happen. If it is not, in step 3 the ACM will send a notice to the individual, asking her to re-authorize the transfer. The individual can choose to agree or not.

This scheme largely reduces the cost to reacquire individuals' consent for data transfer and promote data fluidity and accessibility. But individuals' choices and privacies are still protected. At any time, they can choose to adjust their privacy configuration according to their will or turn off ACM's automated functions.

However, one problem this design may encounter is that when an individual disables all kinds of data transfer in the privacy configuration, the ACM will reject all transfer requests in step 2 and ask for re-authorization every time. The individual might become impatient and just reauthorize all activities, potentially to untrusted data controllers. To prevent this, we will have a more detailed privacy configuration set-up. Individuals should list with less generalization about the specific categories of information that can be transferred. Overall, we encourage individuals to not broadly reject all transfer requests.

Alternatively, data controllers can choose to anonymize the data and transfer it without consent. However, as discussed, there are no complete anonymization and effective ways to stop the recipient data controllers from decrypting data. Under the novel scheme, the Central Technical Experts assigned to data controllers will perform the encryption and decryption over data, strictly following the processing conditions. Data controllers can purchase encrypted information from the Data Trader under coordination with the Data Exchange Bureau.

Diagram 7 Monitored Data Transfer #2

In step 1, data controller #1 will send a request to the data trader specifying the kinds of datasets it intends to purchase. If the data trader has the dataset available, it will set the pricing, and the two parties can engage in transactions. If it is unavailable, the data trader can help locate the dataset from another data controller, who is marked as #2 in

Diagram 7. The data trader will help verify the data controller #2's entity and the quality of the dataset, as well as the privacy risk involved in the transfer. A report might be sent to the Data Exchange Bureau to get signatures if necessary. In step 3, data controller #2 will encrypt the data under the supervision of CTEs and transfer it to the data trader. In step 4, the data trader will distribute the dataset to data controller #1. All parties will keep a record of the transfer and the dataset's meta-data will be updated to track the information flow.

Throughout the process, the Data Exchange Bureau will monitor the data trader to ensure legitimacy and market order. It is impractical to track every transaction and exchange of datasets between individual data controllers, so regulating a centralized free-market platform like the data trader is more desirable.

Notice and Broadcast

In daily maintenance of the data management platform and updating individuals' information, there require several channels to enhance communication. Once in a given period, the ACM should send a RECAP report to the individual detailing the kinds of information that is being processed. RECAP stands for Record, Evaluate, and Compare Alternative Prices ((Thaler, R. H., and Cass R. S., 2009). The idea is first inspired by Professor Thaler and his work *Nudge*. The report serves to increase information availability and help consumers understand complex mortgage payments. Essentially, it will be a fee schedule that breaks down the cost of a mortgage in great detail. Consumers are better educated by the report and can make wiser mortgage plans. Likewise, the ACM can provide reports to display what and by whom the personal information is being

processed. Individuals can check their information and update them timely. If some information needs to be updated, the individual will also be reminded by the ACM.

The Data Notice mechanism can effectively increase the accuracy and timeliness of the data, which will better reflect individuals' circumstances. Data controllers' decisions pertaining to individuals will be based on the most accurate and latest data, which will ultimately benefit individuals from fewer mistakes and biases. Additionally, making individuals update their information frequently will increase the dataset quality, generating more economic values out of them.

The data management platform should also broadcast to individuals about the updates and information related to big data and information technologies to raise awareness. For instance, individuals can be educated to grasp the concept of automated decision-making. They should know under what circumstances they are the target of automated decision-making, so they can better exercise their rights over data if they feel disadvantaged in any way. Without proper education, most individuals will be unaware that they are subject to automated decision-making and might be influenced by the algorithms.

7. Conclusion

This paper explores a novel data governance scheme based on behavioral economics principles. Essentially, this structure serves to increase regulatory efficiency. Current data protections try to identify the data controllers that do not follow regulations and punish them one by one. The paradigm regulates the various groups in a “distributed” fashion, which takes up too much legislative and administrative resources. However, the new scheme aims to utilize a regulatory structure to holistically affect society. The scheme is more dynamic and efficient compared to the traditional distributed regulatory system.

