Università degli Studi di Macerata
In collaborazione con HEART-ITN consortium

Philips Electronics Netherland B.V.
Katholieke Universiteit Leuven

DOTTORATO DI RICERCA IN

Marie Skłodowska-Curie Innovative Training Networks
European Industrial Doctorates
MSCA-ITN-EID 766139

CROSS-BORDER DATA TRANSFER REGULATION: A COMPARATIVE STUDY OF CHINA AND EUROPE

Presentata da: Yuan LI

Coordinatore
Prof. Francesca Spigarelli

Relatore
Prof. Simone Calzolaio
Prof. Laura Vagni
Prof. Alessio Bartolaceli
Prof. Milan Petkovic

Esame finale anno 2021

CROSS-BORDER DATA TRANSFER REGULATION: A COMPARATIVE STUDY OF CHINA AND EUROPE

ACKNOWLEDGMENT

INTRODUCTION

Problem Statement

Data protection law today is like one of those megacities, Shanghai city for example, where construction cranes are everywhere and is keeping sprawling outwards. Walking inside the old city centre surrounded by new skyscrapers and new neighbourhood, even the oldest resident would find it chaotic, yet excited for an effervescent discovery. The rise of the data economy dominates many of the contemporary policy debates. One specific aspect often overlooked is the impact of cross-border data transfer (CBDT) regulation.

With technological progress, the development of economic models and the deepening of global integration, the global economy has transformed from an information-driven model to a data-driven model. The economic value of data is constantly being explored and leveraged. Contemporarily, around half of cross-border trade in global service trade can be carried out through the Internet or other communication means. The development of the Internet has led to the development of cross-border e-commerce that changed the traditional trade model and promoted the interactive exchange of global commodities, resource sharing and business model innovation. Cross-border data flow is the cornerstone of the cross-border e-commerce economy, especially for digital products. Without cross-border data flow, there will be no transactions.

Against this background, it is of practical significance to study the legal issues of cross-border data transfer regulation. Cross-border data flow, as a factual act of intergovernmental cooperation and economic activity, does not necessarily present a positive or negative legal evaluation. However, when data flows are exchanged across

borders, or after entering the storage and digital devices of third countries or other organisations, a series of legal issues will arise.

Large-scale user information leakage incidents have occurred frequently around the world, personal information has been sold everywhere, and national security has been threatened as well. As a result, many countries have introduced data protection policies, such as data localisation, which require citizens' data to be stored in their own borders and prevent them from leaving the country, for the consideration of citizens' privacy and national information security. Although such a data policy protects citizens' privacy and state security to a certain extent, it hinders the flow of cross-border data, thus had a profound negative impact on the global economy and trade.

In the scramble for an adjustable and effective cross-border data transfer mechanism that can successfully tackle the impediments created by the digital economic development and other new technologies, China began exploring various law makings that are in alignment with international conventions and treaties and that is desirably relevant to the mounting demands of the developing Chinese socio-cultural and economic setting. In the frantic search for an unassailable solution, China simply borrowed legislative approaches from developed societies, such as the European Union and the United States, with her unique spice of strict state security requirements added.

The problem of merely legal transplantation is that, when the legislative objectives of each country serve different purposes, there will inevitably arise the issue of incompatibility. After adopted the EU's approach of strict supervision of data flow, it has intensified data export restrictions and increased compliance costs. In this dissertation, the author summarises the legal issues that may hinder the free flow of data

across borders into national security, public interest, and personal information protection. This study focuses on emerging research topics, starting with contemporary public policies on the cross-border data transfer (CBDT).

Background of the Study

General Background Information

Globalisation is one of the main trends in the simultaneous development of the world and information civilisation. As defined by Jürgen Habermas, globalisation is "a structural shift in the world economic system"[1], where economic globalisation is the most important form. It refers to the formation of a unified sphere of market economic activities comprising of intermediate products, final products and service products moving across the geographical boundaries.[2] The importance of cross-border value chains in the international economic cycle is ever increasing and the flow of production factors is profoundly changing the way the global trade operates.

Being the key factor in trade, data flow transferred globally is a typical form of economic globalisation. The trend of data globalisation has been significantly reflected in the development of economic globalisation in the last decade. From 2008 onwards, globalisation has not reversed or stagnated as commodity trade has flattened and cross-border capital flows have declined sharply. Rather, globalisation is entering

[1] Jürgen Habermas, 'Jenseits des Nationalstaats?' in Ulrich Beck (eds), *Politik der Globalisierung* (2nd edn, Suhrkamp Verlag 1998).

[2] Helmut Wagner, *Globalization and Unemployment* (Springer 2000) 19.

a new era due to the soaring flows of cross-border data and information.[3] The cross-border data flow has grown 45 times larger since 2005, while all types of flows acting together have raised world GDP by 10.1 per cent over what would have resulted in a world without any cross-border flows. This value amounted to 7.8 trillion US dollars in 2014 alone, and data flow accounts for 2.8 trillion US dollars of this impact.[4] Global flows of all types of support growth by raising productivity, and data flows are amplifying this effect by broadening participation and creating more efficient markets.

Digital technologies are transforming business and international trade. There has been a massive growth in the complexity and volume of global data flows, accompanied by a change in the nature of such transfers in that they no longer constitute point-to-point transmissions, but occur today as part of a networked series of processes made to deliver a business result.[5] In the 1970s, the term "transborder data flow" was typically understood to refer to data transfers as the "exchange of internal company administrative information, response to requests for service by customers, and maintenance of records concerning or describing customers or subjects". Such flows often only occurred when there was explicit intent to transfer data internationally.[6]

---

[3] Mckinsey, 'Digital Globalization: The New Era of Global Flows' (2016) <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/digital-globalization-the-new-era-of-global-flows> accessed 25 February 2019.

[4] ibid.

[5] Paul Schwartz, 'Managing Global Data Privacy： Cross-Border Information Flows in a Networked Environment' (2009) <http://theprivacyprojects.org/wp-content/uploads/2009/08/The-Privacy-Projects-Paul-Schwartz-Global-Data-Flows-20093.pdf> accessed 3 March 2019.

[6] Christopher Kuner, 'Regulation of Transborder Data Flows under Data Protection and Privacy Law: Past, Present, and Future' (2011) OECD Digital Economy Papers No. 187.

Transborder data flows today involve multiple parties communicating through networks in a distributed fashion, in particular via data-driven artificial intelligence, Internet of Things and cloud computing. The architecture of the Internet and technological solutions mean that even a transfer to a party inside the same country may result in the message or file transiting via other countries. As computing devices are routinely implanted in many varieties of implements used in daily life, great volumes of personal data are collected, processed and transferred internationally even without the direct involvement of a human being. The application of data analytics techniques to large amounts of personal data typically involve the transfer of data from numerous sources without regard to geography, and thus makes the boundaries of physical territories blurring.[7]

Why A Comparative Study of China and the European Union?

As early as 2011, China and Russia, among others, submitted the proposal of the International Code of Conduct on Information Security to the UN General Assembly.[8] In January 2015, the member states of the Shanghai Cooperation Organisation updated the above-mentioned proposal and distributed it as an official document of the General Assembly. In 2015, China enacted National Security Law (NSL) and provided the

[7] Christopher Kuner, Transborder Data Flow Regulation in Data Protection and Privacy Law (OUP 2013) 4.

[8] UN document A/66/359, 'Letter dated 12 September 2011 from the Permanent Representatives of China, the Russian Federation, Tajikistan and Uzbekistan to the United Nations addressed to the Secretary-General' (2011) < https://undocs.org/A/66/359> accessed 25 February 2020.

legislative interpretation of "national security".[9] Article 25 of NSL established the concept of "cyberspace sovereignty" in the form of legislation by ruling the "maintenance of national cyberspace sovereignty, security and development interests". In 2016, the Standing Committee of the National People's Congress promulgated the Cyber Security Law (CSL), in which a security protection framework for critical information infrastructure has been established.[10] China is aware of the importance of cross-border data flow in promoting economic development, and has gradually fastened legislation progress on cross-border data transfer (CBDT). Art 5.4.5 of the Information Security Technology – Guidelines for Personal Information Protection in Public and Commercial Service Information Systems (信息安全技术公共及商用服务信息系统个人信息保护指南) issued by administrative bodies in 2012 for the first time addressed the rules for personal information cross-border flow in China.[11] E-Commerce Law and Personal Information Protection Law, both of which are newly enacted in 2020, provide the scope of "personal information" and users' "information rights". Sector-specific provisions on CBDT focused on, i.e. banking, healthcare and credit investigation industrial rules and regulations.

China's international cooperation in CBDT mainly is discussed through trade agreements. The negotiations of FTAs between Sino-New Zealand, Sino-Peru, and

[9] National Security Law of the People's Republic of China (中华人民共和国国家安全法), as adopted at the 15th session of the Standing Committee of the Twelfth National People's Congress of the People's Republic of China on July 1, 2015.

[10] Cybersecurity Law of People's Republic of China (中华人民共和国网络安全法).

[11] It is ruled that, by default cross-border personal information transferred is not allowed, unless one of three derogations are met, that is, consent, explicit permission in laws, and administrative authorities' approval.

Sino-Chile involve adding e-commerce chapters. Since 2000, the "going global" (走出去) of Chinese enterprises and the rapid development of cross-border e-commerce have brought new opportunities for China's reform and opening up. The development of the "Internet+" (互联网+) strategy is considered the enhancement of international economic competitiveness, the promotion of the supply-side reform (供给侧改革). The Belt and Road Initiative (一带一路倡议), which involves more than 138 countries and 31 international organisations, has brought great challenges of social and legal issues about CBDT.[12]

Significance of the Study

Theoretical Significance

Cross-border data transfer faces conflicts between data sharing and various legitimate interests. Yet, no consensus on how to balance these conflicts of interests is agreed upon. Contemporary literature show a lack of unified or interoperable data cross-border transfer rules at the international level, and the provisions on information security protection and cross-border data flow in the e-commerce chapter of regional trade agreements lack implementation mechanisms. The level of information security protection in different countries is inconsistent. The differences in legal systems make it difficult for data to flow freely across borders on a larger scale. To strengthen legal research on cross-border data flows, a unified comparative study of international and

[12] See Graham Greenleaf, '*Asia's Data Privacy Dilemmas 2014-19: National Divergences, Cross-Border Gridloc*k' (2019) UNSW Law Research Paper No. 19-103, 52.

domestic regulations is needed. The key issue is how to distinguish and define the legal policy objectives that restrict cross-border data flows.

The author categorised the rationales for regulatory measures as the following. Firstly, **national security** in CBDT is essentially in the public international law regime and is a political issue. Hence, the principles and theories of public law should be applied for analysis. The principle of sovereignty equality, which is the most important basic principle in public law, should be considered to be leveraged to cross-border data transfer activities in cyberspace. However, the issue of sovereignty in cyberspace is an emerging challenge for traditional international law principles in the digital era. Particularly, the conflicts existed between developing countries and developed countries due to the differences in technological levels. Therefore, it is needed to understand the principles and legal measures of state security in cross-border data transfer.

Secondly, **public interest** is the important legal basis for CBDT. The first issue is the difficulty of defining public interest, as often it is intersected, overlapped, and indistinguishable from national security. Additionally, to clarify discretion of administrative agencies in defining public interests, including the establishment of legal principles that administrative agencies should observe when performing their duties, are needed. What is more, CBDT involved various sectors such as finance, healthcare, human resource, and custom, which need to collect users' personal information as well. In case of critical infrastructure and sensitive data, the development of e-government, the performance of certain government affairs, and the research of industrial science and technology require the exchange and sharing of overseas data. Yet they often break through restrictions and prohibitions through certain legal cooperation methods.

Thirdly, **personal data protection** is the core of the CBDT study. There are great differences in the rights related to personal data protection and privacy in various countries, including whether the rights in personal information are regarded as fundamental human rights. Similarly, there are issues concerning the legal rights that data subjects should enjoy, the obligations of data service providers, and the issue of legal standards for cross-border transfer of personal data. Personal data also includes various types of data, such as child data, employee data, etc., subject to special regulations. Therefore, the issue of personal data needs to address the general rules and special exceptions of the flow of personal data while securing the protection of personal data.

Practical Significance

With the development of the Internet economy, countries and regions around the world have been pushing for higher data protection standards, while the emergence of new cases has also produced various newly created data subject rights and information protection principles. Studies are conducted to provide legal frameworks and to guarantee economic development and technological progress. At the international level, international cooperation still needs to be expanded, and it is difficult to meet the real needs of fast-developing cross-border data flow. How to accelerate international cooperation and reduce the imbalance of protection levels among different regions and countries on a global scale is the problem that this dissertation attempts to explore.

It is clear from the above research objectives that the European Union and China are of major concern of this dissertation. The reason for this, as has already been partly touched upon in the precious discussion, is as follow: China, the EU, and the U.S. are three regions that digital trade within are most developed with strong market

and social demands. But China and the EU both have adopted authoritarian measures that uphold a harsh restriction to free flow of data cross-border transfer. Especially, China has an inherent censorship system, which gives the problem of transfer data outside the territory of China its unique flavour. In other words, Sino-EU transfer of personal information is a unique sample to show how complex the legal problems of cross-border data transfer could be when political control over multinational digital trade is in place. Compared with the situation in the EU-US data transfer framework, in Sino-EU digital trade, the collision and interaction among the critical issues of trans-border data flow, such as lack of harmonisation, conflicts of laws, unawareness of the historical and legal context and the imbalance of the interests involved are amplified, which could provide observers a better understanding of the transborder-data-related problems and a more vivid and clear picture of the future of our global digital trade regulation system. However, this dissertation is global in scope, and it purposely does not delve into the minutiae of national regulatory requirements, and focuses on the main jurisprudential themes.

Literature Review (TBC)

The objective of this dissertation is to explore a transparent framework of CBDT that is in line with the EU value and China's development priorities. Since the EU has established a set of standards as the legal basis for cross-border personal data transfer using adequacy recognition, it will be used as the reference object. Firstly, the EU CBDT is originated from the rules of personal data protection. The right to information self-determination (Recht auf informationelle Selbstbestimmung) is firstly coined by German scholar Steinmueller in 1971 and was recognized as a form of human right by

the Bundesverfassungsgericht in Volkszahlung case in 1983, and has been in the dominant position in German data protection theory since then. The Court held that, everyone should know and decide by themselves who collects and processes what data from him. The right to self-determination of personal information is formed in Article 2 of German Basic Law and gradually becomes the basis of the right to personal data protection in the EU. Globally, Japan, China and other regions also followed such philosophy. On the contrary, the U.S. has adhered to the basic principles of the privacy laws in cyberspace, that is, the free movement of information and freedom of expression with the privacy as an exception.[13]

The divergence in the understanding of the object of the rights directly leads to two paths in the establishment of CBDT restrictions. Christopher Kuner defines the two as geographically-based approach (adequacy) and organisationally-based approach (accountability). Prior to the GDPR, most of the legal texts were drafted before the vast popularisation of the Internet, including the Sweden data protection law requiring the data recipient to "have similar protection standard" in 1973, the OECD guidelines in 1980 and the EU data protection directive in 1995. Easterbrook's famous *Cyberspace and the Law of the Horse* in 1996 denied the unique nature of cyberspace and started the intensive discussion about whether public authorities should intervene the cyberspace and to what extent. Lessig counterargued Easterbrook's idea in *The Law of the Horse: What Cyberlaw Might Teach* to justify the regulation in cyberspace and proposed the principles for the authorities to govern. It seems that such discussion

[13] Paul Schwartz and Daniel Solve, Information Privacy Law, 6th ed, Wolters Kluwer, 2018.

is only limited inside the U.S.: the 2018 GDPR continued the basic principles listed in the 95 Directive.

The EU personal data protection framework, together with the adequacy approach, does not dynamically modify with the changing of the Internet technologies.[14] It indeed received a large degree of resonance worldwide.[15] The influence not only stems from the leverage of EU impact in legislation, but also benefited from the OECD guidelines.[16] Greenleaf analysed the privacy legislation in Asian countries and studied Japanese and North Korean data protection laws by concluding that the two as the only OECD signatory countries in Asia almost completely transplanted the EU data laws with identical CBDT rules set in the OECD guidelines. Japan received the adequacy recognition from the Commission in 2018 and North Korea is in the waiting list to join. The U.S., however, cannot integrate into the EU system due to its conservative approach that does not support such radical personal data rights. In view of the large amount of data transmission needs in reality, the U.S. and the EU had launched multiple rounds of bilateral negotiations on data transfer, and reached agreements including Safe Habor, Privacy Shield, and Passenger Name Record Agreement.

[14] See, Christopher Kuner, 'Transborder data flow and data privacy law'; Greenleaf, 'The influence of European data privacy standards outside Europe: implications for globalization of Convention 108'.
[15] Greenleaf, *Asian Data Privacy Law*
[16] Michael Kirby, 'The history, achievement and future of the 1980 OECD guidelines on privacy' (2011) International Data Privacy Law 1.

Research Questions

Definitions

The concept of regulation will be broadly construed to include all types of conditions, limitations, and restrictions on the transfer of data across national borders. It also encompasses measures that private actors take, whether or not they have binding legal forces, which limit or constrain the transfer of personal data across national borders. Private sectoral instruments such as contractual clauses, internal company policies, and codes of practices are becoming more widely used to structure and protect international data transfers, and may have binding legal value by contractual obligations or regulatory approvals.[17]

This dissertation focuses mainly on the international flows of personal data, which is defined as data relating to an identified or identifiable natural person.[18] For the most part, it does not examine the flows of non-personal data or data that can only identify a legal person. Determining what are and what are not personal data can create controversial results in this study, and the term can have different meanings in different legal contexts.[19] Additionally, with the development of technology, the variety of data

[17] For example, Schemes whereby instruments used by the private sector are either drafted in advanced by public authorities, e.g. the EU-approved standard contractual clauses, or approved by them, e.g. BCRs in the EU, are becoming increasingly common, resulting in a patchwork of private and public regulation. See Commission Decision (EC) 2010/87/EU of 5 February 2010 on standard contractual clauses; Commission Decision (EC) 2001/497 as regards the introduction of an alternative set of standard contractual clauses.

[18] See General Data Protection Regulation (n), Recital 26.

[19] For example, Article 29 Working Party concludes that IP addresses are protected by EU data protection law, while in *Columbia Pictures, Inc. v. Bunnell*, 245 F.R.D. 443, 69 Fed.R,Serv,3d 173 (C.D. Cal. 2007) the California federal court found that IP addresses were not covered by the term "personal information" contained in the defendants' website privacy policy.

is ever-increasing, and the data which is previously considered as non-personal data or anonymous could lead to an individual given enough time and computing power.[20]

Certain types of data flow carried out by the public sector may rise to special issues, such as those conducted for law enforcement purposes.[21] Law enforcement entities often seek access to personal information processed by the private sector.[22] Given the growing interaction between data processing in the private and public sectors, which routinely involves the transfer of personal data across national borders, this dissertation would not distinguish the two and discuss in a separate way with regard to the regulation of transborder data flows.

Research Question

Digital globalisation is becoming the inevitable trend in global trade, while conflicts and restriction by regulatory measures may directly or indirectly impede the development of global digital trade. How to discover the equilibrium between the demand for trade and the control over transborder data flows remains unknown. Therefore, the present study aims to solve this problem by addressing the following research questions.

---

[20] See, e.g., Paul Ohm, 'Broken promises of Privacy: Responding to the Surprising Failure of Anonymization' (2010) 57 UCLA Law Review 1701.

[21] Great concerns that there should be a sharp boundary drawn between data privacy rules in the public and private sectors arisen since 1970s, since the distinction between activities of the private sector and of the law enforcement sector is blurring. See Opinion of the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions, 'A comprehensive approach on personal data protection in the European Union' (2011) 9.

[22] For example, the EU Data Retention Directive mandates that Internet and telecommunications service providers retain certain types of data generated by their users, and make such data available to law enforcement authorities upon request. The volume and scope of such requests are increasing, in some cases even requiring private sector entities to monitor communications of third persons on an ongoing basis.

This dissertation aims to answer the main research question:

> *Whether a transparent cooperative mechanism can be reconciled between Sino-EU digital trade and the regulatory measures in order to solve the complex problem of transborder data flow regulations, in which latter could have a degree of control over the former, and the activities of the former could be legally justified to a certain degree.*

In order to answer the main research question, this dissertation examines three hypotheses of it:

- The data localisation neck lock and restrictions are caused by advancements in digital technology, and it is very difficult to determine the social desirability of cross-border data transfer, and to enforce data protection laws against commercial activities via traditional means.
- The problems that contemporary measures encounter in a transnational scenario are unique and complex, since they concern the general problems of law enforcement against commercial activities in different countries, and the specific endemic problems multinational cooperation face in China and the EU.
- For foreign entities in the Chinese market, regulatory compliance following a framework, which is designed to reconcile the Sino-EU data transfers, is the only possible solution at the moment.

Research Methodology and Expected Value

Scope of the Study

This dissertation explores the history, rationales, beneficial aspects, and possible future of transborder data regulation between the European Union and China. The objective is to examine whether the policy-makers from both regions could better achieve their goals of promoting digital economy by establishing a mutual understanding with the industrial entities, while maintaining the balance between the protection of personal information and the innovation in digital markets. For that purpose, this research explores the historical development of data transfer regulatory measures in China, the EU and the U.S., studied the specific challenges they are encountering in the data globalisation era.

This dissertation focuses only on legal issues relevant to the regulation of transborder data flows, and those only arise under data protection and privacy law. Such regulation exists in other areas of the law as well, such as export control restrictions, copyright law, tax law, etc., and will not be examined in this research, since they either involve the processing of non-personal data or are only peripherally related to data protection and privacy.

