2019 Colloquium on International Law

15-16 August 2019

Discussion Paper:
Investing in Data – The Related Privacy Risks

by Stephen Kai-yi Wong, Barrister
Privacy Commissioner for Personal Data, Hong Kong, China
# Table of Contents

**Part I – The Emerging Data-driven Economy and the Economic Value of Data** ................................................................. 2

**Part II – Higher Privacy Risks** ................................................................. 4

(1) Ubiquitous and Covert Data Collection, Data Minimisation, Notice and Consent, and Individuals’ Control over Data ................................................................. 5

(2) Unpredictable Analytics, Purpose and Use Limitations ......................................................... 6

(3) Profiling, Risks of Re-identification, “Personal Data” and “Non-Personal Data” ........... 7

(4) Inaccuracy of Inferences, Unfair Discrimination and Other Negative Societal Effects ...... 8

**Part III – Heightened Data-related Risks to Organisations** ......................... 9

(1) Customer Defection .................................................................................. 9

(2) Financial Loss .......................................................................................... 10

(3) Penalties and Enhanced Regulatory Enforcement .......................................... 11

(4) Operations Disruption ............................................................................... 12

**Part IV – Possible Ways to address the Privacy Risks** ............................ 13

(1) Recasting the Scope of Data Protection Law .............................................. 14

   a. Collection of Personal Data ..................................................................... 15

   b. Revisiting the Definition of “Personal Data” for Enhanced Privacy Protection 19

(2) Accountability ......................................................................................... 32

(3) Data Ethics ............................................................................................ 41

(4) Cybersecurity ......................................................................................... 44

**Part V – Concluding Remarks** ................................................................. 45
Part I – The Emerging Data-driven Economy and the Economic Value of Data

We have witnessed parts of metamorphosis of the digital revolution and the evolution of data ecosystems over the last decade, whereby data is collected and analysed through a collection of infrastructure, analytics and applications. Prevalent use and diverse application of Big Data analytics, artificial intelligence (AI) and digital platforms in the modern ICT age have transformed life patterns and business developments. Undoubtedly we have marched into a data-driven economy.

2. In the data-driven economy, data, including personal data, is often considered as new oil\(^1\) and new currency.\(^2\) The analysis and use of data bring tremendous benefits to organisations (i.e. data users) and society as a whole. Jack Ma, the Executive Chairman of Alibaba Group, once said, “The main purpose of Tao Bao is not to sell goods, but to collect retail data and manufacturing data. The main purpose of Ant Financial is to establish the credit scoring system. Our logistics operation is not aimed at delivering goods, but to aggregate the data”.\(^3\)

---

\(^1\) The Economist, “The world’s most valuable resource is no longer oil, but data” (6 May 2017): https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data


\(^3\) People.cn, “Big Data is most valuable; Jack Ma asserts that data is the biggest fortune of Alibaba” (18 April 2015) (in Chinese only): http://media.people.com.cn/n/2015/0418/c14677-26865653.html
3. The economic value of data in this data-driven economy has somehow helped re-arranged the lead table of the market in that the most valuable companies nowadays are predominantly technology companies, in particular, the giants that amass huge amount of data. For example, in Hong Kong, the company with the largest market capitalisation is no longer a bank, but Tencent. In the US, the top 5 most valuable companies on Fortune 500 in 2018 were all technology companies, i.e. Apple, Amazon, Alphabet (parent company of Google), Microsoft and Facebook.

4. The data-driven economy is still emerging and globally, governments and organisations alike continue to invest in the development of data and data-related technologies, such as AI, Internet of Things, autonomous driving, etc. In Hong Kong, the Smart City Blueprint published by the Hong Kong Special Administrative Region Government in December 2017 shows the Government’s commitment to embrace innovation and technology to continuously inspire innovation and sustain economic development. The Smart City Blueprint contains a host of data-related initiatives, such as open data, Big Data analytics platform, eID, smart lampposts, intelligent transportation system, etc.\(^4\) They are all expected to be implemented in the coming few years.

5. As there is substantial value in data especially in mining of data, organisations have generally adopted the practice, if not the policy of “the more the better”. Let’s not forget that original personal data belongs to the

\(^4\) Smart City Blueprint: [https://www.smartcity.gov.hk/](https://www.smartcity.gov.hk/)
individuals (i.e. data subjects). Hence when organisations collect, store and use personal data, the data processing activities should be carried out under the laws are complied with, complemented by practicing the requisite respect for privacy of individuals and meeting their expectation.

6. Against the background, in the remaining parts of this paper, we are going to discuss:

- the major privacy risks in the data-driven economy (Part II);
- the operational, financial and reputational risks to organisations in the event of mismanagement of data (Part III); and
- possible ways to address the risks (Part IV), i.e.-
  - expanding the scope of data protection laws; and
  - promoting privacy accountability and data ethics.

**Part II – Higher Privacy Risks**

7. Revolutionary technologies such as Big Data analytics, AI and the rise of digital platforms in the modern ICT age have transformed not only the world that we live in but specifically the personal data privacy protection landscape.
(1) Ubiquitous and Covert Data Collection, Data Minimisation, Notice and Consent, and Individuals’ Control over Data

8. Powerful computing and low-cost storage enable organisations to collect data ubiquitously, covertly and automatically on a large-scale to feed into the pool of intelligence data. Individuals’ use of smart devices and systems like smartphones, wearable devices and social media networks may lead to release of their information in public domains, and tracking of their online and offline behaviours, sometimes without their knowledge. It also appears that the rise of data brokers facilitates covert circulation of data and undermines individuals’ ability to control how their data is collected, used and transferred.

