Privacy Commissioner Prize in Privacy and Data Protection Law 2021-2022

Topic 9: Other topic

Identify the privacy and data protection challenges posed by the increasing use of Big Data and critically examine how the laws in Hong Kong can meet these challenges, pointing out any inadequacies. Based on your analysis of the Hong Kong situation and learning from the laws and experiences in other jurisdictions, what recommendations would you make to the Hong Kong government as regards the legislative intervention and administrative measures that need to be pursued to secure enhanced privacy and data protection.

(The essay was previously submitted as course assignment for LLAW/JDOC6046 Privacy and Data Protection. This submission is an amended version.)

Name: Man Sum Yi
Word Count: 4707 words

1. Introduction

Data, especially when aggregated, can reveal a lot about a person. Technology advancement facilitates data collection, consolidation, processing, analysis and interpretation. Profiling, big data analytics and artificial intelligence ("**AI**") are frequently used together to discover new patterns and make useful predictions. Big data encompasses enormous value for driving economy, innovation, productivity and efficiency.

This paper first discusses the privacy and data protection challenges posed by the increasing use of big data, and examines whether the law in Hong Kong is adequate in meeting those challenges. Legislative interventions and administrative measures are proposed with reference to overseas experiences.

2. Overview of big data and privacy law

Big data refers to a massive amount of structured and unstructured data[1], generated by people, machines and devices.[2] Big data does not always involve personal data, where identified or identifiable natural persons are not involved[3], such as monitoring natural phenomena or overseeing the manufacturing processes. Nevertheless, under the new big data era, data have become a coveted resource for companies. Protecting online personal information is becoming more challenging than before. Some large corporations, such as Facebook and Google,

[1] Office of the Privacy Commissioner for Personal Data. "Privacy Protection and Data Governance in the Internet of Things" 5 June 2019. https://www.pcpd.org.hk/english/news_events/speech/files/AcademyofLaw_0605.pdf

[2] Gupta, Nirmal Kumar, and Mukesh Kumar Rohil. "Big Data Security Challenges and Preventive Solutions." *Data Management, Analytics and Innovation,* Springer Singapore, Singapore, 2019, p. 285–299.

[3] Definition of personal data under GDPR.

do not only produce data, but also store and manage data.[4] While big data creates a variety of opportunities, one's privacy could be put at stake due to the volume, variety, veracity, value and velocity of data flow.[5]

Big data is more prevalent in our lives than we may imagine. Spotify produces weekly personalized playlists by tracking users' preferences[6], and online retailers keep track of customers' browsing history and purchase records to suggest new items for purchase. Lately, scholars have been using big data to track the spread of COVID-19, by combining real-time aggregated population flow data with number and location of confirmed cases to develop a new risk assessment model which can identify high-risk locales at an early stage.[7]

The Personal Data (Privacy) Ordinance (Cap. 486) ("**PDPO**") establishes the legal framework concerning privacy and data protection in Hong Kong. The enforcement of PDPO is overseen by the Office of the Privacy Commissioner for Personal Data ("**PCPD**"). Data users must comply with the six data protection principles ("**DPPs**"), which represent the core requirements in PDPO. Privacy and data protection challenges will be discussed based on the relevant DPPs.

---

[4] Tene, Omer, and Jules Polonetsky. "Big Data for All: Privacy and User Control in the Age of Analytics." *Northwestern Journal of Technology and Intellectual Property*, vol. 11, no. 5, 2013, p. 239.

[5] Commonly known as the 5V's of big data. Hitzler, Pascal, and Krzysztof Janowicz. "Linked Data, Big Data, and the 4th Paradigm." *Semantic Web*, vol. 4, no. 3, 2013, p. 233-235.

[6] Marr, Bernard. "The Amazing Ways Spotify Uses Big Data, AI And Machine Learning To Drive Business Success" Forbes, 30 October 2017. https://www.forbes.com/sites/bernardmarr/2017/10/30/the-amazing-ways-spotify-uses-big-data-ai-and-machine-learning-to-drive-business-success/?sh=f2b92ae4bd2f Accessed on 13 November 2021.

[7] Jia, J.S., Lu, X., Yuan, Y. et al. "Population flow drives spatio-temporal distribution of COVID-19 in China." *Nature* 582, 2020, p.389–394.

3. Privacy and data protection challenges posed by the increasing use of Big Data

Alan Westin defined privacy as "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others".[8] Maintaining secrecy, anonymity, and solitude of personal data are important elements.[9] Data protection involves securing the data throughout the entire data life cycle.

(a) Excessive collection, retention and privacy invasion

The greater the scope of data collection, the more predictions could be made. The perceived latent opportunities in big data encourage data users to collect as much data as possible from multiple sources and retain them as long as possible for future uses. Retaining data longer than necessary is a contravention of DPP2.

With the advent of the Internet of Things ("**IoT**"), human inputs in devices and data collected by sensors will be shared with the Internet networks. Those devices commonly ask for personal information, such as name, address, age and health information, which may be considered excessive under DPP1. Data collection may be passive, some devices track movement of individuals with sensors, and some deploy facial recognition technology. Therefore, data users can easily develop comprehensive profiles regarding a person's life and preferences.