This dissertation provides (is aiming to provide) a transparent cooperative framework to the transborder data flow between the European Union and China and to solve the problem caused by data protection rules, which can be deprived from different priority of the political objectives in the two regions. It is noteworthy that it is China's sovereign right to determine its own data policies, as long as they are consistent with China's international obligations. The fundamental differences between the various actors, institutions, and norms involved in transborder data flow regulation

mean that it is best classified as a *Midlevel Principle*[23] sphere for which a constitutional or treaty-based framework is unsuited, but individual steps could be taken to reduce the impact and scope of the problems. Therefore, rather than to wait for the Chinese government to change their related laws as requested by foreign entities, the recommendations and proposition resulting from the study offers the foreign companies to cooperate with Chinese companies, thereby lowering the general cost for regulatory compliance and risks of transborder data restrictions. Once that kind of cooperation is established at the industrial level, foreign companies can more effectively benefit from their digital products and services consumers in China, and a framework of cross-border data transfer at the regional level could thus be foreseeable.

Methodology and Outlines of the Dissertation

In questioning whether a cooperative mechanism can be reconciled between the Sino-EU digital trade and data transfer regulatory measures, this dissertation employs comparative law and the midlevel principle theory as the general methodologies.

Chapter 1 is devoted to the tension between data protection in the digital environment and the rising of international digital trade. A theoretical analysis will be conducted to show the different data protection laws regulating data transfers and the general challenges that entities faced when trying to comply with such regulatory measures in the digital era. The result of this analysis would like to show that the rationales for transborder data flow regulation are unique and complex. Although regulatory measures may impede the international digital trade, the cause behind is the

[23] Robert P. Merges, *Justifying Intellectual Property* (Harvard University Press 2011).

development in digital technology. It is the new advancements in technology that urge competent authorities to push for a higher standard of CBDT control and upset the "territoriality principle" in international law, gradually losing the balance between promoting digital economy and guaranteeing data sovereignty, and such impact will be ever-increasing.

Chapter 2 will evaluate the focused jurisdictions' transborder data regulation by examining the policies and purposes for which these regulatory measures were enacted. Transborder data flow regulation involves norms arising from fundamental rights law, economic regulation, law enforcement requirements, and private-sector practices. Legal sources such as national law, regional agreements, and international treaties will be examined via different perspectives of individuals, companies, data protection authorities and national government. The regional cross-border data transfer frameworks will be studied as the comparative analysis for the Sino-EU framework. Furthermore, how the regional data transfer mechanism affects the international digital trade and the protection of personal data, as well as the influence on the potential international agreement will be analysed.

Chapter 3 will employ legal-historical analysis to explain the factors in macro-level (e.g. political culture, laws and regulations concerning the market, pros and cons of enforcement) and micro-level (data protection laws and norms) that constitute the transborder data flow regulation in China. There are endemic obstacles and economic considerations that are in the way to eliminate the negative effect of data control effectively. Vague clauses, regulatory norms in the pipelines, and censorship system in

China create the grey area of the authoritarian competence. Furthermore, how regulatory measures affect the digital market and commercial activities of foreign companies in China will also be analysed. Chapter 4 will further analyse the data export regulation under Chinese laws.

Chapter 5 will introduce the midlevel principle theory, and to propose a Sino-EU framework. The first half of Chapter 5 will study the challenges and potentials arise from the differences to pave the way for the discussion of a potential Sino-EU framework. The fundamental basis for protecting personal data and thus to regulate data transfer varies depending on the legal basis it applies and the objectives the norms would like to achieve. Contemporary theorising about regulating data transfer begins a long way from the protection of fundamental rights. Recasting data protection may help in rebalancing the field at the conceptual level, yet may also risk a thoroughly practical concern. Midlevel principles are the principles upon which actual institutions operate. Decisions made within institutions are made on the basis of midlevel principles without any direct reference to deeper or more foundational principles of comprehensive or general application. By discovering the midlevel principles eminent in data protection laws, it will provide observers a better understanding about the doctrines and practices, therefore lead to a potential framework that may reconcile the transborder data regulations between the EU and China.

Chapter 6 will analyse the challenges that technology brought to transborder data regulation, specifically the issues arise in the development of Internet of Health Technology (IoHT), sensitive data-based distributed machine learning algorithms, and

cloud computing under the data protection laws of the EU and China. Moreover, the proposed Sino-EU framework will be applied at the firm level on a case-by-case basis.

In consideration of the challenges that our transborder data regulation mechanism is facing in the digital era, as discussed in Chapter 1 and 2, and the practical and theoretical difficulties when reconciling a transparent cooperative framework for data cross-border transfers between the EU and China, as identified in Chapter 3, 4 and 5, this chapter would conclude with suggestions for the Sino-EU data transfer based on the theoretical and practical discussion as studied in Chapter 6 and 7.

Expected Values and Limitations

This dissertation discussed the issues arising from cross-border data transfer regulation. Part I studied the evolvement of the CBDT rules. It is pointed out that the CBDT regulation is a technology-led phenomenon yet not novel. It is an emerging threat to privacy posed by the development of technology, thus attracted the scrutiny from the public and the authorities. The CBDT regulation reflects the enforcement of national jurisdiction in the cyberspace, which does not enjoy an indisputable general consensus in the contemporary international law. The rule-making of CBDT cannot avoid the controversial debate over the legitimacy of state supervision of the network. CBDT regulation is originated from the protection of personal data in the EU, yet the disagreement with regard to its philosophy is deprived from the conflict of different legislative values, that is, different legislators have different understandings of the freedom of free flow of information and the right to personal information. The author also questioned the rationale of the EU data transfer rules by discussing the target validity of the current rules, that is, the target validity for data protection.

Part II compared the EU and China's data protection laws as well as the CBDT rules respectively. Challenges that CBDT restriction measures might face are listed, since the data transborder transmission is not a legislative measure by nature. In the process of rule-making and implementation existed dual pressures from domestic and abroad, categorised as technological, international legislative and theoretical challenges. Theoretically, Cyberspace does not have a boundary similar to a physical space, the theoretical premise that the EU CBDT rules ignored is that the state must control the transborder transmission of data by setting the borders. Thus, for China, two aspects must to be addressed: is there an independent cyberspace law, and where is the boundary between the virtual and real world. International legislative challenges arise from the oversea data access of the U.S. government. The EU CBDT framework has limited impact when facing such data access under the cover of FISA and CLOUD Act of the U.S. Particularly, this dissertation discussed the potentials for a free flow of data transfer mechanism between the EU and China. It is worthy exploring the possibility for a region-based bilateral collaboration, such as a free trade zone in China, to seek for the EU Commission's recognition of adequate level of protection of personal information. For general data-intensive entities, binding corporate rules and standard contractual clauses are still a preferrable approach.

Part I A Theoretical Study on Cross-Border Data Transfer Regulation

Chapter 1 Theoretical Basis for Regulating CBDT

1.1 Data as the Essential Element of Digital Economy

The global economy is undergoing an information explosion that can "unlock new sources of economic value, provide fresh insights into science and hold governments to account"[24]. The computerised data and information now circulate freely on an international scale, and the volumes of data crossing borders have reached unprecedented levels. The participants in transborder data flows are diverse, such as commercial and non-commercial organisations, individuals and governments, while the data traded or exchanged across national borders varies, such as data and information related to trading activities, intra-corporate flows, infrastructure communication services and scientific and technological exchanges. The benefits that can be derived from transborder data flows are growing, although the ability of countries to reap such benefits may vary.[25] Although it is widely recognised that countries should have a common interest in facilitating transborder data flows, and in reconciling different policy objectives in this field, the implementation of free flow of transborder data remains vague.

---

[24] The Economist, 'Data, data everywhere – A special report on managing information' (2010) <https://www.economist.com/special-report/2010/02/25/data-data-everywhere> accessed 25 February 2019.

[25] OECD, 'Declaration on Transborder Data Flows' (1985) <http://www.oecd.org/sti/ieconomy/declarationontransborderdataflows.htm> accessed 7 March 2019.

The origin of the above-mentioned problem was rooted in the complex context of data. Data has been considered to constitute a new economic asset class and framed with fancy terms such as "the new oil" or "the new currency".[26] In the digital era, those new classes with rich knowledge and large amounts of data will be formed at an accelerated pace will play an increasingly important role.

- *Protecting data as an asset.* The European Commission stated in the proposal for a Digital Content Directive that, digital content, such as music, can be exchanged with 'a price is to be paid or the consumer actively provides counter-performance other than money in the form of personal data or any other data' in the contract.[27] It demonstrated the opinion that consumers may have the same rights when entering into a contract for the supply of digital content paid with both money and data. However, this approach has already been challenged.[28]

- *Investing data as the capital.* It may increase its value, yet the core difference lays between the two: data is intangible and can be replicated, while capital cannot.

---

[26] World Economic Forum, 'Personal Data: The Emergence of a New Asset Class' (2011) <http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf> accessed 25 October 2020. See the speech of Meglena Kuneva at the Roundtable on Online Data Collection, Targeting and Profiling: '*Personal data is the new oil of the internet and the new currency of the digital world*' (2009) <http://europa.eu/rapid/press-release_SPEECH-09-156_en.htm> accessed 25 October 2020.

[27] Article 2(1), Article 3(1) and Recital 13 of the proposal for a Directive of the European Parliament and of the Council on certain aspects concerning contracts for the supply of digital content (proposal for a Digital Content Directive), (2015), COM(2015) 634 final.

[28] To protect data merely as an asset only presents the contemporary value of the data but overlooked the fact that the value of data would gradually depreciate over time. This is derived from the value of capital. In fact, the natural increase of value in data is rare, but devaluation often occurs. For example, when a secret data is disclosed, its value no longer exists.

- *Protecting data under intellectual property rights*. The propertisation of data triggered intense debates around the world, with significant divergence.[29] Under European copyright and sui generis data protection paradigm, certain limitations are naturally conflict with the development of technologies.

The emerging theory of data as a new factor of production brought new perspectives. Data has been largely recognised as an essential factor of production, together with other four: land, labour, capital and organisation defined in *Principles of Economics* by Alfred Marshall.[30] Especially, the exponential growth of big data – high volumes, high velocity and high variety information assets – provide massive datasets enhanced insights and innovative forms of information processing.

1.2 Data Sovereignty

Digital globalisation is playing an essential role in promoting the economic and social development of all countries, yet producing a realistic dissembling effect on the national sovereign and jurisdiction.[31] To some extent, globalisation could lead to the consequence that national sovereignty being squandered by multinational titans' power, policy orientation, identification and network.[32]

In the 1990s, when the Internet just started the booming, the theory of national sovereignty was not a natural choice for scholars. Back then, David Johnson, David

[29] For examples, the European Union for multiple times demonstrated disagreement against the propertisation of data.

[30] Alfred Marshall, *Principles of Economics* (8th edn, Macmillan 1920).

[31] Lawrence Lessig, *Code 2.0* (2nd edn, Basic Books 2006).

[32] Ulrich Beck, *What Is Globalization?* (Polity 2000).

Post, Lance Rose, Joel Reidenberg and Henry Perrett and other legal scholars had responded to the claims of Internet technology pioneers and supported the theory of Cyberspace as Sovereignty. In this unfamiliar domain, what the scholars firstly touched is not its commonality with the application of the sovereignty principle to the traditional field, but the sense of distance brought about by the lack of rational understanding. With the deepening of the influence of the Internet in the real world, the academia began to describe the Internet from "virtual space" to "heterotopia", and from "heterotopia" to "mass space". Until then, national sovereignty's existence in cyberspace was generally recognised.

Data sovereignty is the inheritance and the expansion of national sovereignty, information sovereignty and cyber sovereignty. There is no conventionally adopted definition of what is data sovereignty. Data sovereignty is believed to be the result of the stimulation of emerging Internet applications and objective features embedded in the Internet. Its most iconic expression is the geographical separation of data owners, users and storage provider and the resulting rights, power identifications and effective enforcement.[33] The concept of data sovereignty can be divided into two categories: *i*) general data sovereignty, which includes national data sovereignty and personal data sovereignty; and *ii*) special data sovereignty, which refers only to national data sovereignty.[34] National data sovereignty is the premise of the enforcement of personal data

[33] Zachary Peterson et al., 'A Position Paper on Data Sovereignty: The Importance of Geolocating Data in the Cloud' (2011) Proceedings of the 3rd USENIX Conference on Hot Topics in Cloud Computing.

[34] Kristina Irion, 'Government Cloud Computing and National Data Sovereignty' (2012) 4 Policy & Internet.

sovereignty while relying on the support and expression of personal sovereignty. Personal data sovereignty can be regarded as the right of personal data, which refers to the data subjects' right to self-determination and self-control over her data.[35] While data sovereignty can also be divided into data management rights and data control rights, it does not mean full control over the transnational flow of data. Rather, it is important to seek a reasonable balance between legitimate rights and free flows. This view clarifies the application scenarios and implementation principles of the data sovereignty concept.

Although some scholars believe that data sovereignty is "only an illusion"[36], from the perspective of the development of sovereignty theory as well as the internal logic of technological change, the emerging of data sovereignty is inevitable. However, to reach data sovereignty is complex. With respect to other types of sovereign object-matter such as territories and population, there exist fundamental differences:

- *The power relations in data sovereignty are interdependent.* Data sovereignty in any country is facing constraints from both vertical and horizontal aspects. Horizontal aspect refers to the power of a country's cyberspace and the power relations between other countries. Vertical aspects refers to the cyberspace power relationships among the state and the supranational, sub-national, and individuals.[37] Unilaterally emphasising the absolute control over the country's

[35] It is believed that data sovereignty discussed by most European scholars refers to the sovereignty of personal data, while Chinese scholars emphasize the sovereignty of national data.

[36] Jonathan Obar, Big Data and the Phantom Public: Walter Lippmann and the Fallacy of Data Privacy Self-Management' (2015) Big Data & Society.

[37] Zachary Peterson et al. (n).

data resources may lead to stagnant data flows and fragmentation of cyberspace, thus erode national data sovereignty.

- *The boundaries of data sovereign jurisdiction are overlapping.* Large multinational data centres are distributed globally. This leads to the data subject, physical platforms and the business operator belonging to different jurisdictions. Due to the lack of international law and the differences between countries, data sovereignty jurisdiction boundaries are overlapping and conflicting.[38]
- *The data sovereign powers are asymmetric*. Advanced technology and innovative business model create ambiguousness in defining the nature and the ownership of data. Massive, heterogeneous, real-time transborder data flows are challenging the jurisdiction of data sovereignty. Powerful data markets dominate cyberspace and the technology industry, forming a realistic suppression upon other countries their capability to guarantee the data sovereignty. The imbalance is still evident and may ever grow.

1.3 Public Interest

Reviewing contemporary CBDT rules and frameworks, it is unexceptionally addressed special restrictions and prohibitions for specific industries or specific types of data that a country's government can formulate based on the protection of domestic public interests. However, the "public interest" in the rules is highly generalised, giving a government great discretion with limited restrictions. On the other hand, with the development of e-government, industry requirements, scientific research progress, and

[38] Imagine the scenario where American company A stored the data collected from German citizen B in the data center C located in Ireland, all three countries may have the jurisdiction over such data.

global welfare improvements, it is in need to circumvent the existing rules that restrict or prohibit data flow in order to practice the sharing and exchange of data.

Recital 4 of the GDPR covers "public interest" under those that are "in relation to its function in society" and other fundamental rights in addition to the right to protect personal data. Recital 10 allows the Member States to maintain or introduce national provisions based on "public interest". Article 6.2 of the APEC Cross Border Privacy Enforcement Arrangement (CPEA) exempts its member economy the 'governmental activities authorised by law when taken to protect security, public safety, sovereignty or other public policy' from obligations.

Public interest is an important part of the daily language in politics, law and economics, but there is no consensus on its context.[39] In legal studies, public interest generally refers to the interests of unspecified members of society. The principle of public interest is presented in various international treaties and national laws. For example, Article XX of the General Agreement on Tariffs and Trade (GATT 1947) provided a general exception of the protection of public morals from obligations. Article XIV of the General Agreement on Trade in Services (GATS) provided a general exception of 'protect public morals or to maintain public order' from obligations. Article 7 of TRIPS recognises the balance between the protection and enforcement of intellectual property rights with public interest considerations, while Article 8 allows members to adopt measures necessary to promote the public interest, including to protect public health.

[39] Mike Feintuck, *Media Regulation, Public Interest and the Law* (Edinburgh University Press 2006) 2nd Edition.

Article 1of China's Cybersecurity Law defined the legislative objective as to "safeguard public interest". Under certain circumstances, public interest often can be interchangeable with terms such as public moral, public order, public welfare and social welfare.

It is noted that, against different societal backgrounds, interpretation of public interest may differ. According to Roscoe Pound's theory of social interests, in order to evaluate the conflicting interests in due order of priority, every society has certain basic assumption upon which its ordering rests, through for most of the time they may be implicit rather than expressly formulated. Three kinds of interests are listed, i.e. individual, public and social. Public interest is referred to as requests, demands, and orders from the standpoint of political life, the life of an organised political society. Simply put, it is in general the national interest. Is there any individual or public interests of greater weight than a conflicting social interest? Where the standpoint applies? These are questions scholars are trying to understand. In China, Pound's concept of public interest is not commonly accepted. Apparently, public interest involves the interests of an unspecified majority of members of such society, rather than all.[40]

The rights citizens and interests of commons are showing an ever-expanding trend. Similar to national security, the context of public order is constantly expanding, demonstrated that the concept has appeared in more and more different types of legislation. In CBDT rules, there are inevitably "public interest" provisions or exception

[40] See Zejian Wang, Civil Code General Principles (民法总则) (2014); Liming Wang, The Hierarchy of Interests in Civil Law and Its Consideration (民法上的利益位阶及其考量) (2014) The Jurist Volum 1.

clauses. Therefore, to define "public interest" in CBDT is a prerequisite for the orderly flow of data.

The author strongly advocates the protection of public interests as the basic legal principle, rather than just a general exception. The processing of personal data should be designed to serve the people. The right to protect personal data is not an absolute right, and shall meet an objective of public interest and be proportionate to the legitimate aim pursued.[41] Since public interest has legal uncertainty, it requires law enforcement agencies to exercise discretion in its application. In order to prevent the abuse of discretion by law enforcement agencies, administrative power should be restricted through the principle of proportionality. For example, Article 4 of the GDPR clearly stipulates that all fundamental rights must be balanced in accordance with the principle of proportionality to ensure the free flow of cross-border data. Among the U.S.-led treaties, the principle of proportionality is more inclined to restrict cross-border data transfer to the lowest extent. This is believed that the U.S. favours commercial freedom and commercial rights when weighing among various social interests based on its own economic and technological advantages. Therefore, different countries and regions have significant differences when applying the proportionality principle. Nevertheless, it is still an effective means to restrict the administrative intervene from abusing their discretion and to safeguard legal authority.

[41] Recital 4 of GDPR.

1.4 Cybersecurity

Restrictions on data flows may not only impede digital trade by entities operating in key overseas markets, but also introduce vulnerabilities that increase the risk of cybercrime and data breaches.[42] These measures are often implemented for security or cybersecurity reasons. However, measures requiring data localisation, source code disclosure, and encryption restrictions may actually increase cybersecurity risks and their associated costs. While there is currently no international framework for cybersecurity law, the Budapest Convention on Cybercrime (2001) established a multilateral standard for national cybercrime laws and enforcement.[43] Signatories agree to a certain level of domestic enforcement, including prosecuting cyber crimes committed in their territories.

Companies employ a series of technical and non-technical controls at the firm level to identify threats, defend against attacks, and respond to network intrusions.[44] Basic technical techniques include deploying firewalls and intrusion detection systems and using cryptography to transmit sensitive information securely and privately. Non-technical controls consist of policies and procedures, such as adherence to a patch management policy. Measures that are often proposed and enacted for cybersecurity or national security reasons may restrict the use of cryptography or require the firms to disclose source code.[45] Cryptography restrictions can impede the transborder data

[42] European Council, ETS 185 - Convention on Cybercrime, 23.XI.2001.
[43] *Ibid*.
[44] See e.g. Information Security Management Systems, ISO/IEC 27001:2013.

[45] USITC, Global Digital Trade I (n).

flows of both encrypted data and physical goods that enable cryptography. Disclosure of source code may lead to trade secret concerns.

1.5 Censorship

The outright blocking or filtering by governments in some countries of their Internet platforms and content is the most direct measure impeding digital trade. Instances of government-mandated disruptions to digital networks or particular digital apps or services are frequently justified on the grounds of maintaining public order, ensuring national security, or protecting local businesses. These have increased sharply in the last decade.[46] Often, developed countries do not block or filter Internet content or applications, although specific exceptions do exist.[47]

Overall, the blocking of certain types of content is acceptable under several international trade agreements. The WTO's General Agreement on Trade in Services (GATS), for example, allows countries to maintain exemptions to certain obligations in order to protect public morals or maintain public order; to protect human, animal, or plant life or health; or to secure compliance with laws or regulations, including measures to prevent deceptive or fraudulent practices.[48] However, incidents of censorship that may fall outside of these exceptions are becoming increasingly common. China, for example, blocks and filters Internet content using a highly advanced censorship apparatus

[46] The countries where the costs to local economy of such Internet shutdowns are likely to have been largest were India, Saudi Arabia, Morocco, and Iraq.

[47] For example, several European countries, including France, Italy, and the UK block websites that promote terrorism or contain certain types of adult content.

[48] World Trade Organization, 'The General Agreement on Trade in Services (GATS)' (1995) <https://www.wto.org/english/docs_e/legal_e/26-gats_01_e.htm> accessed 7 March 2019.

that employs a sophisticated technical platform, the so-called "Great Firewall" of China. Some countries restrict the cross-border transmission of certain data to maintain public order or public morals. For example, Singapore, Lebanon, and Turkey prohibit the import of adult entertainment websites to protect their own public morals. In the negotiation of the General Agreement on Tariffs and Trade, some European countries advocate "cultural exceptions," restricting digital cultural products such as movies, music, and games, in order to protect their own culture from external impact.

