Public Policy Research Funding Scheme

公共政策研究資助計劃

Project Number :
項目編號： 2021.A2.049.21C

Project Title :
項目名稱： Data Literacy of Hong Kong Internet Users
香港互聯網使用者的數據素養

Principal Investigator :
首席研究員： Dr IP Ka Wai
葉家威博士

Institution/Think Tank :
院校 ／智庫： Hong Kong Baptist University
香港浸會大學

Project Duration (Month):
推行期 (月) ： 12

Funding (HK$) :
總金額 (HK$)： 610,914.50

“Data Literacy of Hong Kong Internet Users”

「香港互聯網使用者的數據素養」

Submitted by

Department of Government and International Studies

Hong Kong Baptist University

(Principal Investigator: Dr IP, Ka Wai)

Table of Contents

Acknowledgment

This research project (Project Number: 2021.A2.049.21C) is funded by the Public Policy Research Funding Scheme of the Government of Hong Kong Special Administrative Region. All opinions and analyses expressed in this report are those of the research team and in no way represent those of the HKSAR Government.

Members of the Research Team

Principal Investigator:

-Dr. IP, Ka Wai, Assistant Professor, Department of Government and International Studies, Hong Kong Baptist University

Co-investigator:

-Prof. COLE, Alistair, Head of Department; Professor, Department of Government and International Studies, Hong Kong Baptist University

-Dr. KANG, Yi, Associate Professor, Department of Government and International Studies, Hong Kong Baptist University

-Dr. YUEN Wai Hei Samson, Associate Professor, Department of Government and International Studies, Hong Kong Baptist University

We would like to acknowledge the invaluable assistance of Miss Leung Lok Yi Wendy and Miss Cheng Cagang Crystal during the research process.

Executive Summary

Research Abstract

In a digital society, new forms of literacy are needed for individual citizens to navigate the changing digital landscape. The development of the smart city requires citizen's trust and engagement in the process of digitization. For instance, they are willing to share their personal data, make use of the government's e-services, and engage in informed discussions about policy solutions to regulate data applications. However, recent research suggests that the applications of big data analytics are not always benign. They could lead to loss of privacy, discrimination, social exclusion and facilitate deliberate distributions of misinformation.

Managing the risks that come with increasing datafication requires something more than technical solutions. To function as a proactive member in a digital society requires not only hard skills to use digital technologies but also an extended data literacy in the population as measured by citizens' awareness and critical reflection of big data practices and their implications, as well as the ability to implement such knowledge in their digital attitudes and behavior. Our research project thus focuses on the data literacy of Hong Kong internet users, which consists of data thinking, data doing, and data participation. What is the current level of data literacy of Hong Kong internet users? What are the potential factors influencing a person's data literacy? Are there any demographic groups which are more data literate than others? What kind of intervention, if any, may effectively improve data literacy levels?

To answer these questions, the research team conducted an online survey and a survey experiment. The survey experiment was designed to test the effectiveness of two forms of data literacy intervention—exposure to information and participation in deliberation—on improving data literacy for the participants. In addition, the research team conducted focus group interviews with randomly selected respondents to explore in detail how different interventions have affected the respondents' data thinking, data doing and data participation. Our study found that both forms of intervention were effective in enhancing the participants' data literacy levels, especially in the aspects of data thinking and data doing. Yet, their influence over data participation was rather indeterminate.

Layman summary on policy implications and recommendations

Educational and informational resources about data literacy currently remain scarce. The public lacks the skills and knowledge to navigate society with data, especially for underprivileged groups like low-income, elderly, the disabled, and ethnic minorities. It is thus essential to close the literacy gap by providing access to training and equipping the public to think critically and ethically about data. We believe that the coordination between parties like public institutions, educational institutions, NGOs, and local support groups, can assist the public in nurturing data literacy, improving their experience in the current technology and data-driven smart city.

(1) We recommend that schools incorporate data literacy into the curriculum from the early stages

of education in a cross-module discipline. Young people's development of data literacy can be implemented through continuous exposure to relevant knowledge in applied settings like project learning and visiting activities. Students will not only be able to focus on different facets of data literacy in-depth but also integrate data analysis into their daily lives and illuminate real-world phenomena through data. To facilitate the new curriculum, extensive professional training should be developed and offered for teachers to acquire the requisite data literacy skills that transcend the boundaries of school subjects. Classroom technology also requires updates and should be supported by the adequate provision of hardware to students.

(2) In this research, we observed a strong demand among Hong Kong online users for trustworthy data literacy information. The effectiveness of information exposure in improving data literacy is proven to be riding on the credibility of the sources of information. We recommend that the government further collaborate with educational institutions to launch programs in different forms for promoting data literacy. For example, courses or interest classes organized by NGOs or community centers would help the less privileged groups to become acquainted with the latest online trends and relevant technical knowledge to counter their daily lives.

While the provision of introductory educational materials might not be particularly effective in changing the attitudes and behaviors of recipients who are more literate in data, we recommend that future educational programs be offered in a variety of formats designated for participants with different levels of data literacy. When more in-depth and discipline-focused programs can be launched for advanced data users, fundamental programs with targeted help can be provided to

less literate participants to bridge the gap in skill level.

(3) Participants in deliberation workshops suggested that imposing more regulations on tech companies can hold them more accountable to their users. The General Data Protection Regulation (GDPR), drafted and passed by the European Union (EU), is directly binding and applicable to enhance users' control and rights over their data. Harsh fines are levied against those who violate privacy and security compliance. We believe that the GDPR could be a model for the Hong Kong government to formulate similar laws to regulate organizations' practices in their data collection and management.

(4) To build trust and improve their agency, tech companies should allow users to tailor their digital experiences according to their needs and preferences. We recommend all tech companies implement trustworthy, transparent, and user-friendly designs that keep users informed about how their services work with the collected data and grant them meaningful control over their usage.

行政摘要

報告撮要

在數碼社會中，個體公民需要新的素養形式，以適應不斷變化的數碼環境。智慧城市發展需要數碼化過程中得到公民的信任和參與，例如令他們願意分享自己的個人數據，使用政府的電子服務，並參與討論規範數據應用的政策解決方案。然而，大數據的應用可為社會帶來種種好處，但並非全無風險，例如可能導致失去隱私，歧視，社會排斥，以及虛假信息的傳播。

要應對數據化帶來的不斷增長的風險，除了技術解決方案外，還需要做更多的事情。要成為在數字社會中發揮積極作用的一員，公民不僅需要使用數碼技術和知識，而且還需要有一定的「數據素養」。這要通過公民對大數據的應用及其影響有充份的認知和反省，由此改變他的態度和行為。因此，我們的研究項目側重於香港互聯網用戶的數據素養，包括數據思維，數據處理和數據參與。

研究團隊進行了在線調查和調查實驗。這項調查實驗旨在測試兩種形式的數據素養干預措施（信息接收和參與協商）在提高參與者的數據素養方面的有效性。我們的研究發現兩種干預措施都能夠提升參加者的數據素養，當中在數據知識和數據行為上較為顯著。但對於數據參與的影響則不能確定。

研究項目對政策影響和政策建議的摘要

目前社會上的數據素養教學和資源仍然匱乏，公眾，尤其是低收入人群、老年人、殘疾人、少數民族等弱勢群體，可能缺乏足夠的技能和知識來利用數據驅動社會。為了彌合可見的數字鴻溝，政府必須為公眾提供適當的培訓，讓他們能夠以批判性和道德的方式思考數據。我們相信，公共機構、教育機構、非政府組織和當地團體之間的合作可以幫助公眾培養數據素養，並提升他們在當今科技和數據驅動的智慧城市中的體驗。

（1）我們建議學校將數據素養納入基礎教育課程中。青年的數據素養發展可以通過在項目學習和參觀等應用環境中，持續接觸相關知識來實現。學生不僅能夠深入關注數據素養的不同面向，還能夠將數據分析融入日常生活，並通過數據闡明現實世界的現象。為促進新課程的發展，政府應開發並提供廣泛的專業培訓，使教師能夠獲得跨學科的數據素養知識和技能。課堂技術也需要持續更新，並為學生提供足夠的硬件來支持相關教育。

（2）在這項研究中，我們觀察到香港互聯網使用者對可信的數據素養信息的強烈需求。從研究結果來看，持續接收相關資訊可以提高數據素養，其有效性很大程度上取決於信息來源的可信度。我們建議政府進一步與教育機構合作，推出不同形式的數據素養項目，例如由非政府組織或社區中心舉辦的興趣班，可以幫助弱勢群體熟悉新科技和技術，以應對日常生活。由於介紹性教育材料相對簡單，難以有效改變數據素養較高的使用者的態度和行為，因此我們建議未來的教育項目應針對不同數據素養水平的參與者，定制不

同程度的計劃，例如為數據素養較高的使用者提供更深入和以特定學科為重點的項目，向數據素養較低的使用者提供具有針對性幫助的基礎計劃，以彌合他們技能水平的差距。

（3）審議工作坊的參與者認為，對科技公司進行監管，會使他們對使用者更負責。由歐盟（EU）起草和通過的「一般資料保護規範 (GDPR)」具有直接約束力，適用於增強使用者對其數據的控制和權利，並對違反隱私和安全條例者處以重罰。我們相信 GDPR 可以成為香港政府制定類似法律的參考。

（4）為了建立信任並改善人們的能動性，科技公司應該允許用戶根據他們的需求和偏好去定制他們的數碼體驗。我們建議所有科技公司實施可信賴、透明和對用戶友好的設計，讓用戶了解他們將如何處理所收集的數據，並讓用戶能夠有意義地控制他們的使用。

Research Background

Data, Datafication and the Smart City

As our everyday lives and interactions are increasingly 'datafied' by the growing digitalization of human activities, generating an unprecedented amount of data, it is crucial to include everyone in the digital society in a meaningful way. This 'Big Data Revolution', as some might call it, is powered by the ever-growing, ubiquitous collection, analysis and applications of data about people and their activities by private companies and governments in all areas of society (Mayer-Schönberger and Cukier 2013).

The definition of big data is contested. Earlier designations emphasized the magnitude of the data sets and the volume, variety and velocity of such data (Diebold 2012; Laney 2001). However, recent research has focused on not only the technical characteristics of big data but also its social dimension. Boyd and Crawford (2012), for example, defined big data as a 'socio-technical' phenomenon that rests on the interplay of technology, analysis and mythology. Accordingly, big data refers to the use of computation powers to collect, aggregate and cross-reference large data sets, with a purpose of generating insights into social, political and economic problems. This is associated with a widespread belief that these data sets offer greater intelligence than was previously possible, with a higher level of accuracy and objectivity (Boyd and Crawford 2012, 663).

Like other technological changes, the advance of big data comes with both opportunities and risks. The expansion of big data technologies has played an important role in the development of smart cities. These technologies enable cities to gain valuable insights from a large amount of data collected through various sources, such as connected devices and sensors, and to transform the domains of transportation, energy, health care and education (Hashem et al. 2016). In Hong Kong, for example, the Hospital Authority has launched a big data analytic platform to facilitate health care-related research (Innovation and Technology Bureau 2020, 13-14). Real-time adaptive traffic signal systems with sensors for pedestrians and vehicles can help optimize the green times allocated to vehicles and pedestrians (Innovation and Technology Bureau 2020, 5-6). Multi-functional Smart Lamp posts have been proposed to facilitate the collection of real-time city data to enhance city management and other public services (Innovation and Technology Bureau 2020, 23-24).

However, recent research suggests that the applications of big data analytics are not always benign. They could lead to loss of privacy, discrimination and social exclusion and facilitate deliberate distribution of misinformation. For instance, it has been widely documented that the development of big data technologies, enabled by massive surveillance and data collection, has led to violations of individual and group privacy (see, for example, Turow 2011; Lyon 2014; Schneier 2015; Mittelstadt 2017; Zuboff 2019). In some European countries, the widespread perception that big data infringes privacy has undermined public trust, resulting in lower willingness of citizens to share personal data (Vodafone Institute 2016). Nevertheless, the risks associated with the datafication of society go well beyond the concerns of privacy. Another example of ‘data harm’ is

information distortion. During the COVID-19 pandemic, the use of digital media amplified and exacerbated the spread of misinformation concerning health issues, thus hampering the efforts of medical professionals and government agencies to combat the virus (Vraga, Tully and Bode 2020; Carmi et al. 2020).

Data Literacy

Managing the risks that come with datafication requires more than technical solutions. Citizens need to have the critical abilities to understand the opportunities and risks associated with increasing data collection and engage in informed public debates on these new data practices. Our research project would thus focus on the data literacy of Hong Kong Internet users. The term 'literacy', the ability to read and write, has acquired different meanings in different contexts. For example, 'scientific' literacy means an appreciation of the nature of science, the development of personal attributes and the acquisition of scientific skills and values (Holbrook and Rannikmae 2009). 'Cultural' literacy refers to the skills that enable one to operate efficiently in different cultural and professional contexts (Ochoa, McDonald and Monk 2016). 'Financial' literacy describes the knowledge, skills, confidence and motivation necessary to effectively manage money (Remund 2010).

In a digital society, new forms of literacy are needed for citizens to navigate the changing digital landscape. Gilster (1997) first popularized the term 'digital' literacy as 'the ability to understand and use information in multiple formats from a wide variety of sources when it is presented via

computers'. Eshet-Alkalai (2004) defined digital literacy as the set of 'cognitive, motor, sociological and emotional skills, which users need to function effectively in digital environments'. Hargittai and Hsieh (2012), for example, identify digital literacy with the (self-reported) familiarity with Internet-related terms, which they argue is positively related to a person's actual ability to find accurate information online.

In these formulations, digital literacy is often used as a synonym for digital competence and is also interchangeable with digital skills. Nonetheless, to meaningfully participate in a datafied society, citizens need more than the hard skills to use digital technology. First, one needs to be aware of the potential implications of datafication for their lives and the society in which they live. Second, such awareness should enable more societal and public involvement in the digital realm. This does not presuppose a negative stance towards datafication and big data technology. Nevertheless, it is important for individuals to make informed judgements about the process of datafication. Thus, to be data literate is to be able to critically reflect upon big data collection practices and to implement this knowledge in one's digital engagement (Sander 2020b, 2-3). This conception of data literacy has been called 'critical data literacy' (Sander 2020a; 2020b), a 'data mindset' (D'lgnazia and Bhargava 2018), 'big data literacy' (D'lgnazia and Bhargava 2015) and 'algorithmic literacy' (Gray, Gerlitz and Bounegru 2018).

In our proposed research, however, we adopted the data citizenship framework developed by Yates et al. (2020), which focuses on the skills, thinking and actions needed in a datafied society, to unpack the critical and participatory dimensions of data literacy.

According to Yates et al. (2020, 10), data citizenship is a framework for data literacy that 'outlines the importance of citizens having a critical and active stand, at the time when society's datafication and algorithmically-driven decision making has become normalized'. As a framework for data literacy, it suggests that to function proficiently in a datafied society (i.e. to be data literate), one must be able to carry out individual and collective critical inquiry into the data practices of one's community. The framework consists of three areas (Yates et al. 2020, 12-14; see also Camri 2020, 10):

1. Data Thinking – Citizens' critical understanding of data, such as awareness of data protection rights, the ability to use data for communication (e.g. providing evidence to validate an argument) and the knowledge of different data collection practices.