Big data may reveal details about an individual's intimate life, and bring

[8] Westin, Alan. *Privacy and Freedom*. New York: Atheneum, 1967. p. 7.
[9] Gavison, Ruth. "Privacy and the Limits of Law." *The Yale Law Journal* 89.3. 1980, 421-471, p. 428.

psychological harm such as distress and embarrassment. Researchers found out that by analysing "likes" on Facebook, one can know a person better than his family members and spouse.[10] In the US, Target analysed its customers' purchase records, identified around 25 products that could be used for predicting pregnancy, and started sending pregnancy-related advertisements to customers with high pregnancy scores. A high school girl's pregnancy status was thus revealed to her father.[11]

One of the privacy concerns of data subjects would be mass surveillance. Edward Snowden revealed in 2013 that the US National Security Agency collected millions of telephone metadata records indiscriminately, regardless of whether citizens had any suspected wrongdoing.[12] This illustrated that governments could abuse big data to perform mass surveillance, causing severe privacy invasion into daily lives of ordinary citizens. Meta has recently announced that they will shut down Facebook's face recognition system, in response to growing privacy concerns.[13]

(b) Re-identification of individuals

Traditionally, privacy protection is based on anonymity. Even when anonymisation and de-identification techniques are applied, it is becoming easier to infer a person's identity by linking datasets. People can make use of the publicly available data, such as information on social media, put pieces together

[10] Ahmed, Murad. "Facebook understands you better than your spouse" *Financial Times*, 13 Jan 2015. https://www.ft.com/content/3dfa397c-9a73-11e4-8426-00144feabdc0. Accessed on 20 November 2021.
[11] Duhigg, Charles. "How Companies Learn Your Secrets." *The New York Times Magazine*, 16 Feb 2012. http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html. Accessed on 20 November 2021.
[12] Greenwald, Gleen. "NSA collecting phone records of millions of Verizon customers daily" *The Guardian,* 6 June 2013. https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order. Accessed on 8 November 2021.
[13] Meta. "An Update on Our Use of Face Recognition" 2 November 2021. https://about.fb.com/news/2021/11/update-on-use-of-face-recognition/. Accessed on 8 November 2021.

to "re-identify" individuals. Larger databases and more powerful analysis techniques facilitate the re-identification process.[14]

Earlier in 2008, two US researchers had successfully unveiled people's political orientation, religious views and sexual orientation by applying de-anonymisation technology to the anonymised 500,000-record Netflix's Prize dataset.[15] Professor Sweeney from Harvard University found out that 87% of US individuals could be identified if zip code, gender and date of birth data are provided. She successfully re-identified government-released hospital records by using a voter database.[16]

Although data users may think that anonymised data is safe to be released, one can never know how much data another person holds and the impact if datasets are merged.

(c) Lack of transparency

Opaqueness in big data analytics limits the ability of individuals to exercise effective control and defend their own interests. DPP5 requires data users to ensure openness of personal data policies and practices. However, it seems difficult to explain to data subjects every logic or rationale behind the predictions made with big data. The US White House report published in 2014 indicated that individual autonomy is lost in the "impenetrable set of algorithms".[17] Algorithms

[14] Government Office for Science, UK Government. "Artificial intelligence: opportunities and implications for the future of decision making" 2015. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/566075/gs-16-19-artificial-intelligence-ai-report.pdf. Accessed on 27 October 2021.

[15] Narayanan, Arvind, and Vitaly Shmatikov. "How to break anonymity of the netflix prize dataset." arXiv preprint cs/0610105, 2006.

[16] Perry, Caroline. "You're not so anonymous" *The Harvard Gazette*. 18 Oct 2011. https://news.harvard.edu/gazette/story/2011/10/youre-not-so-anonymous/

[17] Executive Office of the President, the White House, Washington. "Big Data: Seizing Opportunities,

applied to analyse data are often complicated, even some data users may not fully understand the algorithm behind. Nevertheless, it could be argued that the law does not require a technical explanation of all details in an algorithm, but for what purpose.

(d) Profiling, Inaccurate inferences and Discrimination

Profiling refers to "an automated processing of personal data and using data to evaluate personal aspects".[18] DPP1 and DPP3 regulate data collection and use respectively, and profiling can be regarded as a stage in between collection and use.[19] Although PDPO is designed to regulate the entire data life cycle, there is no specific provision regulating profiling. It is uncertain whether the results derived from big data analytics fall under the purview of PDPO. While such data may not directly involve personal data, the results may have a significant impact on individuals.

Inferences drawn from big data may not necessarily reflect causation, but correlations.[20] Google Flu Trends is a prime illustration where big data predictions may not work. In 2013, an article in Nature revealed that Google had tremendously overestimated flu trends.[21] Inaccurate inference of personal data conflicts with the principle of accuracy under DPP2. Hence, unintended discrimination could have resulted. Although big data analytics are derived from

---

Preserving Values" May 2014, p.10.
https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf. Accessed on 1 November 2021.

[18] Article 4(4) of GDPR.

[19] Wong, Stephen Kai-yi. "Engineering Privacy through Accountability" *66th ABA Section of Antitrust Law Spring Meeting* 11 April 2018.
https://www.pcpd.org.hk/english/news_events/media_statements/files/PCPDABA2018.pdf

[20] Rubinstein, I. S. "Big Data: The End of Privacy or a New Beginning?" *International Data Privacy Law*, vol. 3, no. 2, 2013, p. 74–87.

[21] Butler, D. "When Google got flu wrong." *Nature* 494, 2013, p.155–156.

objective algorithms, humans who handle the data are often associated with bias. Examples of bias include availability bias where an analyst tends to rely on the most readily available information, and confirmation bias where an analyst seeks to confirm own hypothesis with data.[22]

Individuals may be the victims of incorrect automated decision-making, including creditworthiness, employment opportunities, insurance coverage or social welfare benefits.[23] Amazon abandoned its AI recruiting programme because it was discovered that the programme was designed based on Amazon's old, biased hiring data, leading to gender bias against women.[24] Contrary to DPP3, bias and discrimination may prejudice the proper use of data. Individuals may be judged on the results of big data analytics rather than the data collected.