The scheme finds deeper resonance with Plato’s Theory of Forms, which narrates “Forms” as the non-physical essences of reality, of which physical matters are the mere projections. In this sense, this data governance is a “Form” that provides the blueprint for regulatory implementations.

“each idea exists

phenomena take their names from the ideas in which they participate”

- Republic, Plato

8. More Scenarios

The scenarios, components, and other features described in this paper is not exhaustive to capture the essence of the governance scheme. More interactions and elements are open to being imagined and discussed by interested individuals. In the following sections, some thought-provoking designs will be introduced, hopefully inspiring more ideas from the readers to further the framework.

Privacy-enhancing Technologies (PETs)

The information technology community lacks the transparency between technology development with users and regulatory bodies. Individuals lack guarantees from the technologies about security and privacy features. An alternative solution is to ask developers to create “privacy-by-design” technologies to embed privacy-preserving features (Danezis, G., Domingo-Ferrer, J., Hansen, M., et al., 2015). Privacy-enhancing Technologies are such that can minimize the processing of personal data and decrease risks of information disclosure. For instance, differential privacy is a technique used to protect privacy by adding randomized noises to a dataset. The perturbation is to the extent that it modifies an individual value but maintains the overall dataset’s attributes and patterns. A practical application is its use by the United States Census Bureau to collect information about individuals (Ruggles, S., Fitch, C., Magnuson, D. and Schroeder, J., 2019). The Bureau will distort some individual’s personal information, like age, sex, but in a way that makes sure on various levels of geography, the statistics are not affected. For example, a state’s gender ratio will remain the same after injecting noises.

Another example of the PET is the Federated Learning System. It is used to train AI models to protect privacy and prevent data leakage (Yang, Q., Liu, Y., Chen, T., et al., 2019). An idea first proposed by Google, the goal is to build machine learning models based on training datasets distributed across multiple edge devices. It will not transmit individual data from their devices to the central server to conduct model training but

localize the model training. A common application is on Apple's Siri system. Each Apple Siri can recognize the device owner's unique voice through training with the local data from the users (Zhao, Y., Zhao, J., Jiang, L., et al., 2020). At the same time, Siri can recognize general human voices, using the models from the central server. Because Federated Learning will only aggregate the trained local model parameters but does not involve any direct personal data transfer to the central server, it will effectively reduce the risks of personal information disclosure and enhance privacy protection. It will also break up data isolation among different data controllers and enhance data fluidity. Under the novel governance scheme, the ACMs can be responsible to manage individual data as well as the local models, conduct local training to implement federated learning. The federated learning system will also become interoperable, so different data controllers can access the local models and generate more values from the system.

According to Gartner's report on Top Strategic Technology Trend for 2021, the PETs will generally feature three kinds of technologies (Gartner., 2020). They can provide a trusted environment or container for data to be analyzed, perform decentralized data processing and analytics, and encrypts data before processing. The three features can well integrate with the design of the data management platform. The ACMs can encrypt an individual's data and the data management platform should provide a solid framework to implement decentralized processing and provide secure interfaces to obtain data from individuals. The platform should also broadcast programs about PETs to raise awareness among developers to integrate privacy features in their applications.

Policymakers can also benefit from gaining a deeper understanding of privacy engineering and develop better audit schemes to verify the trustworthiness of data controllers. The data management framework and the governance scheme can be used to support those audit procedures through providing interfaces to policymakers and facilitate privacy and data protection.

Digital Tax

Digital Tax is the imposition on tech companies that generate revenue by providing digital goods and services (Tax Foundation, 2020). Data constitutes a form of strategic asset which firms can use to build up barriers to market entry and exploit the competitive advantage. As tech giants become increasingly powerful and occupying more market shares, they use their massive amount of data to conduct user profiling and analysis, refine products, and ultimately provide better goods and services (Hovenkamp, E., 2018). Consequently, small digital companies cannot compete with them. Taxation is a common way to discourage the formation of monopolies and foster more competition and innovations.