2. Data Doing – Citizens' everyday engagement with data, such as the ability to use data ethically, the ability to evaluate data quality and credibility (e.g. fact-checking) and the ability to manage data in a safe and secure way.

3. Data Participation – Citizens' use of data to contribute to and shape collective data experiences, such as taking proactive steps to protect their own and others' privacy, the ability to utilize data for societal participation and civic action, the ability to help others with their data literacy and the ability to resist the current hegemonic practices of platforms and data services.

According to this conceptualization, data literacy has cognitive, attitudinal (data thinking), as well as behavioral (data doing and participation) dimensions. In our proposed research, we plan to use surveys to examine the level of data literacy of Hong Kong Internet users.

Citizenship refers to a binding relation between the state and the individuals within its jurisdiction, implying various rights (e.g. to vote, to hold public office and access to public services) and obligations (e.g. taxation, military service and compliance with law) (Tilly 1997). It also entails various degrees of participation in the community's political, social and economic processes (Bellamy 2008). What connects data literacy and the notion of citizenship is the idea that being data literate means being able to participate *reflectively* and *proactively* in data collection practices and make informed decisions in a digital society. This understanding of citizenship focuses on the citizens' agency and the progressive social change that may result from it (Hintz, Dencik and Wahl-Jorgensen 2018, 30). Thus, data citizenship can be compared with the idea of being a data consumer (who is primarily a self-interested chooser of services).

Another reason why the framework of data citizenship is particularly apt for unpacking the concept of data literacy is that it highlights the importance of participation in public life. Accordingly, to function proficiently in a datafied society, one requires not only hard skills and knowledge but also pro-social motivations and behaviors to benefit their community. The classical notion of citizenship gives one rights, but it also prescribes the exercise of civic virtues, such as being able to overcome one's immediate self-interest and become public spirited (Pettit 1997; Dagger 1997).

Therefore, we treat 'data participation' – which is measured by pro-social online behavior, potentially contributing to a better digital environment – as an important dimension of data literacy.

Data Literacy Tools

The next question is whether and how data literacy can be implemented and improved. At present, some Internet users in Hong Kong may not be aware of the potential consequences of sharing personal data online and they may lack the knowledge to participate meaningfully in debates regarding the governance of datafication. The future datafication and smart-city development in Hong Kong needs informed public debate about the implications of data technologies and their applications in various areas of citizens' life. A previous study by Sander (2020a; 2020b) looked into the effectiveness of online data literacy tools, such as multi-media websites, which were designed to improve the users' data literacy. The findings show that those data literacy tools have a positive influence on the users' data literacy, leading to more concern for privacy and more privacy-sensitive Internet usage. Nonetheless, Sander's study was a qualitative study involving only a small sample (N = 10). Our research adopted an experimental study format involving a larger sample (N = 600) to test the effectiveness of interventions that were designed to foster data literacy.

Hong Kong's Digital Readiness

Previous studies on Hong Kong's digital readiness have focused on Hong Kong's digital

infrastructure and the digital attitudes and engagement of Hong Kong residents.

According to the Digital Intelligence Index developed by the Fletcher School at Tufts University, Hong Kong is among the most digitally advanced economies in the world, performing especially well in terms of digital environment and infrastructure. In terms of the digital environment, Hong Kong has a well-developed infrastructure to facilitate digital interactions and transactions, which facilitate innovation in the city's digital economy. Hong Kong is ranked third in digital environment among the 90 economies studied (Chakravorti et al. 2020, 19-23). In addition, Hong Kong residents exhibit a high level of digital engagement, which is measured by the consumers' use of technology, social media, e-commerce and mobile payments (Chakravorti et al. 2020, 33).

A recent study commissioned by Google found that Hong Kong residents in general display a favorable attitude towards digital technology. Hong Kong residents exhibit a high degree of confidence in the ability of Artificial Intelligence (AI) to improve lives and substantial usage of digital (AI-powered) products and services. However, the same study also found that Hong Kong residents have relatively low levels of AI familiarity and knowledge, and willingness to share data (KPMG 2020, 23-36).

What is missing from these studies is the investigation into Hong Kong residents' ability to critically reflect on the opportunities and risks associated with digitalization and to participate proactively in shaping the digital environment. Our research project filled this gap in the literature by applying the data citizenship framework to examine the data literacy of Hong Kong Internet

users. We intend to answer the following research questions:

-What is the current level of data literacy of Hong Kong Internet users?

-What are the potential factors influencing a person's data literacy?

-Are there any demographic groups that are more data literate than others?

-What type of intervention, if any, may effectively improve data literacy levels?

We believe that a highly data literate population is crucial for the future development of Hong Kong as a digitally advanced smart city. As such, a rigorous assessment of the different dimensions – data thinking, doing and participation – of data literacy would be an important first step towards more informed policymaking by the government and more empowered Internet usage on the part of the population.

Research Objectives

1. To assess the level of data literacy of Hong Kong Internet users
2. To determine the factors influencing varying levels of data literacy of Hong Kong Internet users
3. To identify relevant means to improve data literacy and online behavior in the context of Hong Kong
4. To reveal the mechanisms leading to enhanced data literacy
5. To produce new findings that can inform policy-making and public debates about digitalization and smart city initiatives in Hong Kong

Methodology

To answer our research questions, we conducted an online survey (to assess the data literacy of Hong Kong Internet users and to discern the possible factors underlying different levels of data literacy) and a survey experiment (to test the effects of two forms of interventions on data literacy). The PI and Co-Is have previous experience in investigating trust and the smart city, political theory, public opinion, and non-governmental organizations through surveys and interviews.

Online Survey

The research team has designed the questionnaire and monitored its implementation. The target population is Cantonese-speaking Internet users in Hong Kong aged 18 years or above. The Hong Kong Public Opinion Research Institute (HKPORI) was commissioned to implement the online survey. The sample size is 3,279 respondents. 15.7% of the respondents are drawn from the Hong Kong People Representative Panel (probability panel) and 84.3% of the respondents are drawn from the Hong Kong People Volunteer Panel (non-probability panel). Both panels were created and maintained by HKPORI. People would not be able to participate in our survey unless they were invited to do so. This was to ensure the randomness and representativeness of the sample.

We identified two reasons for choosing an online survey format instead of a conventional telephone survey. First, online surveys tend to produce equally accurate results as telephone surveys. Online surveys, as opposed to interviewer-administered polls, can reduce the social

desirability bias of the survey (Chang and Krosnick 2003; Kennedy and Deane 2019). In recent years, some internationally renowned pollsters, such as Pew Research Center, Gallup Poll, and YouGov, have adopted the online survey method. Second, the use of an online survey enables the researchers to reach the intended larger sample at a relatively low cost. A larger sample in the initial survey means a larger pool of potential participants for the survey experiment, which we planned to conduct following the initial survey.

One limitation of using online surveys, however, is that the sample excludes those who do not use or have no access to the Internet. We recognized this limitation but would like to emphasize that the Internet penetration rate in Hong Kong is exceptionally high, with 91.7% of the population aged 10 years or above using the Internet (Census and Statistic Department 2020). Thus, our method would not exclude a significant segment of the Hong Kong population.

The survey results are rim-weighted according to figures provided by the Census and Statistics Department as well as figures from HKPORI's regular telephone tracking surveys. The gender-age distribution of the Hong Kong population came from "Mid-year population for 2021". The educational attainment (highest level attended) distribution and economic activity status distribution came from "Women and Men in Hong Kong-Key Statistics (2021 Edition)." For appraisal of political condition and political inclination, respondents' data came from previous panel surveys, while target distribution came from regular telephone tracking surveys.

Survey Experiment

We recruited 648 respondents from the first online survey to participate in the survey experiment. These participants were randomly assigned into three groups. Groups 1 and 2 were experimental groups that received two different interventions: informational exposure and deliberation. Group 3 served as the control group.

Intervention (1) – Informational Exposure

In the literature on media literacy interventions, it has been repeatedly shown that exposure to information, such as tips on identifying fake news, could help subjects resist false news or other forms of misinformation (see, for example, Guess et al. 2020; Tully, Vraga and Bode 2019; and Vraga, Tully and Bode 2020). To test whether exposure to information would have an influence on data literacy, we invited respondents from Group 1 to join a WhatsApp group in early July. Research team members regularly posted information related to data literacy in the group for a month, and there were a total of 16 posts. (Refer to the Appendix for the full content of the posts).

Intervention (2) – Deliberation Workshop

Another group of respondents were invited to attend the 2.5-hour 'Deliberation Workshop'. Participants would deliberate on issues surrounding data security and the applications of data technologies in society in a small-group discussion format (with 8–10 participants in each group). The research team had prepared the briefing materials and designed the discussion topics for the

workshop. The workshop was held to induce deliberation among the respondents, defined as the weighing of considerations through discussions that were informed, balanced, conscientious, substantive and comprehensive (Fishkin and Luskin, 2005). For instance, before the discussion sessions, the participants were sent balanced briefing materials. (Refer to the Appendix for the full version of the briefing and discussion materials.)

Five trained moderators were invited to lead the small group discussions. They were responsible for maintaining an atmosphere of civility and respect, encouraging different opinions, restraining dominating speakers, and ensuring that all the major views on the issue were expressed and considered (Fishkin and Luskin 2005, 288). Under these conditions, the participants were "effectively motivated to behave a bit more like ideal citizens" in the sense that informed discussions would help them overcome their immediate self-interest and become more public-spirited (Ackerman and Fishkin 2002, 134).

Previous studies of deliberative polling suggested the robustness of its effect across different domains. The findings indicated that participants gain knowledge about an issue through deliberation. In addition, deliberations often lead to changes in policy, attitudes, and opinions (Fishkin and Luskin, 1999). Finally, deliberation can increase political trust and readiness for collective actions (Gronlund, Setälä and Herne, 2010). We hypothesize that an increase in public spirit can foster a stronger sense of data citizenship and induce more pro-social online behavior.

After completing the interventions, the participants answered the phase-two survey with the same

questions as in the initial survey.

Focus Group Interviews

The research team conducted 10 focus group interviews with phase two respondents to supplement the survey experiment in exploring how different interventions—informational exposure and deliberation workshop—have affected the respondents' data literacy. Each focus group comprised 6 to 8 participants. Respondents were asked to recall and interact with one another about their intervention experiences. The PI and Co-Is then focused on the causal chains in each case and then compared them across cases. As the interviews might involve sensitive topics like personal experience or political and policy issues, measures have been taken by the research team to ensure confidentiality.

Survey Finding

Part 1. Demographics

To begin, we present the socio-demographic characteristics of the studies' respondents.

1a. Gender

In the survey, respondents were composed of slightly more females (52.4%) than males (46.5%), with 1.1% of respondents stating themselves as "other" gender. (See Table 1a)

1b. Age

Respondents in the survey were distributed evenly among all age groups under 60. 13.7% of the respondents were 18-29 years old, 15.7% were aged from 30-39 years old, 17.8% were aged from 40-49 years old, and 19.8% were aged from 50-59 years old. 60-69 year-old respondents made up the largest proportion of the survey with 27.4%, and 5.6% were 70 years old or above. (See Table 1b)

1c. Education Level

In general, we observe that most of our respondents were well-literate. 53.1% of respondents have attained upper secondary education (S4-S7 / DSE / YiJin), 10% had completed tertiary education with non-degree programs (including diploma / certificate / sub-degree courses), 19.4% were bachelor's degree holders, and 5.6% achieved postgraduate tertiary education or above. Only 10.2% of the respondents stated that they had a lower secondary education (S1-S3) and 1.7% received primary education or below. (See Table 1c)

1d. Occupation Status

Concerning the occupation status of the respondents, around one third of them were permanent workers (32.8%). Another one-third of the respondents were not the breadwinners for their household, with 22.9% being retired or pensioners, and 10.6% being housekeepers. 12.2% of them were contract workers with a year or longer contract. Freelancers or part-time workers (3.7%), full-time students (3.9%), the unemployed (4.2%), and those with other occupation status (4.6%) took up similar proportions of the respondents. 3.2% of the respondents were employers. Least respondents (1.8%) were short-term contract workers. (See Table 1d)

1e. Household income

Respondents were asked to state their monthly household income in the survey. Respondents with $10,001 to $50,000 household income comprised the principal respondents of the study. More specifically, most of their households (22.6%) had an average income of $10,001 to $20,000,

18.8% have $20,001 to $30,000, and 19.1% have $30,001 to $50,000 per month. The next largest group after the above three was 11.2% of the respondents, whose households made $50,001 to $70,000 monthly. 6% of them earned $70,001 to $100,000 and 2.1% earned more than $100,000. 5.4% of the respondents reported having no household income, 8.3% had less than $10,000, and 6.3% claimed they did not know. (See Table 1e)

1f. Political Inclination

In the survey, we asked respondents to indicate their political inclination from a range of options, including ‘pro-establishment’, ‘centrists’, 'pro-democracy’, ‘localists’, ‘others’, ‘no political inclination/politically neutral/do not belong to any camp’ and ‘don't know/hard to say’.

We observe that the respondents were scattered within the political spectrum. Most of the respondents (30.2%) identified themselves as having no political inclination/ politically neutral/ do not belong to any camp. Pro-democracy was the second most popular political inclination selected by the respondents, taking up 21.5%. Whereas 9.3% representing the localist. 18.5% of them recognized themselves as centrists and 15.4% indicated that they belonged to the pro-establishment camp. Only 5% did not know or found it hard to define their political orientation, while none of them chose "other"’ out of all the options. (See Table 1f)

Part 2. Digital Activities

We asked the respondents about their daily usage of the internet, their online activities, and their usage of social media and instant messaging apps. The four questions asked are adopted from Ofcom's "Adults' media use and attitudes report 2020," a UK-based tracking survey on media literacy.

2a. Daily usage of Internet

The majority of the respondents spent 8 hours or more online (28.0%) or 2-4 hours online (27.7%) on average every day in the last month. 20.4% of the respondents spent 4-6 hours online daily; 18.7% of the respondents spent 6-8 hours online daily. Only a small percentage of respondents spent less than two hours per day online, with 4.4% spending 1-2 hours and 0.6% spending less than one hour. The result indicates that Internet usage is inevitable in most people's daily lives. Around 30% of respondents were heavy users who spend one-third of their day online. (See Table 2a)

2b. Online activities

We included ten online activities in the survey and checked the respondents' participation. More than 90% of the respondents had used instant messengers (91.8%) and search engines (90.5%). More than 70% of the respondents had used social media (85.8%), watched short videos (78.6%) and visited news websites or websites about current affairs and politics (70.6%). Around 60% of the respondents had engaged in online payments (63%) and business transactions (62.1%). 43%

of the respondents had watched videos via online streaming, 33.9% had played games online, and 29.4% had listened to music via online streaming.