(e) Data breaches

Data may be stolen, lost or exposed to unauthorised persons if they are not sufficiently protected. The leakage of big data may cause more devastating consequences due to the quantity of data. A data breach may amount to a contravention of DPP4, which requires data users to take reasonable practical steps to secure data. Where big data is stored physically, in a warehouse or on portable devices, the physical condition of storage may undermine security. Tangible security measures such as access control, verification and authorisation steps should be adopted; and intangible security measures include encryption, firewalls, ant-virus software, and password. Big data allures hackers, because less

---

[22] Legard, Ryan. "How to Manage Big Data Issues in a Project Environment." *Data Analytics in Project Management.* Auerbach Publications, 2018, p. 115-132.
[23] Rubinstein, I. S. "Big Data: The End of Privacy or a New Beginning?" *International Data Privacy Law*, vol. 3, no. 2, 2013, p. 74–87.
[24] Dastin, Jeffery. "Amazon scraps secret AI recruiting tool that showed bias against women" *Reuters*, 11 October 2018. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G. Accessed on 15 November 2021.

cost is required for a single successful hack, and a larger amount of information can be obtained. Malware may be used to steal data and track location information.[25]

HP revealed that IoT devices are extremely vulnerable to attacks. 70% of IoT devices transported data to the internet and local network without encryption, 60% of devices did not use encryption when downloading software updates, and 80% of IoT devices failed to require passwords of sufficient strength.[26]

To prevent data breaches, competency, prudence and honesty are qualities required in persons having access to data. In 2018, an employee of Cambridge Analytica blew the whistle on the company that over 87 million Facebook users' information had been harvested without consent through an external app since 2015.[27] The data was sold to assist Donald Trump's 2016 presidential campaign. The CEO of Facebook subsequently apologised for the scandal and was tested in Congress. In 2019, a US$5 billion fine was imposed by the Federal Trade Commission.

(f) Covert data collection challenges individual control over data

Data subjects could not protect their data without having sufficient control over what and how data is being processed. Big data collection is voluminous and ubiquitous. The increasing use of smart devices such as wearable devices and

[25] Nirmal Kumar Gupta. "Addressing Big Data Security Issues and Challenges." *International Journal of Computer Engineering & Technology* 9.4, 2008, p.229-237.
[26] HP. "HP Study Reveals 70 Percent of Internet of Things Devices Vulnerable to Attack" 29 July 2014. https://www.hp.com/us-en/hp-news/press-release.html?id=1744676#.YY3P-9ZBx-V. Accessed on 10 November 2021.
[27] Ma, Alexandra and Gilbert, Ben. "Facebook understood how dangerous the Trump-linked data firm Cambridge Analytica could be much earlier than it previously said. Here's everything that's happened up until now." Business Insider, 24 Aug 2019. https://www.businessinsider.com/cambridge-analytica-a-guide-to-the-trump-linked-data-firm-that-harvested-50-million-facebook-profiles-2018-3 Accessed on 10 November 2021.

smartphones allows more personal data to go public. The demarcation between public and private spaces becomes blurred. With big data, value of data extends to secondary purposes and uses.[28] Data may be collected, combined and used without the knowledge of data subjects, amounting to the contravention of DPP1 and DPP3 and exceeding a person's reasonable expectation. For example, IoT operates round the clock, and there is no interface to give notice, and no means to accept consent or opt-out. IoT data may not be processed through human beings, but is directly sent to the cloud for processing.

Big data also challenges the notification requirement under DPP1(3). At the time of data collection, organisations may not be aware of the potential use of data. It is a challenge to offer meaningful notice and prevent "notice fatigue". As Professor Ohm pointed out, big data "thrives on surprising correlations and produces inferences and predictions that defy human understanding…how can you provide notice about the unpredictable and unexplainable?". [29] Some suggested that the traditional notice and consent mechanism may not be suitable for the big data era.[30]

4. Inadequacies of Hong Kong Law

PCPD has published different guidelines outlining best practices in handling personal data. In August 2021, PCPD further published a "Guidance on the Ethical Development and Use of Artificial Intelligence" in view of the increasing amount of personal data involved in big data analytics and the use of

[28] Viktor Mayer-Schönberger and Kenneth Cukier. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. John Murray, 2013, p. 1–242.
[29] Ohm, P., "Changing the Rules: General Principles for Data Use and Analysis," *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. 2014, p.100.
[30] Munir, Abu Bakar and Mohd Yasin, Siti Hajar and Muhammad-Sukki, Firdaus. "Big Data: Big Challenges to Privacy and Data Protection". *International Scholarly and Scientific Research & Innovation* 9(1) 2015. May 21, 2015.

AI.[31] Surveying the law in Hong Kong, inadequacies will be discussed.

(a) Unclear definition of "personal data"

Under s.2(1) PDPO, "personal data" is defined as "any data relating directly or indirectly to a living individual from which it is practicable for the identity of the individual to be directly or indirectly ascertained, and in a form in which access to or processing of the data is practicable". In other words, "personal data" needs to satisfy three criteria, attribution, identification and retrievability. The definition under PDPO fall short in protecting personal data due to technological developments. The EU General Data Protection ("**GDPR**") expressly recognises "online identifiers" as "personal data"[32], but there is no equivalent provision in PDPO. For instance, IP address is personal data under GDPR, whereas an Administrative Appeals Board ("**AAB**") case held that IP address did not constitute personal data unless it was coupled with verified personal information.[33] It is noteworthy that the AAB case was decided 15 years ago, hence it might have underestimated the advancement of big data analytics.

Big data compromises both content data and metadata.[34] Content data refers to the subject matter in communication, while metadata provides information about data, such as location, time, and duration of communication. In the Australian case *Privacy Commissioner v Telstra Corporation Ltd*[35], despite

---

[31] Office of the Privacy Commissioner for Personal Data. "Guidance on the Ethical Development and Use of Artificial Intelligence" August 2021. https://www.pcpd.org.hk//english/resources_centre/publications/files/guidance_ethical_e.pdf. Accessed on 2 November 2021.
[32] Article 4 of GDPR.
[33] *Shi Tao v The Privacy Commissioner for Personal Data* (2007), Case No. 16/2007 of the Administrative Appeals Board.
[34] Office of Privacy Commissioner of Canada. "Metadata and Privacy: A technical and legal Overview" October 2014. https://www.priv.gc.ca/media/1786/md_201410_e.pdf. Accessed on 1 November 2021.
[35] *Privacy Commissioner v Telstra Corporation Ltd* [2017] FCAFC 4.

the fact that a journalist's daily routines, work and living locations could be accurately guessed by analysing his telecommunications metadata, it was held that metadata was not data "about" the journalist, but it was about services of Telstra. This case may give rise to debates regarding the scope of personal data, which is not clearly addressed in PDPO. However, such an argument can be said to be purely academic, as it is undeniable that metadata can be extremely intimate and can reveal a great extent of details of personal lives.