However, there are many challenges associated with the digital tax. By which standards we should use to determine the taxation subject and the tax base remain controversial (Olbert, M. and Spengel, C., 2017). We don't know yet if we should tax each transaction about digital services and goods or tax on the entire revenue. Additionally, we cannot prevent double taxation. Data controllers use derived data to generate values and improve their services, yet we have no definite ways to quantify the increased profits generated by data analysis and measure the value of data (Ahmad, N. and Ribarsky, J., 2018). There are also technical difficulties to track transaction records and the flow of data processing.

Through the novel data governance scheme, we hope to find a solution. With the aid of the data management platform and all components and features, there will be increased transparency and tractability in transactions. The platform will provide a technical framework that supports functionalities such as data value estimation to allow digital taxation.

Taking a step further, this platform might inspire more innovative ways to regulate the market and benefit the public. For instance, the U. S. EPA devised a tradable permits

system to control pollutant emissions (Aldy, J.E., 2019). Polluting industries have a permissible emission rate, and for those that manage to bring the emission level down, they can trade the allowances in the market to other industries. Therefore, behaviors to reduce pollutions are encouraged and turned into market behavior without paternalistic regulations. The tradable permit system is also based on the principle of libertarian paternalism. It gives polluting industries more flexibility to set their own polluting limits and incentivize them to reduce pollution through the price mechanism. Likewise, we can use the digital tax and the platform's functionalities to develop new solutions to regulate monopolies. For instance, as analyzed by Carrier in his work, with the current antitrust law, it is less likely to directly disintegrate tech giants and dominant data controllers (Carrier, M. A., 2020). However, since the tech giants are integrating seamlessly with our lives, we can make the tech giants carry more social responsibilities and turn them into a public entity. This should not be done by paternalistic regulations, but through a market mechanism. We can devise a "tradable responsibility system" to let each tech giant carry a certain number of social responsibilities and fulfill public functions. Those who exceed the boundary can trade their contribution to other data controllers. The data management platform and the governance scheme will help quantify and track the performance of these giant data controllers. This might be a novel regulation method that respects the free market and follows relevant legislation.

Ultimately, the mechanisms will introduce more competition and innovations into the market and combat the monopoly and anti-competitive behaviors to benefit the public.

References

- Aaronson, S.A., 2019. Data is different, and that's why the world needs a new approach to governing cross-border data flows. Digital Policy, Regulation and Governance.
- Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L.F., Komanduri, S., Leon, P.G., Sadeh, N., Schaub, F., Sleeper, M. and Wang, Y., 2017. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys (CSUR)*, 50(3), pp.1-41.
- Acquisti, A., Taylor, C. and Wagman, L., 2016. The economics of privacy. *Journal of economic Literature*, 54(2), pp.442-92.
- Ahmad, N. and Ribarsky, J., 2018, September. Towards a framework for measuring the digital economy. In 16th Conference of the International Association of Official Statisticians.
- Albrecht, J.P., 2016. How the GDPR will change the world. *Eur. Data Prot. L. Rev.*, 2, p.287.
- Aldy, J.E., 2019. Evaluating Regulatory Performance: Learning from and Institutionalizing Retrospective Analysis of EPA Regulations. *Case W. Res. L. Rev.*, 70, p.971.
- Allen, A.L., 2016. Protecting one's own privacy in a big data economy. *Harv. L. Rev. F.*, 130, p.71.
- Altorbq, A., Blix, F. and Sörman, S., 2017, December. Data subject rights in the cloud: A grounded study on data protection assurance in the light of GDPR. In 2017 12th International Conference for Internet Technology and Secured Transactions (ICITST) (pp. 305-310). IEEE.
- Ausloos, J., 2018. The Right to Erasure: Safeguard for Informational Self-Determination in a Digital Society?.
- Batista, M., Fernandes, A., Ribeiro, L.P., Alturas, B. and Costa, C.P., 2020, June. Tensions between Privacy and Targeted Advertising: Is the General Data Protection Regulation being violated?. In 2020 15th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-5). IEEE.
- Banterle, F., 2018. The interface between data protection and IP law: the case of trade secrets and the database sui generis right in marketing operations, and the ownership of raw data in big data analysis. In *Personal Data in Competition, Consumer Protection and Intellectual Property Law* (pp. 411-443). Springer, Berlin, Heidelberg.
- Bendiek, A. and Römer, M., 2019. Externalizing Europe: the global effects of European data protection. Digital Policy, Regulation and Governance.