Chapter 2 Global Policy for CBDT Regulation

2.1 CBDT Regulation Mechanisms

With the interaction of data globalisation and data sovereignty, the transborder data flow is becoming the focus of data protection regulation in various countries.[49] Due to the differences which lay in digital economic development levels/models, legal system origins, and data sovereignty objectives, global transborder data regulations show observable divergences and trade-offs. For example, the US-led Tran-Pacific Partnership (TPP) agreement proposes free flow of data between member states[50]; the EU data protection reform further expands the scope of application of the law by revoking the Safe Harbor Agreement that has been implemented for 20 years[51], and implements a

[49] There is a lack of clarity as to the meaning of the term "transborder data flow" even inside one jurisdiction, and often regulatory instruments use different definitions to apply the measures. The EU General Data Protection Regulation (GDPR) refers to "transfer to a third country of personal data" (recital 153) without defining "data transfer"; the APEC Privacy Framework variously uses the terms "international transfer", "information flows across borders", cross-border information flow" and "cross-border data transfer" interchangeably to refer to the movement of personal data across national borders. The OECD Privacy Guidelines refer to "transborder data flows", defining the term as "movements of personal data across national borders" (Section 1(c)), while the Council of Europe Convention 108 refers to "transborder flows of personal data", defined as "the transfer across national borders, by whatever medium, of personal data undergoing automatic processing or collected with a view to their being automatically processed" (Article 12(1)). It is also unclear whether merely making personal data accessible should be considered to result in such a transfer, or whether this requires some active or automatic transmission of the data (*see* Case C-101/01 *Bodil Lindqvist v Åklagarkammaren i Jönköping* [2003] ECR I-12971).

[50] Article 14.11 Trans-Pacific Partnership.

[51] Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECLI:EU:C:2015:650.

new Privacy Shield Framework[52]; Russia enacts the law requiring data to be locally stored[53].

2.1.1 General Principles

In general, the regulations of external cross-border data flows are justified with the following mechanisms and corresponding principles:

(1) *To improve the level of protection of personal data after leaving the country, and to establish a system for transborder data flow regulation.* For example, some countries have adopted the adequacy protection approach to determine the exemption of the restricted cross-border data transfer.[54] Other approaches include standard and ad hoc contractual obligations for the privacy and security control over cross-border data transfer, or self-regulation and self-certificate mechanism operated by enterprises to safeguard the intra-multinational-entity data transfer.[55]

---

[52] Commission Implementing Decision (EU) 2016/1250 of 12 July 2016 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequacy of the protection provided by the EU-U.S. Privacy Shield (notified under document C(2016) 4176) [2016] OJ L 207.

[53] Bryan Cave LLP, 'Russia Data Localization Requirements at a Glance' (2015) <https://www.lexology.com/library/detail.aspx?g=38047e5e-e4e8-4119-8934-db7773011389> accessed 5 March 2019.

[54] The adequacy protection decision refers to the data export region has the power to determine whether the receipt region offers an adequate level of data protection. The effect of such decision is that personal data can flow outside of the original region without any further safeguard being necessary. Often, such decisions are issued in term of a "white list". For example, the European Union has recognized twelve jurisdictions as providing adequate protection.

[55] For example, the European Union provides standard contractual clauses and the Binding Corporate Rules.

(2) *To provide freedom of cross-border data flow by statutory exemptions and strict user authorisation*. Most jurisdictions having a unified personal information protection mechanism also offer exception clauses that are related to the restriction of cross-border data transfers, allowing the flow across the border in specific scenarios. Data users can apply the exception clause for cross-border data transfer without prior approval from the regulatory authorities. In such a case, the consent from data subjects is the basic precondition for a legitimate transfer.[56]

(3) *To drive the development of digital economy with the very priority to give the free flow of data within the region*. With the rise of the data economy, to promote the free flow of data is becoming an important part of bilateral and multilateral international trade negotiations. TPP agreement, as an example, proposes to eliminate the trade barriers and to prevent data localisation from practising "digital Protectionism" among its contracting parties.[57] It is also an effective measure to provide flexible norms for the two regions with different levels of data protection.[58]

---

[56] Many countries have revised their laws to raise the standards of "clear inform" and "explicit consent" of the data subject.

[57] Tran-Pacific Partnership, 'Promoting Digital Trade' (2015) <https://ustr.gov/TPP/#promoting-digital-trade> accessed 7 March 2019.

[58] For example, the validity of EU-U.S. cross-border data transfers turns on whether the transfers from the European Union are conducted pursuant to a privacy regime commensurate with, but not necessarily identical to, those provided in the European Union. However, such bilateral normative measure is unlikely to work if two regions have fundamentally different approaches to data protection laws. Consequently, because transfers are no longer an option, localizing data will become the norm rather than the exception. *See* H. Jacqueline Brehmer, 'Data Localization: the Unintended Consequences of Privacy Litigation' (2018) 67 American University Law Review 3.

(4) *To grasp the extraterritorial influence of domestic laws as a means to dominate the initiative of transborder data regulation*. In the case of the European Union, the EU Data Protection Directive[59] enacted in 1995 applies the principle of territoriality, that is, the establishment of an organisation in the EU or the processing of personal data through the equipment within the EU will fall into the scope of this Directive. Yet, the EU General Data Protection Regulation (GDPR)[60] enacted in 2016 revised this principle and expanded the scope of the application of this regulation.[61]

2.1.2 Applicable Laws

The law governing cross-border data transfers can be categorised into two groups: *i*) commercial data transfers, and *ii*) law enforcement transfers. In the case of EU-U.S. commercial data transfers, the adequacy of the safeguards surrounding the transfer was first challenged in *Schrems I*.[62] In contrast, law enforcement access to user information stored in foreign jurisdictions is governed by Mutual Legal Assistance Treaties

---

[59] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L 281, 23/11/1995.

[60] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119, 4.5.2016.

[61] Some researchers believes that there should be the distinction in data protection law between "extra-territoriality in scope" and "extraterritoriality in effect", and concludes that the provisions of Chapter V of the GDPR are not extraterritorial in scope because they do not apply directly to entities outside the EU. *See* Yves Poullet, 'Transborder Data Flows and Extraterritoriality: the European Position' (2007) 2 Journal of International Commercial Law and Technology 141. However, Chapter V of the GDPR impliedly refer to obligations of data controllers by setting forth the legal bases by which personal data may be transferred to third countries. Even though entities in third countries are not explicitly mentioned as addressees of Chapter V of the GDPR, their processing of personal data transferred from EU still falls within the scope of EU law, since the data transfer mechanisms contained in these articles are based on the application of EU legal protections. *See* Christopher Kuner, 'Extraterritoriality and International Data Transfers in EU Data Protection Law' (2015) University of Cambridge Faculty of Law Legal Studies Research Paper Series No. 49/2015.

[62] *Schrems v, Data Protection Commissioner* (n).

(MLATs). It is the failure of both of these mechanisms to ensure safe, effective, and efficient transfer of user information.

Commercial data transfer

To date, only a few countries have received an adequacy decision from the European Commission based upon the country's domestic laws. Unsurprisingly, the U.S. is not one on the list. Hence, U.S. companies processing EU users' personal information must transfer under an international commitment to use adequate safeguards, contractual clauses, and one of the derogations listed in Article 26 of the Data Protection Directive.

Article 25 of the Data Protection Directive is based entirely on the principle that the transfer of personal information to a third country cannot take place unless that third country guarantees an adequate level of protection of such data.[63] Prior to *Schrems I* decision, U.S. entities under the jurisdiction of the U.S. Federal Trade Commission or the Department of Transportation could transfer EU user data inside or outside the European Union based on the Safe Harbor Framework.[64] The Safe Harbor Framework was believed to be a direct response to the passage of the Directive, and was designed to limit the negative impact of the inherent differences between the EU and the U.S. approaches to personal data protection and international trade.[65]

---

[63] ibid, para 139.

[64] Commission Decision of 26 July 2000 Pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequacy protection provided by the Safe Harbor Privacy Principles and related frequently asked questions issued by the US Department of Commerce [2000] OJ L215 25/08/2000.

[65] The Safe Harbor Framework is an example of an Article 25 international commitment, negotiated by EU-U.S. officials. U.S. companies processing EU user data were able to certify compliance with the principles and transfer data under the Framework. *See* Congressional Research Service, 'U.S.-EU Data

In 2014, being the direct response to the Snowden revelations, Maximillian Schrems filed a case before the Irish Data Protection Agency against Facebook Irish subsidiary arguing that, U.S. law failed to provide adequate protection against U.S. mass surveillance.[66] This lead to the revocation of the Safe Harbor as a valid mechanism for transfers between the EU and the U.S. by the Court of Justice of European Union (CJEU).[67]

Following the invalidation of the Framework in *Schrems I*, the decision had several impacts on EU-US relations. The immediate consequence was that all data transfers from the U.S. to the EU under the Safe Harbor regime were now in violation of the European law. Companies were allowed to use standard contractual clauses or other derogations as an alternative transfer mechanism, until the EU and the U.S. successfully negotiated the Privacy Shield.[68] After receiving wide critics, the EU Commission's adequacy determination for the Privacy Shield was rendered.[69]

Law enforcement access

Privacy: From Safe Harbor to Privacy Shield' (2016) <https://fas.org/sgp/crs/misc/R44257.pdf> accessed 3 March 2019.

[66] *Schrems v, Data Protection Commissioner* (n).

[67] ibid. The CJEU found that the U.S. government permitted generalized access to electronic information and failed to provide redress mechanisms. Therefore, the CJEU determined that the U.S. law did not provide an adequate level of protection essentially equivalent to EU laws.

[68] Communication from the Commission to the European Parliament and the Council on the Transfer of Personal Data from the EU to the United States of America under Directive 95/46/EC following the Judgment by the Court of Justice in Case C-362/14 (Schrems) [2015] COM/2015/0566 final.

[69] Digital Rights Ireland brought the first challenge on 2016, seeking the annulment of the determination on the basis that the Shield failed to provide sufficient substantive changes from the Safe Harbor Framework. This challenge was dismissed for lack of admissibility. French advocacy group La Quadrature du Net also challenged the Commission's decision arguing that the Shield not only continues to violate the Charter, but also fails to provide effective redress mechanisms. This case remains pending.

In criminal matters in general, mutual legal assistance (MLA) instruments are used for cross-border cooperation for the purpose of gathering and exchanging information. In force since May 2017, the European Investigative Order (EIO) Directive is the over-arching EU tool for improving MLA at EU level and simplifying the work of judicial authorities wishing to obtain evidence located in another EU country.[70] As regard to specific areas of cross-border investigations, the European Union is facilitated with multiple legislative instruments to access financial data for anti-money laundering and terrorist financing, to trace and identify the proceeds of crime, and to exchange data with third countries based on bilateral agreements.[71]

2.1.3 Data Localisation

Generally, countries maintain three primary justifications for implementing data localisation regulations. First, some countries view localisation as critical to protecting their respective citizens from foreign surveillance.[72] Second, others justify localisation because it benefits their domestic law enforcement by increasing the accessibility of

[70] EPRS, 'Law Enforcement Access to Financial Data' (2018) <http://www.europarl.europa.eu/RegData/etudes/BRIE/2018/615665/EPRS_BRI(2018)615665_EN.pdf> accessed 3 March 2019.

[71] ibid.

[72] The propensity for the EU to pass data localization laws may have stemmed from the 2013 leak of classified U.S. surveillance documents by National Security Agency (NSA) employee Edward Snowden. See Christopher Kuner, 'Reality and Illusion in EU Data Transfer Regulation Post Schrems' (2017) 18 German Law Journal 4.

user data through local legal processes.[73] Third, data localisation also has a protectionist motive, and countries have used it as a means to bolster domestic markets.[74]

Data localisation measures differ from country to country in terms of industry coverage, geography, types of data covered, complexity, data intensity, and economic impact, among other factors. Accordingly, the measures can be categorised in multiple ways. One observer groups them into four main categories, from most to least stringent: *i*) geographical restrictions on data export; *ii*) geographical restrictions on data location; *iii*) permission-based regulations; and *iv*) standards-based regulations.[75] Another observer places them into two categories: *i*) strict data localisation measures, and *ii*) conditional flow regimes. Strict data localisation measures may require local storage (collecting data on local servers); local storage and processing (collecting and manipulating data to produce meaningful information on local servers); or local storage, processing, and access (thus banning data transfers). Under conditional flow regimes, certain conditions need to be fulfilled for data to leave the implementing jurisdiction, effectively banning the transfer of data. These regimes can be so restrictive as to cause a *de facto* ban on the transfer of specific data.[76]

---

[73] Jennifer Daskal, 'Law Enforcement Access to Data Access Borders: The Evolving Security and Rights Issues' (2016) 8 Journal of National Security Law and Policy.

[74] To some extent, data localization can hinder global markets in favor of local markets by barring foreign services access across borders and inviting reciprocal treatment in return. See Anupam Chander, 'Data Nationalism' (2015) 64 Economy Law Journal.

[75] James Kaplan and Kayvayn Rowshankish, 'Addressing the Impact of Data Location Regulation in Financial Services' (2015) GCIG Paper No. 14.

[76] While a small number of countries (such as Russia and China) have introduced broad and explicit data localization policies, a large number, including the European Union, have introduced narrow data localization policies. Although EU member states currently have competence over data localization

Despite the aforementioned purported benefits, data localisation has several negative consequences. One particularly worrisome consequence is the direct financial burden placed on companies and consumers.[77] In 2013, data localisation was predicted to cost cloud computing services between 21.5 billion and 35 billion U.S. dollars by 2016.[78] Further, a long-term financial impact study of data localisation in seven major countries concluded that recently proposed or implemented data localisation legislation substantially impacted the gross domestic products of all seven countries studied, finding welfare losses of 63 billion U.S. dollars in China and 193 billion U.S. dollars in the European Union.[79] It results in a negative impact on consumer welfare as the companies shift the cost of localisation onto end users.

Another concern rises as such localisation requirements could "threaten the major new advances in information technology - not only cloud computing, but also the promise of data analytics and the Internet of Things (IoT)".[80] In the absence of data localisation measures, Internet data are routed across companies' networks through decisions made autonomously and automatically at local routers, which choose paths based largely on

---

issues, the EU has proposed a "Digital Single Market' which could shift that competence to the supranational level. See John Selby, 'Data Localization Laws: Trade Barriers or Legitimate Responses to Cybersecurity Risks, or Both?' (2017) 25 International Journal of Law and Information Technology.

[77] This cost is derived from many expenses including, but not limited to, building data centers, employing new teams, and complying with local regulations.

[78] Daniel Castro, 'How Much Will PRISM Cost the U.S. Cloud Computing Industry?' (2013) <http://www2.itif.org/2013-cloud-computing-costs.pdf> accessed 5 March 2019.

[79] Matthias Bauer et al., 'The Costs of Data Localization: Friendly Fire on Economic Recovery' (2014) ECIPE Occasional Paper No. 3/2014.

[80] Anupam Chander, 'Breaking the Web: Data Localization vs. The Global Internet' (2014) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2407858>.

efficiency and not on country boundaries.[81] Data localisation dramatically alters this fundamental architecture of the Internet.[82]

2.2 Trends and Challenges for Global CBDT

2.2.1 Policies in Focused Jurisdictions

2.2.1.1 The United States

The United States aim to advocate the free flow of personal data across borders and to use the global leadership of the digital industry to dominate the flow of data. The U.S. has a global lead in the information and communication industry and digital economy, which is the prerequisite for it to dominate the global cross-border data flow. When the U.S. formulated the Trans-Pacific Partnership Agreement (TPP), it proposed that 'on the premise of ensuring that the protection of personal information and other legitimate public policy objectives are protected, the free flow of global information and data will be ensured to drive the Internet and digital economy. The establishment of a data centre will not become a precondition for allowing TPP signatory parties to enter

[81] Data localization affects all Internet communication service providers. However, these measures place small firms at a particular disadvantage, as large companies that operate online often benefit from economies of scale, and thus are better able to craft data policies for individual countries. Companies subject to data localization measures need to rely on country-specific cloud centers and servers, increasing the locations where a company stores data and fragmenting global data into country-specific datasets. Additionally, applying data localization measures to the IoT can reduce data security by forcing providers to create new and previously unnecessary data centers, thus exposing data flows to additional potential breach areas. Further compliance with data localization policies may require detours and inefficient routes, creating latency that reduces IoT functionality. See United States International Trade Commission, 'Global Digital Trade 1: Market Opportunities and Key Foreign Trade Restrictions' (2017) <https://www.usitc.gov/publications/332/pub4716_0.pdf> accessed 5 March 2019.

[82] Matthias Bauer et al., 'Tracing the Economic Impact of Regulations on the Free Flow of Data and Data Localization' (2016) GCIG Paper No. 30.

the market, nor is it required to transfer or obtain software source code.'[83] This proposition is a concentrated expression of the cross-border flow of personal data policies in the United States. In the new round of trade negotiations with other countries, the U.S. has advocated the inclusion of the "free flow of data across borders" into the terms of the agreement, in order to break the market access barriers set up by many countries using the cross-border flow of data.

However, the restrictions of the export of important technical data and foreign investment in specific data fields are expanded to curb the development of strategic competitors such as China, and to ensure the United States' global leadership in the field of science and technology. Since the Trump administration vigorously promoted the "America First" trade protectionist policy, the U.S. has actively deployed such control measures as an important means to contain China and other strategic competitors. The U.S. John McCain National Defense Authorization Act for Fiscal Year 2019 (NDAA) updates and reforms the U.S. foreign investment examination and export restrictions on emerging yet fundamental technologies.[84]

Export Control

In terms of export control, according to the Export Administration Regulations (EAR)[85], together with the Export Control Reform Act (ECRA) of 2018, the U.S. export control is not limited to the export of hardware, but also includes specific technical

[83] TPP

[84] NDAA

[85] EAR

data[86], that is, controlled technical data is "transmitted" to locations outside the United States. Server storage or data processing are required to obtain an export license from the Bureau of Industry and Security (BIS) of the Ministry of Commerce. In November 2018, the U.S. BIS released a list of 14 emerging technologies, planning to develop an export management framework for key technologies and related products, such as biotechnology, artificial intelligence and machine learning technology.[87] On October 29, 2018, the U.S. Department of Commerce also included about 90 Chinese companies on the list of companies that violated U.S. national security and diplomatic interests on the grounds that they "posed a significant threat to U.S. national security interests". It is requested that these companies' export, re-export and transfer of American-origin goods, software and technology must comply with additional licensing requirements. The contents of U.S. export control have a large overlap with the Chinese national strategic plan "Made in China 2025", and its policy is primarily aimed at strengthening the blockade of technology exports to China.

Foreign Investment Review

In terms of foreign investment review, the Committee on Foreign Investment in the United States (CFIUS) has the competence to review and restrict a wide range of investment transactions and export transactions when necessary, and establish multiple

[86] The applicable scope of "Technology" is defined as the 'information necessary for the "development," "production," "use," operation, installation, maintenance, repair, overhaul, or refurbishing (or other terms specified in ECCNs on the CCL that control "technology") of an item.' Technology may be in any tangible or intangible form. Part 772 of the EAR.

[87] Bureau of Industry and Security, Commerce, Review of Controls for Certain Emerging Technologies, https://www.federalregister.gov/documents/2018/11/19/2018-25221/review-of-controls-for-certain-emerging-technologies

mechanisms to identify and protect key emerging technologies to ensure the security of the United States.[88] The reformed "Foreign Investment Risk Review Modernization Act of 2018" (FIRRMA) expands the scope of "covered transactions" by including 'as a result of the transaction, could obtain "critical technologies", "critical infrastructure" and companies that foreigners hold or collect sensitive personal data of U.S. citizens.'[89] Non-controlling and non-passive investments are included in the scope of its review as well. At the same time, CFIUS also requires investors to sign a security protocol, namely Mitigation Measures, which specified detailed content on internal security management systems, localisation of products and services, and rights for government auditing to prevent sensitive information, products and services from leaving the country.[90] Although the bill is aimed at the review of all countries' investment in the core high-tech industries in the U.S., the 35-page bill has mentioned China in 15 places, far more than other countries. It specifically requires the Secretary of Commerce of the U.S. to submit reports on "direct investment by Chinese enterprises in the U.S." and "state-owned enterprises investment in the U.S. transportation industry" to Congress

[88] The proposed 14 emerging technologies and related products on the initial list include robotics, artificial intelligence and machine learning technology, and other advanced surveillance technologies. As regards foundational technologies, the BIS has yet to issue an advance notice of proposed rulemaking. The precise scope of the latter remains unclear, but is expectedly likely to include items that 'enable progress and applications in a variety of problem domains', such as semiconductor technologies. See: Bureau of Industry and Security, Commerce, 'Review of Controls for Certain Emerging Technologies' (Federal Register 2018) Proposed Rule.

[89] Sec. 201 of H.R.5841 - Foreign Investment Risk Review Modernization Act of 2018.

[90] From 2013 to 2015, 10 cases (10%) resulted in the use of legally binding mitigation measures. In 2015, CFIUS mitigation measures were applied to 11 different covered transactions (8% of total 2015 transactions). Latham & Watkins LLP, Overview of the CFIUS Process (2017). <https://www.lw.com/thoughtLeadership/overview-CFIUS-process> accessed 6 February 2021.

every two years. Particularly, when CFIUS is conducting a national security assessment, CFIUS reviews can potentially discriminate among investors from certain countries, that whether the covered transaction involves a country of "special concern", and that country has "demonstrated or declared" the acquisition of a type of critical technology as a "strategic goal", which is clearly directed towards China.[91]

Controlled Unclassified Information

In accordance with the requirements of Executive Order No. 13556 signed by former President Obama in 2010, in order to improve the then situation where the government's controlled non-secret information that are scattered across more than 100 separate departments and agencies, the U.S. Archives took the lead in sorting out and standardising the classification and basis of controlled unclassified information (CUI) backed by U.S. laws, regulations, and government. The CUI program lists in-detail 17 categories, including critical infrastructure, defence, financial, immigration. Intelligence, international agreements, law enforcement, legal, natural and cultural resources, NATO, nuclear, patent, privacy, procurement and acquisition, proprietary business information, provisional, statistical, tax, and transportation information. [92]

The listed categories of data can be regarded as "important data" identified by the U.S. government, and stricter management measures are adopted. Meanwhile, the access of CUI's dissemination is graded into seven categories: no foreign dissemina-

[91] Subtitle A of Title XVII, Section 1702(c)(1) of FIRRMA.