(b) Unclear definition of "collection"

Big data users may easily argue with the *Eastweek Publisher Ltd. & Anor v PCPD* [36] case that there is no collection of personal data. The case laid down conditions for the collection of personal data, (1) the collecting party must be compiling information about an individual, (2) the individual must be the one whom the collector of information has identified or intends or seeks to identify, and (3) the identity of the individual must be of importance to the collecting party. By asserting that big data analytics and AI algorithms are used to predict trends, as opposed to identifying individuals or retrieving their personal information, the absence of "collection" renders PDPO inapplicable. The data users may "remain completely indifferent to and ignorant of an individual's identity" from the beginning till the end of the data cycle, while third parties can still identify individuals directly or indirectly by adopting big data analytics techniques. This case was severely criticised by Berthold and Wacks.[37] As there is no requirement under PDPO that "identification" must be carried out by the data user, there is a loophole in the event of data security incidents and undermine the protection to

[36] *Eastweek Publisher Ltd. & Anor v PCPD* [2000] 2 HKLRD 83.
[37] Berthold, Mark., & Raymond. Wacks. *Hong Kong Data Privacy Law: Territorial Regulation in a Borderless World. 2nd ed*., Sweet & Maxwell Asia, 2003. Chapter 7.

individuals. Although the *Eastweek* case was decided 20 years ago, PDPO has not offered a clarification over the definition of "collection" so far.

(c) Limited extra-territorial applicability

PDPO only regulates data users who "control the collection, holding and processing or use of data" in and from Hong Kong[38], meaning that the Ordinance has a limited reach in the digital age. Multinational corporations whose services are not done in Hong Kong may argue that they are not under the control of PDPO.[39] The European Court of Justice in the case of *Google v Spain*[40] opined that search engines design the algorithm, hence they are controllers of personal data carried out by their search engine. The local case *Dr Yeung Sau Shing Albert v Google Inc*[41] offered similar insights, where Deputy Judge Marlene Ng held that Google designed the algorithms, arguably Google should be treated as publisher, but not a mere facilitator of information. However, the points were made in the interlocutory proceedings, the issue of whether search engine collects personal data is not settled.

Under the current PDPO which came into effect in late 2021, the Office of the Privacy Commissioner for Personal Data ("**PCPD**") may issue cessation notices to request the removal of doxxing-related content regardless of where the disclosure is made (s.66K(1)). The cessation notice can be served on a person in Hong Kong or a non-Hong Kong service provider where an electronic message is made (s.66M (1) and (2)). Nevertheless, the law in Hong Kong is insufficient in protecting personal data of Hong Kong citizens from data breaches happening overseas.

---

[38] S.2(1) of PDPO.
[39] *See* Mak, Elise. "Data privacy: Europe does a check up; Hong Kong just trusts" *Harbour Times*, 6 September 2018. https://harbourtimes.com/2018/09/06/data-privacy-europe-check-hong-kong-just-trusts/ Accessed on 16 November 2021.
[40] *Google Spain SL and Google Inc v AEPD and Mario Costeja González* ECJ Case C-131/12.
[41] *Dr Yeung Sau Shing Albert v Google Inc* [2014] 4 HKLRD 493 and *Dr Yeung Sau Shing Albert v Google Inc (No.2)* [2015] 1 HKLRD 26.

(d) Insufficient control over cross-jurisdiction data transfer

Regarding cross-border data transfer, s.33 of PDPO expressly prohibits the transfer of personal data outside the Hong Kong except under circumstances specified in PDPO. However, s.33 is not yet in operation.[42] S.33(2) specified that one of the conditions must be met for cross-border data transfer: (a) the place must be on the White List; (b) data user has reasonable grounds for believing that there is in force in that place any law which is substantially similar to, or serves the same purposes as PDPO; (c) the data subject has consented in writing to the transfer; (d) the data user has reasonable grounds for believing that the transfer is for the avoidance or mitigation of adverse action against the data subject; it is not practicable to obtain the consent in writing of the data subject to that transfer; but if it was practicable, such consent would be given; (e) the data is exempt from DPP3 by virtue of an exemption under Part 8 of PDPO; or (f) the data user has taken all reasonable precautions and exercised all due diligence to ensure that the data will not, in that place, be collected, held, processed, or used in any manner which, if that place were Hong Kong, would be a contravention of a requirement under PDPO.

Data users transferring data to places outside Hong Kong are also subject to other requirements under PDPO, especially under DPP3 (prescribed consent for a new purpose), DPP2(3) (data users must adopt contractual means to restrict retention period for data processors), DPP4(2) (data user must adopt contract means to ensure data security) and Part VIA of PDPO (data user passing customers' personal data to overseas contractors to make phone calls need to observe requirements).

Despite the Commissioner's complied "White List", so far the HKSAR government has no sign of implementing s.33. PCPD opined that it is not uncommon for overseas jurisdictions not having restrictions on cross-border data transfer, and the existing PDPO provisions have already offered sufficient safeguards by DPPs.[43]

---

[42] PCPD has provided a "Guidance on Personal Data Protection in Cross-border Data Transfer". https://www.pcpd.org.hk/english/resources_centre/publications/files/GN_crossborder_e.pdf

[43] Office of the Privacy Commissioner for Personal Data. "Response to media enquiry on data localization" 15 April 2020. https://www.pcpd.org.hk/english/news_events/media_enquiry/enquiry_20200415.html. Accessed on 10 November 2021.