- Biancotti, C. and Ciocca, P., 2019. Opening Internet Monopolies to Competition with Data Sharing Mandates (No. PB19-3).
- Bozdag, E., 2018. Data Portability Under GDPR: Technical Challenges. Available at SSRN 3111866.
- Brasher, E.A., 2018. Addressing the Failure of Anonymization: Guidance from the European Union's General Data Protection Regulation. *Colum. Bus. L. Rev.*, p.209.
- Bukht, R. and Heeks, R., 2017. Defining, conceptualising and measuring the digital economy. *Development Informatics working paper*, pp. 68.
- Buttarelli, G., 2016. The EU GDPR as a clarion call for a new global digital gold standard.
- Bygrave, L.A., 2002. *Data protection law*. Wolters Kluwer Law & Business.
- Caribou Digital., 2016. *Winners and losers in the global app economy*. Caribou Digital Publishing, Farnham, Surrey
- Carrier, M.A., 2020. Big Tech, Antitrust, and Breakup. *Georgetown Journal of International Affairs* (Jan. 2020).
- Chander, A. and Lê, U.P., 2014. Data nationalism. *Emory LJ*, 64, p.677.
- Clifford, D. and Ausloos, J., 2018. Data protection and the role of fairness. *Yearbook of European Law*, 37, pp.130-187.
- Crampton, J.W., 2015. Collect it all: National security, big data and governance. *GeoJournal*, 80(4), pp.519-531.
- Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J.H., Metayer, D.L., Tirta, R. and Schiffner, S., 2015. Privacy and data protection by design-from policy to engineering. *arXiv preprint arXiv:1501.03726*.
- European Parliament and Council of European Union., 2016. Regulation (EU) 2016/679. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN>
- European Parliament and Council of European Union., 2016. General Data Protection Regulation. *Official Journal of the European Union*. Recital 42.
- Evans, P. and Gawer, A., 2016. The rise of the platform enterprise: A global survey. *The Emerging Platform Economy Series*, 1. The Centre for Global Enterprise, New York, NY

- Forgó, N., Hänold, S. and Schütze, B., 2017. The principle of purpose limitation and big data. In *New technology, big data and the law* (pp. 17-42). Springer, Singapore.
- Gartner., 2020. Gartner Top Strategic Technology Trends for 2021. Available at: <https://www.gartner.com/smarterwithgartner/gartner-top-strategic-technology-trends-for-2021/> (Accessed: November 2020)
- Goddard, M., 2017. The EU General Data Protection Regulation (GDPR): European regulation that has a global impact. *International Journal of Market Research*, 59(6), pp.703-705.
- Gola, Schomerus, *BDSG Kommentar*, page 47, München 2010, ISBN 978-3-406-59834-0
- Graef, I., 2018. When data evolves into market power: Data concentration and data abuse under competition law. In: Moore M and Tambini D, eds. *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*. Oxford University Press, Oxford. pp. 71–97.
- Gruschka, N., Mavroeidis, V., Vishi, K. and Jensen, M., 2018, December. Privacy issues and data protection in big data: a case study analysis under GDPR. In *2018 IEEE International Conference on Big Data (Big Data)* (pp. 5027-5033). IEEE.
- Guanglan, T., 2020. Data Subject Rights in the Era of Big Data and Their Unsolved Issues: The EU GDPR Taken as the Analysis Object. *Journal of Renmin University of China*, 34(6), p.131.
- Hansen, P.G., 2016. The definition of nudge and libertarian paternalism: Does the hand fit the glove?. *European Journal of Risk Regulation*, 7(1), pp.155-174.
- Hovenkamp, E., 2018. Platform Antitrust. *J. Corp. L.*, 44, p.713.
- Isaak, J. and Hanna, M.J., 2018. User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer*, 51(8), pp.56-59.
- Jensen, C., Sarkar, C., and Potts, C., 2007, July. Tracking website data-collection and privacy practices with the iWatch web crawler. In *Proceedings of the 3rd symposium on Usable privacy and security* (pp. 29-40).
- Kim, G.H., Trimi, S. and Chung, J.H., 2014. Big-data applications in the government sector. *Communications of the ACM*, 57(3), pp.78-85.
- Kenney, M. and Zysman, J., 2016. The rise of the platform economy. *Issues in Science and Technology*, 32(3), p. 61-69.
- Khan, L.M., 2018. Its Challenge to Current Antitrust Law. *Digital Dominance: The Power of*