[92] National Archives, 'CUI categories' <https://www.archives.gov/cui/registry/category-list>

tion, federal employees only, federal employees and contractors only, no dissemination to contractors, dissemination list controlled, authorised for release to certain nationals only, and display only.[93]

The CUI program defined the scope of the hotly debated issue of important data, which significantly influenced China's law-making over the important data export.

Extraterritorial Jurisdiction

The Clarifying Lawful Overseas Use of Data Act (CLOUD Act) of 2018 ended the dispute over whether U.S. law enforcement agencies have the right to obtain access to user data stored in overseas servers by U.S. companies in the *United States v. Microsoft Corp*.[94] It is indicated that the entity in the "possession, custody, or control" of data is required to provide it as part of ongoing criminal proceedings, regardless of whether this information is stored in the U.S. or abroad. The law expands the power of U.S. law enforcement agencies to access overseas data, at the same time sets a specific path for the U.S. government to sign bilateral treaties with other countries, and allows law enforcement agencies of "qualifying foreign government" to retrieve data stored in the U.S..

The CLOUD Act requires the U.S. Attorney General, together with the Secretary of State to submit a written report to Congress to determine the conditions for

---

[93] National Archives, 'CUI Registry: Limited Dissemination Controls' <https://www.archives.gov/cui/registry/limited-dissemination>

[94] *United States, Petitioner v. Microsoft Corp.*, 584 U.S. ___, 138 S. Ct. 1186 (2018)

determining the "qualifying foreign government" (QFG), that is, whether the foreign government has a legal system that institutes “robust substantive and procedural protections for privacy and civil liberties". The determination is based on credible information and expert opinions, considering several factors:

(1) whether the QFG has sufficient substantive and procedural laws in terms of cybercrime and electronic evidence; whether it has joined the Budapest Cybercrime Convention; or its domestic law is consistent with the basic rules of the Budapest Convention;

(2) whether the QFG demonstrates respect for the rule of law and the principle of nondiscrimination;

(3) observes international human rights obligations or demonstrates respect for international basic human rights, including protecting privacy from illegal interference, right to a fair trial, freedom of speech;

(4) whether there exist the clear legal mandates that how the agencies can collect, retain, use, and share data, as well as effective oversight for the data activities as mentioned above;

(5) whether the QFG demonstrates commitment to promote and protect the global free flow of information and the open, distributed, and interconnected nature of the Internet.

The CLOUD Act sets aside traditional bilateral or multilateral judicial assistance treaties, intensifying current data-related judicial sovereignty conflicts between countries. Its effective implementation depends on the strong international economic and political position of the United States and cooperation with relevant countries.

2.2.1.2 Singapore

Singapore is the fourth largest Internet data centre in the Asia-Pacific region, following Japan, China and India. Through the "Smart Nation" initiatives, Singapore has modernised its information infrastructure and promoted investment in the telecommunication industry and data centres.[95] In the Cloud Readiness Index (CRI) of 2018 released by the Asian Cloud Computing Association (ACCA), Singapore ranked the first. It is ranked at the leading position in the regimes of broadband quality, network security, privacy protection, government supervision, and intellectual property protection, showing its superiority in infrastructure and supervision.[96]

Geographically, Singapore is close to the developed public cloud service markets in the Asia-Pacific region such as Australia, Japan, and South Korea. This is also an important factor in promoting Singapore's strategic goal of building a data centre in the Asia-Pacific region. Singapore has established cross-border data transfer requirements similar to those of the European Union, prohibiting the transfer of data to countries or regions with a lower level of data protection than Singapore. However, under special circumstances, companies can apply for an exemption from the Personal Data Protection Commission.[97] In addition, the commission also provides guidance

[95] In 2016, Singapore's data center revenue was US$730 million, while India's US$1 billion, China's US$2.9 billion, and Japan's US$6.5 billion. According to Broad Group's report 'Datacenter Markets SE Asia 2018', Singapore currently has 22 data center operators managing 46 facilities.

[96] ACCA, Cloud Readiness Index2018 (2018) <http://asiacloudcomputing.org/studies/cri2018/results/> accessed at 7 January 2020.

[97] Article 26(3), Personal Data Protection Act 2012, Republic of Singapore.

for data cross-border transfer contractual clause as a supplement.[98] The flexibilities of the mechanisms have made Singapore a priority for multinational companies to set up data centers in the Asia-Pacific region.

In February 2018, Singapore joined the Cross-Border Privacy Rules system endorsed by the Asia-Pacific Economic Cooperation, and is one of the most important and active participants in APEC-CBPRs, as well as the advocate of regional data free flow. Based on CBPR rules, an APEC economy must demonstrate its compliance with APEC privacy framework, including privacy protection law, privacy protection enforcement authority, privacy certification agencies, among others. Singaporean Personal Data Protection Commission actively explores the certification mechanism to bridge with CBPR. With the said certification, entities operating inside Singapore may be able to freely transfer data with other entities being certified in CBPRs member states.

2.2.1.3 Japan

Japan is the only country that enjoys the recognitions from both the EU and APEC frameworks. Being the most advanced economy in data protection regime in Asia, Japan rectified its Act on the Protection of Personal Information (APPI) as early as 2003, with an amendment in 2015. The globalisation of the digital economy has also prompted Japan to introduce regulations on cross-border data transfers when revising

[98] The Singaporean data protection commission has released data transfer agreements, for which includes model clauses. The Amendment Bill of 2020 provides a new right of data portability on electronic data. Individuals may request a "porting organization" to transmit certain data about them to another organization. The porting organization must have an ongoing relationship with the individual, and have collected or generated such data. The commission has published guidance for data sharing, includes intragroup and third party sharing, with practical nonbinding clauses.

the APPI. Three legal basis for transferring personal data outside of Japan are addressed:

(1) obtained an individual's consent prior to the transfer;

(2) the recipient country has the same level of data protection as Japan and is recognised by the Japanese Personal Information Protection Commission (white-list countries); and

(3) the recipient companies outside of Japan have established a comprehensive system for data protection in accordance with the requirements of the Personal Information Protection Commission (which are consistent with APEC-CBPRs).

Although the cross-border data transfer rules in Japan is very much GDPR alike, its interpretation of the rules is more flexible, providing space for the free flow of data across borders. For example, according to the guidelines issued by the Japanese Personal Information Protection Commission (PIPC), if the recipient of Japan's data transfer to a foreign country can ensure that "appropriate and reasonable methods" are adopted, such data transfer can be allowed. The PIPC guidelines provide some examples, such as business operators in Japan entrust the processing of personal data to foreign business operators through legally binding contractual clauses between the data sender and receiver; or personal information are transferred within the same group in accordance with cooperation rules or privacy policies.

On the one hand, Japan actively follows the US policy of free cross-border data flow by participating in the US-led Trans-Pacific Partnership Agreement (TPP) and

APEC's CBPR rule system. It has become the dominant "comprehensive and advanced" member of the Trans-Pacific Partnership Agreement (CPTPP) after the US withdraws from the TPP. Meanwhile, being a CBPR member, Japan has established a certification authority to provide certificates for companies that are compliance with CBPR rules. Japan's Ministry of Economy, Trade and Industry (METI) is also actively promoting the CBPR system, designating the Japan Association for the Advancement of Information Economy and Information (JIPDEC) as an independent accountability agency required by CBPR to review and certify cross-border data transfer activities of companies.

On the other hand, Japan published Supplementary Rules on the protection of personal data to mitigate the difference from the European Union standard. Specifically, the protection of sensitive data, the data subjects' rights, as well as the further transfer of personal data originated from the EU are emphasised. In return, the EU Commission officially recognised Japan's adequate protection of personal data in 2019, hence a bilateral recognition framework is completed. Further initiatives include a cross-border data free flow framework among the U.S., the EU and Japan, which is on the schedule for negotiation at the G20 summit. It is considered, however, that Japan's commitment to the free flow of data is mainly due to the existed various forms of protectionist policies in the Japanese market.[99] Hence, no other data protectionism policies are needed, such as data localisation.

[99] Japan’s economic policy relies heavily on large companies and corporate alliances (Keiretsu). For example, the Japanese Taxi Association attempted to block Uber from entering Japan. Japanese car market has been the most protected car market, of which over 96% of the market is controlled by the domestic auto spare parts market. See Helen Lui, ‘Japan: Policy Commitment to Free Data Flow with

2.2.1.4 India

India adopted data localisation policy to promote the development of the domestic digital economy. According to the prediction of the Indian technology business community, India's digital economy will reach 4 trillion U.S. dollar by the year 2025. 40 The Indian government's funding for the "Digital India" program reached 48 billion U.S. dollar in 2018-2019. India is seeking to transform the country into a connected economy and conduct extensive research, training and skill development in the fields of robotics, artificial intelligence, digital manufacturing, big data intelligence, and quantum communications, to establish India's global knowledge and digital society dominance. The purpose of implementing data localisation in India is mainly to promote the development of domestic data economy. The preface of the "Draft National Policy Framework for E-commerce in India" clearly stated that India would gradually promote the data localisation policy and require the establishment of data centres. India as well claimed that it is not to implement strict "data protectionism", but free flow of data will not be allowed. Therefore, on the one hand, its data localisation strategy aims to integrate into the trend of data globalisation, and on the other hand, it wants to stimulate the development of India's digital economy. The Draft E-Commerce Framework lists a series of exemptions for data localisation, such as data transfer by start-up companies, internal data transfer by multinational companies, and data transfer based on contracts.

Informal Restrictions' (2018) <https://jsis.washington.edu/news/japan-policy-commitment-free-data-flow-informal-restrictions/#_ftnref16> accessed 30 January 2021.

India enforced the classification scheme of personal data with respectively different data localisation requirements. In the Draft Personal Data Protection Bill of 2018, personal data is categorised into three types: general personal data; sensitive personal data and critical personal data. General personal data and sensitive personal data are required to be stored at least in copies in India, without the restriction to be transferred abroad. Exemptions for data localisation are also provided. Critical personal data can be processed only in a servers or data centre located in India, and it is absolutely prohibited to leave the country.[100] However, the draft did not specify the specific context and scope of "critical personal data".

Indian particularly concerns about finance data control. Payment data is required to be locally stored to promote the banking and finance industries development. The Central Bank of India ruled that all payment companies in India to compulsorily store their data locally before October 2018, despite the wide criticism from the EU and the U.S. companies. Studies believed that such sector-specific compulsory data localisation policy is large because of its low bank penetration rate. Not only do Indian regulators attempt to promote the domestic banking industry, but also seek access to such data for law and tax enforcement.

2.1.2.5 Russia

In 2006, Russia passed the Federal Personal Data Law, but the law has not been strictly enforced. In 2014, Russia released personal data localisation rules, requiring all oper-

[100] Article 40(2) of the Personal Data Protection Bill, 2018, India.

ators that collect and process personal data of Russian citizens to use data centres located in Russia. The law does not restrict the export of personal data, but rules that the data must be stored on a server in Russia at the time of the initial storage. The Russian data localisation policy is mainly serving economic and law enforcement purposes. From an economic point of view, Russia's increasingly weak economy hinders the development of its IT industry. Especially in recent years, the economic sanctions imposed by the U.S. and Europe have negatively influenced the development of Russia's data industry, leading to an oversupply of data centres in Russia.[101] The implementation of data localisation rule has enabled Russia to rapidly develop the big data market and the construction of a large number of data centres by multinational companies.[102] For law enforcement, Russia also attempts to strengthen the government's law enforcement power and control over data through localised data storage. This is reflected in its amendment of anti-terrorism law "Yarovaya's Law." The law requires the organisers of information dissemination on the Internet to retain Russian users' Internet communication data, user data and certain user activity data. Such data shall be stored in Russia for a minimum period of 6 months, and to be disclosed to Russian authorities upon request.

However, Russia is the signatory country of the Convention on the Protection of Individuals in the Automatic Processing of Personal Data (the Convention 108), which gives it a special role in CBDT international collaborations. The Federal Data Protection Law recognized the Convention 108 signatories had provided sufficient protection

[101] Alexey Danilyants, 'State of the Russian Infrastructure Market,' (2017) Datacenter Dynamics.

[102] Ibid.

of personal data. Thus, in principle, data can be freely transferred. Additionally, the Russian supervisory authority Roskomnadzor established a white list of 23 countries that have been recognized as having an adequate level of protection for personal data.

2.2.2 International Collaborations

The benefits that can be derived from cross-border data flows are growing, while the ability of countries to reap such benefits may vary.[103] Although it is widely recognized that countries should have a common interest in facilitating cross-border data flows and reconciling different policy objectives in this field, the implementation of the free flow of cross-border data remains vague. Due to the differences lay in digital economic development, legal systems, and data sovereignty objectives, it is difficult for countries to impose effective regulations on cross-border data transfer through one's own. In contemporary legislations, a trend of preference for establishing one data flow model inside a region within a given group of countries is emerging.

2.2.2.1 Multilateral International Agreement

The Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108) adopted by the Council of Europe in 1981 is the first and up-to-date only binding multilateral international agreement to set standards for the transborder data flows. The early version of Convention 108 provides general principles that require signatory countries not to restrict or impose any special authorisations to prevent the flow of personal data among the member states and aims to

[103] See OECD, 'Declaration on Transborder Data Flows' (1985) <http://www.oecd.org/sti/ieconomy/declarationontransborderdataflows.htm> accessed 7 March 2019.

achieve a greater unity between its members[104]. The Convention 108 was further developed in the Additional Protocol in 2001 to introduce the concept of an "adequate level of protection" for the intended data recipient countries that are not the signatories to Convention 108.[105] Such exporting party is also subject to exceptions where the transfer is in need of individuals' legitimate interests and public interest, or is based on authority-approved contractual clauses.

The Convention 108 is the result of the implementation of the European Convention on Human Rights with regard to privacy protection. It attempts to build consistent data protection principles to safeguard individuals' rights while keeping active exchanges of such personal information across the borders. Be great as it may, the significance of Convention 108 is limited.[106] Although international agreement as an instrument for dealing with modern societal and legal topics is advantageous in terms applicable scope of the rules, enforcement and guidance, its complex and lengthy establishment procedures have slowed down the reaction time to the emerging issues in international community, especially in areas where international consensus has not yet been reached.

---

[104] See Preambles, 'Details of Treaty No. 108: Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data', Council of Europe, www.coe.int/en/web/conventions/full-list/-/conventions/treaty/108.

[105] See 'Details of Treaty No. 181: Additional Protocol to the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, regarding supervisory authorities and transborder data flows', Council of Europe, http://conventions.coe.int/Treaty/EN/Treaties/Html/181.htm.

[106] Limited signatory countries, overbroad content and free applicable scope eliminate the practical performance of the Convention 108. Additionally, International Law Commission listed "protection of personal data in the transborder flow of information" in its long-term working programs as early as 2006, yet fruitless so far. See Report of the International Law Commission Fifty-eighth session (1 May-9 June and 3 July-11August 2006) p.489.

2.2.2.2 Bilateral International Agreement

In view of the latency of the international community's cooperation in the field of cross-border personal data transfer, multiple emerging countries in the digital economy have actively launched bilateral negotiations based on their own development needs. By reaching the bilateral agreement, it is provided with a legal basis for the personal data exchanges between the signatory countries. The EU-U.S. Privacy Shield Framework is an example. In 2014, being the direct response to the Snowden revelations, *Schrems I* case led to the revocation of the Safe Harbor as a valid mechanism for transfers between the EU and the U.S. by the Court of Justice of European Union (CJEU).[107] The EU and the U.S. successfully reached the Privacy Shield Framework as the alternative, putting forward more stringent and descriptive data transfer requirements for data controllers.[108] After receiving wide critics, the EU Commission's adequacy determination for the Privacy Shield was rendered.[109] American companies may be permitted to acquire personal data from a total of 28 European countries after being registered under the Privacy Shield program and demonstrated that they fulfil the "ad-

[107] The CJEU found that the U.S. government permitted generalized access to electronic information and failed to provide redress mechanisms. Therefore, the CJEU determined that the U.S. law did not provide an adequate level of protection essentially equivalent to EU laws. See *Schrems v, Data Protection Commissioner*.

[108] Communication from the Commission to the European Parliament and the Council on the Transfer of Personal Data from the EU to the United States of America under Directive 95/46/EC following the Judgment by the Court of Justice in Case C-362/14 (Schrems) [2015] COM/2015/0566 final.

[109] Digital Rights Ireland brought the first challenge on 2016, seeking the annulment of the determination on the basis that the Shield failed to provide sufficient substantive changes from the Safe Harbor Framework. This challenge was dismissed for lack of admissibility. French advocacy group La Quadrature du Net also challenged the Commission’s decision arguing that the Shield not only continues to violate the Charter, but also fails to provide effective redress mechanisms. This case remains pending.

equacy protection" requirement by self-certifying procedures. Privacy Shield Framework additionally includes verification, assessment and supervision mechanisms, as well as special rules related to arbitration procedures.[110] The bilateral agreement allows two countries to make more detailed arrangements for cross-border data transfer issues. It is advantageous in terms of negotiation efficiency and enforcement, as well as the flexibility of contents. Yet, its scope of application is limited to the jurisdictions of the two countries. For the establishment of a regional framework of personal data cross-border transfer, a bilateral agreement has very limited effect on bridging different legal standards.

2.2.2.3 Soft Laws

Soft laws often play important roles in encouraging reluctant states to consider and eventually agree upon policies and strategies in areas where serious differences exist. Many international organisations have issued soft laws to regulate cross-border transfer of personal data, which has given certain guidance to the national legislation and implementation. The OECD Privacy Guidelines released in 1980 serve as the first internationally agreed upon set of personal information protection principles and focus on balancing between the needs for digital economy and the protection of individual's rights. It addressed the needs for greater efforts to tackle the global dimension of privacy through improved interoperability and provided the member states with a basic framework of free flow of personal data for further negotiations. The APEC framework, published by the Asia-Pacific Economic Cooperation in 2004, is a framework

[110] Similarly, the U.S. also reached Swiss-U.S. Privacy Shield Framework with Switzerland.

to protect privacy while enable regional personal information transfers to promote consumer trust and business confidence, to lighten compliance burdens and booster digital economy. The data controllers' obligations are particularly emphasised as data subject's consent is mandatory prior to the transfer of the personal information, and the adequate level of data protection shall be guaranteed. This framework is used as a basis for the APEC Cross-Border Privacy Rules ("CBPR"). The U.S.-led CBPR system comprises Privacy Enforcement Authority, privacy certification institutions and recognized entities operating upon nine general privacy principles and a bundle of practical requirements. A joint APEC-EU working team attempts to discover more opportunities for "double compliance" via EU BCR and APEC CBPR referential.[111]

2.2.2.4 Non-Binding Guidance

The level of economic and social development of the ASEAN member states varies, but they reached a landmark regional declaration on data protection in 2012. The ASEAN Ministerial Conference adopted the ASEAN Framework on Personal Data Protection, which established a series of principles to guide the data protection practices of the member states and regions.[112] The ASEAN Data Protection Framework

[111] The Referential for Requirements for Binding Corporate Rules (BCR) and APEC Cross Border Privacy Rules system serves as an informal checklist for companies to apply certifications under the BCR and CBPR system. The referential outlines common compliance requirements and ad hoc requirements for each of the systems. Although the referential was superseded after the enactment of the GDPR in 2018, EU representatives have continued to express a strong interest in developing a work plan for future efforts. See Article 29 Data Protection Working Party, Opinion 02/2014 on a referential for requirements for Binding Corporate Rules submitted to national Data Protection Authorities in the EU and Cross Border Privacy Rules submitted to APEC CBPR Accountability Agents.

[112] ASEAN, 'Framework on Personal Data Protection' (2016) <http://asean.org/storage/2012/05/10-ASEAN-Framework-on-PDP.pdf> accessed 30 January 2021.

aims to promote regional integration and cooperation, and to establish a safe, sustainable, transformative and upgraded digital economy in ASEAN. To achieve this goal, strengthening personal data protection and promoting digital trade and information flow among ASEAN member states is essential. The ASEAN data protection framework is designed with great flexibilities to adapt to the different levels of member states in data and privacy protection supervision. Economies that apply this framework at the level of member states can take exceptional measures applicable to their national conditions. Notably, the framework is not domestically and internationally binding.

The Southern African Development Community (SADC) developed the Model Law on Data Protection in 2010 containing general data protection principles for cross-border data transfer, as another example. Notwithstanding the efforts, many African countries are still struggling with enacting laws to regulate the collecting and processing of personal information. The organisations' practices stopped at proposing a broad framework of guidance. Further discussions over effective solutions to the conflicts of applicable laws of personal data transborder regulation are needed. However, international negotiations and corporations are worthy of recognition.

2.2.3 Trends in CBDT Policy Makings

2.2.3.1 Important Data based on National Security Concerns is the Core of CBDT

The influence of geopolitics on cross-border data flow policies will further increase, and important and sensitive data with national security concerns at the core will become the focus of restrictions on cross-border flows. After the Snowden incident, the destruction of the rules of the international trading system by national security in the cyberspace has been spreading globally. Concepts such as data localisation policies,

cybersecurity evaluation, and autonomous governance have emerged. As the competition between China and the United States in the high-tech field has evolved into a technological cold war, "important sensitive data" with "national security" concerns at the core has also become the focus of cross-border flow restrictions.