Respondents actively engaged in socializing and entertainment on online platforms. They relied significantly on online platforms to obtain information and news. A notable number of respondents practiced online transactions. Respondents' usage of various online streaming services to obtain entertainment associated with the increasing popularity of these services worldwide. (See Table 2b)

2c. Social Media

We asked the respondents if they had registered for a personal account with 12 selected social media sites or apps to explore their social media usage. The three most common social media platforms are Facebook (91.4%), YouTube (76.4%), and Instagram (65.1%). Other social media platforms are less popular among respondents; their usage ranged from 27.4% (Twitter) to less than 1%. Only 1.1% of the respondents indicated that they never used any social media platforms. The result indicates that almost all respondents had a social media profile. However, not all social media platforms require their users to register a personal account, like YouTube and online forums. Instead, their users can browse the platforms as guests or without an account. (See table 2c)

2d. Instant Messenger

Respondents were asked if they had ever used the 12 listed instant messaging websites or apps for communication. With 95.9% of respondents using it, WhatsApp was the most popular instant messenger. It was followed successively by Facebook Messenger (66%) and WeChat (62.4%). More than 50% of the respondents (55.4%) have used Zoom. In early 2020, Zoom's usage has increased remarkably since quarantine measures were adopted around the globe in response to the COVID-19 pandemic. It is now one of the most common video communications platforms for remote learning and working. 55% of the respondents had used Signal and 40.2% had used Telegram. In early 2021, WhatsApp issued a new privacy policy suggesting sharing access of users' data on calls, photographs, texts, videos, and documents, etc. with its parent company, Facebook. It sparked a fierce backlash over privacy issues and sensitive personal data violations, with many users threatening to quit the service. Some have migrated to Signal and Telegram, messaging platforms that are recognized as more secure. For example, their privacy capabilities can protect their users from third-party monitoring. While Signal shares very similar features with WhatsApp, Telegram is known for mass texting by allowing a mega chat group size, anonymity (phone number is not required to register an account), and no default limit on message forwarding. It has become a popular instant messenger since the Anti-Extradition Law movement in Hong Kong in 2019. Besides the above messengers, Line (31.8%), Instagram Direct Message (24.8%), and Skype (22.2%) were also common platforms that respondents used. (See table 2d)

3. Data Technology and Confidence

In part 3, we investigated two aspects of confidence in data security: acquired confidence and self-

confidence. For acquired confidence, we asked the respondents to indicate their confidence in keeping their personal information safe via various external measures. For self-confidence, we asked the respondents if they were confident in securing their personal privacy and information online.

3a. Acquired Confidence

Most respondents (71%) agreed (somewhat agree: 45.5%; strongly agree: 25.5%) that changing privacy settings can keep their information safe (M = 3.99, SD = 0.78). A similar number of respondents (68.3%) agreed (somewhat agree: 42.2%; strongly agree: 26.2%) that enabling multi-factor authentication can keep their information safe (M = 4.03, SD = 0.81). 63.6% of the respondents agreed (somewhat agree: 38.6%; strongly agree: 25%) that using a firewall or other related antivirus software can keep their information safe (M = 3.91, SD = 0.85). 55.6% of the respondents agreed (somewhat agree: 31.5%; strongly agree: 24.1%) that using websites or apps that are more secure or sensitive to privacy (like Signal, Telegram, and Duckduckgo) can keep their information safe (M = 3.83, SD = 0.90). Finally, around half of the respondents (49.9%) agreed (somewhat agree: 35.8%; strongly agree: 14.1%) that using a Virtual Private Network (VPN) can keep their information safe (M = 3.73, SD = 0.84). (See Table 3a-b)

3b. Self-confidence

Respondents were asked to indicate their confidence in securing their personal privacy and

information online. We observe a consistent view across respondents, indicating that they were impartially confident in securing their online information safety. Among the suggested options, they were most confident in determining whether online information is true or not (M = 3.3, SD = 0.83) and they were least confident in understanding the terms and conditions offered by websites or apps (M = 2.91, SD = 0.95). The remaining options are listed as follows, according to descending order: confident in using online public information and data for their problem solving (M =3.14, SD = 0.84); confident in keeping their information safe (M = 3.05, SD = 0.88); confident in managing who can see the information they shared (M = 3.05, SD = 0.94); confident in understanding data analysis which they encountered online (M = 3.04, SD = 0.87); confident in securing their personal privacy online (M = 3.04, SD = 0.86); and confident in managing who can use their personal information and data online (M = 2.92, SD = 0.96).

When respondents were quite confident that digital measures can protect their data security, they were less confident that "they" can protect their own data and information safely. Such neutral views might imply that individuals can hardly rely on their own (like their capabilities) in protecting their data and information online. The supplementary external measures can act as a fortification tool to strengthen their confidence. Also, it is quite difficult for individuals to verify whether they have successfully secured their data and information due to the lack of substantial evidence. Thus, respondents were inclined to offer conservative answers. (See Table 3c-e)

Part 4. Data Technology and Attitude

We researched on six perspectives to understand respondents' attitudes towards data technology.

4a. Data and Society

We asked respondents about their attitude towards data technology's impact on society. The question was adopted from Doteveryone's "People, Power, and Technology: The 2020 Digital Attitudes Report". A vast majority (78.3%; much better: 16.4%, somewhat better 61.9%) of the respondents believed that data technology had improved their personal lives (M = 3.93, SD = 0.74). However, respondents (59.5%; much better: 12.5%, somewhat better 47.0%) were less convinced that data technology had improved their society (M = 3.53, SD = 1.09). The perceived gap may imply that: while most respondents can benefit from data technology, they are also aware of the risk of data technology abuse in society, like the exacerbation of social inequalities and potential mass surveillance. (See Table 4a)

4b. Data and Disempowerment

Respondents were asked to indicate their perceived disempowerment while using data technology. The question was adopted from Doteveryone's "People, Power, and Technology: The 2020 Digital Attitudes Report". We discovered a significant perceived power imbalance between respondents and technology companies. The lack of transparency in how tech companies manage their users' data has greatly frustrated the respondents, leading to a severe sense of powerlessness. Respondents desired opportunities to tailor their digital experience to their needs and preferences

outside of the constraints imposed by tech companies.

First, most respondents felt powerless about how their data was used by third parties. 82.6 % of respondents (somewhat agree: 46.3%, strongly agree: 36.3%) recognized that they have no control over how their personal information and data are collected and used by tech companies (M = 4.15, SD = 0.82). Besides the lack of control, the result implies that the respondents might lack relevant knowledge about how tech companies managed their data. (See Table 4b(i))

Respondents presented a strong feeling of resignation towards websites and apps' terms and conditions (T&C). 74.6% of the respondents (somewhat agree: 46.5%; strongly agree: 28.2%) usually accepted the T&C without reading or fully understanding them when they used a website or app (M = 3.96, SD = 0.91). (See Table 4b(ix)). Furthermore, 81.3% of them (somewhat agree: 33.3%, strongly agree: 48%) expressed that it is meaningless to read the terms and conditions on websites or applications as they are obligated to accept them in exchange for the service provisions (M = 4.26, SD = 0.85). Tech companies are, in a way, forcing their users into a dilemma: they can only choose between accepting the T&C or quitting the service. Thus, users are highly discouraged from reading and understanding the T&C if they would like to consume the services. Furthermore, participants from the deliberate workshops presented the common impression that most T&Cs are lengthy, repetitive, intricate, and sometimes in unreadable font size. (See Table 4b(i))

More than half of the respondents (61.9%; somewhat agree: 36%, strongly agree: 25.9%) found reporting inappropriate or harmful content to a platform ineffectual, as they believed the platform

would not take it seriously (M = 3.88, SD = 0.9). This result echoed focus group participants' previous encounters with unresponsiveness from tech companies when they reported their concerns and problems to the corresponding platforms. (See Table 4b(ii))

65.7% of respondents (somewhat agree: 42%, strongly agree: 23.9%) perceived that they have no say over the issue even if there were negative impacts on society created by tech companies' websites or apps (M = 3.80, SD = 0.96). This question seeked to elicit respondents' perceived influence in voicing their concerns towards tech companies. Incessant scandals of data breaches and misuse have come to light in recent years. Individual users have limited feasible countermeasures, with quitting the services as the ultimate solution. (See Table 4b(iii))

Half of the respondents (52%, somewhat agree: 32.9%, strongly agree: 19.1%) felt changing privacy settings did not have an impact on protecting their personal privacy (M = 3.63, SD = 0.92). 59.6% of them (somewhat agree: 36.4%, strongly agree: 23.2%) also found the related settings user-unfriendly, taking too much time and effort (M = 3.7, SD = 1.02). Customizing personal privacy settings requires extra effort from respondents to identify and understand their own needs, rights, and potential subsequent restrictions. They considered the effort of adjusting respective settings disproportionate since the outcomes were deemed uncertain. The default privacy setting (e.g., on cookies collection) suggested by the tech companies was regarded as a more convenient option for the respondents, which further discouraged them from making privacy setting changes. (See Table 4b(i) and 4b(ii))

While tech companies are expected to prioritize the benefit of their users, only 30.0% of the respondents (somewhat agree: 17.5%; strongly agree: 12.5%) recognized that tech companies would take their users' best interests into account when they design and operate their websites and apps (M = 2.99, SD = 1.18). This result suggested a strong distrust of tech companies among respondents. They believed that tech companies value their own profits more than their users' interests. Several participants attended the deliberate workshops even expressed their speculation of tech companies' total ignorance of users' interests, with the occurrence of companies sharing access or selling users' information to third parties. (See Table 4b(ii))

Respondents (41.5%, somewhat agree: 29.3%; strongly agree: 12.2%) tended to believe that they were not safe online and that they would be deceived or harmed in some way (M = 3.38, SD = 0.97). In both the survey and focus groups, we are aware that respondents presented high levels of concern about the potential harm that they could encounter online and most of them believed that harmful experiences online were inevitable. (See Table 4b(iii))

67.5% of the respondents (somewhat agree: 37.6%; strongly agree: 29.8%) believed that their personal information and data online was being collected excessively (M = 3.92, SD = 0.96). Respondents were not only aware that their data is being collected but also consider the collection excessive. (See Table 4b(x)). Several participants of the deliberate workshop recalled their experiences of filling in unreasonably excessive personal information when registering accounts online. They expressed dissatisfaction towards tech companies' requirements of collecting unnecessary private information, especially sensitive ones (e.g., ID numbers) that mismatches with

their acquired services. They were concerned about how the tech companies store and use their information. Some even suspected that tech companies were selling their data to third parties without their authorization. Along with such an accusation, only one-fifth of the respondents (21.8%, somewhat agree: 12.3%; strongly agree 9.5%) agreed that tech companies were protecting their users' personal data (M = 2.87, SD = 1.12). Instead, 59.1% of respondents (somewhat agree: 34%; strongly agree: 25%) argued that tech companies were violating their privacy by collecting their personal information and data online (M = 3.69, SD = 1.07). It is thus not surprising that respondents (11.8%; somewhat agree: 9.7%; strongly agree 2.1%) barely welcomed tech companies' collection of their personal information and data for service improvement (M = 2.28, SD = 1.06). (See Table 4b (ix), (x) and (xi))

Most respondents (90.1%; somewhat agree: 38.2%, strongly agree: 51.9%) firmly believed that their personal data should be protected (M = 4.4, SD = 0.734). While they expressed disappointment towards tech companies' insensibilities on personal data protection, what about their expectations of the government? We found that only 30% of respondents (somewhat agree: 16.2%; strongly agree: 13.8%) agreed that the government was protecting Internet users' personal data (M = 2.91, SD = 1.32), and more than half of the respondents (52.4%; somewhat disagree: 16.9%, strongly disagree 35.5%) refused the government to collect their data to improve public services (M = 2.51, SD = 1.43). The above results highly resemble our observation in deliberate workshops and focus group interviews that participants manifested their concern over third parties' mismanagement of users' personal data. They were highly aware of their own responsibilities for protecting their information and data. (See Table 4b(ix) and (xi))

4c. Data Right

In this part, we explored respondents' perceived rights online. They upheld anonymity when engaging in online activities. 59.6% of respondents (somewhat agree: 28.4%; strongly agree: 31.2%) believed that citizens should have the right to access the Internet anonymously (M = 3.71, SD = 1.20). 58.6% of respondents (somewhat agree: 27.6%; strongly agree: 31.1%) agreed that citizens should have the right to anonymously express their opinions online (M = 3.67, SD = 1.21). They believed they should be able to articulate their opinions and ideas online freely. Around half of the respondents (49%; somewhat agree: 21.2%; strongly agree: 27.8%) believed that they should have the right to speak freely online, even if the speech is harmful or controversial (M = 3.43, SD = 1.34). 43% of the respondents (somewhat agree: 12.7%; strongly agree: 30.3%) deemed that citizens should have the right to speak freely online without being subject to criminal charges (M = 3.27, SD = 1.49). While most of the respondents upheld anonymity online and against suppression in free expression, participants from the workshop and focus groups also expressed their concerns over hate speech and cyber-bullying. They were thus not unaware of the potential limitations or boundaries of free speech. (See Table 4c(i) and (ii))

4d. Data Usage

From the previous sections, we observed that respondents were in general reluctant about offering their data and information to tech companies yet stuck in an inevitable situation of having to

comply as they hoped to use the services. Nevertheless, from the perspective of tech companies, information collection is essential for their operation and the customization of services for their users. Thus, we hoped to further explore in which circumstances the respondents were more willing to provide their information. Questions below are adopted from Ofcom's "Adults' media use and attitudes report 2020".

Respondents presented a tendency toward unwillingness to disclose personal information in exchange for extra services or benefits. More than half of respondents (53.6%; somewhat: 25.8%; strongly unwilling: 27.8%) would not give out their personal information for free services like public Wi-Fi networks (M = 2.35, SD = 1.09). 55.9% of them (somewhat unwilling: 29.2%; strongly unwilling: 26.7%) would not offer their information for more relevant advertisements or information, such as discount offers (M = 2.37, SD = 1.13). 45.9% of the respondents (somewhat unwilling: 27.2%; strongly unwilling: 18.2%) were unwilling to offer their personal information for personalized services like weather service and video recommendations (M = 2.61, SD = 1.10). Among the options, respondents were the least reluctant to offer their information for creating personal accounts (35.3%; somewhat unwilling: 23.3%; strongly unwilling: 12%) (M = 2.8, SD = 1.02), as personal accounts are a prerequisite to many websites or apps' services. Respondents expressed moderate unwillingness to provide personal information to improve the service quality of the website or app (M = 2.62, SD = 1.02). (See Table 4d(i) and (ii))

However, the negative scene had shifted when the data usage is clearly stated on the website or app (M = 3.02, SD = 1.04). It helped improve users' understanding and thus reduce perceived

ambiguity and distrust. Furthermore, when respondents could exercise control over how much personal information and data they share with tech companies, their unwillingness to share personal information receded significantly (18%; somewhat unwilling: 12.2%; strongly unwilling: 5.9%) (M = 3.35, SD = 1.03). The result implies that respondents were not completely averse to sharing personal information with tech companies, especially in circumstances where they are sufficiently informed about how their data is managed and have the agency to select what data to disclose. (See Table 4d(iii))

Aside from simply concluding that respondents were aware that tech companies collect information about them, we would like to delve deeper into respondents' perspectives on the ways that their information is collected. This question was adopted from Doteveryone's "People, Power, and Technology: The 2020 Digital Attitudes Report".