9. Big Data analytics and AI thrive on enormous amount of data being collected to produce profound insights, contrary to the principle of data minimisation. Ubiquitous and covert data collection blurs the demarcation between private and public spaces, and challenges principles of informed consent, adequate notification and data transparency, diminishing individuals’ control over their data.
(2) Unpredictable Analytics, Purpose and Use Limitations

10. Big Data analytics “rely not on causation but on correlations⁵”, they may generate surprising revelations. All too accurate predictions may reveal sensitive and “core private information⁶”, causing embarrassment, psychological and reputational harms to individuals. In the classic privacy incident of Target, a U.S. retailer, analysis of purchase history had revealed a teenager’s pregnancy status well before her father knew it.⁷

11. Organisations may find themselves confused about the data they may potentially reveal, attenuating their ability to inform individuals of the risks during data collection. Individuals are unable to understand the full extent of privacy consequences and give informed consent.⁸

12. Unpredictable analytics may also challenge the principle of purpose limitation⁹ as organisations are unable to specify how they will use the data. Individuals are stripped of the ability to control “when, how and to what

---

extent information about them is communicated" and adopt privacy protections accordingly.

(3) Profiling, Risks of Re-identification, “Personal Data” and “Non-Personal Data”

Profiling often involves aggregating and combining datasets, at the expense of destroying individuals’ anonymity or revealing their personal information. In another classic privacy case, Netflix released anonymised movie ratings data of their subscribers to contenders in a competition to help improve their movie-recommendation algorithms; little did they know that the contenders were able to combine the anonymised ratings with other publicised data to re-identify the subscribers and to reveal the subscribers’ political views and religious beliefs.

13. Risk of re-identification undermines the fundamental distinction between “personal data” and “non-personal data”, shedding doubts on when data protection comes in for businesses to take up obligations and

---

11 There is no universal definition for “profiling”. Under Article 4(4) of the GDPR, “profiling” is defined as “any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person’s performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements.”
reciprocally, for data subjects to be entitled to their fundamental human right.

(4) **Inaccuracy of Inferences, Unfair Discrimination and Other Negative Societal Effects**

14. The reliability of Big Data analytics depends on the accuracy and representativeness of the inputted data, and the robustness of the algorithms, among others. Inaccurate inferences drawn and predictions made by Big Data analytics may conflict with the principle of data accuracy under data protection laws.\(^{15}\) When these outcomes are used for automated decision making in vital areas such as credit ratings, job prospects, eligibility for insurance coverage and welfare benefits\(^ {16}\), individuals’ socioeconomic status may be adversely affected, causing the individuals irreparable harms and impacting their fundamental human rights beyond privacy intrusions. An extreme reality is that people will be judged not on the basis of their actions, but on the basis of what all the data about them indicate their probable actions may be.\(^ {17}\)


15. Inputs of inaccurate or prejudicial data in algorithms may reinforce social prejudices and lead to discriminatory decisions.\(^\text{18}\) When algorithms and decision making processes are opaque and not transparent, individuals cannot exercise effective control over their own data, not to mention defence of their own interests and rights.

**Part III – Heightened Data-related Risks to Organisations**

16. The amplified privacy risks in the data-driven economy have drawn the concerns of the public and regulators, which if not addressed properly can cost organisations a great deal in their operations.

(1) **Customer Defection**

17. Various surveys and studies showed that data breaches and privacy scandals might well impair consumers’ trust and lead to “customer defection”. According to a recent survey by Symantec, 90% of people in Hong Kong were “more alarmed than ever” about their privacy.\(^\text{19}\) Another recent study by Microsoft revealed that in the event of negative trust


experience (e.g. breach of security and privacy), 53% of the consumers in the Asia Pacific region would either switch to another organisation, reduce the usage of the digital service (36%) or stop using the digital service altogether (34%).

(2) Financial Loss

18. Mishandling of personal data and loss of customers’ trust come at a high price, as the ensuing examples show. On 19 March 2018, the share price of Facebook dropped by 6% (or US$35 billion in value) when the Cambridge Analytica incident was revealed, in which the UK data analytics firm was accused of misusing the data of 87 million Facebook users to influence the 2016 U.S. Presidential Election. On 25 July 2018, Facebook announced along with its quarterly results that more than three million users in Europe had left the social media platform since the revelation of the Cambridge Analytica incident. The share price of Facebook dropped by 19% (or US$119 billion in value) on the following day.

---

19. In addition to the negative impact in the market, data mishandling could attract hefty monetary fines imposed by regulators under data protection laws, albeit fragmented. In October 2018, the Information Commissioner’s Office of the UK imposed a £500,000 fine to Facebook for failing to protect users’ personal data in the Cambridge Analytica incident\(^{23}\), being the maximum fine the UK regulator might impose under the Data Protection Act 1998. Under the new Data Protection Act 2018, which almost mirror-images the General Data Protection Regulation (GDPR) of the EU, the administrative fine that may be imposed by the UK regulator can be up to 4% of the global annual turnover of a company. In the event of Facebook, this is equivalent to US$2.2 billion based on its published revenue in 2018. It was recently reported that the Federal Trade Commission of the U.S. might impose civil penalty between US$30 to 50 billion on Facebook for breaching a consent agreement reached in 2011 with FTC in light of the Cambridge Analytica incident. Facebook is not the only company embroiled in regulatory enforcement. Indeed, the French data protection authority tested its new penalising power under GDPR by imposing a €50 million fine on Google in January 2019 for its lack of transparency, inadequate notice and lack of valid consent regarding the ads personalization.\(^{24}\)


20. Compliance with the data protection law of a single jurisdiction is a heavy one. Organisations having operation bases in multiple jurisdictions would find complying with the currently fragmented data protection laws an added burden. The updating, enhancement and alignment of data protection regulation across the globe would make their operations no less easier. Besides, as personal data is often associated with crimes and public order / security, whether and how to disclose individuals’ personal data for criminal investigations as and when requested by laws enforcement authorities is becoming a more trying decision than ever.

(4) Operations Disruption

21. Poor data governance including lack of adequate and sufficient data security may not only harm individuals’ privacy right and breach data protection laws, but also disrupt the operations of organisations, and may even put the sustainability of the operations into doubt.

22. In May 2017, the WannaCry ransomware attack broke out and spread all over the world quickly. The ransomware encrypted the data in the affected computers and asked for ransom from those impacted to decrypt. It was reported that at least 75,000 computers in 99 countries were been affected in a few days.\textsuperscript{25} In the UK, dozens of public hospitals and clinics were also affected, leading to cancellation of some operations and

appointments. Some infections were also reported in Hong Kong, although they did not appear to cause significant damage.