- Google, Amazon, Facebook, and Apple, 710, p.805.
- Koning, M.E., 2020. The purpose and limitations of purpose limitation (Doctoral dissertation, [Sl: sn]).
- Koops, B.-J. 2014. "The Trouble with European Data Protection Law." *International Data Privacy Law* 4 (4): 250–61. <https://doi.org/10.1093/idpl/ipu023>.
- Labrinidis, A. and Jagadish, H.V., 2012. Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, 5(12), pp.2032-2033.
- Maguire, S., Friedberg, J., Nguyen, M.H.C. and Haynes, P., 2015. A metadata-based architecture for user-centered data accountability. *Electronic Markets*, 25(2), pp.155-160.
- Malgieri, G. and Custers, B., 2018. Pricing privacy—the right to know the value of your personal data. *Computer Law & Security Review*, 34(2), pp.289-303.
- Mantelero, A., 2017. From group privacy to collective privacy: towards a new dimension of privacy and data protection in the big data era. In *Group Privacy* (pp. 139-158). Springer, Cham.
- McDermott, Y., 2017. Conceptualising the right to data protection in an era of Big Data. *Big Data & Society*, 4(1), p.2053951716686994.
- Medium, 2017. *Five loopholes in the GDPR*. Available at: <https://medium.com/mydata/five-loopholes-in-the-gdpr-367443c4248b> (Accessed: September 2020).
- Mitchell, A.D. and Hepburn, J., 2017. Don't Fence Me In: Reforming Trade and Investment Law to Better Facilitate Cross-Border Data Transfer. *Yale JL & Tech.*, 19, p.182.
- Mittal, S., 2017. Old Wine with a New Label: Rights of Data Subjects Under GDPR. *International Journal of Advanced Research in Computer Science*, ISSN, (0976-5697).
- Mueller, M. and Grindal, K., 2019. Data flows and the digital economy: information as a mobile factor of production. *Digital Policy, Regulation and Governance*.
- Muravyeva, E., Janssen, J., Dirkx, K. and Specht, M., 2018, December. Students' Attitudes Towards Personal Data Sharing in the Context of e-Assessment: Informed Consent or Privacy Paradox?. In *International Conference on Technology Enhanced Assessment* (pp. 16-26). Springer, Cham.
- Olbert, M. and Spengel, C., 2017. International taxation in the digital economy: challenge accepted. *World tax journal*, 9(1), pp.3-46.