(e) Absence of mandatory data breach notification mechanism

A breach of the DPPs is not by itself an offence stated in s.64A(2)(a) PDPO. The Commissioner can serve enforcement notices on data users (s.50), non-compliance with enforcement notices may attract penalties after the judicial process (s.50A). For a first conviction, the data user is subject to a fine at level 5 and to imprisonment for 2 years; for the second or subsequent conviction, the data subject is subject to a fine at level 6 and to imprisonment for 2 years.

Despite PCPD's "Guidance on Data Breach Handling and the Giving of Breach Notifications"[44] calls for notification on major data breaches, there is no mandatory requirement under PDPO. This indirectly encourages data users to adopt a "wait-and-see" approach, putting faith in luck that the data breach will remain undiscovered. Oftentimes, PCPD and the affected individuals are notified only when the data breach hits the headline, and this may hinder timely follow-up actions. There is a significant rising trend in the number of individuals being affected by major data breaches in recent years. The average number of individuals being affected by each incident increases from 184 million in 2018 to 308 million in 2020.[45] The government has proposed legislating the notification requirement for data breaches involving "a real risk of significant harm"[46], albeit it has not been adopted yet.

5. Recommendations

A major challenge to protect personal data is the evolving nature of technology and business operations. As there is a global trend towards enacting

---

[44] Office of the Privacy Commissioner for Personal Data. "Guidance on Data Breach Handling and the Giving of Breach Notifications" Jan 2019. https://www.pcpd.org.hk/english/resources_centre/publications/files/DataBreachHandling2015_e.pdf. Accessed on 10 November 2021.

[45] Office of the Privacy Commissioner for Personal Data. "Protection of Personal Data and Cyber Security Challenges in Healthcare Sector" https://www.ehealth.gov.hk/filemanager/content/pdf/common/training/2021/10/08/protection-of-personal-data-and-cyber-security-challenges-in-healthcare-sector.pdf. Accessed on 10 November 2021.

[46] Constitutional and Mainland Affairs Bureau, HKSAR. "Review of the Personal Data (Privacy) Ordinance" 20 January 2020. https://www.legco.gov.hk/yr19-20/english/panels/ca/papers/ca20200120cb2-512-3-e.pdf. Accessed 2 Nov 2021.

more comprehensive privacy and data protection laws, Hong Kong must not lag behind. Recommendations on legislative intervention and administrative measures are offered.

5.1. Legislative Intervention

Making amendments to the current PDPO would be a straightforward solution to enhance big data privacy and data protection.

Expand the definition of personal data and extra-territorial applicability

The definition of personal data has remained unchanged since the PDPO came into force in 1996. The Hong Kong government recognises the need for expanding "personal data" definition, and has proposed to cover information relating to "identifiable" natural persons in the definition.[47] This amended definition would align with definitions in EU, Canada, Australia and New Zealand.[48] GDPR adopts a prescriptive approach for the definition of "personal data", which explicitly includes location data, online identifier and other sensitive data, such as genetic data and biometric data. Hong Kong should adopt a similar approach to avoid ambiguity. In particular, "metadata" should be specifically stated to circumvent the uncertainty brought by the *Telstra* case, because metadata could be used to identify an individual. The expanded definition would be able to reflect the reality that big data analytics technology could be used to re-identify individuals.

[47] Constitutional and Mainland Affairs Bureau, HKSAR. "Review of the Personal Data (Privacy) Ordinance" 20 January 2020. https://www.legco.gov.hk/yr19-20/english/panels/ca/papers/ca20200120cb2-512-3-e.pdf. Accessed on 2 Nov 2021.
[48] *Ibid*.

It is not uncommon that privacy and data protection laws have extra-territorial reach, such as the GDPR, Singapore's Personal Data Protection Act and Australia's Privacy Act 1998. The laws apply to data users and processors processing personal data of the individuals of that region regardless of the processing location. The newly enacted Personal Information Protection Law ("**PIPL**") in China also has the same extra-territorial effect. PDPO should have a similar extra-territorial reach. In case there is a big data breach incident involving a massive number of Hong Kong people, the government will no longer take it lying down.

Regarding the uncertainty of data "collection", it is crucial that the government clarifies in PDPO that the three conditions in *Eastweek* case do not apply. The laws in Singapore[49], European Union[50], Canada[51] and Australia[52] do not require any "identification" threshold in the data collection process. In those jurisdictions, data protection laws come into play whenever personal data is involved.

More stringent regulation on profiling

Some jurisdictions have extended protection to regulating profiling and automated decision-making. Under the GDPR, Article 13 requires disclosure to individuals of the existence, logic, significance and envisaged consequences of profiling. Article 22 secures the right of an individual to avoid being subject to a decision that is based solely on profiling and produces legal effects. Similarly, article 24 of PIPL grants individuals the right to require an explanation from data

[49] Personal Data Protection Commission, Singapore. *Advisory Guidelines On Key Concepts In The Personal Data Protection Act.* 2013, rev. 2017. https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Advisory-Guidelines/important-terms-in-pdpa---ch-3-9-(270717).pdf Accessed on 17 November 2021.
[50] Article 4(2) of GDPR.
[51] Office of the Privacy Commissioner of Canada. *The Personal Information Protection and Electronic Documents Act (PIPEDA).* http://laws-lois.justice.gc.ca/eng/acts/P-8.6/index.html
[52] *Australia Privacy Act 1998*. https://www.legislation.gov.au/Details/C2017C00283

processors and object to the handling of personal data when automated decision-making processes are adopted.

In view of the privacy and data protection challenges, it is pertinent to make it clear that PDPO regulates the use of profiling for automated decision-making. The Australian Government reminded organisations in the "Guide to Data Analytics" that opinions and inferences drawn about individuals from other data are personal data.[53] PDPO should expressly include the right to object to processing and being subject to a decision solely based on profiling.

Mandatory data breach notification

A mandatory breach notification is an international norm. It would allow PCPD to effectively oversee and monitor data breach incidents, especially if there are sufficient follow-up actions.