23. From time to time, organisations (social platforms that amass data capable of mobilizing the mass in particular) have experienced cyberattacks (including DDoS attacks) hampering their services.

24. In 2018, Hong Kong Computer Emergency Response Team Coordination Centre (HKCERT) received 10,081 security incident reports, a 55% increase year-on-year. HKCERT estimated that cyberattacks would continue to rise in 2019, with financially motivated cybercrimes becoming more targeted.

**Part IV – Possible Ways to address the Privacy Risks**

25. There is no silver bullet that can resolve all the privacy issues brought by new technology. Concerted efforts by policy makers, regulators, organisations and individuals are required. Below we will focus on

---


27 A distributed denial of service (DDoS) attack overwhelms a website or an online service with more traffic than the server or network can accommodate. The goal of a DDoS attack is to render the website or the online service inoperable. A DDoS attacked is primarily accomplished by remotely controlling a network of hacked computers or connected devices known as “botnet”. The botnet is then used to flood targeted servers or networks with more data than they can accommodate.

discussing two possible ways which are considered to be of significance in Hong Kong:

- recasting the scope of data protection law to cover appropriate data and data processing activities; and
- practising privacy accountability and data ethics to improve data governance of organisations.

26. Other possible ways which are outside the ambit of this discussion include: (i) enhancing the privacy right of individuals; (ii) imposing appropriate restrictions on the collection, processing, use and retention of personal data by organisations; and (iii) increasing the deterrent effects against non-compliance with data protection laws.

(1) Recasting the Scope of Data Protection Law

27. Most, if not all, data protection laws only apply to processing of “personal data”. Therefore, to expand the scope of data protection laws, one direct way is to expand the definition of “personal data”. However, before we deal with the definition of personal data, we will first discuss a related issue held by the Hong Kong Court of Appeal in Eastweek Publisher Limited & Another v Privacy Commissioner for Personal Data [2000] 2 HKLRD 83 (Eastweek case).
a. *Collection of Personal Data*

**Background of Eastweek Case**

28. The Eastweek case is the first landmark case of the data protection law of Hong Kong, i.e. the Personal Data (Privacy) Ordinance, Chapter 486 of the Laws of Hong Kong (PDPO) after its implementation in 1996.

29. The Eastweek case arose from a complaint lodged with the Privacy Commissioner for Personal Data (PCPD) back in September 1997. The complainant, while walking on the street, was photographed by a magazine photographer without her knowledge or consent. The photograph was subsequently published in the magazine accompanied by unflattering and critical comments on her dressing style. The publication caused embarrassment and inconvenience to the complainant amongst her clients and colleagues.

30. After investigation, the PCPD decided that the publisher of the magazine contravened Data Protection Principle (DPP) 1(2)(b) in Schedule 1 to the PDPO, i.e. unfair collection of personal data as the complainant’s photograph was taken without her knowledge or consent. The publisher appealed against this decision and the case eventually went all the way up to the Court of Appeal. In March 2000, by 2-1 majority, the Court of Appeal quashed the decision of the PCPD.
Implications of Eastweek case

31. The Court of Appeal held, inter alia, that the PDPO would not come into play at all if there was no “collection” of personal data at the first place. The Court laid down the necessary conditions for a “collection” of personal data:

(i) the collecting party must be compiling information about an individual; and
(ii) the individual must be the one whom the collector of information has identified or intends or seeks to identify.

32. The Court of Appeal held that there was no collection of personal data by the publisher because of “the complainant’s anonymity and the irrelevance of her identity so far as the photographer, the reporter and Eastweek were concerned”, and that the publisher “remained completely indifferent to and ignorant of her identity right up to and after publication of the offending issue of the magazine.”

33. The Court of Appeal also discussed a few provisions of the PDPO that could “only operate sensibly on the premise that the data collected relates to a subject whose identity is known or sought to be known by the data user as

__29 Eastweek Publisher Limited & Another v Privacy Commissioner for Personal Data [2000] 2 HKLRD 83, paragraph 14__
an important item of information”,\textsuperscript{30} such as an individual’s right to request for access to his own personal data (i.e. section 18 of the PDPO) and the use limitation of personal data (DPP 3 in Schedule 1 to the PDPO).

34. By interpreting the above judgement and obiter dicta, it appears that there is a third condition for “collection” of personal data which is closely related to the second condition, i.e.

(iii) the identity of the individual must be an important or relevant item of information to the collecting party.

Meaning of “Collect” revisited in the Digital Age

35. The incident of the Eastweek case occurred more than 20 years ago. It is indisputable that ICT developments and the ways of processing data (including personal data) have undergone revolutionary changes. The following examples would illustrate the need to have the law updated.

(i) CCTVs are widely used in Hong Kong and many other parts of the world for security purposes. Pursuant to Eastweek case, if a person installs CCTVs for monitoring the surroundings, as opposed to identifying individuals from the CCTVs images, is not considered to “collect” personal data. The PDPO will not come into play. Yet,

\textsuperscript{30} Eastweek Publisher Limited & Another v Privacy Commissioner for Personal Data [2000] 2 HKLRD 83, paragraphs 23-28
it is undeniable that storage of CCTVs’ images increases risks of privacy harms from data mishandling or breaches.

(ii) While data users getting hold of data may be completely indifferent to and ignorant of the identities of the individuals being tracked initially, it is nonetheless possible for them, or other parties, to subsequently identify individuals directly or indirectly and reveal details of their intimate lives by applying techniques of Big Data analytics and profiling. It is therefore desirable for the PDPO to come into play in the first place to ensure adequate protection of the data by the data users.