Trump administration successively issued a series of documents such as the National Security Strategy, the 301 Report, the Foreign Investment Risk Review Modernization Act, and the Reform of the U.S. Foreign Investment Review Board, to implement the "strategic encirclement and suppression" of China's high-tech development. Recently, it has imposed export restrictions on more than 40 state-owned enterprises in China. Together with a series of export restrictions on high-tech products originally set on China, the U.S. created a barrier for China from both investment and export. The Sino-US trade war has provided a united domestic environment for the U.S. to introduce the above-mentioned bills and policies. The two parties and American society have reached a high degree of consensus on this issue. Regardless of how the trade war will be ultimately resolved, the U.S. exercises control over China in the field of cutting-edge and basic technologies. It restricts the cross-border transfer of large amounts of technical data and sensitive personal data through long-arm jurisdiction and strong intelligence and law enforcement capabilities. Additionally, the United States' strong propositions in the domain will inevitably affect its strategic allies' technology transfer and data cross-border flow strategies to China. The strengthened value orientation of the data cross-border flow policy with national security as the main consideration will further damage the existing business and trade rules and hinder the development of global digital trade.

2.2.3.2 Digital Economic Competitiveness Decide CBDT Strategy

In addition to the consideration of national security and geopolitical factors, the choice of policy paths for cross-border data flow in various countries is also greatly restricted by whether their industrial capabilities and economic development status can control data flow.

One is an aggressive strategy represented by the United States. From the above analysis of the global digital economy development and industrial competition situation, the U.S. is a global leader in the competitiveness of the digital economy industry. Its strategy is to actively advocate the free flow of data to prevent and eliminate digital trade barriers. It emphasises the importance of data flow to economic growth, the value of the free flow of data across borders, the cost of data localisation, and the avoidance of unnecessary security measures. Therefore, the U.S. has actively participated in a series of trade negotiations (TPP, TIPP, NAFTA) to promote the free flow of data in e-commerce.

The second is a regulatory strategy represented by the European Union. The EU's overall industrial capabilities in the digital economy are weaker than those of the U.S. and China. Although it recognises many liberal objectives, it is subject to the reality of industrial competitiveness, and it relies more on its ability to make rules and pass high-level data protection requirements. Strong regulatory capabilities will increase the cost of cross-border data flow for foreign companies, and build protectionism barriers for the development of the digital industry in the region.

The third is the export restriction strategy represented by Russia. Due to the lack of competitiveness of the country's digital economy industry and concerns about

data loss that will damage their industrial development and national security, Russia, Indonesia, India and other countries tend to adopt protectionist measures such as restrictions on cross-border data flow and localised storage of data. Although the requirements for data localisation in different countries vary, such measures are largely used as a market entering barrier to protect domestic industries' development, especially cloud computing and other infrastructure. As a result, it damaged the monopoly of the U.S.-led cloud computing industry to a certain extent.

2.2.3.3 Different Mechanisms for General Personal Information and Important Data

The regulation of the cross-border flow of personal data is based on corporate self-discipline, backed by government supervision as the guarantee for implementation. The main purpose of the regulation of the cross-border flow of personal data is based on the protection of personal information, including the consent of the data subject, the protection of the rights and interests of the data subject, the contract between the domestic data transferer and the overseas data recipient, the country and region where the data recipient is located, and the adequacy of data protection mechanisms. The regulatory measures mainly include the following two types:

(1) Certificate. The regulatory agency or a third-party agency recognised by the regulatory agency is the subject of certification, applying a combination of substantive review and formal review to conduct assessment and certification while relying on the industry associations and the market's self-discipline. For example, The EU's "white list" is evaluated and certified by the EU's 31st Working Group to protect data subjects' rights and interests in the applicant country and region. Various factors are considered,

including the effectiveness of the legislation and implementation of personal information protection, the establishment of regulatory agencies, and the status of accession to international conventions. The transfer of personal data from within the EU to countries and regions that have passed certification is exempted from review. The EU's "Binding Corporate Rules" (BCR) stipulates that multinational group companies shall submit applications to data protection agencies in relevant countries within the EU. Data protection agencies shall evaluate and authenticate their applications in accordance with the requirements of the BCR. The multinational group organisation can legally export data within the scope of BCR requirements without reassessment. The "code of conduct" certification proposed by the GDPR relies on industry associations, and it can take effect through a binding commitment after it is recognised by the regulatory agencies of the Member States or the EU Data Protection Commission. Additionally, the recognised market certification can also be used as a legal mechanism for cross-border data transfer. After the APEC CBPR certification is approved (annual evaluation is required), cross-border transfers of personal information and data can be carried out within the scope of CBPR regulations.

(2) Contractual obligation. The EU, Australia, and other government agencies have formulated and implemented data export contract models. The contract specifies the obligations of relevant entities, to regulate data recipients' behaviour and manage the outbound flow of personal information. For example, the European Commission adopted the Standard Contractual Clauses (SCC) based on the Data Protection Directive in 2001, 2004, and 2010, respectively. The two regulatory models can be used in parallel, based on corporate self-discipline and government supervision.

Important and sensitive data are by default prohibited from exporting, with exceptions based on data classification and export review. Depending on the data attributes and risk levels, as well as national conditions and political and cultural differences, export restrictions are commonly imposed on critical infrastructure and important industry data, such as government, banking, finance, credit investigation, health, and taxation information. For example, France requires data for taxation, management, and commercial development to be stored locally. Australia prohibits the transfer of health records outside Australia. India rules that payment data is prohibited from leaving the country. The U.S. stipulates that data belonging to security classifications cannot be stored in any connected public cloud database.

Administrative review on a case-by-case analysis is the most common supervision from the authorities. Before certain types of important data leave the country, the data exporter shall submit the export declaration documents to the government department. After a review of the corresponding export activity, the data may leave the country upon approval. For example, the U.S. Department of Commerce reviews whether controlled technologies have obtained export licenses for overseas storage and processing. South Korea has established an export application negotiation mechanism for map data. The National Institute of Geosciences, the Ministry of Future Creation and Science, the Ministry of Foreign Affairs and other departments jointly assess risks and determine whether the export is allowed.

2.2.3.4 Data Sovereignty Increases Conflicts of Jurisdictions

As data has become an important strategic resource of the country, the accumulation, processing and governance of data have become important factors that determine the

lifeline of the country's economy. The thirst for data resources is reflected in the expansionary data sovereignty strategies of major countries and the expansion of jurisdiction.

The data sovereignty strategies of the United States and the European Union are quite “aggressive”, expanding their cross-border data enforcement powers through extraterritorial jurisdiction. For example, the U.S. CLOUD Act empowers U.S. law enforcement agencies to "control" the data of U.S. companies, regardless of whether they are in the U.S. or abroad. For U.S. citizens’ data and personal data in the U.S., foreign governments must go through the domestic judicial process in the U.S. This kind of long-arm jurisdiction enables the expansion of US data sovereignty to the global market where US companies are located. The EU's GDPR also applies to all companies that provide products and services to EU residents, regardless of whether the company is located in the EU. The extraterritoriality of the jurisdictions will undoubtedly increase sovereign conflicts with countries where data is stored.

Relatively, the data sovereignty strategies of China, Russia and other countries are “passive”, by solving the problems of applicable laws and domestic law enforcement through data localisation. In addition, the slow progress of the traditional judicial assistance treaty (MLAT) between countries indirectly encouraged governments to be more willing to choose data localisation policies. Data local storage is at least convenient for law enforcement, and it is also a strong defence in the application of laws.

The Internet is global, but legislation and regulation are local. For a long time, the issue of Internet jurisdiction and applicable laws has not reached consensus. At present, the expansion of data sovereignty has led to an increase in the number of connection

points for the applicable laws in various countries, and jurisdictional conflicts have brought difficult conflicts of obligations to cross-border service companies.

2.3 Dilemma in the development of cross-border data transfer regulation (TBC)

3.3.1 Different objectives 1

3.3.2 State security 1

3.3.3 Personal information protection 1

3.3.4 Different interests among developing and developed countries 1

Part II A Comparative Study of China and the EU

Chapter 3 China's Data Protection Laws from the European Perspective

3.1 Introduction

China's data protection law is an evolving project that still under development, with various administrative regulations and department rules mushrooming. The Personal Information Protection Law has been incorporated into the law-making plan of the 13th Standing Committee of National People's Congress, and was released with the draft for public comment on October 21, 2020. The legislators especially emphasised the protection of public interest and state security, taking into account the needs of the protection of data subject's rights, and took a reluctant position on the regulation of cross-border data transfer. *The Cybersecurity Law* (enacted in 2017) for the first time addressed data localisation and security assessment of data export requirement for Critical Information Infrastructure providers.[113] *The Civil Code of China* (adopted May 28, 2020) newly introduced greater protection of privacy rights and personal information.[114] It clarified that (i) the rights and interests of natural persons over their personal information are civil rights and private rights; (ii) the natural persons' rights to their personal information belong to personality rights; and (iii) the distinction is

[113] The Cybersecurity Law 2017.

[114] "The personal information of a natural person shall be protected by law. Any organization or individual that needs to acquire the personal information of an individual shall obtain such information in accordance with law and guarantee the safety of such information. Any illegal collection, usage, processing, and transfer of the individual's personal information, or illegal trade, making available or disclosure of other's personal information is the violation of law." Article 111 Civil Code of the People's Republic of China.

made between privacy and personal information. These three pieces of legislations constitute the foundation of China's personal information protection laws.

The Measures on Personal Information and Important Data Export Security Assessment (draft for comments) was released in 2017 by the Cyberspace Administration of China (CAC). It was planned to contain elements in the scope of the security assessment, such as the consent of data subject, the security protection status of data recipient, and the risk of data leaving China. Upon receiving constructive criticism, the CAC updated its second version of *The Measures on Personal Information Export Security Export* (draft for comments) in 2019. One essential element – the important data – was removed while one important element – the standard contractual clauses – was introduced.

This chapter aims to provide a comprehensive analysis of China's data protection laws, following the GDPR as a frame to organise and systematise the most important Chinese regulations.

3.1.1 The Evolution of China's Personal Data Protection Laws

Chinese concepts of privacy and personal data protection vary through different historical periods. Most of them are rooted in Chinese traditional ethics or moral standards, and partially integrated with socialism ideology.[115] Since the economic transmission from centrally planned market to free market in the 1990s, Chinese communities began to experience a greater variety of roles in participating economic, societal and political activities. Although traditional predominant values still hold a deep influence

[115] See e.g., B.S. McDougall and A. Hanson (Eds.), *Chinese Concepts of Privacy* (2002) Brill, 8.

on people's behaviours, individualism and subjectivity have dramatically promoted in their social life. Scrutiny and concerns over the importance of individuals' privacy and the protection of emerging personal data processing are ever-growing. Baidu, the largest Chinese search engine provider, was sued by a consumer rights protection association for illegally collecting user data without consent.[116] Alibaba, another internet giant, was challenged by Chinese users for the misuse of their digital transaction records and social media presence on Zhima Credit (an online credit service that offers loans based on users' digital activities).[117] The consciousness of privacy in contemporary China has been gradually expanded and individuals have raised their expectations for the right to be let alone.

Prior to the CSL, China's personal data protection policy was integrated into a number of laws and administrative rules through the protection of personal dignity and reputation. Article 28 of the Chinese Constitution provides citizens with an inviolable personal dignity from "insult, defamation or false charge." Article 252 of Criminal Law (1997) prohibits any violation of the freedom of citizen's communication rights by hiding, destructing or illegally opening other's letters. Article 101 of General Principles of Civil Law (1986) confers natural person and legal person the right of reputation. The Supreme People's Court in 2001 for the first time confirmed the legal ground for

[116] See M. Jing, China consumer group accuses Baidu of snooping on users of its smartphone apps (2018) <https://www.scmp.com/tech/china-tech/article/2127045/baidu-sued-china-consumer-watchdog-snooping-users-its-smartphone> accessed 30 January 2021.

[117] See X. Wang, Zhima Credit apologizes for its annual report's 'mistake' (2018) <https://news.cgtn.com/news/78637a4e35637a6333566d54/share_p.html> accessed 30 January 2021.

claiming remedies for the damages caused by the violation of one's privacy or other personal rights.

Personal Information was firstly defined in the *Notice concerning Punishing Criminal Activities of Infringement of Citizen's Personal Information* in 2013, stating that "personal information includes any information that can identify the citizen's personal identity or information and data involving the citizen's personal privacy, such as name, age, ID number, and so on." In response to the rapid development of technology, Chinese authorities released over 200 pieces of laws, administrative regulations and sector-specific rules regulating the collecting and processing of personal information across domains like banking, healthcare, medical record or disease control.[118] A comprehensive framework for personal data protection laws is urgently in need.

3.1.2 Personal Information Protection

There is no chapter entitled "personal information protection" in the CSL, yet provisions related to the protection of personal information are scattered through this law. Chapter 4 Network Information Security covered most of the personal information protection provisions. “Network operator” is the core subject-matter that most of the obligations imposed upon. Data subject's rights have been conferred passively through the legal obligations for network operators, i.e. network operator shall correct or delete

[118] China provides direct protection of personal information through The Seventh Amendment of Criminal Law, Tort Law, Telecommunication Law, Junior Protection Law, Consumer Protection Law, etc. Indirect protection of personal information is provided though Constitution Law and Civil Law. For example, the Ministry of Industry and Information Technology is in charge of regulating the ISPs via *Measures on Protecting Personal Information of Telecommunication and Internet Users*, *Measures on SMS service management*, etc.

on the request of the data subject when the personal information are incorrect or wrongly processed.

The structure comprises basic principles for processing, legal grounds for processing and a non-exhaustive example list of prohibited conduct. Personal information can only be collected when the data subject is informed and agree to the purpose and scope of the collection. The processing of personal information must follow basic principles listed in Article 40 - 42, 47, 49 which share substantive similarities with the APEC privacy framework. Consent is the ONLY legal ground for the processing of personal information.[119] This is to ensure that the data subject has sufficient autonomy to decide the way his or her personal data will be collected, processed and distributed. Such autonomy is endorsed by the sufficient informing requirement, meaning that only after data subject is informed of the purpose, scope and means of processing the personal data can he or she be capable of giving the genuine consent. The network operator has to perform the information obligation before collecting the individual's personal data.

On October 21, 2020, the full text of the "*Personal Information Protection Law (Draft)*" (the PIPL) was finally unveiled under the attention of the public. The draft deepens the personal information protection system in all aspects, reflecting the legislative thinking that focuses on the protection of personal information while taking into account the complexity of economic and social life. The 8,000-character draft represents the first attempt of China to systematically define, construct and organically integrate the protection and regulations of personal information at the law level. It is a

[119] Article 41 Cybersecurity Law (n12).

condensed version of the CSL, with various international data protection legislative experiences implanted, such as the GDPR.

3.1.3 Enforcement and Authorities

The CSL's provisions relating to data privacy formed the most comprehensive and broadly applicable set of privacy rules. It acts as an umbrella that covers a bundle of administrative regulations and numerous normative texts scattered across most of the industries. To date, there is no independent authority for data protection. Multiple competent authorities or supervisory authorities are in charge of the implementation and enforcement of the rules.

3.1.3.1 Regulatory Framework

Various types of documents have the force of law in China. Among all the legal instruments, the Constitution Law enjoys the highest primacy yet is rarely applied directly. The law made by the National People's Congress or the Standing Committee of NPC has the highest legal effect in the respective regime, such as the Cybersecurity Law, and the Personal Information Protection Law (Draft).

Administrative regulations are rules promulgated by the State Council. Its legal effect is lower than the Law but higher than the Department rules. To date, two administrative regulations were issued: the *Regulation on Critical Information Infrastructure Security Protection* and the *Regulation on Cybersecurity Multi-level Protection Scheme*. Additionally, sector-specific administrative regulations also affect China's personal data export study, such as the *Regulation on Computer Information Security Protection* and the *Regulation on Human Genetic Resources Information Management*.

Department Rules are legal documents issued by the ministries and commissions under the State Council, along with other agencies with administrative functions directly under the State Council. The applicable scope is determined by the competence of the issuing government department. For example, the aforementioned *Measures on Personal Information Export Security Assessment* is a department rule issued by the CAC. To date, around 30 department rules were issued by various authorities in the field of security, data protection and export.

Judicial interpretations are the explanations to specific legal questions made by the State Supreme judicial institutions during the application of the laws. Both the Supreme People's Court and the Supreme People's Procuratorate had released interpretations relating to cases that infringe personal information.

Standards (no legal effect) are mandatory or voluntary technical standards published by the Standardisation Association of China (SAC). In Cybersecurity and Data protection fields, TC260 group under the SAC is responsible for a series of standards titled Information Security Technology that covers methodologies, definitions or scopes of the norms. Within China, national standards play an important role in implementing laws and regulations. Despite the non-compulsory nature, they are better understood as a quasi-regulation rather than a technical specification typically presented in the Western context. Since 2010, over 240 national standards in this field have been published. It remains debatable with the necessity of such a big amount of technical standards in force.

Additionally, local regulations are directly applied within the scope of the provinces, autonomous regions and municipalities directly under the Central Government.

3.1.3.2 Competent Authorities

Under the CSL, different parties are in charge of specific area of works. The *State* is to (i) make cybersecurity strategies; (ii) clarify fundamental requirements and objectives of cybersecurity; (iii) guide key area cybersecurity policies and measures. Additionally, the State shall adopt measures to guarantee the cyberspace free from attacks, interferences and crimes. The *network-related industrial associations* shall provide guidance for entities' self-regulation and promote the healthy development of the industries. The *network operators* are required to fulfil obligations addressed in the CSL and to uphold societal responsibilities.

Respectively, the *Congress* is responsible for determining the scope of CII and key areas. The *Cyberspace Administration of China*, an administrative agency directly under the State Council, is in charge of the coordination and management of all cybersecurity-related issues. The Ministry of Industry and Information Technology and the Ministry of Public Security are responsible for supervising and managing affairs within the scope of their competence.[120] The Standardisation Association of China publishes national and sectoral technical standards.

The CAC, also framed as an agency directly under the Chinese Communist Party, inherently carries a heavy stroke of political colour. It is the most important supervisory authority of cybersecurity and directly reports to the State Council for managing Internet information and contents. It works independently from the Ministries of information, public security or commerce. The CAC also leads the drafting of department

[120] Article 8, Cybersecurity Law (n12).

rules implementing the CSL. Its branches at the province level are the main enforcement institutions that supervise, investigate, and impose administrative fines.

(a structure chart will be presented here)

3.1.3.3 Enforcement

Enforcement of the CSL and related rules in China follows a typical bottom-top approach. Supervisory authorities have broad discretionary powers as well as the competence to impose administrative fines upon entities. Overlapping areas of jurisdictions often pop up among different authorities. The CAC is responsible for coordinating all issues that arise through the enforcement. Although not legally binding, the competent authorities often refer to the Information Security Technology standards when performing assessments or issuing certifications.

The supervisory authorities are actively performing their duties since the year 2015. Means of enforcement include communication with the operator, supervising the modification of business, or administrative fines and termination of the operation. A special operation targeting at illegally collecting and processing personal information through mobile applications is jointly conducted by the CAC, MIIT and SPS.

It is rebuttable that the CAC has the competence in imposing administrative fines. According to the Organic Law of the State Council, the CAC is not one of the departments under the State Council. The legal ground for the CAC should be Article 11 of the Organic Law of the State Council ruling that "the State Council can establish agencies directly under the Council for managing specific affairs or assisting the Primer to handle specific affairs". However, it is not explicitly informed that which agency the

CAC is established for. The official documents issued by the later agencies are categorised as "other kind of administrative documents" which cannot be enforced as the basis for administrative fines.[121]

According to the CSL, it is clear that the responsibilities of the CAC are coordination and supervision. Therefore, the rules and measures issued for imposing fines might not be legitimate, even their legal effect could be challenged (emphasis mine). Such gap originated from the boost of cybersecurity legislations, and shall be bridged in the future law makings. With the working-in-progress Personal Information Protection Law, the CAC is expected to (i) remain as an agency under the CCP for supervising the Internet affairs, and the national independent Data Protection Authority is formed for data protection regulation; or (ii) be conferred the legitimacy under the new law.

3.2 General Provisions

3.2.1 Objectives

Prior to the PIPL, Chinese data protection law is mostly viewed as being fragmentary, insufficient, ineffective and difficult to understand. There exists a vast amount of relevant national, local and sector-specific regulations that affect the comprehensibility of this regime. Even with the releasement of the PIPL, one shall not view China's data protection laws solely from one legislation – supplementary department rules and national or industry standards must be taken into consideration.

[121] Article 14, Law of the People's Republic of China on Administrative Penalty 2018.

As the European GDPR is considered to be the most comprehensive and modern data protection framework, it strives to align the laws of the European Union Member States with certain discretion of national legislations. It aims to balance the protection of personal data with the free flow of data inside the region.[122] Derived from Article 8 of the Charter of Fundamental Rights of the European Union, Article 16 of the TFEU and the Convention 108, the GDPR particularly considers the protection of personal data as the protection of fundamental rights.

The protection of personal information as defined in the PIPL and the CSL, however, are not considered a fundamental right. The Chinese Constitution only indirectly recognised the right to privacy, which has strict preconditions to be referred to. In various cases, access to users' information, such as the user's contact information or social media friends' information, are not considered a violation of the right to privacy by the courts.[123] In practice, the Constitutional Law is rarely referred to, since there is no constitutional court as well as the possibility to assert constitutional rights.[124]

The Chinese legislators demonstrate a clear preference towards stronger data protection. Data protection rules have been increasingly included when amending or creating high-level laws, such as the Tort Liability Law, the Consumer Protection Law, the Criminal Law, the E-Commerce Law, and the General Provisions of Civil Law. After

[122] Article 1 of the GDPR.

[123] *Wang v. Tencent*, Nanshan People's Court Civil Judgment (2020) yue 0305 min chu 825 (粤 0305 民初 825 号)

[124] The first case based on a constitutional right was in *Qi Yuling v. Chen Xiaoqi et al*. of 2001, where the right to receive education was quoted as the applicable provision for the basis of the lawsuit. The *Qi* case was officially withdrawn in 2008 since it was "no longer in use".

the enactment of Civil Code, Tort Liability Law and the General Provisions were abolished. Instead, the right to privacy and the protection of personal information are clarified for the first time under Chapter 6 Personality Rights, Volume 4 of Civil Code.