Different jurisdictions have different thresholds for notification. GDPR requires data controllers to report data breaches to a relevant supervisory authority when there is a "risk to the rights and freedoms" of individuals.[54] Canada requires notification where there is a "real risk of significant harm".[55] Australia demands notification where there is "unauthorised access to or unauthorised disclosure of personal information".[56]

GDPR requires notification to be made to the supervisory authority within 72 hours after having knowledge of the breach, reasons should be provided in

[53] Office of the Australian Information Commissioner, Australian Government. "Guide to data analytics and the Australian Privacy Principles" 21 March 2018. https://www.oaic.gov.au/privacy/guidance-and-advice/guide-to-data-analytics-and-the-australian-privacy-principles. Accessed on 20 November 2021.
[54] Article 33 of GDPR.
[55] n49 above.
[56] n50 above.

case of a delay.[57] Canada, Australia and New Zealand only require notification to both supervisory authority and individuals to be done "as soon as feasible".

PDPO should give a clear definition of data breach, state the notification threshold, list out what information should be provided to PCPD and impacted individuals, impose a reasonable timeframe, and specify a mode of notification. The number of affected individuals could be used to determine the threshold of notification. The government's proposed timeframe of five business days is deemed appropriate, so that a contravention could be objectively determined.[58] The same notification requirement should equally apply to data processors, given that a high volume of data is entrusted to them.

Heavier sanctions

The consequences of big data breaches could be farfetched. Assuming that the reformed PDPO has an extra-territorial effect, levying fines is more appropriate than custodial sentences. Breaching provisions in PDPO generally attracts fines of HK$50,000, which is wholly inadequate as a deterrent, especially to large organisations. GDPR's two-tier fine approach is linked to the company's annual global turnover. In serious breaches, the maximum fine could be 4% annual global turnover or €20 million, whichever is greater. Some commented that the large amount of fine is a "tit for tat" between the US and EU.[59] Hong Kong may refer to Singapore's level of penalty, which is up to 10% of the organisation's annual turnover in Singapore, or SGD 1 million (equivalent to HK$5.7 million), whichever is higher.

---

[57] Article 33 of GDPR.
[58] n48 above.
[59] Bullock, Peter & McCormack, Urszula. "LegCo's Review of the Personal Data (Privacy) Ordinance ("PDPO")". https://www.kwm.com/en/hk/knowledge/insights/legco-reviews-pdpo-20200203. Accessed on 20 November 2021.

5.2. Administrative Measures

Besides amending PDPO to expand its scope of privacy and data protection, it is important to implement administrative measures that embrace "contingency, flexibility and openness to the new"[60], as technology is always more fast-changing than the law.

Imposing industrial standards

One of the significant features of GDPR is its built-in accountability principle. In response to that, PCPD has published the "Privacy Management Programme ("**PMP**"): A best practice guide" to encourage organisations to treat personal protection as an accountability issue.[61] The government can publish industry standards for big data upon consultation. The Hong Kong Money Authority has been providing updates in respect of the use of big data analytics and AI to their authorised institutions, and encouraging an ethical accountability framework in the collection and use of personal data.[62]

Although there is no one-size-fits-all approach, the type of adverse consequences in a data breach, level of sensitivity of data and business operation models may be highly similar within the same industry. PCPD can set maximum retention limit of different categories of personal data. For instance, PCPD recommends employers not to retain employment-related personal data of former employees for more than seven years.[63]

---

[60] Vermeulen, Erik, et al. "Regulation Tomorrow: What Happens When Technology Is Faster than the Law." *American University Business Law Review*, vol. 6, no. 3, 2017, p. 561–594.
[61] Office of the Privacy Commissioner for Personal Data. "Privacy Management Programme: A best practice guide" August 2018. https://www.pcpd.org.hk/pmp/files/pmp_guide2018.pdf. Accessed on 18 November 2021.
[62] Hong Kong Monetary Authority. https://www.hkma.gov.hk/eng. Accessed on 10 November 2021.
[63] Office of the Privacy Commissioner for Personal Data. "Code on Human Resource Management" Clause 1.3.3.2, April 2016.

The Australian government is planning to offer industry standards for AI technologies in sectors including agriculture, human services, financial services, transport and logistics and mining, oil and gas.[64] PIPL has specific rules for processing sensitive personal information, including biometrics, medical health data, religious belief, financial accounts and information about children under 14.[65] The government can make reference to these types of industries and consider implementing industry standards for sensitive data as their first priority. The government could guide how to conduct privacy impact assessments, and review "privacy-by-design" approaches for organisations. Moreover, learning about previous data breaches in the industry is important because organisations can avoid repeating mistakes.

Step up education efforts

Citizens have a crucial role to play in privacy and data protection. On one hand, big data has significant impacts on their lives, on the other, they are well placed for spotting any abuses. Individuals are often tempted by free services without putting much thought to the consequences, such as free Wi-Fi services. Enabling them to understand new technologies can avoid personal data being collected and used unwillingly. For example, individuals could switch off unnecessary tracking functions on their wearables and smart devices, anonymise their data when necessary, and minimise the publication of personal information on social media. Playing the role of an educator, the government should hold public education campaigns regularly to raise public awareness of the risks in big

https://www.pcpd.org.hk/english/data_privacy_law/code_of_practices/files/PCPD_HR_Booklet_Eng_AW07_Web.pdf. Accessed on 20 November 2021.

[64]Standards Australia. "Developing standards for artificial intelligence: Hearing Australia's Voice" June 2019. https://www.standards.org.au/getmedia/aeaa5d9e-8911-4536-8c36-76733a3950d1/Artificial-Intelligence-Discussion-Paper-(004).pdf.aspx. Accessed on 19 November 2021.

[65] Article 28-32 of PIPL.

data. Although PCPD has been running the student ambassador programme annually, its university privacy campaign, mass media campaign and public roadshow have ceased since 2016.[66] CNIL, the Data Protection Authority in France has been promoting "new digital literacy" from primary school to university level, which allows citizens to be familiarised with data algorithms.[67] It is suggested that PCPD could organise public awareness campaigns every two years, and the Education Bureau could introduce big data ethics to primary and secondary education.