(iii) Big Data analytics and AI algorithms mostly function to predict trends to inform businesses’ decisions, as opposed to identifying individuals, though the operations necessarily involve use of personal data. If general privacy principles (e.g. transparency, data minimisation, use limitation, data security) do not apply as a result of not meeting the definition of “collect” under the PDPO, this may undermine individuals’ knowledge and control on the use of their personal data, and reduce the protection to the individuals.
b. Revisiting the Definition of “Personal Data” for Enhanced Privacy Protection

36. It has been accepted that a telephone number *per se* does not qualify as “personal data” as it does not enable the identity of individual to be ascertained. Nowadays, this interpretation seems to run contrary to widely held expectations that a telephone number, in particular a mobile telephone number, is more “personal” than the Identity Card number. Moreover, given powerful search engines and the prevalence of telephone number look-up services, it is possible to track down an individual by using his telephone number alone.\(^{31}\)

37. Email addresses with brief identifying information as usernames such as <skywong@gmail.com> has not been accepted as “personal data” either on the ground that the username *skywong* alone may not enable the identity of an individual to be ascertained, and the domain name *gmail.com* does not reveal any affiliations. Nonetheless, nowadays each email address is unique, and may be used to link up with a lot of activities of an individual. By the use of search engines and/or profiling techniques, it is likely that the identity of the owners of such email addresses can be uncovered.

---

38. Metadata\textsuperscript{32} and IP address\textsuperscript{33} do not relate directly to an individual, and hence may not be considered as personal data under the PDPO. However, nowadays, individuals’ interests, preferences and professions may be deduced from tracking their online browsing histories and their interactions with different social networking websites. Such tracking and inference are facilitated by the use of IP addresses and metadata.

**Metadata and IP address**

39. Metadata is generated as and when an individual uses technology, and enables one to know the who, what, where, when and how of a variety of activities. Metadata provides certain details about the creation, transmission and distribution of a message. Those details may include the IP address of the person sending an e-mail, time and location information, the subject, the addressee, etc. The analysis of metadata may reveal intimate or behavioral information about an individual.

40. An IP address is a numerical label assigned to an electronic device (e.g. computer, smartphone) that uses the Internet Protocol\textsuperscript{34} for communication. Online trackers may use metadata, including IP addresses, to track online activities of certain individuals.

\textsuperscript{32} Metadata is a set of data which describes and gives information about other data, such as time and duration of a phone call.

\textsuperscript{33} An IP address is a numerical label assigned to an electronic device, such as computer and smartphone, that uses the Internet Protocol for communication.

\textsuperscript{34} The Internet Protocol (IP) is the principal system of rules that allow two or more entities of a communications system to transmit information across network boundaries.
41. Nowadays, individuals’ interests, preferences and professions may be deduced from tracking their online browsing histories and their interactions with different social networking websites. By analysing the geolocation data derived from individuals’ mobile phones, the trackers may even get insight into some of the habits of the tracked individuals. It is now well accepted that these metadata and IP address under modern technology will help build up a detailed personal profile about an individual and has caused privacy concern. Not all the information collected for profiling directly identifies individuals. It then raises the issue about whether metadata constitutes “personal data” and whether there is sufficient protection to information privacy about an individual in light of the widespread tracking and profiling activities conducted.

Anonymisation and Pseudonymisation

42. When considering the issue of “personal data”, it is important to consider anonymisation and pseudonymisation as these measures would impact on the application of data protection laws to the underlying data.

43. Anonymisation and pseudonymisation are processes commonly used to get rid of or obscure “personal data” from information. It is trite that fully anonymised data (to the extent that a data user or anyone else is unable to re-establish the identity of any individual with other existing or future information on the individual) is not personal data.
Existing issue - Identifiability

44. The current definition of “personal data” under the PDPO is limited to information about an “identified” person. When an individual is “identified”, it means the individual is distinguished from other individuals. By contrast, pursuant to Recital 26 of the GDPR, an individual is “identifiable” when he/she can be identified, taking into account all the means reasonably likely to be used by a data user or a third party. In this regard, the concept of “personal data” under the PDPO has been challenged by the ever-changing and growing technology and brings loopholes for certain online monitoring or tracking activities which remain blatantly unregulated under the PDPO. For example, the existence of metadata and IP address might lead to unwanted compilation of a detailed personal profile about an individual and cause privacy concerns.

45. Concerns about the definition and scope of “personal data” under the PDPO were expressed in several decisions of the Administrative Appeals Board (AAB). In AAB no. 16 of 2007, the PCPD received a complaint relating to disclosure of an internet subscriber’s information, including the IP address of the computer that disseminated the information. The PCPD expressed the view that an IP address was information about an inanimate computer (i.e. not an individual) and the IP address alone could not reveal the identity of the relevant computer user. However, in certain circumstances, an IP address can constitute “personal data” when it is read together with other information (such as name, identity card number and
address), provided that the identity of an individual can be ascertained. While the AAB in this case ruled that the IP address together with other information collected (namely, the business address and email account which did not on their own reveal the identity of the internet subscriber) did not amount to “personal data”, the AAB expressed the view that when an IP address was coupled with verified personal information such as name, identity card number and address, it would constitute personal data.

46. In AAB no. 25 of 2008, it was decided that an email address in some circumstances could be information from which the identity of an individual might be directly or indirectly ascertained. However, the AAB did not accept that an email address comprising the initials of the complainant was, without any further information, sufficient to lead to the conclusion that the complainant’s identity would become reasonably ascertainable from such email address. In this case, the AAB decided that the email address in question was not the complainant’s personal data.

47. As noted from the above AAB decisions, it is not uncommon in our regulatory experience of the PCPD that disputes arose from the restrictive meaning of the definition of “personal data”. Given the rapid development in technology and Big Data analytics, the PCPD envisages that there will be increasing uncertainties about what constitutes “personal data” within the current definition.
Overseas position

48. In the last decade, changes have been observed in overseas data protection legislation. Necessitated by the rapid technological enhancement, overseas judicial authorities, especially those in EU, also support the application of data protection regime be extended to cover IP address or online identifiers.

European Union judicial decision and GDPR position

49. With regard to the application of EU Directive to IP addresses, the Court of Justice of European Union (CJEU) ruled on the position in two cases, namely Scarlet Extended (C-70/10) and Breyer (C-582/14).