Most respondents (89.7%; likely: 37.7%; definitely: 52%) agreed that traces of their internet activities, like browsing and search history, are collected by website or app providers (M = 4.38, SD = 0.81). And 80.3% of them (likely: 39.7%; definitely: 40.6%) believed that information about them on social media, like their posts and photos, are collected by website or app providers (M = 4.19, SD = 0.84). The results suggest that respondents strongly believed that website or app providers were collecting records of their online activities and their information online.

63% of respondents (likely: 38.4%; definitely: 24.5%) believed the website or app providers would collect the information that they stored on their electronic devices, e.g., mobile phones and

computers (M = 3.74, SD = 1.05). And more than half of them (54.5%, likely: 38.4%; definitely: 24.5%) assumed that their biometric data records, such as fingerprints and facial features, would be collected (M = 3.62, SD = 1.12). It's worth noting that respondents suspected websites or app providers who were able to collect both kinds of information that were ostensibly stored offline without authorized access. The above result may reflect respondents' sense of insecurity in the online world, with a belief that their digital data would be collected.

75.5% of respondents (likely: 36.7%; definitely: 38.8%) believed their habits in real life, including hobbies and frequently visited places, would also be collected (M = 4.06, SD = 0.97). When the use of the internet has highly penetrated into people's daily lives, traces of their offline habits are inevitably collected. Participants in deliberate workshops, for example, believed that their mobile phones' GPS functions collect data about where they go; and their browsing records, such as videos watched and information searched, could easily point to their hobbies. (See Table 4d(ix) and (x))

4e. Data Truthfulness

We observed that the respondents were highly aware of data truthfulness. Only 4.5% of them (somewhat agree: 4.4%; strongly agree: 0.1%) believed that online information is always true (M = 2.32, SD = 0.88). And only 8.4% of the respondents (somewhat agree: 0.5%; strongly agree: 7.9%) indicated that they did not care about the truthfulness of the information they searched online (M = 2.07, SD = 0.96). It shows that the circulation of misinformation has been a concern among respondents. They were skeptical about the authenticity of the online information obtained.

(See Table 4e(i))

Most respondents (77.5%, somewhat agree: 59%; strongly agree: 18.5%) were likely to acquire information that resonates with their interests when searching the internet (M = 3.92, SD = 0.74). In Part 2, we observed a prevalence (90.5%) among respondents of using search engines to look up information online. 59.9% (somewhat agree: 59%; strongly agree: 18.5%) of them presented a tendency to look for online information that aligns with their views (M = 3.61, SD = 0.90). When respondents only encountered information that matches their views, their exposure to diverse perspectives and opinions would be limited due to the echo chamber effect. Participation in an echo chamber amplifies and reinforces people's existing beliefs and lowers their capacity to consider opposing viewpoints and discuss complicated issues. While they are fueled by confirmation bias that only favors information that shares their beliefs, they may result in spreading misinformation and hate speech to defend themselves and distort others. In the long term, echo chambers may lead to political polarization and extremism. (See Table 4e(i) and (ii))

Part 5. Data Literacy

5a. Data Thinking

Online users view and analyze the world through data in their everyday lives. We hoped to explore citizens' critical understanding of data and their abilities to deal with data. (See Table 5a)

(i). Critical data analysis (16. DT3）

"Critical data analysis" refers to "the ability to consider, examine, and discuss data bias, methodological errors, and inaccurate data visualization" (Pawluczuk, Alicja, et al. 2020). In question 16, we presented the respondents with four examples of data visualization and asked them to identify the data visualization that is free from visualization mistakes. Data visualizations synthesize and unmask the meaning of complex raw data into insightful takeaways with visual means of data patterns, trends, and correlations. Deceptive data visualization may lead to false interpretation and biased decision making. For option A, the pie chart is the wrong visualization method for the presented data. For option B, reverse scale on the Y-axis may lead to a wrong interpretation of the trend. For option C, omitting the baseline and starting the Y-axis above zero creates a misleading impression of a pronounced difference. 50.8% of the respondents picked the right answer, option D, the only chart without misleading components. The result implies that around half of the respondents might potentially be misled due to their inability to understand visualized data. Increased awareness and further education on data interpretation, like criteria of effective data visualization and common tricks concealed in manipulated data visualization, can help improve their data analysis skills and ability to conduct data communication with integrity.

(ii) Understanding data society (17 DT7)

"Understanding data society" refers to "the ability to understand the way the data economy works (e.g., how platforms are funded, what cookies are, broadly what algorithms do)" (Pawluczuk,

Alicja, et al. 2020). In question 17, respondents were asked to identify the correct description of "cookies". "Cookies" are small text files that store information about users' activities for later processing on a website in a web browser. 53.9% of the respondents chose the correct option B. It is worth noting that 15.9% of the respondents mistakenly believed that they do not have the right to clear, enable, and manage cookies. In recent years, privacy laws, like the EU's ePrivacy Directive and General Data Protection Regulation (GDPR), have addressed and governed the use of cookies. Many websites are therefore displaying cookie banners to inform their users about the cookie use and, if necessary, obtain users' explicit consent through affirmative action, e.g., by clicking "Accept" for the corresponding use. Some websites allow their users to view and indicate their consent for respective cookie processing purposes on the banners. We suggest that online users pay attention to the mechanism of data society to better protect their privacy online.

(iii) Privacy (18, DT5)

"Privacy" refers to "the ability to consider and implement privacy-protective behavior when using data (e.g., using avatars, deleting tweets every couple of weeks)" (Pawluczuk, Alicja, et al. 2020). In question 18, we asked respondents about the approach to be taken when using public Wi-Fi services to avoid potential security risks like personal information leakage. 61% of the respondents correctly identified using a Virtual Private Network (VPN) service when using public Wi-Fi is the best option given. VPN is a privacy protection tool that hides users' Internet Protocol (IP) addresses and encrypts their internet traffic. It prevents users' identities, digital traces, and personal information from being tracked and recorded by tech companies, internet service providers, and

cybercriminals. VPN is one of the most common online security tools available on the market for online users who seek to strengthen their online privacy protection. However, VPNs are not necessarily the perfect tool for every occasion. Online users can also protect themselves from online privacy violations by employing various measures, including changing social media privacy settings.

(iv) Problem solving using data (19 DT6)

"Problem solving using data" refers to "the ability to search for, identify, and use data to solve problems (e.g., open data projects)" (Pawluczuk, Alicja, et al. 2020). Many people now turn to search engines to find solutions to their daily problems. In question 19, we asked respondents to analyze search engine results, testing their awareness of targeted advertising. 61.8% of the respondents spotted that the result quoted was an advertisement or sponsored link. Search engine marketing allows enterprises to reach their target audiences via search engines. For example, when users search for keywords related to a product or service on the Google search engine, Google search partners' advertisements popped up as the users' first result. Pop up ads are generated chiefly by keywords but can also be affected by the user's IP address, browsing history, and search history. Thus, the critical ability to recognize advertisements when aided online users in better sourcing and identifying relevant information, allowing them to solve their problems more efficiently. Furthermore, some participants from the deliberate workshops expressed their privacy concerns and discomfort with how companies targeted them with ads by using their data. To prevent data tracking, they proactively avoided advertisements by using ad blockers or paying for

ad-free experience.

(v) Data Safety (20, 23 DT4)

"Data Safety" refers to "the ability to consider and implement data protective steps when using data (e.g., using private browsing features or more secured browsers, using multiple secured search engines, strong passwords)" (Pawluczuk, Alicja, et al. 2020). In question 20, we asked the respondents about the meaning of Hypertext Transfer Protocol Secure (https). 31.5% of the respondents were able to identify the correct description of option D, which means that the information input to the website will be encrypted on a website with a Uniform Resource Locator (URL) starting with "https://" instead of "http://". HTTP is not encrypted and thus is vulnerable to attacks that will give the attackers access to website accounts and sensitive information like passwords, credit card information, and browsing history, and modify web pages to inject malware or advertisements. HTTPS is designed to withstand such attacks and is considered secure against them with encryption. 54.6% of the respondents answered "Don't Know" to this question. The result is not surprising, as a lock logo is usually presented at the front of the address bar in browsers to indicate a website connection is "secured" instead of showing "https" nowadays. Furthermore, most browsers also display a warning to the user when visiting a site that contains a mixture of encrypted and unencrypted content in a dialogue box or by showing it across the entire window.

Data security has long been a concern for online users in terms of practical data management. Most websites and apps require users to create passwords in order to protect personal information and

data in their accounts. However, password protection alone may not be sufficient for some websites and apps that involve sensitive personal information, like online banking, as passwords can be stolen, guessed, and hacked on the sly. And it would be too late for the users to find out their accounts were being sneaked into. In question 23, we asked respondents to identify an example of multi-factor authentication. Multi-factor authentication is an electronic authentication method in which a user is granted access to a website or application only after successfully presenting two or more pieces of evidence or factors to authenticate their identities when they log in. 57.4% of the respondents chose the correct option D, which illustrates an SMS-based verification, one of the multi-factor authentication methods. To authenticate, users can use their personal access codes to the SMS-receiving device, typically password or biometric features, plus a one-time-valid, dynamic passcode, usually consisting of 4 to 6 digits, sent to them by SMS. The function adds extra security to the user by preventing third parties from breaking into the account even if the single password is leaked. For other options, option B illustrates the function of security questions, and options A and C are forms of “Completely Automated Public Turing Test to Tell Computers and Humans Apart (CAPTCHAs)." They are computing tests to distinguish humans from robots and computers.

(vi) Understanding of data collection (21 DT8)

“Understanding of data collection” refers to “the ability to understand the different data-collection practices of different institutions (e.g., governments, advertising organizations, data brokers) as well as different databases (e.g., National Health Service, local government voter registers, data

brokers)" (Pawluczuk, Alicja, et al. 2020). Most common web browsers are offering private browsing options, allowing their users to go "incognito". With the function on, browsers will not record users' search and browsing history, cookies, and credentials like passwords. However, users are not completely "anonymous," meaning that their digital traces can still be tracked by the sites visited, Internet service providers, and authorities that manage the network. In question 21, respondents were asked to pick the right description of the private browsing function. 45.3% of the respondents answered "Don't Know" and 37.5% of the respondents picked the right option: the feature prevents web browsers from storing information like browsing history. The result may imply that some respondents were not aware of the provision of private browsing features. It is worth noting that institutions collect data through various practices; the private browsing mode may only be partially effective in preventing implicit data collection. Improving the understanding of data collection mechanisms would help online users take relevant measures, like tracker blocking tools, to strengthen their information and data protection.

(vii) Awareness of data protection rights (22 DT1)

"Awareness of data protection rights" refers to "being aware of local or regional data protection policies and laws" (Pawluczuk, Alicja, et al. 2020). In question 22, respondents were asked about the principles of the Hong Kong Personal Data (Privacy) Ordinance (PDPO). The ordinance was passed in 1995 and took effect in December 1996. It aims to protect people's privacy rights in relation to their personal data. It is one of Asia's longest-standing comprehensive data protection laws. The ordinance consists of six data protection principles. 66.3% of the respondents picked the

right description of the principles, which data users are required to collect personal data of data subjects in a fair and lawful way. Option B is incorrect as the Data Collection Principles clearly specify that personal data must be collected for a purpose directly related to the data user's activity and data subjects must be notified of the above purpose. Option C is incorrect as the Data Access and Correction Principle states that a data subject must be given access to his/her personal data and allowed to make corrections if it is inaccurate. Finally, option D is incorrect as the Data Security Principle explains that a data user needs to take practical steps to safeguard the personal data they collect. While the Ordinance is set up to better safeguard citizens' privacy and personal data, participants from the deliberate workshop had mixed views on the effectiveness of the Ordinance in protecting their data. Some found the Ordinance ambiguous and disconnected with common online activities in real life and concerned about the limitations of the ordinance, like its obsolescence and lack of deterrent effect. Numbers of participants have shown high reluctance towards the government's control of online activities. They believed that the regulations would probably restrict the publishing of online content and opinions, thereby interfering with online users' freedom of expression.

(b) Data Doing

"Data Doing" refers to the practical, critical, and ethical skills in data handling and management, particularly in response to challenges associated with data usage (Pawluczuk, Alicja, et al. 2020).

(i) Data Creation (DD4)

"Data Creation" refers to "the ability to create data in different formats (e.g., creation of a blog post, social media post/ hashtag, presentation)" (Pawluczuk, Alicja, et al. 2020). In question 24, we asked respondents about their social media usage, particularly in information sharing. We observed that respondents (46.8%; somewhat disagree: 26.1%; strongly disagree: 20.7%) were quite reluctant to share information about themselves on social media (M = 2.49, SD =1.05). They also expressed low interest (39.4%, somewhat disagree: 23.7%; strongly disagree: 15.7%) in sharing information about community or public affairs on social media (M = 2.67 and SD = 1.05). Echoing the survey result, participants from the deliberate workshop and focus group interviews expressed a strong sense of resignation on social media. A number of participants stated that they were more hesitant to share information about themselves and political affairs on social media today than they were a few years ago. Instead of revealing their real names, some participants had set up social media accounts with fake identities when touching on public affairs so as to keep themselves away from the literary inquisition. Although ordinary citizens have not been getting any explicit warnings, they have been really cautious not to actively advocate or spread their own political agenda since the National Security Law came into effect. It has consequently had a chilling effect across the territory. People are drawing red lines and censoring themselves more lately to avoid potentially falling afoul of authorities. (See Table 5b(i))

(ii) Accessing and Assessing (DD1 & DD2)

"Accessing" refers to "the ability to search for, identify and access services, websites and data"

(Pawluczuk, Alicja, et al. 2020). And "assessing" refers to "the ability to evaluate data quality and credibility (e.g., fact-checking, checking sources of social media posts)" (Pawluczuk, Alicja, et al. 2020). Half of the respondents (50%; somewhat agree: 40.4%; strongly agree: 9.6%) indicated that they often collect information from various sources, even if it differs from their position or point of view (M = 3.47, SD = 0.88). Collecting information from sources with different points of view is one way to prevent the echo chamber effect. 57.2% of the respondents (somewhat agree: 43.5%; strongly agree: 13.7%) reported that they often collect information from various sources to verify information (M = 3.63, SD = 0.86). Verifying information with multiple sources is an effective method of fact-checking, promoting the veracity and correctness of reporting. Many focus group participants have expressed their concerns about the circulation of fake news. According to the Proceedings of the National Academy of Sciences, fact-checking curtails belief in misinformation and leaves a more enduring mental imprint than false claims (Porter and Wood, 2021). Through analyzing various sources of information, people are able to restore the full original context as well as identify misplaced or manipulated information and ascertain reliable information from credible sources for countering misinformation. Furthermore, assessing the credibility of information sources based on their reputation, stance, credit, and motivation can assist audiences in distinguishing incorrect or misleading information. (See Table 5b(ii))

(iii) Data Management and Data Deletion (DD6 & DD8)

"Data management" refers to "the ability to store, encrypt, and manage data in a safe and secure way" (Pawluczuk, Alicja, et al. 2020). And "data deletion" refers to "the ability to delete data (e.g.,

deletion of cookies, browsing history)" (Pawluczuk, Alicja, et al. 2020). In question 26, respondents were asked to indicate if they had taken the listed measures to protect their information and data. 60.6% of the respondents (likely: 45.4%; definitely: 15.2%) were prone to using a firewall or other antivirus software (M = 3.63, SD = 0.96). 50.8% of the respondents (likely: 40.8%; definitely: 10%) were inclined to use filtering software to filter out spam, calls, or advertisements (M = 3.47, SD = 0.92). 27.7% of the respondents would about half the time delete cookies or browsing history in their web browser regularly and 25.6% indicated that they would not do so (M = 3.12, SD = 1.12). 34.9% of the respondents would about half the time modify their privacy settings on websites or apps when it is needed (M = 3.09, SD = 1.00). 28.2% of the respondents likely prefered websites or apps that are more secure, like Telegram and Signal (M = 3.34, SD = 1.11). 35.9% of the respondents likely used multi-factor authentication feature on websites or apps (M = 3.46, SD = 0.95). 42.6% of the respondents likely created strong passwords on their devices, websites or apps (M = 3.8, SD = 0.93). 32.4% of the respondents likely did software updates when newer versions are available (M = 3.65, SD = 1.02). Among the above measures, respondents were more inclined to use strong passwords, firewalls, and software updates to protect their data. Filtering software and the use of multi-factor authentication features were also common. As mentioned above, respondents were unwilling to spend too much time and effort modifying privacy settings on websites or apps given the uncertain effectiveness of data protection and handiness of default settings. It also explains the reason that respondents expressed comparatively lower likelihood among the options to delete cookies or browsing history in web browsers, as they were not aware of the necessity to do so.