Conclusion

While the astute use of big data offers ample opportunities for improving human lives, privacy and data protection should remain as the core principles in data use. This paper discussed major deficiencies in Hong Kong law, including the unclear definition of personal data and collection, limited extra-territorial applicability, insufficient control over cross-jurisdiction data transfer, and absence of mandatory data breach notification mechanism.

The PDPO should be amended so as to align with the international standards, and the proposed administrative measures serve as an additional layer of protection. All in all, the government plays a dominant role in protecting individuals, meeting privacy and data protection challenges, and maintaining Hong Kong's status as an international data hub.

---

[66] Office of the Privacy Commissioner for Personal Data. "Events and programmes" https://www.pcpd.org.hk/english/news_events/events_programmes/roadshow/index.html. Accessed on 18 November 2021.

[67] CNIL. "How can humans keep the upper hand?" December 2017. https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf. Accessed on 20 November 2021.

Bibliography

Books and Journal Articles

Berthold, Mark., & Raymond. Wacks. *Hong Kong Data Privacy Law: Territorial Regulation in a Borderless World. 2nd ed*., Sweet & Maxwell Asia, 2003. Chapter 7.

Butler, D. "When Google got flu wrong." *Nature* 494, 2013, p.155–156.

Gavison, Ruth. "Privacy and the Limits of Law." *The Yale Law Journal* 89.3 (1980): 421-471, p. 428.

Gupta, Nirmal Kumar, and Mukesh Kumar Rohil. "Big Data Security Challenges and Preventive Solutions." *Data Management, Analytics and Innovation,* Springer Singapore, Singapore, 2019, p. 285–299.

Hitzler, Pascal, and Krzysztof Janowicz. "Linked Data, Big Data, and the 4th Paradigm." *Semantic Web*, vol. 4, no. 3, 2013, p.233–235.

Jia, J.S., Lu, X., Yuan, Y. et al. "Population flow drives spatio-temporal distribution of COVID-19 in China." *Nature* 582, 2020, p.389–394.

Legard, Ryan. "How to Manage Big Data Issues in a Project Environment." *Data Analytics in Project Management.* Auerbach Publications, 2018, p. 115-132.

Munir, Abu Bakar and Mohd Yasin, Siti Hajar and Muhammad-Sukki, Firdaus. "Big Data: Big Challenges to Privacy and Data Protection". *International Scholarly and Scientific Research & Innovation* 9(1) 2015. May 21, 2015.

Narayanan, Arvind, and Vitaly Shmatikov. "How to break anonymity of the netflix prize dataset." *arXiv preprint cs/0610105*, 2006.

Nirmal Kumar Gupta. "Addressing Big Data Security Issues and Challenges." *International Journal of Computer Engineering & Technology* 9.4, 2018, p.229-237.

Ohm, P., "Changing the Rules: General Principles for Data Use and Analysis," *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. 2014, p.100.

Rubinstein, I. S. "Big Data: The End of Privacy or a New Beginning?" *International Data Privacy Law*, vol. 3, no. 2, 2013, p. 74–87.

Tene, Omer, and Jules Polonetsky. "Big Data for All: Privacy and User Control in the Age of Analytics." *Northwestern Journal of Technology and Intellectual Property*, vol. 11, no. 5, 2013, p. 239.

Vermeulen, Erik, et al. "Regulation Tomorrow: What Happens When Technology Is Faster than the Law." *American University Business Law Review*, vol. 6, no. 3, 2017, p. 561–594.

Viktor Mayer-Schönberger and Kenneth Cukier. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. John Murray, 2013, p. 1–242.

Westin, Alan. *Privacy and Freedom.* New York: Atheneum, 1967. p. 7.

Cases

Eastweek Publisher Ltd. & Anor v PCPD [2000] 2 HKLRD 83.

Privacy Commissioner v Telstra Corporation Ltd [2017] FCAFC 4.

Shi Tao v The Privacy Commissioner for Personal Data (2007), Case No. 16/2007 of the Administrative Appeals Board.

Government publications

Australia Privacy Act 1998. https://www.legislation.gov.au/Details/C2017C00283

CNIL. "How can humans keep the upper hand?" December 2017. https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf. Accessed on 20 November 2021.

Constitutional and Mainland Affairs Bureau, HKSAR. "Review of the Personal Data (Privacy) Ordinance" 20 January 2020. https://www.legco.gov.hk/yr19-20/english/panels/ca/papers/ca20200120cb2-512-3-e.pdf. Accessed 2 Nov 2021.

Executive Office of the President, the White House, Washington. "Big Data: Seizing Opportunities, Preserving Values" May 2014, p.10. https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf. Accessed on 1 November 2021.

Government Office for Science, UK Government. "Artificial intelligence: opportunities and implications for the future of decision making" 2015. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/566075/gs-16-19-artificial-intelligence-ai-report.pdf. Accessed on 27 October 2021.

Office of Privacy Commissioner of Canada. "Metadata and Privacy: A technical and legal Overview" October 2014. https://www.priv.gc.ca/media/1786/md_201410_e.pdf. Accessed on 1 November 2021

Office of the Australian Information Commissioner, Australian Government. "Guide to data analytics and the Australian Privacy Principles" 21 March 2018. https://www.oaic.gov.au/privacy/guidance-and-advice/guide-to-data-analytics-and-the-australian-privacy-principles. Accessed on 20 November 2021.

Office of the Privacy Commissioner of Canada. *The Personal Information Protection and Electronic Documents Act (PIPEDA).* http://laws-lois.justice.gc.ca/eng/acts/P-8.6/index.html

Office of the Privacy Commissioner for Personal Data. "Code on Human Resource Management" Clause 1.3.3.2, April 2016. https://www.pcpd.org.hk/english/data_privacy_law/code_of_practices/files/PCPD_HR_Booklet_Eng_AW07_Web.pdf. Accessed on 20 November 2021.