50. In Scarlet Extended, the CJEU held that the IP addresses of internet users were protected personal data because they allowed users to be precisely identified. More elaborations as to when IP address will constitute personal data can be found in the Breyer’s case.\footnote{Mr Breyer (a visitor to a website) brought an action to restrain a website operator from storing his dynamic IP address claiming that the IP address constitutes his personal data. Although the dynamic IP address did not directly identify Mr Breyer, the operator was able to indirectly identify him from combining such information with the records about him (e.g. account details) held by his internet service provider (“ISP”). One of the issues was whether the IP address stored by the website operator constituted personal data when the identifying information was in the hands of ISP, a third party. The CJEU contrasted this with the Scarlet Extended case which concerned the collection and identification of users’ IP addresses by the ISP themselves (as opposed to website operators).} In Breyer, the CJEU held that a dynamic IP address recorded by a website operator (in this case a German government body) when an individual visited the website constituted personal data of that individual, provided that the website
operator had the legal means enabling it to identify the individual with the additional information in possession of the internet service provider. In effect, the CJEU accepted that in some circumstances, IP addresses would not necessarily be personal data (e.g. where it is illegal to identify the individual). This position has not been changed by the GDPR which only expressly refers to an online identifier as “personal information” if the online identifier can directly or indirectly identify an individual.

51. As compared with the previous definition under EU Directive, the GDPR introduced location data and online identifier as examples of personal information. Article 4(1) of the GDPR defines “personal information” as-

“any information relating to an identified or identifiable natural person; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number, location data, an online identifier, or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity of that natural person.”

52. To ascertain whether a natural person is identifiable, one should take into account all the means reasonably likely to be used by a data controller or another person to identify the natural person directly or indirectly. All objective factors such as the costs of identification, the time required for identification, the available technology at the time of processing and the
technology development will be relevant for determining whether certain means are reasonably likely to be used.\(^{36}\)

53. The recent GDPR has tackled this issue head-on by explicitly including identifiers such as location data and an online identifier as “personal data” if such identifiers can directly or indirectly identify an individual.\(^{37}\)

54. The GDPR explicitly states that data protection principles do not apply to anonymous information.\(^{38}\) On the other hand, personal data, which has undergone pseudonymisation, would still be regarded as information about an identifiable natural person and hence regulated.\(^{39}\)

55. The broadening of the scope of “personal information” under the GDPR has the effect of giving better clarity and certainty. Certain data which was not considered as “personal information” may now fall within the regulatory net of the GDPR with the advent of the digital age.

United Kingdom

56. The definition of “personal data” under section 3(2) of the Data Protection Act 2018 is the same as that under the GDPR, that is:-

\(^{36}\) Recital 26 of the GDPR  
\(^{37}\) Article 4(1) of GDPR  
\(^{38}\) Recital 26 of the GDPR  
\(^{39}\) Recital 26 of the GDPR
““Personal data” means any information relating to an identified or identifiable living individual. “Identifiable living individual” means a living individual who can be identified, directly or indirectly, in particular by reference to –

(a) an identifier such as name, location data or an online identifier, or

(b) one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of the individual.”

57. Furthermore, section 171 of Data Protection Act 2018 makes it an offence for a person knowingly or recklessly to re-identify information that is de-identified personal data without the consent of the controller (which is responsible for de-identifying the personal data). The person, who re-identified the data, has a statutory defence if he or she proves that:

(a) the re-identification is necessary for preventing or detecting crime;

(b) he or she had acted in the reasonable belief that either the data subject or the data controller concerned would have had consented to the re-identification; or

(c) the re-identification, in the particular circumstances, was justified as being in the public interest.
Comparision of the current regime under the PDPO with overseas’ regimes

58. The current definition of “personal data” under the PDPO includes information that relates to an “identified” person. In many other jurisdictions (e.g. Australia, Canada, EU, New Zealand and UK), the definition of “personal data” also includes information that relates to an “identifiable” natural person. The Canadian Court adopted the following test when deciding what constitutes identifiable information:

“Information will be about an identifiable individual where there is a serious possibility that an individual could be identified through the use of that information, alone or in combination with other available information.”  

59. Thus, an individual is “identifiable” if it is reasonable to expect that he or she can be identified from the information in issue or when it is combined with other information whereas an “identified” person refers to a person who is distinguished from other members of the group from the data in issue. So the definition of “personal data” under Australia, Canada, EU, New Zealand and UK has a broader scope than that under the PDPO.

60. To illustrate the difference between the terms “identified” and “identifiable”, the Breyer’s case under paragraph 50 is a good example to

---

40 Mike Gordon v Minister of Health and Privacy Commissioner of Canada 2008 FC 258
41 Abu Baker Munir, Siti Hajar Mohd Yasin and Md. Ershadul Karim, “Data Protection Law In Asia” on page 240
42 Supra note 19
show when a person is “identifiable” from certain information. Mr Breyer is identifiable by the operator of the website (which was a government body) by using the dynamic IP address as the website operator had the legal means enabling it to identify the individual with the additional information in possession of the internet service provider. Hence, the CJEU ruled that the dynamic IP address in the possession of the government body was Mr Breyer’s personal data. The UK’s Information Commissioner’s Office has discussed what identifiability is. For example, the name “John Smith”, by itself, may not always be personal data because there are many individuals bearing this name. However, if one is saying “John Smith with blonde hair and green eyes with a tattoo on his right arm, who works at the Post Office in Wilmslow”, there will be sufficient information to single out the particular individual John Smith. Hence he is “identified”, and John Smith as a piece of information is his personal data.

61. The phrase “by reference to” in the GDPR’s definition, followed by some examples including online identifiers and location data, provides a wider coverage than the existing wording “from which it is practicable for the identity of the individual to be directly or indirectly ascertained” as currently provided for under the PDPO.

62. The express inclusion of location data and online identifier as references in the definition of “personal information” under the EU’s GDPR and the UK’s Data Protection Act 2018 will have the effect of catching the types of data that is provided (e.g. through applications, social media
channels, etc.), observed (e.g. cookies, location services, etc.), derived (e.g. credit ratios, average purchase per visit) and inferred (e.g. credit score, behaviour predictions, etc.) in this digital age that enables tracking and profiling of an individual through his / her online activities.

Justifications for amendments

63. As noted from above, the definition of “personal data” under the PDPO conveys narrower meaning than the corresponding definitions under overseas privacy laws. The definition under the PDPO has not been revised since the enactment of the PDPO in 1996 when it was not envisaged that personal data of a person could take various forms as it is today, with which any other person would be capable of identifying the person.