The CSL stated in Article 1 that, the objectives for the law is:

> *[to safeguard cybersecurity; to maintain cyberspace sovereignty, national security and societal public interests; to protect legitimate interests of citizens, legal persons and other legal organisations; to promote the healthy development of economic and social informationalisation.]*

The PIPL Draft stated in Article 1 that, the objectives is to:

> *[protect personal information rights and interests; standardize personal information processing activities; safeguard the lawful, orderly and free flow of personal information, and stimulate the reasonable use of personal information.]*

This is aligned with the special aspect in terms of multiple objectives in Chinese law makings, particularly those areas where face most of the challenges brought by emerging issues. As this provision suggests, the objectives are to govern everything within the country's cyberspace infrastructure, ranging from internet activities to data export.

The downside is, however, observable. It is not unusual that such generality and flexibility, sometimes excessive omissions, can be found in Chinese law drafting. Coupled with a wide discretionary power conferred on lower-level competent authorities in order to implement the law, predictability and certainty of law often are compromised.

Furthermore, in order to identify a complete set of independent objectives and to prioritize them, the law makers are required to hold clear concepts, logical foundations and thought-provoking procedures.[125] In China, most of the data protection rules were made in response to an existing problem. Despite insufficient experiences in data protection law makings and "rent-seeking" among various authorities, one essential aspect is the missing of a unified value for the protection of personal information. It is yet not crystal clear in other jurisdictions as technology and law in this regime are significantly inter-dependent. Without the clear value set ahead, multiple objectives would affect the fundamental principles as well as the conceptual framework of data protection. The immediate consequence is the vague defining of rights and obligations for stakeholders involved. This echoes the lack of legal predictability and certainty.

3.2.2 Material Scope and Territorial Scope

3.2.2.1 Material Scope

The GDPR applies to the 'processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of [it]'. Article 3(1) of PIPL Draft, employs almost the same wording, provided with exemptions of family private matters and administrative authorities' activities related to statistic and archive management.[126] The CSL and its supplementary administrative regulations apply to inter-

[125] See R. L. Keeney, 'Identifying, prioritizing, and using multiple objectives' (2013) EURO J Decis Process 1, 45-67.

[126] Article 68 of PIPL Draft.

net information service providers, which are understood broadly and include e-commerce service providers, social media service providers, online advertising and mobile services. The Civil Code, specifically the previous Tor Liability law and general provisions also addressed network user.[127] Not very much surprisingly, prior to CSL public processors are excluded from the processing regulations. The CSL, PIPL Draft and Civil Code started drawing a line for public agencies to lawfully process personal information. Criminal Law also applies to certain violations of the processing of personal information, for which is even stricter towards state offenders.[128] In general, China's data protection laws, despite the enactment of CSL and PIPL Draft, still are very much sector-focused.

3.2.2.2 Territorial Scope

Prior to PIPL Draft, the Chinese legislation rarely touched extraterritorial applicability of the rules. They mostly concentrated on domestic compliance and courts do not tend to apply domestic laws to internet companies whose server is located outside of China. Due to the strict Internet censorship rules and data localisation requirements, the information flows are restricted, especially the personal data going abroad. However, Article 3 of PIPL Draft introduced extraterritorial jurisdiction provision, ruling that

[127] Article 36, 62 of Tort Liability Law; Article 1194 of Civil Code.

[128] Crime of infringing citizens' personal information. According to the currently effective Article 253A, as revised by the Ninth Amendment, whoever sells or provides citizen's personal information to others in violation of relevant state provisions or steals or otherwise illegally obtains a citizen's personal information will be sentenced to imprisonment of not more than three years or criminal detention, and/or a fine when the circumstances are serious. If the circumstances are especially serious, the offense is punishable by three to seven years' imprisonment and a fine. The article, however, does not provide a definition of personal information or which circumstances may be deemed "serious" or "especially serious" in sentencing. (Ninth Amendment, art. 253A.)

any organisation or individuals handling personal information activities of natural persons within the borders of the People's Republic of China shall comply with the law. Where the personal information processing activities are carried outside of China, but (i) the purpose of the processing is to provide products or services to natural persons in China; (ii) the processing involves conducting analysis or assessment of activities of natural persons in China; and (iii) other purposes provided in laws and administrative regulations, the PIPL shall apply. This is just like Article 3 GDPR.

3.2.3 Definitions

Personal Information vs Personal Data

The definition of personal data in Article 4(1) of the GDPR and of personal information in Article 4 of the PIPL Draft overlap considerably, both are information identifiable to the data subjects. Anonymised data are exempted, yet the understandings of "anonymous" are quite different.

The core of the difference between personal data and personal information lies in the difference between "data" and "information." Some scholars believe that "data" has a broader scope than "information", since information is a kind of data after being analysed and processed, while data is the carrier of information. Information systems domain uses "data" as a fairly common term to measure any real-world phenomenon. "Information" is data related to a specific decision, so data only becomes information in a specific situation. This distinction lies in the important concept of "relevance".

Therefore, it is more appropriate for most of the data protection regulators to use "data" instead of "information", especially when it comes to matters related to use and disclosure. For example, Singapore has also adopted this view that the scope of “data” is broader than “information”. Hence, it decided to use the term “personal data” when formulating the “Model Rules on Data Protection for Private Organizations”, instead of "personal information". In contrast, “information” has been considered a broader connotation than “data” when studying Japan’s Personal Information Protection Law. "Personal information" is defined as the information of a living individual, which can identify a specific individual by name, date of birth or other description contained in this information, including information that allows easy reference to other information, thereby enabling specific individuals to be identified. This definition is so broad that it even includes public information, such as information in telephone books, public journals, and personnel lists.

In China, the term of "personal data" was earlier used by scholars in the introduction to German data protection laws. It is believed that "personal data" includes any personal data recorded by computers, but no distinction between "personal information" and "personal data" was mentioned.[129] Others use "personal information” and "per-

[129] Xiaohui Li, ‘German Advantages in Protecting Personal Data’ (个人数据保护德国有长处) (1997) China National Conditions and Strength, volume 12.

sonal data" interchangeably. It is believed that "personal data" refers to "personal information" that can directly or indirectly identify citizens' personal identity information.[130]

In summary, "personal data" and "personal information" cannot be equivalent under certain context. "Personal data" is a carrier, or a kind of raw material, while "personal information" needs to be refined. In most legal documents and research papers, however, they are interchangeable. This dissertation does not intend to make a distinction between the two.

Consents

Article 14 of PIPL Draft defined consent as the 'voluntary, explicit expression of agreement, with sufficient acknowledgement as the precondition.' No further explanation is provided regarding what constitutes the "sufficient acknowledgement". Consent can be either oral or written, collective or separate. Different obligations are required in different situations. When essential elements of a personal information processing activity changed, such as the purpose and the means of the processing, or the types of information being processed, a new consent shall be obtained.

De-identification and Anonymisation

Article 19(3) defines de-identification as the process in which the personal information can not identify the specific natural person without additional information. Article 69(4) defines anonymisation as the process in which the personal information can no

[130] Aimin Qi, 'International Comparative Study on Personal Information Protection Laws in Big Data Era' (大数据时代个人信息保护法国际比较研究) (2015) Law Press China.

longer identify the specific natural person, and such a process cannot be recovered. The GDPR has also provided the guidance for the determination of the threshold of the possibility of identification in Recital 26, indicating that a hypothetical possibility of identification is not sufficient to make information identifiable, there must be a reasonable likelihood. When the possibility of singling out an individual does not exist or is negligible, the personal information should not be considered as identifiable. The threshold for identifiability under Chinese context is considered lower than the GDPR. In practice, data processors follow the technic standard of de-identification to demonstrate compliance. Most of the commonly adopted standards are recognised by the authorities, including standards from ISO, NIST, and SAC.

3.3 Principles

3.3.1 Principles Relating to the Processing of Personal Data

Article 5 of the GDPR enumerated lawfulness, fairness and transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity, confidentiality and accountability as the processing principles. Article 5 – 9 of the PIPL Draft listed lawfulness, fairness, purpose limitation, data minimisation, transparency, accuracy and security as the general principles. Notably, the principles are listed in such a general level, and no further explanations are provided in the laws. In contrast, the majority of the non-binding administrative regulations and agency-backed technic standards contain almost all of the GDPR principles and identical explanations.

Lawfulness and Fairness

Lawfulness means that personal information must only be processed when data processors have a legal ground for processing the data. Fairness is the core principle embedded in civil laws, which is linked to the idea that data subjects must be aware of the fact that their personal information will be processed. No fraud or misleading are allowed.

Fairness needs to be assessed on how the processing will affect the data subject. When the processing negatively affects the individuals, such detriment is not justified, thus is unfair. For example, the travel agencies are frequently reported to "kill old customers" (大数据杀熟) using the data they collected. Many online travel agencies offered unreasonably disadvantageous prices to existing customers using big data analysis.[131] This is undoubtedly unfair processing of personal information. In contrast, a different situation may occur, when the processing negatively affected individuals but such detriment is justified. In *Ren Jiayu v. Baidu*, the plaintiff discovered that his name was shown in Baidu search results together with a controversial education company, which is his former employer.[132] Ren sued Baidu for the infringement of his name and reputation and requested to delete the information. The Court rejected Ren's claims, ruling that the search result is merely an objective result where the search terms are used in a certain period and shown based on algorithms. The search result reveals Ren's former working experience which is associated with a controversial education company.

---

[131] Discriminatorily disadvantageous pricing against different customers are reported, including data-based service providers such as Amazon, Orbits, Uber, Ctrip, DiDi, among others. Winnie Lee, '盘点"大数据杀熟"案例' (2019) < https://t.qianzhan.com/caijing/detail/190315-3d15cfc7.html> accessed 2 January 2021.

[132] Ren Jiayu v. Baidu, Beijing No.1 Intermediate People's Court Civil Judgment (2015) Yi Zhong Min Zhong Zi Di 09558 Hao (一中民终字第 09558 号).

Such information is an important factor that the client or students will take into consideration. Therefore, such information is necessary for the public, and the detriment is justified.

Purpose Limitation and Data Minimisation

Purpose limitation means that data processors must only collect and process personal information to accomplish specified and reasonable purposes and not process personal information beyond such purposes. Secondary processing, though not explicitly addressed in the laws, could only be lawfully carried out when such processing is considered compatible with the original purpose for which the personal data was collected.

Data minimisation means that data processors must only collect and process personal data that is relevant and necessary to accomplish the purpose for which it is processed. Similar to the GDPR, two concepts shall be considered in the practical implementation of this principle: necessity and proportionality. The personal information processor must assess whether the personal information to be collected is suitable and reasonable to accomplish the specific purpose. Additionally, one must assess whether the amount of data to be collected is excessive in relation to the purpose that the processor aims to accomplish.

3.3.2 Lawfulness of Processing

Under the GDPR, the processing of personal data will be considered lawful only when and to the extent one of the six legal grounds is met: consent; contract performance; legal obligation; vital interest of individuals; public interest; and legitimate interest. Interestingly, prior to the PIPL Draft, consent is the only legal basis for processing

personal information. Article 13 of PIPL, a GEPR-alike set of circumstances is listed as the legal basis:

- where consent is obtained from the individuals;

- where necessary to conclude or fulfil a contract in which the individual is an interested party;

- where necessary to respond to public health incidents or protect natural persons' lives and health, or the security of their property, under emergency conditions;

- where the processing of personal information, within a reasonable scope, is to implement news reporting, public opinion supervision, and other activities for the public interests;

- other circumstances provided in laws and administrative regulations.

Responding to public health incidents as a legal basis is obviously a criterion set for the Covid-19 pandemic in the year 2020, providing great flexibility for multiple government agencies to operate in controlling the disease expansion. Limitations for such entitlement, however, has not yet been clarified. Concerns are addressed regarding the misuse or abuse of such legal ground in the name of controlling public emergency incidents.

Public interest as a legal basis is largely limited within the scope of news reporting or public opinion supervision. However, Internet information violence is ever-growing. During the Covid-19 pandemic, details of infected persons' information, including identify number, home address, telephone number, employers' information, etc. had been disclosed and spread all over the social media platforms. Occasionally, personal

information of individuals are disclosed on the Internet, for the purpose of sentencing the individuals the so-called "societal death", referring to the situation where a person is believed to conduct certain behaviours which are not a violation of the law, but are considered not align with the moral standards, and therefore will be banished by the communities. Leaving the discussion regarding the justification of "societal death" aside, public interests being interpreted in the provision is problematic.

The GDPR allows the data processing within the limits of the applicable laws, which may include data protection laws and also other applicable rules and codes that deal with area such as employment, competition, health, tax or any other objectives of general public interest, depending on the particular case. To seek harmonisation within the EU, the GDPR prescribes a high level of protection for the rights and freedoms of individuals. The EU national legislation also needs to determine what tasks are carried out in the public interest under this criterion. Under the PIPL Draft, no such requirements are ruled. Data subjects have not opportunity or right to object the processing of their personal data, since once the information is disclosed on the media, it is impossible to stop the spreading of such information. Particularly, the vague wording of "reasonable scope" creates uncertainties. As commonly recognised, when revoking public interest as the legal basis for the processing of personal data, the ground must be sufficient to 'override the interest, rights of freedoms of the data subject or for the exercise or defence of legal claims.'[133]

3.3.3 Consents

[133] Article 21 GDPR.

As afore-mentioned in 4.1.2, prior in the CSL, consent is the only legitimate legal ground for processing, which at first glance leads to a higher degree of protection. Many administrative regulations allow implied consent without adding specifications or requirements apart from information duties,[134] in part only requiring explicit consent for sensitive personal information.[135] Echoes to fairness principle, deception and misleading are forbidden as a means of obtaining consents.

Chapter 2 Article 13 – 19 of PIPL Draft are devoted to the personal information processing rules based on consent. The processing of personal information shall be allowed when the data subject's consent is:

- sufficiently informed;

- freely given;

- explicitly expressed;

- unambiguous indications of willingness.

The responsibility lies with the processor to demonstrate that the data subject has consented to the processing. Where a declaration of consent is pre-formulated by the processor, which will be the case in most circumstances, the consent should be provided in an easily accessible form and using clear language. The processor shall not refuse to provide products or services on the basis that the data subject does not consent the processing of personal information, unless the processing of personal information is needed for the functioning of the products or services. Unlike the GDPR, where freely-

[134] Article 29 Consumer Protection Law ; Article 1035 Civil Code ; Article 41 CSL.
[135] Article 14 PIPL Draft.

given consent is always rebuttable, especially when there is a clear imbalance between the data subject and controller, Chinese laws do not provide such guidance on how to perceive the "genuity" of the freedom to consent.

Prior to the processing of personal information, the personal information processing notice shall be provided, in a clear and plain language, in an easily visible manner. The following information must be provided to the data subject:

- identity and contact details of the processor;

- the purpose and means of the processing, the types and retention period of the personal information being processed;

- information regarding the rights of the data subjects entitled under the PIPL and the procedure to enforce them;

- other issues obligated to inform data subject under laws and regulations.

There are exemptions from the notification obligation. Where the laws or administrative regulations provide that secrecy shall be preserved or notification is not needed, the processors need not to inform the data subjects regarding the personal information processing. Under the emergency circumstance, where it is impossible to notify individuals in a timely manner in order to protect data subjects' vital interests, such as life, health, and property, the processors are allowed to inform the data subjects afterwards.

When the personal information is shared with a third party, a personal information processing notice about the third party shall be provided to the data subject. Specific and separate consent is required for the sharing. According to Article 24 PIPL

Draft, a third party processor bares the equal obligations as the processor in terms of obtaining consent and providing processing information.

Unlike the GDPR, the data subject's consent is not required for automatic decision-making using personal information. Yet, the data subject has the right to reject the processing of the personal information when (i) the legal ground for processing is based on the consent;[136] or (ii) the personal information is used for automatic decision-making.[137] In other words, the consent for the collection of necessary personal information and the consent for the collection of other additional information should be separated.

3.3.4 Processing of Special Categories of Personal Data

Sensitive personal information is defined as the personal information that, once leaked or illegally used, may lead to personal discrimination or serious harm to personal and property safety. Sensitive personal information includes race, ethnicity, religious belief, personal biological characteristics, medical health, financial accounts, personal location and other information.[138] The scope of sensitive personal information defined in the PIPL Draft is broader than GDPR, especially information such as financial accounts and personal location data. Together with the Personal Information Security Specification, communication records, address books, web browsing records, and accommodation information, are included as sensitive personal information.

[136] Article 16 PIPL Draft.
[137] Article 25 PIPL Draft.
[138] Article 29 PIPL Draft.

The processing of sensitive personal information shall be based on the individual's separate consent, and the written consent shall be obtained when necessary. The data subject shall be informed of the necessity of processing sensitive personal information and the impact on the individual.

It is believed that, when the consent requirements for personal information are to be lowered, a high level of protection must to be provided for sensitive personal information. Consequently, a clear differentiation between the general personal information and sensitive personal information is advocated. All laws and administrative regulations have special rules regarding sensitive personal information, of all are defined non-exhaustive and more general. The Specification 2020 gives a long explanation and enumerates many examples of sensitive personal information, many of which go beyond the GDPR. One opinion regarding this phenomena is, unlike the fundamental rights protection oriented rules in the GDPR, the Chinese regulations mostly evolved within the security context. To ensure the safety of persons and property is the main criterion. Article 15 of the Management Measures added that a record shall be filed at the competent supervisory authorities when processing sensitive personal information for a commercial purpose. In summary, strict rules are imposed on the processing of sensitive personal information.

3.3.5 Conditional Consents of Child

The "Measures on the Protection of Children's Personal Information on the Internet" is the first specific regulation for the protection of personal information of children. The Specification of 2018 requires personal information processors to treat the child’s

information as the sensitive personal information. The explicit consent of their guardians is required when processing children's personal information.[139] The special protection measures for sensitive personal information stipulated in the Specification also apply to children's personal information. These measures include obtaining item-by-item authorisation for each function of the service or product, encrypting the storage and transfer of information, applying for and obtaining authorisation for internal information access, and obtaining clear consent before sharing and transferring personal information. The "Privacy Policy Template" attached to the Specification also contains paragraphs on the processing of children's personal information.

Consistent with the child's information protection measures, the PIPL Draft ruled that children under the age of 14 are minors. Where the personal data of children below the age of 16 are concerned, Article 8(1) of the GDPR requires the consent of the holder of parental responsibility.

3.4 Rights of the Data Subject

Compared to the Chinese data protection laws, the GDPR is considerably more complex and far-reaching in terms of the extensive set of rights. This is in part because bolstering individual's rights was one of the main objectives of the Commission in proposing the new data protection framework. The PIPL Draft provides a larger scope of data subjects' rights compare to the CSL, set forth in Article 44 – 49 of the Law.

[139] Of the binding and adopted regulations, it was the new Civil Code that first introduced rules for this case. In principle it requires the consent of the guardians to the processing of any personal information of those under their guardianship, which means minors under the age of 18 and adults with no or limited civil capacity. Article 17, 27, 1035(10, 1037(1) of Civil Code.

3.4.1 Transparency

Transparency is fundamental to any data protection system, as individuals' right to privacy cannot be assured if they are not properly informed about the processing activities. Article 12(1)(5) GDPR obliges controllers to provide information in a 'concise, transparent, intelligible and easily accessible form' as well as free of charge, unless requests are unfounded or excessive. Articles(13)(f) GDPR differentiate information obligations depending on whether information was collected from the data subject or not. In addition, the controller is required to 'facilitate the exercise of data subject rights' by providing 'mechanisms to request' and responding 'without undue delay and at the latest within one month'.

Article 13 of PIPL Draft requires that data subjects have all the information they need in order to understand the nature of the processing and to exercise their further statutory rights. Consequently, Article 44 requires that data subjects have the right to information, the right to decide, and the right to restrict or reject others' processing of their personal information. Article 29 of the Consumer Protection Law requires that controller to disclose their rules regarding the processing of personal information. Article 41(1) of the CSL imposed obligation to inform the data subject the manner, content and purpose of the processing. Data subjects can get access to a complaint and reporting system where the complaint is required to be handled in a timely manner. In the administrative regulations, a time of 15 days is set for responding to the complaints. Furthermore, the controllers should make transparent the channels for accessing and correcting information and the consequences of refusing to provide personal information. Article 18 PIPL Draft and other Measures give a more detailed list concerning

the information, and the later in particular largely cover the requirements of the GDPR. In addition, Article 14 Draft Administrative Measures and Article 5.4 Standard 2020 undertake a differentiation as in Article (13)(f) of the GDPR. Going beyond the GDPR, Article 5.5 and Appendix D Standard 2020 explain the functions of the privacy policy and provide a very detailed and long template combined with writing requirements, which contain very comprehensive rules that are clear and easy to understand and could also inspire the European legislator.

The PIPL Draft non-exhaustively exempts controllers from the obligation to respond to requests or disclose information in Article 19(1). The exemptions include, but not limited to: (i) if they are directly related to national security, public safety, major public interests or criminal prosecution; (ii) if the data subject is abusing his or her rights; (iii) if it will cause serious damage to the legitimate rights and interests of the data subject or others; and (iv) if trade secrets are involved. The uncertainty concerning, as discussed before, 'public interests' and the relative openness of these exceptions could lead to substantive limitations to the information obligations. Nevertheless, the set of personal information protection laws would establish comprehensive information obligations and meet sufficient demands of Chinese scholars and practitioners. In particular, the statutory obligations could alleviate issues such as the lack of privacy policies in many companies or the absence of contact.

3.4.2 Right of Access

Compare to the right of information in Article 13 and 14 GDPR, the right of access set out in Article 15 is in a sense the active counterpart to the more passive right. Any data subject that requests to know must be told about the personal data the controller or

processors holds about them, and why and how it is handled. In comparison to the CSL and PIPL Draft, the GDPR has a considerably larger scope of the mandatory categories of information that the entity must provide.