Using a VPN was the least likely action that the respondents would take. 31.3% of them indicated that they use a VPN about half the time, and 27.5% of them said they were unlikely to use a VPN (M = 2.81, SD = 1.08). Top notch VPNs are usually chargeable. The result shows that the respondents might be resistant to investing monetarily to obtain extra personal data protection. (See Table 5c(i), (ii) and (iii))

5c. Data Participation

Given the data society's collective and interconnected nature, "data participation" refers to citizens' proactive social engagement with data as well as their literacy networks (Pawluczuk, Alicja, et al. 2020). We investigated how people use their existing skills to participate in their increasingly digitized society, how they exercise their rights; and how they shape their collective data experience. While we were aware of how societal factors and circumstances influence people's data participation, we had designed our questions based on the Hong Kong context. For example, we are aware that Hong Kong citizens' approaches to data activism have become more reserved, particularly since the implementation of the National Security Law.

(i) Participation in society using data

“Participation in society using data” refers to “the ability to use data for societal participation and civic action” (Pawluczuk, Alicja, et al. 2020). We divided data participation in society into two dimensions. First, we asked respondents about their usage of online government services. Second,

we asked respondents about their online socio-political participation.

35.8% of the respondents indicated that they rarely look up public service information online via government websites, such as GovHK, and 32% of them sometimes do so (M = 2.86, SD = 1.02). 31.1% of the respondents sometimes used government online services, and 25.4% of them often did so (M = 3.20, SD = 1.09). 39.5% of the respondents had never participated in public consultations launched by the government nor expressed their opinions to relevant bodies online, while 34.3% of them indicated that they rarely did so (M = 1.96, SD = 1.00). We observed that most respondents had experience of looking up public service information online (94%) and using government online services (95.5%). Most government services are available both online and offline. Nonetheless, since the pandemic, we have seen the government promote digital public service in the suspension of face-to-face services during the adoption of targeted social distancing and infection control measures. Electronic consumption vouchers and the “leavehomesafe” app are examples of prevalent e-government measures. Some participants in the deliberate workshop expressed concern about society's digitization, believing that the government should provide offline options for citizens who are less data literate or choose to opt out of digitalization. (See Table 5c(i))

Regarding respondents’ online socio-political participation, we asked the respondents if they had participated in joint signature campaigns online or participated in instant message or social media groups that were related to public affairs or community affairs. 41.5% of the respondents sometimes participated in joint signature campaigns and 25.9% of them rarely did so (M = 2.67,

SD = 1.02). 36.8% of the participants rarely participated in social media groups related to public affairs or community affairs, and 28% of them sometimes did so (M = 2.24, SD = 0.98). Many of the respondents had experience of participating in joint signature campaigns (84.3%) and social media groups related to public affairs or community affairs (73.6%). (See Table 5c(ii))

Most participants in deliberate workshops and focus groups had implicitly presented a proclivity for keeping a low profile in socio-political participation. The retroactive nature of the NSL is one of their concerns. Since the National Security Law came into force, at least 58 civil society organizations, including unions, churches, media groups, and political parties, have disbanded (Hong Kong Free Press 2022). The law has been accused of silencing dissent and suppressing civil society (Li 2021). We observed that the prevalence of community affairs groups on social media is a sign of a thriving local economy or ecosystem, as well as an alternative form of social participation.

(ii) Data Activism

"Data Activism" refers to "the ability to take proactive steps to protect individuals' and collective privacy and well-being in the data society" (Pawluczuk, Alicja, et al. 2020). Respondents presented a strong inclination (56.1%, likely: 31.6%; definitely: 24.2%) to report to the platforms when they receive suspected spam messages on instant messenger or in social media groups (M = 3.61, SD = 1.13). 47.4% of the respondents (likely: 37.3%; definitely: 10.2%) stated that they would clarify with others in the group when they received untrue information in an instant message

or social media group (M = 3.35, SD = 1.00). They were also likely (43.4%, likely: 29.2%; definitely: 14.2%) to report to the platforms when they see inappropriate or harmful content online (M = 3.27, SD = 1.10). The findings indicate that respondents were taking proactive steps to help construct a desirable online environment. For example, their reporting of spam messages and clarification of untrue information to others could stop the further circulation of misinformation and its potential negative outcomes. They expressed the least likeliness (33.1%, likely: 23.8%; definitely: 9.2%) to report to the platforms when they find misinformation circulating on social media (M = 3.01, SD = 1.10). The abundance and high visibility of misinformation online makes reporting every single one of them difficult and impractical. They might also see misinformation posing less direct harm than that from inappropriate or harmful content or suspected spam messages. (See Table 5c(iii) and (ix))

(iii) Supporting others with their data literacy

"Supporting others with their data literacy" refers to "the ability to help others with their data literacy (e.g. helping others with their privacy settings)" (Pawluczuk, Alicja, et al. 2020). 53% of the respondents (likely: 33.7%, definitely: 19.3%) had helped others to keep their data safe, like helping others to set up secure passwords or change privacy settings (M = 3.66, SD = 1.03). The result points to respondents' belief that maintaining cyber security is a shared responsibility; individuals cannot stay aloof or immune from it. Collective effort can create synergy in safeguarding data privacy and security. (See Table 5c(x))

Many participants from the deliberate workshop shared their experience of helping others with their data literacy. The most common case was teaching their elderly significant others to use the "leavehomesafe" app. Most of them found the teaching experience difficult but inevitable as it is stipulated by the government. They believed that people who are less digitally capable due to insufficient technological resources, knowledge, and ability, like the elderly and low-income groups, are gradually overlooked in the increasingly digitalized society. It is more difficult for these people to keep up with technological trends and are more likely to fall into digital crises such as data leakage and internet fraud. The fully digitized systems of virus risk notification, vaccination appointments, and consumption voucher schemes all illustrate the inevitability of the digitization process. If citizens do not follow, they may lose their freedom, rights, and even personal identity.

Summary of Survey Finding

Internet use is indispensable and unavoidable in people's daily lives. Around 30% of the respondents were heavy users who spend one-third of their day online. Respondents actively engaged in socialization and entertainment online. Almost all respondents had a social media profile and were keen to socialize with others via instant messenger. They were also inclined to obtain news and information online.

Respondents were more confident that digital measures, like using multi-factor authentication, changing privacy settings, and the use of antivirus software, could keep their information safe. Yet they were more conservative in claiming that they (themselves) are confident in protecting their

own data and information. They were more certain about their ability to determine the truthfulness of information and solve problems with online information but find the terms and conditions (T&C) of websites and apps dubious and managing who can access their data difficult.

Most respondents believed that data technology has improved their personal lives but were less convinced that data technology has improved their society. We observed a significant perceived power imbalance between respondents and technology companies, and respondents presented a severe sense of powerlessness. For example, most respondents (82.6%) believed that they had no control over how their personal information and data were collected and used by tech companies. And they (65.7%) had no say even though tech companies are creating negative impacts on society. Reporting inappropriate or harmful content to online platforms was deemed ineffective, as respondents believed that the platforms would not take their reports seriously. In general, respondents showed a strong distrust towards tech companies, believing that companies value their own profits more than their users' interests.

Although respondents hoped to personalize their digital experiences based on their needs and preferences, constraints imposed by technology companies significantly hindered them from doing so. Respondents presented a strong feeling of resignation towards websites and apps' terms and conditions (T&C). The lengthy, repetitive, and intricate nature of the T&C, as well as the requirement to accept T&C in order to use the service, strongly discourage users from reading and comprehending the T&C. Respondents were also highly discouraged from customizing their personal privacy settings as they considered the time and effort of adjusting respective settings too

high and not particularly effective in protecting their personal privacy. Following the default privacy setting suggested by the tech companies is a more convenient option for the respondents.

Respondents were generally reluctant to offer their data and information to tech companies but saw it as inevitable to comply as they hoped to consume the corresponding services. They were dissatisfied with technology companies requiring their private and sensitive information excessively, which mismatched with their acquired services. While respondents presented a tendency toward unwillingness to disclose personal information in exchange for extra services or benefits, they were not completely averse to sharing personal information with tech companies, especially in circumstances where they are sufficiently informed about how their data is managed and have the agency to select what data to disclose.

Respondents were feeling insecure in the online world. Most of them believed that harmful experiences online were unavoidable and were highly concerned regarding potential risks. On top of collecting records of their online activities and their information online, they strongly believed that website or app providers were collecting digital data that they stored offline even without explicit authorized access, such as their biometric data saved on devices. And when the internet has penetrated people's daily lives, data tracking on their offline habits are becoming inevitable.

Most respondents upheld anonymity, believing that they have the right to access the Internet and express their opinions anonymously. And they were against the suppression of free expression or literary inquisition of all forms.

Data truthfulness had been a concern among respondents. They were skeptical about the authenticity of the online information obtained. We observed that respondents exhibited a tendency to look for online information that aligns with their views. When respondents only encounter information that matches their views, their exposure to diverse perspectives and opinions would potentially be limited due to the echo chamber effect.

"Data Thinking" explores online users' critical understanding of data and their abilities to deal with data. We asked the respondents eight "data thinking" questions that covered seven aspects of critical ability. Respondents answered 4.3 questions correctly on average. While we could hardly claim that we could fully assess respondents' "data thinking" ability with only eight questions, the result implied that online users' inability in "data thinking" would result in vulnerability to misleading and manipulated online information. They are advised to acquire sufficient knowledge of the mechanisms of data society to better protect their privacy and data online, like implementing relevant digital measures.

"Data Doing" refers to online users' practices in data handling and management, particularly in response to challenges associated with data usage. Respondents were more reluctant to share information about themselves and political affairs on social media, presenting a strong sense of resignation on social media and a tendency towards self-censorship.

Respondents concerned about the circulation of misinformation and acknowledge the importance

of fact-checking. Analyzing information from various sources helps online users in distinguishing incorrect or misleading information that are misplaced or manipulated.

Respondents' concerns and distrust towards third parties' data management were not limited to tech companies but also the government. They were highly aware of their own responsibilities for protecting their information and data. Respondents were more inclined to use strong passwords, firewalls, and software updates to protect their data. They were comparatively unwilling to spend time and effort modifying privacy settings and deleting cookies or browsing history on websites or apps, given the uncertain effectiveness of data protection and the handiness of default settings.

"Data Participation" refers to online users' proactive engagement with data society. We observed that most respondents had been involved in E-government. They had experience of looking up public service information online (94%) and using government online services (95.5%). Respondents were proactively keeping a low profile in socio-political participation. The launch of the National Security Law was the major concern. Yet the prevalence of community affairs groups on social media is a sign of a thriving local civil society as well as an alternative form of social participation.

Majority of the respondents were taking proactive steps to help construct a desirable online environment. For example, their reporting of spam messages and clarification of untrue information to others could stop the spread of misinformation and the potential derived negative outcomes. Respondents also presented a strong tendency to help others in safeguarding their data

and information. The above result points to respondents' belief that safeguarding data privacy and security is a shared responsibility that requires collective effort; individuals cannot stay aloof or immune from it.

Experiment Findings

Compared with the control group, participants of both intervention groups, information exposure (group A) and deliberation workshop (group B), improved in data literacy. The result illustrates a positive educational effect for both treatments. The interventions were designed based on two different ways of communication. Participants of the information exposure group received one-way short posts via WhatsApp for a month. There were a total of 16 short posts. Deliberation workshops' participants took part in a one-time 2.5-hour discussion session led by trained moderators. Each group consisted of 8-10 participants. In the following sections, we first compared the results of two interventions, and then we discussed their corresponding implications.

Data Thinking

Participants of both intervention groups demonstrated a better critical understanding of data and an improvement in their abilities to manage data (See Table E1). We observed a general increased trend in the correct rates of question 16-23 and the respective proportion of participants answering "don't know" had reduced. We will illustrate the statistically significant results of each question in the following.

Q16 aimed to assess participants' ability in critical data analysis via identifying inaccurate data visualization. For group A, number of participants answering wrong answers decreased from 20% to 14% (p<.05*). For group B, the correct rate increased from 71% to 79% (p<.05*). Q17 aimed

to assess participants' understanding of data society, particularly on their understanding of cookies. For group B, the correct rate increased from 78% to 88% (p<.01**), participants answering "don't know" decreased from 12% to 6%. Q18 assessed participants' awareness of privacy through the scenario of using public Wi-Fi and their understanding of VPN. For group B, the correct rate increased from 83% to 92% (p<.01**), the rate of don't know decreased from 14% to 5% (p<.01**). Q20 assessed participants' awareness of data safety based on their knowledge of internet protocol. For group A, the correct rate increased from 60% to 72% (p<.01**), the rate of "don't know" decreased from 32% to 18% (p<.01**). For group B, the correct rate increased from 61% to 71% (p<.05*), the rate of "don't know" decreased from 29% to 19% (p<.05*). Q21 assessed participants' understanding of data collection via their knowledge about private browsing mode. For group A, the rate of "don't know" decreased from 17% to 10%. Q22 assessed participants' awareness of data protection rights, particularly on Hong Kong Personal Data (Privacy) Ordinance. For group A, the correct rate increased from 75% to 91% (p<.01**), the rate of "don't know" decreased from 19% to 6% (p<.01**). For group B, the correct rate increased from 75% to 85% (p<.01**), the rate of "don't know" decreased from 19% to 9% (p<.01**). Q23 assessed participants' awareness of data safety via their understanding of multi-factor authentication. For group A, the number of participants picking the wrong options decreased from 27% to 17% (p<.05*) and the correct rate increased from 69% to 80% (p<.05*). In Q19, we see no statistically significant changes across both groups. We believe this is because we did not specifically touch on the relevant issue in both interventions.