64. In view of the rapid technological development and the closer human interaction with online activities, a more updated, wider scope of protection and clarity in the definition of “personal data” will be preferred to reflect what is already prevailing and will entrench in the digital age.

65. Explicit references to location data and online identifier such as IP address in the definition of “personal data” could address the privacy risks of possible surveillance and tracking of an individual’s activities through information collected from one’s online activities. In addition, the online world is borderless. The GDPR has already regulated websites that monitor behaviours of EU individuals. Including location data and online identifier
as references in the definition of “identifiable person” will bring Hong Kong’s data protection law in alignment with international standards.

66. The broadening of the scope of “personal data” can also give better clarity and certainty so that certain data which might not hitherto be considered as “personal data” can also fall within the regulatory net after the reform. With higher certainty of the definition of “personal data”, we expect arguments or debates on the scope of “personal data” would be minimised. This would hopefully facilitate compliance in the modern age.

67. It is also important that privacy risks arising from re-identification of personal data are adequately managed, mitigated and backed by legislative sanctions. Penalising unjustifiable act of re-identification (as the case in the UK) also provides a balance between legitimate processing of data and free flow of information.

68. The PCPD therefore sees merits in (a) expanding the meaning of the definition of “personal data” to include (i) information from which it is practicable to ascertain the identity of a person directly or indirectly, and (ii) information relating to an identifiable person, (b) providing a definition for “identifiable person” who must be a living individual and can be identified, directly or indirectly, by reference to an identifier such as name, location data or an online identifier, and (c) putting in place an additional measure to tackle individuals and organisations that circumvent anonymisation.
(2) Accountability

69. The PDPO does not contain explicit provisions on the accountability principle, although Data Protection Principles (DPPs) 2 and 4 require a data user to “take all practicable steps” to ensure compliance with data accuracy, data retention and data security requirements.

70. According to the European Data Protection Supervisor 43, accountability is “a common principle for organisations across many disciplines; the principle embodies that organisations live up to expectations for instance in the delivery of their products and their behaviour towards those they interact with.” The core requirements of accountability principle are to (a) put in place appropriate technical and organisational measures to ensure compliance with data protection laws and (b) be able to demonstrate the effectiveness of the measures when requested. 44

71. In February 2014, the PCPD published the “Best Practice Guide on Privacy Management Programme 45”, which manifested the accountability principle. The Hong Kong Government, and a number of public and private organisations, pledged to implement the privacy management programme (PMP).

---

43 The European Data Protection Supervisor is EU’s independent data protection authority, responsible for monitoring and ensuring protection of personal data and privacy when EU institutions and bodies process personal information of individuals.
45 The Best Practice Guide on PMP was revised in August 2018: https://www.pcpd.org.hk/pmp/guide.html
72. The PMP encourages data users to shift their paradigm in data protection from compliance to accountability and to embrace data protection as part of their corporate governance. It also encourages data users to apply the PMP as business imperative throughout the organisations. The PMP helps organisations manage compliance with the PDPO and build trust among customers and employees. Organisations implementing the PMP must, among other things:-

(a) solicit top management support to foster a privacy respectful culture and integrate it into the organisation’s governance structure;
(b) establish policies and programme controls giving effect to the legal requirements under the PDPO;
(c) conduct risk assessment (i.e. privacy impact assessment) when there are material changes in their data processing operations or the regulations;
(d) provide sufficient staff training in respect of data protection;
(e) make sufficient privacy communications to individuals, including staff members and customers;
(f) devise plans for responding to data breach and incident; and
(g) incorporate internal oversight and review mechanisms to keep the PMP on track and up to date.
73. As part of the PMP, an organisation should also appoint a data protection officer (DPO). The roles and responsibilities of the DPO should include:

(a) establishing and implementing programme controls;
(b) coordinating with other appropriate persons responsible for related disciplines and functions within the organisation;
(c) ongoing assessment and revision of programme controls;
(d) representing the organisation in the event of an enquiry, an inspection or in any investigation by the PCPD; and
(e) advocating personal data protection within the organisation itself.

Existing issues

74. As discussed above, increase in the use of ICT, such as Internet of Things, Big Data analytics and AI results poses risks to personal data privacy rights. It is well recognised that in the data-driven economy, development of laws lags behind the ever developing technologies and changing business models. Thus, it becomes crucial for data users in all sectors to adopt accountability principle in personal data protection in order to minimise risks, build and maintain a good reputation, and ensure the trust of citizens and consumers.
75. From the enforcement experience of the PCPD, many organisations in Hong Kong still take data protection as an afterthought rather than embed the data protection principles in their business processes. For example, among the investigation cases handled by the PCPD in the last three fiscal years (i.e. 2016-17 to 2018-19), about 40-50% of the cases were resolved as a result of remedial actions taken by the data users upon the advice of the PCPD. These actions could have been taken earlier and the incidents leading to the PCPD’s investigations could have been prevented if the data users had implemented accountability principle and embedded personal data protection in their business processes.

76. For example, in a recent incident involving intrusion into a customer database of a telecommunications company, the PCPD found that the company did not conduct a comprehensive and prudent review after system migration, leading to the failure to delete personal data in the customer database.\textsuperscript{46} In another incident involving loss of two notebook computers containing personal data of members and electors of an election committee\textsuperscript{47}, the data user was found to lack the requisite awareness and vigilance expected of it in protecting personal data, and lack clear data protection policies, procedures and measures to prevent the incident. These two incidents demonstrated the inadequacy of accountability in some Hong Kong organisations in personal data protection. Although the data users


concerned in the two incidents agreed to comply with the enforcement notices issued by the PCPD and take remedial actions, damage had already been done and the loss of personal data in both incidents could not be recovered.

77. Although the PCPD has been promoting the PMP since 2014, the number of entities (i.e. data users) pledged to implement the PMP is limited.⁴⁸ The figure seemingly indicates that only a limited number of data users voluntarily adopt the PMP.