- Pankowska, M., 2018. Privacy Awareness in the GDPR Implementation Circumstances. In B. Andersson, B. Johansson, S. Carlsson, C. Barry, M. Lang, H. Linger, & C. Schneider (Eds.), *Designing Digitalization (ISD2018 Proceedings)*. Lund, Sweden: Lund University. ISBN: 978-91-7753-876-9. <http://aisel.aisnet.org/isd2014/proceedings2018/Transforming/5>.
- Pagallo, U., 2017. The legal challenges of big data: putting secondary rules first in the field of EU data protection. *Eur. Data Prot. L. Rev.*, 3, p.36.
- Politou, E., Alepis, E. and Patsakis, C., 2018. Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions. *Journal of Cybersecurity*, 4(1), p.tyy001.
- Posner, R.A., 1977. The right of privacy. *Ga. L. Rev.*, 12, p.393.
- Richards, N.M. and King, J.H., 2014. Big data ethics. *Wake Forest L. Rev.*, 49, p.393.
- Ruggles, S., Fitch, C., Magnuson, D. and Schroeder, J., 2019, May. Differential privacy and census data: Implications for social and economic research. In *AEA papers and proceedings* (Vol. 109, pp. 403-08).
- Sagers, C., 2019. Antitrust and Tech Monopoly: A General Introduction to Competition Problems in Big Data Platforms: Testimony Before the Committee on the Judiciary of the Ohio Senate. Available at SSRN 3471823.
- Sagiroglu, S. and Sinanc, D., 2013, May. Big data: A review. In 2013 international conference on collaboration technologies and systems (CTS), pp. 42-47.
- Schermer, Bart W., Bart C., and Simone, H., 2014. The Crisis of Consent: How Stronger Legal Protection May Lead to Weaker Consent in Data Protection. *Ethics and Information Technology*, March. <https://doi.org/10.1007/s10676-014-9343-8>
- Sepúlveda Carmona, M.M., 2018. Is biometric technology in social protection programmes illegal or arbitrary? an analysis of privacy and data protection (No. 994989393402676). International Labour Organization.
- Shah, A., Banakar, V., Shastri, S., Wasserman, M. and Chidambaram, V., 2019. Analyzing the Impact of GDPR on Storage Systems. In 11th USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage 19).
- Solove, D. J., 2013. Privacy Self-Management and the Consent Dilemma. *Harvard Law Review* 126: 1880–1903.
- Soh, S.Y., 2019. Privacy nudges: An alternative regulatory mechanism to informed consent for online data protection behaviour. *Eur. Data Prot. L. Rev.*, 5, p.65.

- Stallkamp, M. and Schotter, A.P., 2019. Platforms without borders? The international strategies of digital platform firms. *Global Strategy Journal*.
- Tax Foundation. 2020. Digital Taxation Around the World. Available at: <https://taxfoundation.org/digital-tax> (Accessed: November 2020)
- Thaler, R. H., and Cass R. S., 2009. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. United States of America: Yale University Press.
- Tikkinen-Piri, C., Rohunen, A. and Markkula, J., 2018. EU General Data Protection Regulation: Changes and implications for personal data collecting companies. *Computer Law & Security Review*, 34(1), pp.134-153.
- United Nations. General Assembly, 1949. Universal declaration of human rights (Vol. 3381). Department of State, United States of America.
- United Nations. 2019. Digital Economy Report 2019. United Nations Conference on Trade and Development, pp. 27-33.
- Van Dijck, J., 2014. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & society*, 12(2), pp.197-208.
- Warren, S.D. and Brandeis, L.D., 1890. The right to privacy. *Harvard law review*, pp.193-220.
- Wachter, S., Mittelstadt, B. and Russell, C., 2017. Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harv. JL & Tech.*, 31, p.841.
- Willoughby, A., 2017. Biometric Surveillance and the Right to Privacy [Commentary]. *IEEE Technology and Society Magazine*, 36(3), pp.41-45.
- Wired., 2018. The Strava Heat Map and the End of Secrets. Available at: <https://www.wired.com/story/strava-heat-map-military-bases-fitness-trackers-privacy/> (Accessed: July 2020).
- Yang, Q., Liu, Y., Chen, T. and Tong, Y., 2019. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(2), pp.1-19.
- Yook, K.C. and Everett, R., 2003. Assessing risk tolerance: Questioning the questionnaire method. *Journal of Financial Planning*, 16(8), p.48.
- Zeyu, L., 2018. The Interpretation and Application of Purpose Limitation Principle in Personal Data Protection. *Journal of Comparative Law*, (5), p.2.

Zhang, M.L. and Chen, M.S., 2019. China's digital economy: Opportunities and risks. International Monetary Fund.

Zhao, Y., Zhao, J., Jiang, L., Tan, R., Niyato, D., Li, Z., Lyu, L. and Liu, Y., 2020. Privacy-preserving blockchain-based federated learning for IoT devices.