Article 45 of the PIPL Draft created the right of access for the first time in China, ruling that individuals have the right to acquire and copy the personal information that a personal information processor has. This provision echoes Article 1036(1) Civil Code. The wording "copy" is confusing though. In the combination of other norms and consideration under the context, the intent would be to acquire a copy of the personal information the processor has, as well as the processing activities the personal information are involved. Exemptions are listed according to Article 19 PIPL Draft. Personal information processors are required to provide the responding in a timely manner after receiving the request. The Measures added that, the information the processors provide shall be "truthfully and free of charge". More specifically, Article 5.3.7 of the Guidelines suggests the personal information processor to inform data subjects:

- whether or not the personal information is processed;

- the content and the status of the processing;

- unless the cost or frequency of the request is beyond a reasonable level.

Article 8.6 of the Specifications limits the content of copies to basic personal information, personal identity information, and health/psychological/education/employment information.

Compared with the broad coverage in Article 15 GDPR, Chinese laws failed to address many essential aspects of the processing, such as the purpose of the processing, the categories of the information, the information of the recipient, the source of the information if not collected from the data subjects, the retention period of the information, and most importantly, the existence of automated decision-making.

3.4.3 Rectification and Erasure

Right of rectification

Article 16 of the GDPR provides data subjects the right to rectification of their personal data. Article 17 of the GDPR provides the so-called right to be forgotten that becomes the most actively scrutinised aspects of the original proposal by the Commission. In comparison, Article 43 CSL, Article 46 PIPL Draft, among other regulatory rules, allow data subjects to request the processor for information completion or rectification. Such right is, however, very limited. Previously, the Regulation limits the information that is uploaded by the data subjects only. In PIPL Draft, processors are obliged to correct, update, and complete the information in accordance with the data subjects' request within a reasonable time. Such right of rectification is considered to be significantly weakened, since the corresponding right to access is limited.

Right to be Forgotten

Right to be forgotten is not recognised by the courts in China. In *Ren Jiayu v. Baidu*,[140] the court dismissed the plaintiff's claims from two aspects: the right of name and right

[140] Plaintiff Ren worked in education sector at Wuxi Taoshi Company, and terminated the employment contract since November 26, 2014. Since February 2015 the Plaintiff discovered website links titled "Taoshi education Ren Jiayu" and "Wuxi Taoshi education Ren Jiayu. Since Taoshi education is quite

of reputation. The name was used neutrally, without misappropriation or counterfeiting. The reputation was not damaged due to the search terms. As to the right to be forgotten, the court reasoned as below:

[*The right to be forgotten is a concept established by the European Court of Justice. There is no legal provision for the right to be forgotten in the contemporary laws in China. Ren claims that his right to be forgotten should be a personal interest based on his general personality rights. If one's personal interest should be protected, Ren must prove its legitimacy and the need for the protection in his case.*

… such information is necessary for the public. Therefore, the interest that Ren claimed to delete based on the right to be forgotten cannot be justified and is unnecessary to be protected under laws. The Court does not support his claim.]

Apparently, the Chinese court does not like the idea set in *Google Spain*, largely due to the fact that China is not suitable for introducing such a right. Take a review back at Article 17 GDPR, the right to be forgotten is strictly limited by other fundamental rights. Such a right can be declined to the extent that processing is needed for exercising other rights. The balance between freedom of expression and information is carefully considered. In the absence of other rights and balances in Chinese legislation, the rash transplantation of the right to be forgotten may lead to the most extreme consequences: any user can request any network information service provider for any information or data related to him/her without reason the delete processing.

---

controversial in the debate, the Plaintiff believed that such titles and links infringed and significantly harmed his reputation. The Plaintiff sent multiple emails to Baidu requesting for deletion of the information, yet Baidu failed to delete.

Right of Deletion

Article 47 PIPL Draft gives data subjects the right to obtain the erasure of their personal information without undue delay, namely:

- the original purpose for processing is accomplished;

- the agreed storage period is expired;

- the personal information processors terminated the supply of products or services;

- data subject withdraws the consent;

In contrast, Article 43 SL, Article 1036(2) Civil Code, and similar Article 8.3 Specification 2020 formed the right on the basis of the violation of legal norms - the processor violates the laws or breaches the agreement. Only Article 21 of the Measure provides a general right to request deletion of the personal information and a corresponding obligation for the processors. In summary, the right to deletion is presented in various forms in Chinese data protection laws.

3.5 Special Requirements

3.5.1 General Obligations

Article 21 of the CSL requires all network operators to be obliged with different security measures according to the *cyberspace multi-level protection scheme* ("MLPS"*)*. Under the MLPS, network operators shall safeguard the cyberspace from interference, destruction or unauthorised access, and to protect the internet data from leak or fraud.

Security obligations include but not limited to (i) the establishment of the internal security management protocol; (ii) the appointment of a person in charge of security affairs; (iii) the deployment of technical measures for cyber attacks; (iv) the record of internet operation activities no shorter than six months and the response plan for security incidence; and (v) the classification of data and the backup and encryption of the important data.

The MLPS was born from the demands of the national computer system security in 1994 and thus falls under the competence scope of the Ministry of Public Security. After a series development of administrative regulations, the updated draft of *Regulation on Cybersecurity Multi-level Protection Scheme* as a milestone was released in 2018. Together with a bundle of supplementary national technical standards, the so-called MLPS 2.0 framework of cybersecurity in China is finalised.[141] The MLPS Regulation as a supporting document of Article 21 CSL defines descriptive obligations and requirements for the network operators fell under different levels of MLPS. Eleven general obligations are listed to clearly allocate the liability and to set technical and organisational security measures. Specific obligations need to be met according to the level of the network operator's activities that would affect the state and public security, scaled from 1 the least risky to 5 the most risky.[142] After being classified, which is based upon a self-assessment, the network operators are required to deploy special

[141] The three newly released national standards are: (1) GB/T 22239-2019 Information Security Technology-Basic Requirements for the Multi-level Protection, (2) GB/T 25070-2019 Information Security Technology- Cybersecurity Multi-level Protection Security Design Technical Requirements, and (3) GB/T 28448-2019 Information Security Technology-Cybersecurity Multi-level Protection Assessment Requirements, which was into force on 1 December 2019. Another national standard titled GB/T 25058-2019 Information Security Technology-Implementation Guide for Cybersecurity Classified Protection comes into effect on 1 March 2020.

[142] For the description of the security levels, see Table 1.

security measures such as personnel management, datasets backup and encryption to protect important data.

The compliance with the MLPS 2.0 will be essential for understanding the personal data export regulation in China. Not only because such compliance is mandatory, but also the second pillar of the CSL, critical information infrastructure protection, is based on the classification within MLPS.

3.5.2 Critical Information Infrastructure

Critical Information Infrastructure ("CII") is a major challenge in implementing China's cybersecurity strategy and had been recurred at top-level national cybersecurity meetings. On the basis of the cybersecurity MLPS, the state implements key protections to CII which, "if destroyed, suffering a loss of function, or experiencing leakage of data, might seriously damage national security, social welfare, and public interests". A non-exhaustive example list (including public telecommunication and information service, energy, transportation, water resources, finance, public service and e-governmental information) is given in Article 31 CSL showing the broad scope of the application of CII requirement. In principle, any network operators that being graded above level III (including level III) under the MLPS shall be regarded as CII operators.

CII operators are imposed stricter security requirements due to the nature of the data being processed. More importantly, Article 37 CSL rules that:

[Personal information or important data that CII operator collected or generated during its operations within the territory of the People's Republic of China shall be stored within the territory of China.]

Transferring CII information outside of China is only allowed under exceptional circumstances where actual needs for business are in place and a security assessment is approved by competent authorities. Under the CSL, CII operator is the only subject-matter that is required to comply with the data localisation policy and security assessment for cross-border data transfer. However, the definitions of CII and other key concepts such as important data remain unclear.

CII is in essence a network facility, information system, digital asset, or a collection of such elements.[143] In the early stages of informationalisation, CII was considered to be part of Critical Information ("CI") that was scoped clearly. With the changing of the technical landscape, sources of risks are far beyond the scope of CI, such as the attacks coming from virtual entities, i.e. ICT or Operation Technology domain.[144] At present, large-scale network destruction of CII is a high-risk yet low-probability incident that very limited examples of CII being damaged from cyber-attacks or data leakage can be provided. Therefore, the assessment of security and risks of CII mainly reply on the experts in the domain, instead of evidences or case studies. This brought inconsistency in determining the scope of CII and eventually made it difficult to implement relevant policies. Generally, all ICT service providers fall within the scope of CII operators according to the laws, which is not efficient in the digital economic community.

---

[143] CAC, National Cyberspace Security Strategy, 2016 (unofficial English translation available at https://chinacopyrightandmedia.wordpress.com/2016/12/27/national-cyberspace-security-strategy/). See also, Title VII, the USA PATRIOT Act, 2001; Framework for Improving Critical Infrastructure Cybersecurity Version 1.1, National Institute of Standards and Technology of the U.S., 2018.

[144] For example, some malware target industrial operation system in electricity, gas, or chemical plants, while some cyber attacks target the control or tampering of information and data.

Chapter 4 Cross-Border Regulation in China

4.1 Critical Information Infrastructure Data Export

The starting point for the study of personal data export is to define the CII operator. Quoting the data localisation requirement discussed in 4.5.2, any personal information or important data that are involved in CII shall not be transferred abroad unless a security assessment is conducted with the supervisory authorities' approval.

- Defining CII

On 11 July 2017, the CAC released the draft for comments *Regulation on Critical Information Infrastructure Security Protection* (CII Regulation). Aligned with the CSL, the scope of the CII shall be determined by a two-step test: (i) if business falls within the industry or sector listed in the Regulation; and (ii) if the business is graded security level 3 or above as demonstrated in Table 1.

Additionally, the CAC's Guidelines on State Cybersecurity Inspection (no legal effect) proposed three aspects to help self-evaluating the CII:

1, key business domain, e.g. data centre cloud service, domain name resolution service, or voice data internet basic network and hub in Telecommunication sector;

2, information system or industrial control system that supports the key business, e.g. generator set control system or information management system;

3, quantity of CII device, e.g. registered users above 10 million, or active users above 1 million, or daily transaction exceeds 10 million RMB for a platform service.

- Defining CII operator

The rules apply to registered entities operating inside the territory of the PRC, as well as those which do not register inside China but offer business and services to Chinese customers. The criteria to determine whether the entity provides business or service in China include: (i) using RMB as currency; (ii) using Chinese as the language; and (iii) delivering goods to China. Any of the abovementioned criteria is sufficient to lead multinational companies to store the collected personal information and important data inside China.

4.2 Personal Information Export

4.2.1 Personal information

Personal Information is defined as "any information that is recorded, electronically or by other means, can be used or in combination with other information to identify the identity of a natural person" (Art. 76(5) CSL; Art. 4 Personal Information Protection Law (draft)). It is a commonly adopted "capacity to identity" methodology.

In the Information Security Technology – Personal Information Security Specification 2017, based on the definition given in the CSL, this standard enlarged the scope by using a very expansive wording: "any information recorded electronically or by other means". This targets all operators from both public and private sector, as well as all collecting and processing activities of personal data they conduct. Furthermore, the standard added that "personal information is … or any information that can reflect a specific natural person's activities". This may be consistent with the broad interpretation of personal data held by the CJEU.

Important Data has been repeatedly addressed in the CSL. It is of crucial importance for assessing CII and CII data export requirement, yet surprisingly not defined in the law. The draft of Information Security Technology – Data Export Security Assessment Guidelines (the Guidelines) defines important data as "raw data and inferred data collected or generated by entities, organisations and individuals inside of China, that do not involve national secrecy, but are closely related to state security, economic development or public interests". Publicly accessible government information is excluded from the scope of important data. An index for determining important data is attached with this standard, comprising of 27 main categories and 223 sub-categories. The categorisation is similar to the U.S. controlled unclassified information (CUI） system.

4.2.2 Measures on Personal Information and Important Data Export Security Assessment 2017

On 11 April 2017, the CAC circulated the draft for public comments entitled "Measures on Personal Information and Important Data Export Security Assessment" (the 17' Measures). Unlike the CSL, the 17' Measures expands the subject-matter of Article 37 CSL from "CII providers" to "network operators". Under the CSL, any owners or managers of networks and network service providers are defined as network operators. It is disappointing since the main issue that practitioners were expecting from the 17' Measures is to distinguish between the CII operator and the ordinary network operators. A clear definition of important data is also missing, only stated that "data closely related to state security, economic development and societal public interests." It further cited the Guidelines as the reference.

Being the first legislation concerning data export regulation of China, the 17' Measures provided guidance to assess the necessity of the export and data that are prohibited from exporting. Security assessment is classified into self-conducted and authority conducted. Data that do not exceed the benchmark (500,000 pieces of personal information/1,000 GB data/important domains) can be exempted from administrative procedures of approval. Unfortunately, all essential issues were kept untouched, or otherwise worded vaguely, making it very difficult to comment.

4.2.3 Measures on Personal Information Export Security Assessment 2019

After receiving a large number of public comments, the CAC published the second draft titled "Measures on Personal Information Export Security Assessment" (the 19' Measures). As its name suggests, the 19' Measure only applies to personal information. The legal requirements set out in the 19' Measures are significantly more onerous than the 17' Measures. Within two-year considerations, the legislators demonstrated observable preference in data export regulation approach.

- Data localisation

The 19' Measure requires all personal information to be stored domestically for security assessment before being provided to recipients outside of China.[145] Two aspects are implied: all personal information need to be locally stored; and all personal information exports need to go through security assessment.

While data localisation is gradually adopted in international data regulation standards, one shall notice that data localisation does not necessarily mean the restrictions over

[145] Article 2 of the 19' Measures.

cross-border data flows. Either the EU GDPR or the U.S. CUI system both emphasise that data localisation, backed with transparent regulatory rules, can reconcile the objectives of safeguard state security and personal rights and free flow of data across borders, which are of equal importance. The 19' Measures itself aims to functioning as a precise and predictable mechanism for cross-border personal data transfer.

- Security assessment

Network operators shall submit the applications for a clearance for the personal information export to the province-level Cyberspace Administrations after a transfer contract is signed with the recipient. The supervisory authority after received the application shall conduct security assessment based on the submitted documents, and to complete it within 15 working days, with the possibility of extensions depending on the complexity of the export.

The security assessment focuses on (i) legal compliance; (ii) protection of data subject's rights; (iii) enforceability of the transfer contract; and (iv) the recipient's record on whether it had infringed data subject's rights or had security incidence. When serious data leakage or data misuses occur, the data subjects are unable to protect their legitimate interests, or the parties are unable to provide protection of the personal information, the authority can request the network operator to pause or terminate the transfer. The security assessment shall be performed at least once per two years. When the substantive factors, such as the purpose of transfer or the retention period, have changed, a new application of assessment shall be submitted.

- Standard contractual clauses

The requirement of the legally-binding contractual agreement between the network operator and the recipient is probably the biggest surprise in the 19' Measures. This so-called transfer contract is the EU Standard Contractual Clauses alike, taking into consideration the limitation of territorial jurisdiction, recognises *inter partes* autonomy.

The contractual clauses are required to include: (i) the purpose, type and retention period of the personal information export; (ii) the data subject is the beneficiary of the clauses involving data subject's interests; (iii) the legal ground for the data subject to claim for remedies when infringement occurs; (iv) when the recipient is unable to perform the contract due to its state's legal environment changed, the contract shall be terminated or re-assessed; and (v) the termination of the contract shall not exempt the obligations involving the legislative interests of the data subject, unless the personal information is destroyed or anonymised. The 19' Measures further clarifies the contractual obligations of network operator and recipient, respectively.

The adoption of standard contractual clauses integrates the regulatory requirements into contract autonomy. It is expected to indirectly abide off-shore entities by the China's standard. This approach largely depends on the supervision of the post-transfer performance of the parties. Considering that China is still waiting for her own Personal Information Protection Law, it is more likely that China's personal data protection and cross-border transfer regulation will be tilted towards the European standard. On the other hand, there is no clear line between personal information and important data. Important data naturally could contain a large amount of personal information. The

regulation on important data and important data export is waiting for the other boot to drop.

4.2.4 Personal Information Protection Law (draft)

On 21 October 2020 the Legislative Affairs Commission of the Standing Committee of the National People's Congress released the draft of Personal Information Protection Law (the PIPL) and invited for public comments. Different from the 19' Measures, the PIPL draft does not require all kinds of personal information transborder activities to be examined through the security assessment.

- Derogations

Cross-border transfer of personal information is by default not allowed, unless at least one of the derogations is granted:

1, Where the amount of personal information being processed reaches the threshold for CAC security assessment, the personal information processor shall firstly store the personal information inside China. Such personal information can only be transferred outside of China after the security assessment being conducted and approved by the CAC.[146]

2, Prior to the cross-border transfer, the processor shall provide the data subject with information including the identify and contact of the recipient, purpose and means of processing, types of personal information, and means for data subject to implement the rights. The transfer is allowed when the individual's consent is obtained.[147]

---

[146] Article 38(1) and Article 40, Personal Information Protection Law (draft).
[147] Article 39, Personal Information Protection Law (draft).

3, A personal information protection certificate issued by a CAC-recognised organisation.[148]

4, Contractual obligations over the recipient with regard to the personal information protection[149] (similar to the contractual clauses described in Sec. 3.2.2).

- Restrictions

For the concerns of the protection of China's data subjects and data sovereignty in the global data governance, as well as to achieve a delicate balance in international relations, the PIPL draft for the first time introduced restrictions and countermeasure clauses over personal data. The measures embodied a "black list", on which the subjects to the restrictions or countermeasures will be included in the list that personal information transfer is restricted or prohibited. The applicable conditions of restrictions and countermeasures have also been strictly limited. The subjects to the restrictions include foreign institutions or individuals engaged in personal information processing activities that (i) damage the rights of Chinese data subjects; and (ii) endanger China's national security and public interests.[150] The subjects to the countermeasures are countries or regions that impose discriminatory restrictions, prohibitions or similar measures on China[151].

- DPIA requirement

[148] Article 38(2), Personal Information Protection Law (draft).
[149] Article 38(3), Personal Information Protection Law (draft).
[150] Article 42, Personal Information Protection Law (draft).
[151] Article 43, Personal Information Protection Law (draft).

Data protection impact assessment (DPIA) is one of the most important means for the continuous and autonomous operation of the compliance operations that personal information processors shall demonstrate and/or self-certify. Prior to the Personal Information Protection Law, DPIA is recommended via non-mandatory technic standards. For the first time DPIA is ruled as a legal compliance that more stringent requirements have been put forward for the establishment of an organisation's internal compliance system. Specifically, the DPIA is required when personal information are transferred to a recipient that is located outside of China. A period of minimum three years has been proposed as the retention time for keeping the result of the DPIA and the record of the processing.[152]

- Transfer by national agencies

The access and transfer of personal information are possible based on the request for international judicial assistance. Where national agencies need to transfer personal information abroad, special laws and regulations shall be complied with.[153]

4.3 Conclusion

With the increasing participation of emerging countries in the global data governance, the traditional legislative paradigm dominated by the European Union and the United States is constantly being broken and reshaped. It is particularly important for China

[152] Article 54, Personal Information Protection Law (draft).
[153] Article 41, Personal Information Protection Law (draft).

to establish the regulatory framework of cross-border data transfer, for not only it involves the rights of Chinese citizens and entities, but also the cyber sovereignty and national security, as well as the framing of global cyberspace rules.

China keeps leveraging the data sovereignty to fasten the law makings to support the development of critical technology in digital domains and the infrastructure construction. The cross-border data transfer regulation prefers a strict unidirectional data flow administration that focuses on controlling the flow of the data being transferred outside of China. The regulation is largely orientated by the CAC agencies, which weakens the autonomy for individuals and entities in terms of self-governance and enforcement. It is better to objectively value the importance of efficiency in digital economy and to avoid the excessive rigid adherence to traditional sovereignty, of which, the data localisation requirement as the strongest manifestation of data sovereignty is imposed.

In practice, either the "common European data space" proposed by the European Data Strategy, or the "certified governments" recognized by APEC CBPR system are both an attempt to establish cross-border judicial corporation frameworks among trusted entities for the application of rules and efficient enforcement. However, China has not established a mutual trusted mechanism for transborder data flow with other countries. The proposed initiatives largely remain at the conceptual level without practical operability.

Despite the limitations, there are various positive dynamic developments in the framing of China's cross-border data regulation. The CSL, together with Civil Code and Personal Information Protection Law demonstrate great willingness towards a stronger

data protection regime and more flexible regulatory mechanism. By introducing contractual obligations and statutory derogations while strengthening domestic personal data protection standard, it is observable that China's legislation is continually moving towards the European approach. Given the fact that countries are unlikely to form a corporation framework in a short period, cross-border data transfer between China and the EU would be profoundly rooted in bilateral and multilateral trade and investment negotiations.

Chapter 5 Challenges and Potentials for a Sino-EU Collaborative Framework

Part III Bridging the Gap: Industrial Solutions

Chapter 6 Towards the Health Data Cross-Border Transfer Compliance: A Case Study of HEART-ITN Project

Chapter 7 Operationalisation of eHealth Privacy Requirements in the Context of the GDPR and CSL

Chapter 8 Federated Machine Learning in Data Protection

Federated machine learning is in essence a machine learning technique where models are learned on datasets distributed across multiple devices without accessing them directly. This privacy-preserving property of federated learning triggers the scrutiny of studies to explore this new paradigm of data storage and processing - privacy could be finely secured and protected while minimizing the compromise on accuracy of trained models. It is yet unclear how privacy-protecting statutes would interpret this new technology and to which extent it may be deployed. This chapter examines data privacy and security in federated learning from a multi-disciplinary perspective. By unpacking the technology of federated learning and analyzing through lens of data protection laws, it is concluded by finding that federated learning *per se* cannot be GDPR complied. Rather, it could be compatible if properly designed so that to enable the better governance of her data by data subject. The roles for federated learning in legal contexts in

relation to data minimisation; storage limitation; data protection by design and by default; transparency and interpretability are proposed. Additional guidance to apply privacy enhancing technologies on federated learning to facilitate assessment of privacy risks is also provided.