Overseas position

78. Accountability is not a new concept. Back in 1980, the Organisation for Economic Co-operation and Development (or OECD) already set forth the principle of accountability in its privacy guideline, placing obligations on data controllers to adopt measures to give effect to the data protection principles.⁴⁹ Rapid technological development in the past decades calls into question the ability of data protection regulations to catch up and to effectively protect personal data privacy. In 2000, Canada became the first country to include the accountability principle in its data protection law. EU (as well as UK, which implements the GDPR) further developed the accountability principle in the GDPR by introducing the requirements of

---

⁴⁸ As of May 2019, 39 organisations have publicly pledged to implement the PMP, in addition to all the bureaux and departments of the Hong Kong SAR Government: [https://www.pcpd.org.hk/pmp/pledging.html](https://www.pcpd.org.hk/pmp/pledging.html)

data protection by design and by default, and data protection impact assessment, etc. In Australia and Singapore, although the word “accountability” is not mentioned in their data protection laws, various elements of the accountability principle are enshrined in the laws, which require data users to implement policies, practices and/or systems to ensure compliance with obligations under the data protection laws.

European Union

79. The GDPR expressly introduces the accountability principle.

80. Article 5(2) of the GDPR states that:-

“[t]he controller shall be responsible for, and be able to demonstrate compliance with [the principles relating to processing of personal data] (‘accountability’).”

81. Article 24(1) of the GDPR provides that:-

“the controller shall implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing is performed in accordance with this Regulation. Those measures shall be reviewed and updated where necessary.”
82. Article 25(1) of the GDPR also requires a data controller to adopt data protection by design and by default through a risk-based approach, “[t]aking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing.”

83. The GDPR also contains other provisions that specify measures to give effect to accountability. For example:-

(a) Article 35(1) of the GDPR requires a controller to conduct data protection impact assessment where processing “is likely to result in a high risk to the rights and freedoms of natural persons”;

(b) Article 37(1) of the GDPR requires a controller to designate a DPO under certain circumstances\(^{50}\);

(c) Articles 38 and 39 of the GDPR make further provisions on the position and functions of the DPOs. For example, (i) a DPO shall directly report to the highest management level, and shall not be dismissed or penalised for performing his or her tasks,\(^{51}\)

---

\(^{50}\) Appointment of DPOs is required pursuant to Article 37(1) of the GDPR where:-

“(a) the processing is carried out by a public authority or body, except for courts acting in their judicial capacity;
(b) the core activities of the controller or the processor consist of processing operations which, by virtue of their nature, their scope and/or their purposes, require regular and systematic monitoring of data subjects on a large scale; or
(c) the core activities of the controller or the processor consist of processing on a large scale of special categories of data pursuant to Article 9 and personal data relating to criminal convictions and offences referred to in Article 10.”

\(^{51}\) Article 38(3) of the GDPR
and (ii) the DPO shall advise the controller and monitor compliance with the GDPR.\textsuperscript{52}

\textbf{United Kingdom}

84. The UK’s Data Protection Act 2018 does not contain provisions that derogate from the requirements of the GDPR with regard to accountability. Therefore, the accountability principle under the GDPR equally applies in the UK.

\textbf{Justifications for amendments}

85. In 2010, Article 29 Working Party (\textbf{WP29}\textsuperscript{53}) noticed that the now-repealed EU Directive was not fully successful in ensuring that data protection requirements were translated into effective mechanisms for delivering genuine protection, especially if personal data protection was not part of an organisation’s shared values. WP29 recommended introducing accountability-based mechanisms as a way of encouraging data controllers to implement practical policies and procedures to protect personal data.\textsuperscript{54}

86. In Singapore, the Personal Data Protection Commission recognised the importance of data for innovation and economic growth. Meanwhile, the

\begin{flushleft}
\textsuperscript{52} Article 39(1) of the GDPR \\
\textsuperscript{53} The Article 29 Working Party was an independent European working party under the European Data Protection Board. It dealt with issues relating to the protection of privacy and personal data until 25 May 2018 (the date on which the GDPR entered into force): \url{https://edpb.europa.eu/our-work-tools/article-29-working-party_en}

\end{flushleft}
Commission also acknowledged that ubiquitous computing in the digital age had changed the nature of data collection from active interaction to passive one, rendering consent-based data protection laws ineffective in striking a balance between innovation and personal data protection.\textsuperscript{55} As a result, the accountability principle was called for.\textsuperscript{56}

87. As noted above, accountability in personal data protection is not a new concept. It is enshrined in the OECD Guidelines. Data protection laws in Australia, Canada, EU, Singapore and UK all have provided for some requirements of accountability.

88. The PCPD considers that accountability is a preferred and effective approach for strengthening personal data protection, in particular in the digital age, where data may be collected, processed and used in ways beyond the envisioning of lawmakers and regulators, rendering prescriptive rules or traditional data protection principles ineffective or incompatible.

89. The beauty of the accountability principle is that it is scalable. Data users are provided with the flexibility to identify and evaluate the privacy risks of their data processing activities, as well as to develop appropriate measures to address the particular risks, without significantly compromising their business objectives and legitimate interests. Moreover, the


\textsuperscript{56} The Personal Data Protection Commission, “Response to Feedback on the Public Consultation on Approaches to Managing Personal Data in the Digital Age”, 1 February 2018: https://www.pdpc.gov.sg/Legislation-and-Guidelines/Public-Consultations#ACTR1
accountability principle emphasises development and implementation of policies and measures to ensure compliance with data protection laws and prevent (rather than remediate) data breach incidents. Therefore, the accountability principle is effective in protecting personal data privacy of individuals.

(3) Data Ethics

90. Globalisation of commercial and data processing activities means that businesses now have to comply with multiple regulatory regimes. Given the uneven data protection landscapes across the globe, due diligence has to be exercised by businesses to ensure that they do not fall short of the legal requirements in the jurisdictions in which they have operation. This mission is proved to be challenging to businesses by a few high-profile data breaches in recent years in which individuals in multiple jurisdictions were affected, leading to the probes by multiple regulators at the same time.