Group A participants made improvements in 5 questions (Q16, 20-23). The result is not surprising,

as the WhatsApp short posts covered topics about Q16 (data visualization), Q17 (cookies), Q18 (public Wi-Fi and VPN), Q22 (Personal Data Privacy Ordinance), and Q23 (multi-factor authentication). Participants answered 6.13 and 6.75 questions correctly on average in phase 1 and phase 2 respectively, the difference is 0.62.

Group B participants also made improvements in five questions (Q16-18, 20, 22). We had included information about the Personal Data (Privacy) Ordinance in the pre-workshop material. As the discussion sessions were semi-structured, respondents were free to discuss different issues about data literacy. Thus, the above topics might have been brought up during discussions. Participants improved from answering 6.03 questions correctly on average in phase 1 to 6.67 of that in phase 2. The difference is 0.64.

We identified some limitations for the survey experiment. First, we aimed to test the effectiveness of the two interventions by setting the "data-thinking" questions. We are aware that we cannot assess participants' ability in data thinking thoroughly with only eight questions. Second, we noticed that participants would likely make improvements in "data thinking" in phase 2, as it is their second attempt to do the same set of questions. We had an untreated control group in which participants did not receive any intervention and the correct rates of Q17 and Q22 also improved. For Q17, the correct rate increased from 72% to 80% *(*p<.05*), rate of wrong answers decreased from 16% to 7% (p<.01**). A similar trend is also observed in Q22, where the correct rate rose from 67% to 75% *(*p<.05*). In phase 1, participants answered 5.56 questions correctly on average, and the result improved to 5.94 questions in phase 2. The difference is 0.38, much less than those

in the two intervention groups. As a result, we believe that the two interventions were effective in improving the participants' ability in data thinking.

Data Doing

While participants of both intervention groups made improvements in "data thinking", they would react more proactively in protecting their data and privacy (See Table E2). For group A, participants were more willing to do software updates when newer versions are available (M1 = 3.91, M2 = 4.10, $p<.05$*). For group B, participants were more inclined to use multi-factor authentication feature on websites or apps (M1 = 3.78, M2 = 3.95, $p<.05$*); and create strong passwords on devices, websites or apps (M1 = 3.91, M2 = 4.09, $p<.05$*). We observed no statistically significant change in the control group. From the above result, we conclude that both interventions influenced participants' behaviors. It implies that the interventions had raised participants' awareness towards privacy and data safety, participants thus took the practical recommendations or tips shared via short posts or from other discussants to better safeguard their data and privacy.

Data Participation

There was no statistically significant change in data participation for both intervention groups. We still observed a strong resignation attitude towards data participation among participants. While concerns regarding the suppression of civil and political freedoms are deemed the major cause of

participants' resignations, it was difficult for deliberation workshops and exposure to data literacy information to effectively alleviate their hesitation (or fear) in data participation.

Confidence

Compared with phase 1, group A participants' acquired confidence in firewall or other related antivirus softwares (M1 = 3.90, M2 = 4.11 $p<.01$**) and multi-factor authentication (M1 = 4.30, M2 = 4.45 $p<.05$*) had improved since the invention (Table E3). In the short posts shared, installing antivirus software was suggested to protect data and privacy. And one of the short posts was dedicated to multi-factor authentication and its functions. As a result, it is not surprising that participants were more confident that the two digital measures mentioned above could protect their privacy and data (See Table E3).

However, group B participants were less confident that using a VPN could keep their information safe (M1 = 4.03, M2 = 3.87, $p<.05$*) since the intervention. We noticed that VPN is frequently mentioned in deliberate workshops as a digital measure taken by participants to protect themselves in the online world. Some participants had discussed the drawbacks of VPN, like how it had prevented them from accessing some Hong Kong websites and how the use of VPN had reduced their Internet speed. Other participants shared their uncertainty about the effect of VPN in protecting their data and privacy. Due to the above reasons and the cost of the VPN subscription, several participants mentioned that they had stopped using VPN. The sharing of "VPN experiences" among participants might influence some people's perceptions of VPN, leading to a loss of

confidence.

Participants of group A presented an overall self-confidence growth in protecting their own data and information (See Table E4). After intervention, they were notably more confident in securing their personal privacy online (M1 = 3.06, M2 = 3.35, p<.01**), keeping their information safe online (M1 = 3.10, M2 = 3.33 p<.01**), managing who can see the information they share online (M1 = 2.92, M2 = 3.25 p<.01**) and understanding data analysis that they encounter online (M1 = 3.21, M2 = 3.46 p<.01**). Their confidence was also elevated in managing who can use their personal information and data online (M1 = 2.75, M2 = 2.97 p<.05*) and using online public information and data to solve their problems (M1 = 3.24, M2 = 3.47 p<.05*). By understanding the importance of data protection and getting to know the methods to do so, the participants were empowered to recognize the potential threats and risks more easily. It would add to their awareness and make them feel more prepared and capable of better reacting to them.

However, participants from group B were less confident about their abilities to manage who can see the information they share online (M1 = 3.03, M2 = 2.72 p<.01**) and use their personal information and data online (M1 = 2.81, M2 = 2.51 p<.01**). In the deliberate workshops, there was a case study about how commercial companies collect and use their users' data and information online. Many participants expressed their unwillingness to provide sensitive personal information to commercial companies, believing that such data collection is unnecessary and poses risks to them. They mentioned various types of risk, such as (1) commercial companies selling their data to third parties for profit (the incident of Octopus Card Holdings selling personal data of

customers in 2010); (2) data breaches of technology companies (Facebook data leakage in 2021); (3) common examples of credit card frauds and online scams; and (4) skepticism of tech companies' surveillance (spying activities like recording their daily conversations offline). We believed that the sharing of negative experiences, skepticisms, and feelings like fear, anxiety, and anger in deliberate workshops adversely affected some participants' confidence in safeguarding their data, particularly in managing who can see and use their data.

Disempowerment

After the intervention, group A participants felt less disempowered when using data technology. Participants believed that they had slightly more control over how tech companies collect and use their personal information and data (M1 = 4.24, M2 = 4.01 p<.01**). Their defiance toward reading terms and conditions (M1 = 4.43, M2 = 4.19 p<.05*) had reduced. But the highly inclined result shows that participants still hold the conviction that reading terms and conditions is meaningless if they must use the website or app. Their negative perception that changing privacy setting would help protect personal privacy mitigated (M1 = 3.46, M2 = 3.05 p<.01**).Their reluctance towards changing privacy settings because it takes too much time and effort also scaled down (M1 = 3.55, M2 = 3.24 p<.01**), participants were still not convinced that the platforms would take their report on inappropriate or harmful content seriously (M1 = 4.00, M2 = 3.76 p<.01**). And they would remain powerless if the tech company's website or app is creating negative impacts on society (M1 = 3.78, M2 = 3.57 p<.05*). They believed that there are still deception and harm prone in the unsafe online world (M1 = 3.32, M2 = 3.11 p<.05*). Even though

a generally declining trend was observed in the extent of distrust in the mentioned areas, participants' sense of powerlessness was alleviated, while remaining strong. They expected themselves to exert more control over their personal data protection after the intervention but were still frustrated by the beliefs that they could not determine the outcomes or reinforcements that they seeked, and the problems were not solely solvable with their own efforts.

In group B, participants' disbelief towards the tech companies' upholding their users' best interests in mind during websites and apps design and operation intensified (M1 = 2.73, M2 = 2.40 $p<.01$**). We observed that participants of the deliberation workshop were inclined to share their negative opinions and skepticism towards tech companies and certain government policies, focusing on previous scandals and alleged misbehavior. Disbelief and distrust were further intensified in the discussions, leading to the potential "demonization" of tech companies. (See Table E5)

Implication of experiment findings

Participants in both intervention groups, information exposure (Group A) and deliberation workshop (Group B), showed improvement in data literacy, illustrating the positive educational effect of both treatments. However, the two treatments led to diverse influences on participants' level of confidence and sense of disempowerment. We observed that the prolonged and frequent exposure to data literacy information in the WhatsApp Group was more effective in raising participants' confidence in safeguarding their data and information, as well as alleviating their sense of disempowerment in the online world. However, participating in the deliberation workshop

lowered participants' confidence in data management, increasing their sense of insecurity and powerlessness.

The practice of constantly releasing short posts about data literacy in the WhatsApp group to participants manifested a repetition learning effect. It is an efficient way to enhance memory performance (Ebbinghaus, 1964). Participants could thus recollect memories and associate past experiences with new stimuli, leading to a better and longer memory retention. Furthermore, participants might regard the research team, who was responsible for releasing the educational materials, as a more credible and reliable source of information. It was mainly due to the research team's solid academic background and political neutrality. Participants were thus more likely to find the information shared trustworthy, even if they were from the news outlets or the government publication. Such a result further implies and acknowledges the importance of credible institutions, like universities, acting as a bridge between citizens and the government in knowledge transfer.

It was common for participants to share their negative online experiences and sense of insecurity during the deliberation workshop. Some had experiences of encountering online risks, like leakages of credit card information and online scams, that unfortunately led to terrible outcomes. And some participants were particularly skeptical towards the potential government and tech companies' surveillance, acquiring a strong sense of powerlessness and disempowerment. We believe that the occurrences of the spread of fear and anxiety in the discussion were grave but inadvertent. It was manifested and intensified by Hong Kong's political context especially after the NSL rolled out. A number of participants in the deliberation workshops expressed their worries

towards the vagueness of what may constitute an offense of sedition or “endangering national security” and, hence put themselves at imminent risk of criminal prosecution.

While presenting a positive finding on the effectiveness of both interventions, we acknowledge that the interventions had certain limitations. For example, the interventions were not particularly effective in raising participants’ data participation. Focus group participants, who self-identified as being interested in, familiar with, or aware of the topic of data literacy, found the information shared on WhatsApp or in the deliberation workshop diversified but most of it was commonly known.

Focus Group Findings

We conducted ten focus group interviews in October 2022 to further explore survey experiment participants' views and supplement the experiment's findings. Interviewees were recruited from the pool of participants who had taken part in the survey experiment.

Deliberation Workshop

Most workshop participants joined the activity because they were concerned about data security and privacy. They were aware of how the issue affected their daily lives. The workshop allowed them to discuss with participants from diverse backgrounds and listen to different perspectives. A participant appreciated that another member of her group shared his views from an IT expert perspective, which were significantly different from her user or consumer perspective.

Most participants enjoyed their deliberation workshop experience. They believed that data literacy was not a common topic that they would discuss with friends or relatives. The deliberation workshop offered a platform for people concerned with data literacy to exchange ideas and conduct in-depth discussions about specific issues with each other. Some participants believed that the discussion had pushed them out of their comfort zone, enabling them to acquire new knowledge. A participant further explained how his experience in the workshop "consolidated" his existing data practices. Through confirming his habits with other participants, we "verified" that what he was doing was right and appropriate. He found the support he gained from others was valuable, as

he knew he was not the only one doing so.

While workshop participants are asked to share impressive views of others that have surprised them or are notably distinct from their own, several participants shared about others' "extreme" practices to opt out of certain online behaviors due to potential risks. For example, a participant shared about a member's experience of stopping using credit cards after being hacked; some participants were against online shopping because they were afraid of data abuse. Participants found the choice of “exit” difficult and non-practical, especially when they needed the online services provided.

Participants instead would act proactively to negotiate, manage, and control the potential online risks. For example, one person regularly checked his credit card transaction record to ensure no misuse; another used biometric passwords to reduce the risk of password breaches. It was not uncommon to find participants sharing “successful experiences” to counter online risk, like a participant successfully reduced junk messages on instant messenger by adjusting the app’s privacy settings. Another participant presented how she offered fake information during account registration to protect her personal data.

Participants were asked to pinpoint moments when they felt they were at risk. Receiving calls from strangers (or unknown numbers) was a shared experience among participants. Most would simply ignore those calls or use software to filter them, believing they were mainly advertisements or scams. However, the experience frustrated them as they realized that their phone number (as a

piece of private information) was leaked to an unknown third party.

Surveillance was also an issue mentioned by different participants. Several participants believed their devices would record everything they said due to the coincidence of encountering online advertisements that matched their offline conversations with others. The experience acted as a foundation for their skepticism towards tech companies. Several participants discussed their resignations from social media. They no longer posted photos with their appearances and set up fake accounts that could not chase their real identities. Self-censorship was common among participants, who believed that they should be very careful when expressing online opinions. Some explicitly mentioned that the launch of the national security law was the primary reason.

Participants expressed a sense of fear and anxiety while discussing relevant experiences. The situation applied to both the deliberation workshop and the follow-up focus groups. We believe the workshop and focus group offered a safe, open, and free space for participants to express their opinions and feelings. The expression and communication of negative emotions, like fear, anxiety, anger, and powerlessness, would spread and affect others. For example, a few participants mentioned realizing they did too little to protect their data after joining the workshop. It was the sense of insecurity that sparked the reflections.

While the sharing of negative experiences was prevalent among focus groups, we are not surprised that participants of the deliberation workshop presented lower self-confidence and more disempowerment. The finding of our survey experiment implies that, in a social context where the

room for political participation is significantly curtailed, deliberative communications will not serve to enhance people's sense of empowerment.

Information exposure

Most focus group participants found the information shared in the WhatsApp group basic yet diverse. According to the participants, the divergence of topics allowed them to understand the concept of data literacy from various perspectives, and the use of real-life examples helped them to associate the concepts with their everyday experience. Several participants mentioned that they would search for further information when they found the topics discussed interesting.

For participants who self-identified as aware of data literacy, the content of the WhatsApp posts was deemed too simple. They were familiar with the concepts and had long used the practices listed as tips to protect their data security. They thus found the impact of the posts to be minimal. The above comments demonstrated the drawbacks of one-way information exposure. As participants are prevented from discussing with other participants or the research team, the learning effect is limited by the posts' content. Furthermore, participants' views implied a diversified demand for data literacy information at various levels. Future education programs should avoid the "one-size-fits-all" approach.

Participants found topics and information that could be associated with their daily lives most impressive. For example, a participant mentioned that the post about data collection of CCTV

reminded him of his school's policy to post warning stickers about the installation of CCTV. The association of data literacy information with real-life incidents allowed him to identify and reflect the reason behind the policy.

Several participants expressed that consistent and interactive information exposure can increase people's awareness of data literacy. A participant's company's IT department would occasionally send "fake" scam emails to their employees. When someone presses the scam email link, educational material that warns the person not to access scam emails will appear. The participants found such practice interesting and useful to increase people's awareness of online risks. Future educational programs can take this example as a reference to come up with more interactive learning formats.

To conclude, participants from both intervention groups expressed improved awareness of data literacy after the intervention. And some even stressed that the awareness has led to specific behavioral changes. While we believe that the communication forms of the two interventions are major factors leading to distinct results, we should not ignore the impact of the socio-political context.