8.1 Introduction

Recent advances in computing technologies and the explosion of Big Data have empowered machine learning (ML) and artificial intelligence (AI) applications to learn rich insights from large datasets. These applications have enabled automation of knowledge-based tasks, improved decision making and predictive intelligence benefiting consumers and service providers alike. The accuracy as well as the degree of personalisation of these applications improve with collection of more personal data. However, large scale collection of images, audio and textual data to train ML algorithms are framed risky and dangerous for privacy, equality and autonomy.

Typically in centralized ML approaches, datasets are logged on a central- ised/distributed server to train a model. Such data is both willingly shared by the users and or implicitly collected through interaction with users. The downside of large scale personal data collection for centralised ML techniques is that (i) this data may also contain accidentally captured private information about the users like ambient noise and conversations, unintended faces, license plates, etc and (ii) once the data is collected for learning, the user loses control over their data while service providers store the data indefinitely, and also keep ownership over the learned model.

Contemporarily, the widespread acceptance of anonymisation techniques has led to a de facto standard for extracting information from a dataset while protecting the confidentiality of individuals whose data are processed. Anonymisation techniques are heavily researched, ranging from the commonly adopted k-anonymity, to the differential privacy, and to the latest highly acclaimed synthetic datasets. But the basic trade-off between usability and privacy still remains.

The aforementioned concerns have prompted a proactive approach towards privacy-preserving ML with federated machine learning. Federated machine learning (FML) aims to tackle these problems by training models over datasets stored in devices directly. This is performed in a two tier-approach, (i) train local models over on-device datasets and (ii) disseminate updates on the local model to one or more servers to train the global model. The privacy gains made with FML amount to "decoupling of model training from the need for direct access to the raw training data".

While most of the recent studies on FML focus on improving the communication overhead, model accuracy, security and privacy aspects of FML, there is a lack of existing studies that analyse the likelihood for FML to be regulatory compliant under privacy-protecting laws. Existing literature addresses the challenges and the approaches of traditional ML techniques to meet their compliance obligations and protect the privacy of individuals under European data protection laws. However, FML techniques significantly varies from traditional ML as highlighted in Table. This entails the need for a study on FML from both legal and privacy engineering perspectives to interpret how FML can contribute in achieving cross-disciplinary GDPR compliance.

8.2 Background

Google coined the term federated learning, a form of privacy-preserving de- centralized machine learning, to leverage shared models learned on rich and contextual on-device data without the requirement of disclosing the raw user data to a centralised server. We shortly outline the operational model of federated learning and common privacy-related considerations below.

Data

Stages: In federated learning, data is processed in three distinct stages as presented in Fig. : (i) raw data is collected and stored on client devices in a local data store, (ii) the local model learned from the data is transformed into an update for the server, and (iii) the global model stored on the server is updated accordingly and the update is distributed to all clients. The privacy-preserving nature of federated learning derives from the fact that personal data is processed locally on the client device, and the personal data is reduced through the principle of data processing inequality.

Feature and sample spaces: Federated learning enables learning over data with diverse feature and sample spaces. In horizontal federated learning, models are learned from data with similar features, collected from a diverse set of users, i.e. a shared feature space but diverse sample space. One example is the learning of article preferences of users belonging to different countries and language groups from clickstream data. On the other hand, in vertical federated learning, the models are learned from the same user space, over a diverse set of features, i.e. a shared sample space but diverse feature space. For example, the creation of a recommendation model to predict user interests in products from articles they read and the purchase history of the user collected by an

e-commerce application. Liu et al. proposed a transfer learning based approach to learn models on data with minimal overlap on both sample and feature space. Learning of user preferences from text and image inputs from multiple countries is one such example. In this article, we mainly focus on horizontal and vertical federated learning systems.

Client

Role: Client devices execute the federated learning application and are primarily responsible for (i) collecting and storing data obtained from sensors/peripherals to be used for training the local model, (ii) extracting features from the data and training of the model locally, and (iii) applying necessary security- and privacy- enhancing measures before sharing model updates with the server.

Security and privacy issues and measures: Even though in federated learn- ing, local datasets are not disseminated to the server, the problem is shifted towards protecting the dataset on the device itself. Bonawitz et al. present best practices for client devices to periodically remove stale datasets and to ensure that the data is protected at rest from malware or disassembly attacks through encryption. In federated learning, clients are not only raw data providers but also contribute to the global model. Hence, they are more powerful than clients in traditional centralized machine learning and can launch attacks such as model poisoning. An adversarial third party can eavesdrop on the model updates sent to the server and perform model inversion attacks. The goal of such attack would be to extract private information about the client from the updates. To prevent such attacks, the exchange of models between the client devices and the server should be encrypted in an end-to-end fashion.

Server

Role: In federated learning, the server acts as a coordinator of the federation of devices. The server chooses the participating devices in each round and com- putes an update on the global model in each step from the aggregated device updates.

Security and privacy issues and measures: An honest but curious server can attempt to maximize the information extracted from clients, i.e. by inferring private information individually from the client updates. Measures against such a server can be adopted from a client perspective, e.g. by adding noise to the models to achieve data-level or user-level differential privacy.

The server can also gain additional information on client devices by targeting specific devices in each round. By design, we can mitiagate this attack by ensuring the server can (i) only pick a subset of devices advertising participation randomly and (ii) access the aggregated update after the local updates have been processed. This can be achieved with the server treating the updates as a collective federated value such that the individual contents of the federated value remain opaque to the server. Moreover, since the server executes simple functions such as averaging over the aggregated models, these operations can be performed over the encrypted data models with approaches such as homomorphic encryption.

Model Federated learning models in the existing literature are primarily of two types: statistical models and neural network models. **Statistical models** such as logistic and linear regression are typically used for classification and prediction applications. These models are smaller in size, easy to interpret and can be trained on devices with constrained resources. Neural networks on the other hand are used to model more

complex structures in datasets with the use of highly parameterized and flexible functions. However, neural network models have stringent resource requirements and have larger model sizes. As such, interpretability is traded in for more predictive power.

Local and global model updates of both types comprise either the models entirely or their delta from the model in the previous iteration. Models can also be compressed to reduce communication overhead, especially for deep learning models.

Security and privacy issues and measures: Local and global model updates can be exploited for model inversion attacks and for information retrieval. To minimize the vulnerabilities stemming from such attacks, the model updates should be ephemeral and not stored on the global datastore.

8.3 Case-Study

Among the various domains to which FML techniques are applied, recommender systems is a domain where these techniques connect the end-user directly to the service provider. The service provider, on one hand, aims to offer increased visibility of items the end-user is more likely to consume while the end-user benefits by having an improved user experience. The performance of recommender systems on the other hand is fuelled by data; more contextual data the service provider accumulates on the end-user, more likely is the ability to offer a suitable recommendation. Furthermore, a high degree of personalization is also a key desirable feature for end-users. However, as the name suggests, personalization entails the need to access the user's personal data. Thus, a trade-off between the desirable performance of recommender systems and the need to access personal data arises.

To this end, the analysis is around a FML based collaborative filtering approach to an article recommender system. In particular, it focuses on the collaborative filtering method for recommender systems due to the following reasons, (i) the models learned from the data of each user can be easily interpreted as personal or non-personal data facilitating our analysis, and (ii) the models themselves are explainable in comparison to other models like neural networks.

8.3.1 Data

The objective of recommender systems is to recommend items to users based on their explicit ratings or implicit feedback of previous items they have interacted with. These ratings are specified as a matrix R where each term rij corresponds to rating of item j by user i. The objective of collaborative filtering is to compre- hend certain character- istic preferences in end-users and find items with similar characteristics to recommend to the end-users. These characteristics are termed latent factors. For example, if we consider music as an item, latent factors can represent the genres of music. The latent factors can be obtained by factorizing the user-item rating matrix R into (i) user affinity matrix (U) and (ii) article affinity matrix (A) which we refer to as model parameters for the recommender system. Each row in these matrices specify the affinity of an article/user towards the latent factors. The data flows in this application is highlighted in a data flow diagram (DFD) in Figure .

8.3.2 Actors

The key actors in a federated collaborative filtering application are (i) end-user: users of the application who consume the articles recommended by the application, (ii) server admin: control over hyper-parameters for the training phase of the application,

(iii) model engineer and analyst: and (iv) content creator: the entity that produces the articles and is interested in improving consumption of the corresponding articles.

8.3.3 Processing

The processing of the aforementioned data is performed in two phases, (i) local processing of data on-device and (ii) aggregation of local updates on the server. The overall goal of the training phase is to minimize the following function denoting the error between the actual ratings and predicted ratings, where, λ is a regularization parameter.

$$f(u,a)^{(t)} = \sum_{(r_{ij} \in R)} (r_{ij} - u_i.a_j)^2 + \lambda(||U||^2 + ||A||^2)$$

On-device processing: Each device stores the history of the interaction of a user i and article a which comprises of reading time Tij and an explicit opinion Oij. The implicit rating of the article is thus considered as a function of these two values. The user-affinity matrix is computed for each user on-device as Ri. Furthermore, when a client device is chosen for training, it receives the latest version of the article affinity matrix. The training of the matrices occur as follows on each selected device. Each user updates the latent factors over k iterations of training.

$$u_i \leftarrow u_i + \alpha.a_j(r_{ij} - u_i.a_j) - \lambda.a_j \quad (2)$$

Similarly, the latent factors for each article are updated on-device as follows to compute the matrix A for user i.

$$a_j \leftarrow a_j + \alpha.u_i(r_{ij} - u_i.a_j) - \lambda.u_i \quad (3)$$

After the conclusion of k iterations on device, the locally updated article affinity matrix Ai is shared with the server.

On-server processing: The server is responsible for selecting client-devices satisfying certain criteria like (i) processing resource availability and (ii) device on-charge and (iii) Internet connectivity status for each training round. Once each training round is concluded, the server aggregates the local updates received from each participating client as follows. Assuming d devices participated in the training round, the update is computed as follows,

$$A = (\sum A_i)/d$$

8.4 Methodology

The research methodology it follows in this chapter is to study the outline of processing and storage in the context of FML applications in order to identify the critical aspects of GDPR that require attention to be complied with. To conduct the study, the author choose a case-study of a FML collaborative filtering application to deduce the implications of GDPR principles on this application. To this end, Privacy Impact Assessment (PIA) methodology proposed by CNIL is used to carry out a comprehensive study of the above application as described below.

8.4.1 Study of FML techniques and case study definition

In the first step, the general concept of FML techniques in light of the principles these techniques are based on including focused data collection and data minimization are studied. Furthermore, it concentrates on data distribution among clients, local and

global model parameters for FML algorithms as well as processing on-device and the server to identify areas that invoke attention. Taking into consideration the aforementioned areas, it defines a use-case of a FML based recommender system.

8.4.2 Privacy Impact Assessment

The PIA methodology is broadly classified into four phases, (i) context definition, (ii) study of fundamental rights and principles, (iii) risk assessment related to the security and privacy of the data and (iv) validation of the PIA. It is defined the context of the application and discuss the processing of personal data. Furthermore, the specific steps necessary to ensure compliance with the fundamental principles of GDPR are highlighted. The likelihood and severity of privacy risks are assessed to steer the research towards solutions to ensure the risks are treated adequately.

8.4.3 Identification of attention points and validation

Based on the privacy risks discussed in the previous step along with proposed mitigation of the risks , it is concluded that if the PIA is validated or conditional on improvement. Furthermore, it is needed to identify action required in the next iteration of the PIA study to further address the areas where residual risks might be unacceptable.

8.5 Analysis

The study on the background of FML applications as described in section 2, has led to the identification of the aspects that require attention in the context of GDPR. In this section, we analyze the use-case comprehensively with a PIA study in accordance to the second step of the methodology.

8.5.1 Study of context

The study of the application context is aimed at obtaining a reasonable understanding of data processing operations to be taken into consideration for our case study.

Outline of processing In accordance with the terminology described in the definition of the case study in 8.3, the outline of data processing is defined in the following steps and is illustrated in figure.

Step 1: Local data accumulation: For each user i, the time spent on interact- ing with each article j is accumulated as Tij which is measured in seconds. The opinion of each user i on an article already interacted with is also accumulated as Oij. This step is performed locally on each user device.

Step 2: Affinity matrices computation: The latent factor is received by each device as a hyperparameter from the server. For each user i, from the above variables, the user-article rating vector is computed denoted as Ri. This is computed as a function of Oij and Tij. Thereafter, the vector Rij is factorised into vector Ui and matrix Ai.

Step 3: Device participation: Each device periodically updates the models stored locally on the device with the updated global model on the server. The device sends a request to the server to download the updated model and initiate a training round. The server chooses suitable devices for the training round according to the following criteria (i) charging status of the device and (ii) connectivity to a non-metered WiFi connection.

Step 4: On-device training and local updates: On being chosen for a training round, the device undergoes training for iterations equal to the preset number of epochs, defined as a hyperparameter. This training phase results in computation of the updated

values of Ui and Ai on each device based on equations (2) and (3). Once computed the devices transmit the updated matrix Ai to the server.

Step 5: Global model update: The server can follow different approaches to accumulate the local updates from the devices. In this case-study we consider d to be a hyperparameter which signifies the minimum number of local updates required by the server to compute the global update. Once d updates are received by the server, the server computes the global update and stores the update into memory as the updated version of the global article affinity matrix model.

Step 6: Model verification: Following the update computed on the global model, a model analysts verifies the performance of the recommender system using this model. If the performance is satisfactory then the particular global model is chosen as the deployed model (Adeployed). Otherwise, the global model is discarded and removed from the server. This step can be performed for each global model update or more infrequently with a batch of global updates.

Data and processes. The processed data highlighted in the outline is classified into personal and non-personal data to define which data and corresponding pro- cesses fall under the scope of GDPR and which data and corresponding processes can be exempted.

Data. The data structures used in the data collection and processing phases are illustrated in table 2 along with the data store in which they are stored.

Processes. The processes involved in the recommender system applications are defined as follows:

P1: User input. Users can voice their opinions on the articles with a like/dislike for each article through the application interface.

P2: User interaction. User interaction process records the clickstream data from the user along with interaction time of the user, swipes and clicks on articles, among others.

P3: Application interface. The application interface offers the recommended articles to the user for interacting with the article. The application interface receives IDs of recommended articles from the recommendation engine (P20) which are then fetched from the article repository DS4 aided by the article re- quest handler P14.

P4: Data splitter. The data collected from the user-interface is split into train- ing and testing datasets and further shared with the data store access helper. This data is stored for further training and testing purposes.

P5: Data Store access helper. This data store access helper allows interaction with DS2 for updating and removing the training and testing datasets stored on-device.

P6: Time Normalization. The interaction times of users vary according to various factors, like the language of the article, their reading proficiency, age, among others. This process normalizes the interaction times of the articles for each user.

P7: Relevance calculation. In this process the normalized time and user opinion (if available) is used to calculate the relevance of a user i to an article j. P8: Data Store access helper This data store access helper is responsible for storing the model parameters, i.e. the user and article affinity matrices from P7 and to update them from P10.

P9: Step Update. The process P9 calculates the gradient of the error function from equation (1) for a given U and A. This gradient is the direction in which the next step towards the minima of the error function is taken.

P10: Step update. The gradient calculated by P9 is multiplied by the step size α and the values of A and U are updated accordingly through P8.

P11: Model updater. The model updater is used by the device to send the local model update computed over the latest training round to the server. There can be two variants of the algorithm, (i) where both the affinity matrices are transmitted in model updates or (ii) only the article affinity matrix is transmitted. The user device can also invoke participation in a training round through this process by requesting the updated global model.

P12: Article submission interface. The article creator uses the process to upload or submit content to the content provider's repository.

P13: Data store access helper. This data store access helper allows access the to the article repository to update the articles as well as fetch articles from the repository.

P14: Article request handler. This process handles 2 types of requests, (i) to handle the requests to fetch articles from the repository based on article ID and (ii) poll the repository for updates on the repository itself.

P15: Update aggregator. The local updates transmitted by the model updater are aggregated in this step. Since the aggregation can occur asynschronously or synchronously this process is separated from the federated averaging process.

P16. Federated Averaging. In this process the federated averaging of the ac- cumulated local updates is performed according to equation (4).

P17. Global Model Updater. The global model is updated with the output value from P16. This model is stored using the data store access helper separately from the de- ployed model.

P18. Data Store Access. Helper This datastore access helper updates the global and deployed models on the server and also offers read access to P19. This process is used by both the server admin to update hyperparameters and the model analyst to update the deployed model.

P19. Device Model Updater. If the deployed model changes after the federated aver- aging is computer or a model update is requested by a device to participate in a training round, this process transmits the deployed global model to the user device.

P20. Article Recommender. The article recommender process computes the ordered list of articles by their predicted ratings and recommends the user the top articles through P3, the user interface.

Data controller and processor.

According to the definition of data controller, the role of data controller is to determine the purposes and means of personal data. In this use-case thus, the role of the data controller is performed by the content provider since the goal of the content provider is to offer a better user- experience to the users to access their content as well as rec- ommend suitable articles to users to increase their revenue. The data processor on the

other hand performs the data processing according the purposes stated by the controller on its behalf. In this use-case, two scenarios can arise; (i) if the content provider opts for their own recommendation service developed in-house, there is no involvement of a data processor; (ii) the content provider outsources the recommendation service to another entity which then acts as the data processor.

Analysis of personal data.

In accordance with the types of data defined in table, the ones that fall under the category of personal data are as follows, (i) article interaction time (Tij), (ii) user opinion (Oij), (iii) user-affinity matrix (Ui) and (iv) user-article relevance matrix (Ri). The interaction statistics lie on- device throughout the lifecycle of the application. Similarly, the relevance matrix is also computed on device and stored on device to update the affinity matrices in each training round.

In the first variant mentioned in process P11, when the user-affinity matrix (Ui) is also transmitted along the article affinity matrix (Ai), then the update is considered as personal information and it is transmitted from the client device to the server. The user-affinity matrix can be either stored directly on the server or averaged among the users participating in the training round. The purpose of storing the user-affinity matrix in this variant to produce content based on the popular latent factors among users.

In both variants, the article affinity matrix which is non-personal data by itself, derived from personal data, is averaged using P16 and stored on the server until it is tested for deployment. If it is suitable to be deployed it is stored for a longer period as Adeployed or otherwise is discarded and removed. The averaged user-affinity matrix

can be stored and updated in each round to offer new users a baseline to recommend popular articles to address the cold start problem.

8.5.2 Study of fundamental principles

Legal basis It is important that all parties involved shall have a justification of lawfulness. Besides the criteria mentioned in Article 6 GDPR, all parties involved need to also have an exception defined under Article 9 GDPR for the processing of special categories of personal data (i.e., sensitive data).

For the recommender system, Article 6(f) GDPR can be invoked:

> "*processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which re- quire protection of personal data, in particular where the data subject is a child*."

Article 9(j) can be invoked:

> "*processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject."*

Purpose limitation. The recommender system will be used for research and statistics in patients' healthcare monitoring. For each new survey a data application will be filed as described in the process P3.

Assessment of controls protecting data subject rights. The recommender system will be used for research and statistics on pregnancy plus. Maxima Medical Centre Veldhoven PIA is applicable. Right to access; right to rectification and supplementation; right to object; right to the restriction of processing; right to data portability; right to be forgotten are listed.

8.5.3 Study of risks related to security of data

This step of the PIA methodology is aimed at understanding the measures that are already put in place by the data controller to "ensure a level of data security appropriate for the level of risk presented by processing personal data" according to Art. 32 of GDPR. In particular, the controls must ascertain that while processing the personal data, a natural person does not get access to or process personal data in a way that is not instructed by the controller or in means beyond the purpose specified.

Assessment of existing controls. The existing controls for assessment are primarily of three types, (i) controls bearing specifically on the data being processed, (ii) general controls regarding the system, (iii) organizational controls. In this paper we will primarily focus on the controls specific to the data being processed in the case study, since the information on the other two types of controls are not entirely available and are also not relevant to the study of the FML techniques. Furthermore archiving and paper document security are excluded since they are not applicable.

8.6 Conclusion

This section have evaluated the approach of federated learning in the specific context of the GPDR, which is generally applicable in the EU and is the most stringent set of data protection regulations that in practice applies to many – also non-EU organizations. Federated learning is well-placed to help enhancing the security and privacy measures and regulatory compliance of the GDPR, particularly in relations to data minimization and data protection by design requirements.

Firstly, FML per se is not immune from the GDPR application. Although federated learning avoids communication of large collections of raw data to the back-end server, we argue that the derived information (i.e. as contained within model updates) is to be considered pseudonymous data and therefore can not be excluded as such from GDPR application.

Secondly, the complexity of FML structure leads to complicate allocation of liabilities. It is important to identify the controllership before the processing of personal data starts. Requirements of fair processing information policy and processing agreement differ from jurisdiction to jurisdiction. The ubiquitousness of the end devices further amplifies this problem. Yet, FML is unlikely to exempt the data controller from such obligations.

Thirdly, based on our in-depth analysis of the applicable GDPR principles, the author argues that the GDPR requirements of visibility and transparency remain the most problematic to attain due to the lack of interpretability of the learned models in general, but this is also true for more centralized machine learning techniques.

Fourthly, new technology can lead to new openings for attacks. Concerns have already emerged regarding the robustness to attacks and failures, as well as the bias and fairness of the training data. It is important for policy makers to be sceptical towards the deployment of the FML. This does not mean unnecessary hindering of the technological progress. Rather, further observations should be invested.

The analysis demonstrates how legal certainties are blurring when operation difficulties lay between the practice and the rules on paper, particularly when an abrupt technology is emerging. The GDPR in general was not designed to significantly govern the development of data-driven innovations. Privacy preserving techniques like federated machine learning are encouraged for further deployment, for it might provide data subjects more control over their data while promoting data sharing in a secured way. Many properties of federated learning present its technological advances compared to the state of art. The follow-up experimentation of federated learning with a regulatory focus can provide insights for both computer scientists and policy-makers.