91. Meanwhile, despite relentless attempts to revamp data protection laws, the development of technological innovation invariably outpaces regulatory efforts. As a result, meeting regulatory requirements alone would not be effective enough to adequately protect and live up to individuals’ expectations in privacy protection, especially in jurisdictions which lack robust deterrent sanctions.
92. Organisations in general that amass and derive benefits from personal data should be held to a higher ethical standard that meets the stakeholders’ expectations alongside the requirements of laws and regulations. Reiteration by enterprises about their compliance with regulatory requirements does not spare them from the devastating damage to their hard-earned corporate reputation and consumers’ trust. In this regard, data ethics could bridge the gap between legal requirements and the stakeholders’ expectations. It is time for data users, as well as regulators, to promote and practise data governance and ethics.

93. The PCPD commissioned a study on data ethics in 2018, with a view to drawing up recommendations on what an Ethical Data Stewardship framework should look like, and providing tools for organisations to achieve fair and ethical processing of personal data. The study report\(^{57}\) was published in October 2018, recommending that organisations who conduct advanced data processing activities should implement ethical data stewardship by adhering to the three core ethical values – respectful, beneficial and fair – and conducting ethical impact assessments, amongst other things. The objective of ethical data stewardship is to ensure that the impact on the interests, rights and freedoms of all stakeholders are duly considered and addressed in data processing activities.

\(^{57}\) See the study report “Ethical Accountability Framework for Hong Kong, China (2018)” and “Data Stewardship Accountability, Data Impact Assessments and Oversight Models - Detailed Support for an Ethical Accountability Framework (2018)” on the PCPD’s website: https://www.pcpd.org.hk/english/resources_centre/publications/surveys/surveys.html
94. Data ethical values focus on fairness, respect and mutual benefits. In practical terms, they involve genuine choices, meaningful consent, absent of bias or discrimination, and fair value exchange between individuals and organisations.

95. It is encouraging that in early May 2019, the Hong Kong Monetary Authority issued a circular to all authorised institutions in Hong Kong, urging them to adopt and implement the Ethical Data Stewardship framework of the PCPD in the collection and use of personal data in Fintech development. 58

96. At the 40th International Conference of Data Protection and Privacy Commissioner held in Brussels in October 2018, a Declaration on Ethics and Data Protection in Artificial Intelligence was passed, of which the PCPD was one of the co-sponsors. The Declaration sets out six guiding principles to preserve human rights in the development of artificial intelligence, i.e. (i) fair; (ii) continued attention and vigilance; (iii) transparency and intelligibility; (iv) ethics by design; (v) empowerment of every individual; and (vi) reducing and mitigating biases or discriminations. A new permanent working group has been set up pursuant to the Declaration to further promote and develop the six guiding principles across the globe. Being one of the co-chairs of the permanent working group, the PCPD will continue to work closely with multi-stakeholders, both at home and abroad, to nourish a culture and environment that respects privacy. It is hoped that a

proper balance would be struck between privacy protection and free flow of information that will facilitate and not hinder technological innovation.

(4) Cybersecurity

97. The privacy, financial, operational and other risks discussed at the beginning of this paper are brought by the use of data and data technologies. Therefore, cybersecurity is something that cannot be ignored when addressing those risks. Nowadays, a lot of personal data are stored in information systems or cyber spaces, the absence of cybersecurity will certainly undermine privacy protection. Although the PCPD is no expert of cybersecurity, organisations are strongly advised to follow the guidelines issued by other relevant authorities to strengthen their cybersecurity. In December 2018, the National Institute of Standards and Technology (NIST) of the USA updated its “Risk Management Framework for Information Systems and Organizations”\(^\text{59}\) to help organisations more easily meet this goal. Steps for risk management as suggested by NIST include:

(a) Prepare tasks – To carry out essential activities to prepare the organisation to manage its security and privacy risks; examples of the essential activities include (1) identifying and assigning key roles for executing the Risk Management Framework to individuals; (2) identifying the types of data that will be

processed; and (3) identifying the stakeholders having an interest in the system.

(b) Categorise tasks – To inform organisational about the relevant risk management processes and tasks by determining the adverse impact with respect to data security incidents.

(c) Select tasks – To select, tailor, and document the controls necessary to protect the information systems.

(d) Implement – To implement the controls in the security and privacy plans.

(e) Assess – To determine if the controls selected for implementation are implemented correctly, operating as intended, and producing the desired outcome.

(f) Authorise – To provide organisational accountability by requiring a senior management official to determine if the security and privacy risk is properly addressed.

(g) Monitor – To maintain an ongoing situational awareness about the security and privacy posture of the information system and the organisation in support of risk management decisions.

**Part V – Concluding Remarks**

98. Risk assessment is part and parcel of the investment game. Privacy risk management controls the data investment game.
99. Revolutionary and innovative developments in ICT like Big Data analytics, AI and Fintech present challenges to the regulatory strengths and effectiveness of existing data protection laws. These developments set forth races by data protection regulators to come up with novel regulatory solutions, including expansion of the scope of data protection laws.

100. To give effect to the legal requirements, there is also an expectation of comprehensive, effective and evidenced privacy compliance policies and programmes being put in place, relevant and scalable for the businesses concerned, as well as demonstrable internally and externally. This legitimate expectation comes from both the customers, who are the data subjects, and the regulators.

101. In face of rapid technological developments threatening to annihilate robust law amendments efforts, it is suggested that data protection laws should embody privacy accountability. Comprehensive, flexible and responsibility-based, accountability is the crucial framework to strike a balance between data protection and facilitation of businesses and innovation. Regulators should further explore the possibility of data ethics as an essential solution to bridge the expectation gaps between organisations and individuals.

102. The idea of good data stewardship and governance, or accountability has also been incorporated in the new laws and regulations of many jurisdictions, notably the EU GDPR implemented in May 2018.
Notwithstanding that similar principle of accountability is yet to be provided for in the law of Hong Kong, businesses in Hong Kong should be well poised to adopt proactive data management as corporate digital values, ethics and responsibilities in this era of data driven economy, translating legal requirements into risk-based, verifiable and enforceable corporate practices and controls, to address regulatory changes worldwide; enable updated business models, digitalisation, globalisation and ensure data protection, sustainability and trust.

103. For organisations, given the privacy, financial and other operational risks that may arise from the use of modern ICT, proper policies and measures are expected to be implemented to identify, assess and mitigate privacy and data security risks.