Office of the Privacy Commissioner for Personal Data. "Events and programmes" https://www.pcpd.org.hk/english/news_events/events_programmes/roadshow/index.html. Accessed on 18 November 2021.

Office of the Privacy Commissioner for Personal Data. "Guidance on Data Breach Handling and the Giving of Breach Notifications" Jan 2019. https://www.pcpd.org.hk/english/resources_centre/publications/files/DataBreachHandling2015_e.pdf. Accessed on 10 November 2021.

Office of the Privacy Commissioner for Personal Data. "Guidance on the Ethical Development and Use of Artificial Intelligence" August 2021. https://www.pcpd.org.hk//english/resources_centre/publications/files/guidance_ethical_e.pdf. Accessed on 2 November 2021.

Office of the Privacy Commissioner for Personal Data. "Guidance on Personal Data Protection in Cross-border Data Transfer". https://www.pcpd.org.hk/english/resources_centre/publications/files/GN_crossborder_e.pdf. Accessed on 20 November 2021.

Office of the Privacy Commissioner for Personal Data. "Privacy Management Programme: A best practice guide" August 2018. https://www.pcpd.org.hk/pmp/files/pmp_guide2018.pdf. Accessed on 18 November 2021.

Office of the Privacy Commissioner for Personal Data. "Privacy Protection and Data Governance in the Internet of Things" 5 June 2019. https://www.pcpd.org.hk/english/news_events/speech/files/AcademyofLaw_0605.pdf

Office of the Privacy Commissioner for Personal Data. "Protection of Personal Data and Cyber Security Challenges in Healthcare Sector" https://www.ehealth.gov.hk/filemanager/content/pdf/common/training/2021/10/08/protection-of-personal-data-and-cyber-security-challenges-in-healthcare-sector.pdf. Accessed on 10 November 2021.

Office of the Privacy Commissioner for Personal Data. "Response to media enquiry on data localization" 15 April 2020. https://www.pcpd.org.hk/english/news_events/media_enquiry/enquiry_20200415.html. Accessed on 10 November 2021.

Personal Data Protection Commission, Singapore. *Advisory Guidelines On Key Concepts In The Personal Data Protection Act.* 2013, rev. 2017. https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Advisory-Guidelines/important-terms-in-pdpa---ch-3-9-(270717).pdf Accessed on 17 November 2021.

Standards Australia. "Developing standards for artificial intelligence: Hearing Australia's Voice" June 2019. https://www.standards.org.au/getmedia/aeaa5d9e-8911-4536-8c36-76733a3950d1/Artificial-Intelligence-Discussion-Paper-(004).pdf.aspx. Accessed on 19 November 2021.

News articles

Ahmed, Murad. "Facebook understands you better than your spouse" *Financial Times*, 13 Jan 2015. https://www.ft.com/content/3dfa397c-9a73-11e4-8426-00144feabdc0. Accessed on 20 November 2021.

Dastin, Jeffery. "Amazon scraps secret AI recruiting tool that showed bias against women" *Reuters*, 11 October 2018. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G. Accessed on 15 November 2021.

Duhigg, Charles. "How Companies Learn Your Secrets." *The New York Times Magazine*, 16 Feb 2012. http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html. Accessed on 20 November 2021.

Greenwald, Gleen. "NSA collecting phone records of millions of Verizon customers daily" *The Guardian,* 6 June 2013. https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order. Accessed on 8 November 2021.

Ma, Alexandra and Gilbert, Ben. "Facebook understood how dangerous the Trump-linked data firm Cambridge Analytica could be much earlier than it previously said. Here's everything that's happened up until now." *Business Insider*, 24 Aug 2019. https://www.businessinsider.com/cambridge-analytica-a-guide-to-the-trump-linked-data-firm-that-harvested-50-million-facebook-profiles-2018-3b Accessed on 10 November 2021.

Mak, Elise. "Data privacy: Europe does a check up; Hong Kong just trusts" *Harbour Times*, 6 September 2018. https://harbourtimes.com/2018/09/06/data-privacy-europe-check-hong-kong-just-trusts/ Accessed on 16 November 2021.

Marr, Bernard. "The Amazing Ways Spotify Uses Big Data, AI And Machine Learning To Drive Business Success" Forbes, 30 October 2017. https://www.forbes.com/sites/bernardmarr/2017/10/30/the-amazing-ways-spotify-uses-big-data-ai-and-machine-learning-to-drive-business-success/?sh=f2b92ae4bd2f Accessed on 13 November 2021. Accessed on 16 November 2021.

Perry, Caroline. "You're not so anonymous" *The Harvard Gazette*. 18 Oct 2011. https://news.harvard.edu/gazette/story/2011/10/youre-not-so-anonymous/ Accessed on 16 November 2021.

Other references

Bullock, Peter & McCormack, Urszula. "LegCo's Review of the Personal Data (Privacy) Ordinance ("PDPO")". https://www.kwm.com/en/hk/knowledge/insights/legco-reviews-pdpo-20200203. Accessed on 20 November 2021.

Hong Kong Monetary Authority. https://www.hkma.gov.hk/eng. Accessed on 10 November 2021.

HP. "HP Study Reveals 70 Percent of Internet of Things Devices Vulnerable to Attack" 29 July 2014. https://www.hp.com/us-en/hp-news/press-release.html?id=1744676#.YY3P-9ZBx-V. Accessed on 10 November 2021.

Meta. "An Update on Our Use of Face Recognition" 2 November 2021. https://about.fb.com/news/2021/11/update-on-use-of-face-recognition/. Accessed on 8 November 2021.

Wong, Stephen Kai-yi. "Engineering Privacy through Accountability" *66th ABA Section of Antitrust Law Spring Meeting* 11 April 2018. https://www.pcpd.org.hk/english/news_events/media_statements/files/PCPDABA2018.pdf Accessed on 16