Conclusion

To conclude, the research team accomplished the research objectives according to the proposed research plan. First, the research team formulated an online survey that suited the Hong Kong context and utilized the survey to assess Hong Kong Internet users' level of data literacy in an impartial and rigorous manner. The survey results clearly illustrated how different factors influence varying levels of data literacy. To be "data literate", citizens need not only hard skills to use digital technologies, but also awareness and critical reflection of data practices and ability to implement data knowledge in digital attitudes and behaviors. Second, the research team proposed two means, information exposure and deliberation workshop, to improve data literacy. The survey experiment proved that both means were effective in promoting internet users' data literacy. To further investigate the underlying mechanisms leading to enhanced data literacy, the research team conducted ten focus group interviews with the survey experiment participants. The qualitative findings of the focus group did not only supplement the experiment findings, but also explained how contextual factors contribute to different outcomes of the two interventions. The research team believed that the fruitful findings and corresponding policy recommendations were useful and essential to inform future policy making and public debates, particularly in the area of digitalization and the development of smart city.

Policy Recommendations

Like many cities in advanced economies, Hong Kong has embraced the "smart city" agenda, which involves using technological solutions to improve the management and efficiency of the urban environment. In *Hong Kong Smart City Blueprint 2.0* (henceforth the *Blueprint*) published in 2020, the SAR Government set its vision to "embrace innovation and technology to build a world-famed Smart Hong Kong characterized by a strong economy and high quality of living." The smart city initiatives in Hong Kong involve various government bureaux and departments. Against this background, data literacy should be a concern for policy-makers. The Education Bureau has been providing various supports for teachers, students and parents in the implementation of information literacy education in schools.

Education

While large institutions make use of individuals' data to drive decision-making, small organizations and civic hackers are tapping into the availability of open and public data to navigate the data-rich society to interface with data directly for social innovation. Notwithstanding, the educational and informational resources which enable individuals to understand this data remain scarce, making the public at large lack the skills and knowledge to take advantage of this opportunity, especially the underprivileged groups, such as low-income, elders, disabled persons, and ethnic minorities. It becomes essential to develop data literacy and close the literacy gap by providing access to training and equipping the public to think critically and ethically about data.

We recommend that schools incorporate data literacy in the curriculum from the early stage of education in a cross-module discipline. Like the highly emphasized and heavily invested STEM education in Hong Kong curricula, young people's development of data literacy can be invested through continuous exposure to relevant knowledge in applied settings. "Information Literacy for Hong Kong Students" Learning Framework was first introduced in 2018. The EDB puts the emphasis on developing students' media and information literacy (MIL) and develop and apply generic skills such as critical thinking, mathematical skills and IT skills in an integrative manner. However, data literacy (as conceptualized in this report) is a much broader concept than MIL, and it includes an element of proactive social engagement in the digital landscape. Instead of passively being instructed with data-related skills within the context of an independent subject, it can be adopted with a student-centered pedagogy, such as project learning and visiting activities outside the classroom. Students will have the opportunities to focus on different facets of data literacy in depth, including mathematics, science, social science and humanities courses. They will be able to learn integrating data analysis into their daily life and to illuminate real world phenomena through data.

Meanwhile, we recognize the potential challenges inherent in curriculum change. Teachers need extensive professional training to develop the requisite skills as it transcends boundaries of school subjects. Classroom technology also requires updates and should be supported by adequate provision of hardware to students. This may especially be an issue for schools with less resources and such capacity for teacher training.

Governmental Support

In the experiment above, we observed a strong demand among Hong Kong online users for trustworthy data literacy information. The effectiveness of information exposure in improving data literacy is proven to be riding on the credibility of the sources of information. We recommend that the government further collaborate with educational institutions to launch programs in different forms for promoting data literacy. As mentioned by numbers of participants in the deliberation workshop, courses or interest classes organised by NGO or community centres would help the less privileged groups become acquainted with the latest online trends and technical knowledge when browsing the net. They also suggested the provision of introductory educational materials is not particularly effective in changing the attitudes and behaviors of the recipients who are with more background knowledge and identifying themselves as data literate. A one-size-fits-all class therefore may not account for the users from varying backgrounds and skills. We recommend that future educational programs be offered in a variety designated for participants with different levels of data literacy. To foster usability, more specific and discipline-focused programs can be launched for advanced data users. Fundamental programs with targeted help can be provided to less literate participants to bridge the gap of skill level.

We therefore call on the Government for financial support to unlock the expertise of civil society in order to support people to address the backfire of the growth of technology-dependent lifestyles. Coordinated action between charities, educational institutions and support groups can help people to seek redress and encourage improved understanding of data management.

Regulation

While more and more people are entrusting their data with cloud services, data breaches are often their concerns. Participants in deliberation workshops suggested that imposing more regulations on data (controlling or processing) organizations can hold them more accountable. Drafted and passed by the European Union (EU), the General Data Protection Regulation (GDPR) put into effect on May 25, 2018. GDPR is directly binding and applicable to enhance users' control and rights over their personal data and to simplify the regulatory environment for international business within the EU. Harsh fines are levied against those who violate the privacy and security compliance. Numbers of workshop participants recognised that it could be the model for the local government to formulate similar laws and practices to regulate organizations that target or collect data related to people.

Building Transparent Data Ethics

If tech companies want to build and maintain trust with their users, they must get their users informed and involved. They should allow their users to tailor their digital experiences according to their own needs and preferences. We recommend all tech companies implement trustworthy, transparent, and user-friendly design that shows their users how their services work with the collected data and give them meaningful control over their digital experiences. The Government should empower the public by providing support to enhance data literacy, creating motivation to

improve online experiences.

The respondents generally had high levels of concern about the potential harms that technology can cause. But their experience of reporting problems was mostly pitiful, and many believed that harmful experiences are just part and parcel of life online. We recommend tech companies, websites and social media platforms create accessible and straightforward ways for people to report concerns and provide clear information about the corresponding actions they would take. Legislation is a promising means to regulate how companies within scope handle complaints. If they are done robustly, we believe it can directly and tangibly benefit the public and build confidence that regulation is proving effective.

Details of the Public Dissemination Held

The research team proposed to hold public dissemination from three perspectives.

Academic Conference

We presented the preliminary findings of this research at「第十屆 香港文化與社會研討會」. It is a Hong Kong based academic conference held on 5th November 2022, focusing on Hong Kong society and its culture. In particular, the research team explained how the complex Hong Kong context affected the research findings, and the research's implication to future Hong Kong's digitalization. The research team is open to join more academic conferences in the coming future.

Publication

On top of presenting research findings at academic conferences, the research team also planned to publish research articles in peer-review journals. Given the relatively short research period, the team will focus more on publication after the submission of the research report and is going to submit at least one academic article within 2023.

Public Dissemination

After the completion of the project, the research team will hold a press conference to present and disseminate the findings of this research. We hope to invite policymaker, such as the Innovation, Technology and Industry Bureau; and representatives from relevant professional organizations, like the Hong Kong Information Technology Joint Council (HKITJC), the Hong Kong Association for Computer Education (HKACE), the Internet Professional Association (iProA) and the Internet Society Hong Kong (ISOC HK). Media, academics, teachers, students and the general public are also welcomed to join.

References

Ackerman, Bruce, and James S. Fishkin. 2002 “Deliberation Day.*” The Journal of Political Philosophy* 10, no. 2: 129–52.

Bellamy, Richard. 2008. *Citizenship: A Very Short Introduction.* New York: Oxford University Press.

Carmi, Elinor, Simeon J. Yates, Eleanor Lockley, and Alicja Pawluczuk. 2020. "Data citizenship: Rethinking data literacy in the age of disinformation, misinformation, and malinformation." *Internet Policy Review* 9, no. 2: 1-22.

Chakravorti, Bhaskar, Ravi S Chaturvedi, Christina Filipovic, and Griffin Brewer. 2020. “Digital in the Time of Covid: Trust in the Digital Economy and Its Evolution Across 90 Economies as the Planet Paused for a Pandemic.” The Fletcher School Tufts University. https://sites.tufts.edu/digitalplanet/files/2020/12/digital-intelligence-index.pdf

Chang, Linchiat, and Jon A. Krosnick. 2009. "National surveys via RDD telephone interviewing versus the Internet: Comparing sample representativeness and response quality." *Public Opinion Quarterly* 73, no. 4: 641-678.

Census and Statistics Department. 2020. “Thematic Household Survey Report No.69.” Accessed March 3 2021. https://www.statistics.gov.hk/pub/B11302692020XXXXB0100.pdf

Dagger, Richard. 1997. *Civic Virtues: Rights, Citizenship, and Republican Liberalism.* New York: Oxford University Press.

Diebold, Francis.2012. "*On the Origin(s) and Development of the Term 'Big Data'*". PIER Working Paper No. 12-037. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2152421.

D'Ignazio, Catherine, and Rahul Bhargava. 2015. "Approaches to Building Big Data Literacy". Bloomberg Data for Good Exchange Conference, New York. http://rahul-beta.connectionlab.org/wp-content/uploads/2011/11/Edu_DIgnazio_52.pdf

D'Ignazio, Catherine, and Rahul Bhargava. 2018. "Cultivating a Data Mindset in the Arts and Humanities". Public, 4(2). https://public.imaginingamerica.org/blog/article/cultivating-a-data-mindset-in-the-arts-and-humanities/

Ebbinghaus, E. E. (1964). Memory (Trans. H. A. Ruger and C. E. Bussenius). New York: Dover (Original work published 1885).

Eshet, Yoram. 2004. "Digital literacy: A conceptual framework for survival skills in the digital era." *Journal of Educational Multimedia and Hypermedia* 13, no. 1: 93-106.

Fishkin, James S., and Robert C. Luskin.2005. "Experimenting with a democratic ideal: Deliberative polling and public opinion." *Acta politica* 40, no. 3: 284-98.

Fishkin, James S., and Robert C. Luskin. 1995. "Bringing deliberation to the democratic dialogue." In *The Poll with a Human Face: The National Issues Convention Experiment in Political Communication*, edited by Maxwell Mccombs, and Amy Reynolds, 3-38. Mahwah, London: Routledge.

García Ochoa, Gabriel, Sarah McDonald, and Nicholas Monk. 2016. "Embedding cultural literacy in higher education: a new approach." *Intercultural Education* 27, no. 6: 546-59.

Glister, Paul. 1997. *Digital Literacy*. New York: Wiley Computer Pub.

Guess, Andrew M., Michael Lerner, Benjamin Lyons, Jacob M. Montgomery, Brendan Nyhan, Jason Reifler, and Neelanjan Sircar. 2020. "A digital media literacy intervention increases discernment between mainstream and false news in the United States and India." Proceedings of the National Academy of Sciences 117, no. 27: 15536-45.

Grönlund, Kimmo, Maija Setälä, and Kaisa Herne.2010. "Deliberation and civic virtue: Lessons from a citizen deliberation experiment." European Political Science Review: EPSR 2, no. 1: 95.

Hargittai, Eszter, and Yuli Patrick Hsieh. 2012. "Succinct survey measures of web-use skills." Social Science Computer Review 30, no. 1: 95-107.

Hashem, Ibrahim Abaker Targio, Victor Chang, Nor Badrul Anuar, Kayode Adewole, Ibrar Yaqoob, Abdullah Gani, Ejaz Ahmed, and Haruna Chiroma. 2016. "The role of big data in smart city." International Journal of Information Management 36, no. 5: 748-58.

Hintz, Arne Lina Dencik and Karin Wahl-Jorgensen. 2018. Digital Citizenship in a Datafied Society. Cambridge: Polity Press.

Holbrook, Jack, and Miia Rannikmae. 2009. "The meaning of scientific literacy." *International Journal of Environmental and Science Education* 4, no. 3: 275-88.

Hong Kong Free Press. “Timeline: 58 Hong Kong Civil Society Groups Disband Following the Onset of the Security Law.” Hong Kong Free Press HKFP, 30 June 2022, https://hongkongfp.com/2022/06/30/explainer-over-50-groups-gone-in-11-months-how-hong-kongs-pro-democracy-forces-crumbled/.

Innovation and Technology Bureau. 2020. “Hong Kong Smart City Blueprint 2.0.” Accessed March 3 2021. https://www.smartcity.gov.hk/modules/custom/custom_global_js_css/assets/files/HKSmartCityBlueprint(ENG)v2.pdf

Kenney, Courtney and Claudia Deane. 2019. “What our transition to online polling means for decades of phone survey trends.” Pew Research Center, February 27, 2019. https://www.pewresearch.org/fact-tank/2019/02/27/what-our-transition-to-online-polling-means-for-decades-of-phone-survey-trends/

KPMG. 2020. “AI for Everyone.” Accessed February 5, 2021.https://services.google.com/fh/files/misc/google_smarter_digital_city_4_ai_for_everyone_whitepaper.pdf

Laney, David. 2001. *3D data management: controlling data volume, velocity and variety.* META Group Research Note.

Li, Cho-kiu. 2021. "Raw fear in Hong Kong." *HAU: Journal of Ethnographic Theory* 11, no.3: 1045-1059.

Lyon, David. 2014. "Surveillance, Snowden, and big data: Capacities, consequences, critique." *Big Data & Society* 1, no. 2:1-13.

Mayer-Schönberger, Viktor and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston: Houghton Mifflin Harcourt.

Mittelstadt, Brent. 2017."From individual to group privacy in big data analytics." *Philosophy & Technology* 30, no. 4: 475-94.

García Ochoa, Gabriel, Sarah McDonald, and Nicholas Monk. 2016."Embedding cultural literacy in higher education: a new approach." *Intercultural Education* 27, no. 6: 546-59.

Ohme, Jakob. 2019. "Updating citizenship? The effects of digital media use on citizenship understanding and political participation." *Information, Communication & Society* 22, no. 13: 1903-28.

Pettit, Philip. 1997. *Republicanism: A Theory of Freedom and Government*. Oxford: Oxford University Press.

Porter, Ethan, and Thomas J. Wood. 2021. "The global effectiveness of fact-checking: Evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom." *Proceedings of the National Academy of Sciences,* 118, 37: e2104235118.

Remund, David L. 2010. "Financial literacy explicated: The case for a clearer definition in an increasingly complex economy." *Journal of Consumer Affairs* 44, no. 2: 276-95.

Tilly, Charles.1997. "A primer on citizenship." *Theory and Society* 26, no. 4: 599-602.

Tully, Melissa, Emily K. Vraga, and Leticia Bode. 2020. "Designing and testing news literacy messages for social media." *Mass Communication and Society* 23, no. 1: 22-46.

Turow, Joseph. 2011. *The Daily You: How the New Advertising Industry is Defining Your Identity and Your Worth*. New Haven, CT: Yale University Press.

Schneier, Bruce. 2015. *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World.* New York: W. W. Norton & Company.

Vodafone Institute for Society and Communications. 2016. "Big Data: A European Survey on the Opportunities and Risks of Data Analysis." https://www.vodafone-institut.de/wp-content/uploads/2016/01/VodafoneInstitute-Survey-BigData-en.pdf

Vraga, Emily K, Tully, Melissa, and Leticia Bode. 2020. "Empowering Users to Respond to Misinformation About Covid-19." *Media and Communication (Lisboa)* 8, no. 2 (2020): 475–79.

Yates, Simon, Elinor Camri, Alicja Pawluczk, Eleanor Lockley, Bridgette Wessels, and Justin Gangneux. "Me and My Big Data Report 2020: Understanding Citizens' Data Literacies: Thinking, Doing & Participating with Our Data." Accessed March 3 2021. https://www.liverpool.ac.uk/media/livacuk/research/heroimages/Me-and-My-Big-Data-Report-1.pdf

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power.* New York: Profile Books.

Appendix

- Table
- Question Book
- WhatsApp Short Post
- Workshop
    - Moderator guide
    - Pre-workshop material
    - Discussion Material
- Focus Group
    - Question Guide

List of Tables (Survey)

*"Refuse to answer" are voided as missing data and not including valid results as shown below. Sample base of a total 3261.

Part 1. Demographics

Table 1a. Respondents' Gender

Gender	Male	46.5%
	Female	52.4%
	Other	1.1%
Total		100% (3,220)

Table 1b. Respondents' Age

Age	18 - 29	13.70%
	30 - 39	15.7%
	40 - 49	17.8%
	50 - 59	19.8%
	60 - 69	27.4%
	70 or above	5.6%
Total		100% (3,242)

Table 1c. Respondents' Education attainment

Educational Level	Primary or below	1.7%
	Lower secondary (S1-S3)	10.2%
	Upper secondary (S4-S7 / DSE / YiJin)	53.1%
	Tertiary: Non-degree	10.0%
	Tertiary: Bachelor degree	19.4%
	Tertiary: Postgraduate or Above	5.6%
Total		100% (3,237)

Table 1d. Respondents' occupation status

Occupation Status		

	Employer	3.2%
	Permanent Worker (No contract renewal required)	32.8%
	Contract worker (1 year or more)	12.2%
	Short-term contract worker (Less than 1 year)	1.8%
	Freelancer / Part-Time work	3.7%
	Retired / Pensioner	22.9%
	Housekeeper	10.6%
	Full-time student	3.9%
	Unemployed	4.2%
	Others	4.6%
		100% (3,244)

Table 1e. Respondents' income

Monthly household income		
	No income	5.4%
	Below $10,000	8.3%
	$10,001 to $20,000	22.6%
	$20,001 to $30,000	18.8%
	$30,001 to $50,000	19.1%
	$50,001 to $70,000	11.2%
	$70,001 to $100,000	6.0%
	8 Over $100,000	2.1%
	Don't Know	6.3%
		100% (3,235)

Table 1f. Respondents' political inclination

		Total
Political inclination		

	Localist	9.3%
	Pro-democracy camp	21.5%
	Centrist	18.5%
	Pro-establishment camp	15.4%
	Other	0%
	No political inclination / politically neutral / do not belong to any camp	30.2%
	Don't know / hard to say	5%
		100% (2,981)

Part 2 Respondents' digital activities

Table 2a. Respondents' daily usage of the Internet

		Total
In the last month, how much time on average do you spend online every day?		
	Less than 1 hour	0.6%
	1 to less than 2 hour(s)	4.4%
	2 to less than 4 hours	27.7%
	4 to less than 6 hours	20.4%
	6 to less than 8 hours	18.7%
	8 hours or more	28.0%
	Don't know/ hard to say	0.3%
Total		100% (3,261)

Table 2b. Type of online activities respondents engaged in

		Percentage of cases
In the last month, have you done the following online activities?		
	Use instant messenger (like WhatsApp, FB	91.8%

	Messenger, Telegram, Signal, WeChat)	
	Use search engines to search for information (like Google, Yahoo!, Bing)	90.5%
	Use social media (like Facebook, Instagram, Twitter, LIHKG)	85.8%
	Watch short videos online (like YouTube, Instagram, Facebook)	78.6%
	Visit news websites or websites about current affairs and politics (like MingPao, NowTV, Initium, HK01)	70.6%
	Use online payment software (like PayMe, WeChat Pay, Alipay)	63.0%
	Buy and sell online (include goods and services)	62.1%
	Watch TV show or movie via online streaming (like Netflix, Disney+, Apple TV)	43.0%
	Play games online	33.9%
	Listen to music via online streaming (like Spotify, KKbox, Joox)	29.4%
	Others	2.1%
Total		100% (3,261)

Table 2c. Social media respondents registered

		Percentage of cases
Have you ever registered for a personal account with the following social media sites or apps?		
	Facebook	91.4%
	YouTube	76.4%
	Instagram	65.1%
	Twitter	27.4%
	Online forums (like LIHKG, Discuss.com.hk, Baby Kingdom)	21.2%

	LinkedIn	19.6%
	Pinterest	15.3%
	Snapchat	12.3%
	TikTok	8.3%
	Reddit	6.5%
	Tumblr	4.7%
	Twitch	4.1%
	Other	3.8%
	I do not use any social media websites or apps	1.1%
	Don’t know / hard to say	0.9%
Total		100% (3,232)

Table 2d. Instant messengers respondents used

		Percentage of cases
Have you ever used the following instant messaging websites or apps?		
	WhatsApp	95.9%
	Facebook Messenger	66.0%
	WeChat	62.4%
	Zoom	55.4%
	Signal	55.0%
	Telegram	40.2%
	Line	31.8%
	Instagram Direct Message	24.8%
	Skype	22.2%
	Discord	12.1%
	Google Hangouts	10.1%
	Viber	4.1%
	Other	1.4%
	I do not use any instant messaging	0.5%

	websites or apps	
	Don’t know / hard to say	0%
Total		100% (3,257)

Part 3. Data Technology and Confidence

Table 3a. Respondents’ acquired confidence on technological tools

How much do you agree with the following statements?		Using a firewall or other related antivirus software keeps my information safe	Using Virtual Private Network (VPN) keeps my information safe	Changing privacy settings keeps my information safe
	Strongly agree	25.0%	14.1%	25.5%
	Somewhat agree	38.6%	35.8%	45.5%
	Neither agree nor disagree	27.0%	26.4%	20.3%
	Somewhat disagree	1.0%	3.1%	2.7%
	Strongly disagree	1.4%	1.0%	0.1%
	Do not know / Hard to say	7.0%	19.6%	5.9%
Total		100% (3,261)	100% (3,261)	100% (3,261)
	Mean	3.91	3.73	3.99

Table 3b. Respondents’ acquired confidence on technological tools(Cont’)

How much do you agree with the following statements?		Enabling multi-factor authentication keeps my information safe	Using websites or apps that are more secure or sensitive to privacy (like Signal, Telegram and Duckduckgo) keeps my information safe

	Strongly agree	26.2%	24.1%
	Somewhat agree	42.2%	31.5%
	Neither agree nor disagree	16.8%	30.0%
	Somewhat disagree	3.1%	3.7%
	Strongly disagree	0.23%	0.6%
	Do not know / Hard to say	11.6%	10.0%
Total		100% (3,260)	100% (3,239)
	Mean	4.03	3.83

Table 3c. Respondents' self confidence on securing personal privacy and information online

How much do you agree with the following statements?		I am confident in securing my personal privacy online	I am confident in keeping my information safe online	I am confident in managing who can use my personal information and data online
	Strongly agree	4.8%	3.6%	4.6%
	Somewhat agree	16.3%	21.1%	16.6%
	Neither agree nor disagree	53.5%	49.5%	45.8%
	Somewhat disagree	12.3%	10.8%	16.2%
	Strongly disagree	5.0%	6.2%	8.6%
	Do not know / Hard to say	8.1%	8.8%	8.3%

Total		100% (3,261)	100% (3,259)	100% (3,253)
	Mean	3.04	3.05	2.92

Table 3d. Respondents' self confidence on securing personal privacy and information online (Con't 1)

How much do you agree with the following statements?		I am confident in managing who can see the information I share online	I am confident in determining whether online information is true or not	I am confident in understanding the terms and conditions offered by websites or apps
	Strongly agree	5.2%	4.3%	3.7%
	Somewhat agree	20.3%	34.8%	18.4%
	Neither agree nor disagree	47.1%	41.5%	43.7%
	Somewhat disagree	12.2%	9.7%	18.1%
	Strongly disagree	7.0%	2.7%	8.2%
	Do not know / Hard to say	8.2%	7.0%	7.9%
Total		100% (3,261)	100% (3,261)	100% (3,261)
	Mean	3.05	3.3	2.91

Table 3e. Respondents' self confidence on securing personal privacy and information online

(Con't 2)

How much do you agree with the following statements?		I am confident in understanding data analysis which I encounter online	I am confident in using online public information and data to solve my problems

	Strongly agree	2.1%	2.3%
	Somewhat agree	24.9%	28.1%
	Neither agree nor disagree	43.7%	43.4%
	Somewhat disagree	14.3%	12.2%
	Strongly disagree	5.3%	3.8%
	Do not know / Hard to say	9.7%	10.1%
Total		100% (3,234)	100% (3,258)
	Mean	3.04	3.14

Part 4. Data Technology and Attitude

Table 4a. Respondents' views on impact of data technology on personal life and society

Overall speaking, do you think data technology makes the following better or worse?		Personal life	Society
	Much better	16.4%	12.5%
	Somewhat better	61.9%	47.0%
	No change	12.7%	10.9%
	Somewhat worse	4.1%	13.9%
	Much worse	0.8%	5.3%

	Don't know / Hard to say	4.1%	10.4%
Total		100% (3,225)	100% (3,169)
	Mean	3.93	3.53

Table 4b(i). Respondents' perceived disempowerment while using data technology

How much do you agree with the following statements?		I have no control over how tech companies collect and use my personal information and data	It is meaningless to read terms and conditions because I have to accept them in order to use the website or app	Changing privacy settings does not help to protect personal privacy
	Strongly agree	36.3%	48.0%	19.1%
	Somewhat agree	46.3%	33.3%	32.9%
	Neither agree nor disagree	13.0%	14.6%	34.3%
	Somewhat disagree	3.2%	3.6%	9.6%
	Strongly disagree	0.7%	0.1%	0.3%
	Do not know / Hard to say	0.6%	0.5%	3.8%
Total		100% (3,261)	100% (3,261)	100% (3,260)
	Mean	4.15	4.26	3.63

Table 4b(ii). Respondents' perceived disempowerment while using data technology (Con't 1)

How much do you agree with the following statements?		Changing privacy settings takes too much time and effort	Generally, even if I report inappropriate or harmful content to a platform, the platform will not take my report	Tech companies keep their users' best interests in mind when they design and operate their websites

			seriously	and apps
	Strongly agree	23.2%	25.9%	12.5%
	Somewhat agree	36.4%	36.0%	17.5%
	Neither agree nor disagree	24.7%	26.5%	27.8%
	Somewhat disagree	10.5%	3.3%	26.4%
	Strongly disagree	2.2%	1.2%	8.4%
	Do not know / Hard to say	3.0%	7.2%	7.4%
Total		100% (3,258)	100% (3,260)	100% (3,237)
	Mean	3.70	3.88	2.99

Table 4b(iii). Respondents' perceived disempowerment while using data technology (Con't 2)

How much do you agree with the following statements?		Even if I think that a tech company's website or app is creating negative impacts on society, I will have no say over the issue.	Being online is not safe and there will always be someone trying to deceive or harm me in some way
	Strongly agree	23.7%	12.2%
	Somewhat agree	42.0%	29.3%
	Neither agree nor disagree	22.2%	40.2%
	Somewhat disagree	8.3%	10.6%
	Strongly disagree	1.5%	3.6%

	Do not know / Hard to say	2.3%	4.1%
Total		100% (3,239)	100% (3,258)
	Mean	3.80	3.38

Table 4b(ix). Respondents' views on data collection

How much do you agree with the following statements?		When I use a website or app, I usually accept the terms and conditions without reading or fully understanding them	I welcome the government to collect my personal information and data to improve their services	I welcome tech companies to collect my personal information and data to improve their services
	Strongly agree	28.2%	11.4%	2.1%
	Somewhat agree	46.5%	17.9%	9.7%
	Neither agree nor disagree	19.0%	15.6%	31.7%
	Somewhat disagree	3.2%	16.9%	25.0%
	Strongly disagree	2.4%	35.5%	29.8%
	Do not know / Hard to say	0.8%	2.8%	1.7%
Total		100% (3,239)	100% (3,261)	100% (3,261)
	Mean	3.96	2.51	2.28

Table 4b(x). Respondents' views on data collection (Cont')

How much do you agree with the		I believe that tech companies are violating my privacy by collecting my	I believe that my personal information and data online is being collected

following statements?		personal information and data online	excessively
	Strongly agree	25.0%	29.8%
	Somewhat agree	34.0%	37.6%
	Neither agree nor disagree	25.5%	22.3%
	Somewhat disagree	9.5%	4.1%
	Strongly disagree	3.6%	2.2%
	Do not know / Hard to say	2.3%	3.9%
Total		100% (3,261)	100% (3,261)
	Mean	3.69	3.92

Table 4b(xi). Respondents' views on data protection

How much do you agree with the following statements?		Internet users' personal data should be protected	The government is protecting Internet users' personal data	Tech companies are protecting Internet users' personal data
	Strongly agree	51.9%	13.8%	9.5%
	Somewhat agree	38.2%	16.2%	12.3%
	Neither agree nor disagree	8.1%	25.9%	38.6%
	Somewhat disagree	0.7%	16.9%	20.5%
	Strongly disagree	0.8%	17.6%	11.4%

	Do not know / Hard to say	0.3%	9.7%	7.7%
Total		100% (3,261)	100% (3,261)	100% (3,261)
	Mean	4.40	2.91	2.87

Table 4c(i). Respondents' perceived rights online

How much do you agree with the following statements?		Citizens should have the right to access the Internet anonymously	Citizens should have the right to anonymously express their opinions online
	Strongly agree	31.2%	31.1%
	Somewhat agree	28.4%	27.6%
	Neither agree nor disagree	24.0%	24.6%
	Somewhat disagree	6.4%	7.3%
	Strongly disagree	7.6%	8.0%
	Do not know / Hard to say	2.4%	1.4%
Total		100% (3,261)	100% (3,261)
	Mean	3.71	3.67

Table 4c(ii). Respondents' perceived rights online (Cont')

How much do you agree with the following statements?		Citizens should have the right to speak freely online, even if the speech is harmful or controversial	Citizens should have the right to speak freely online without being subject to criminal charges

	Strongly agree	27.8%	30.3%
	Somewhat agree	21.2%	12.7%
	Neither agree nor disagree	23.6%	24.5%
	Somewhat disagree	12.5%	10.2%
	Strongly disagree	11.3%	18.7%
	Do not know / Hard to say	3.6%	3.6%
Total		100% (3,261)	100% (3,261)
	Mean	3.43	3.27

Table 4d(i). Respondents' willingness to provide personal information

How much are you willing to provide your personal information to websites or apps for the following reasons/purposes?		To create personal accounts (like social media accounts)	To get free services (like public Wi-Fi network)	To get personalized services (like weather service, video recommendations)
	Strongly willing	3.4%	2.6%	4.8%
	Somewhat willing	21%	11.8%	15.6%
	Half-half	38.4%	30.5%	32.4%
	Somewhat unwilling	23.3%	25.8%	27.7%

	Strongly unwilling	12%	27.8%	18.2%
	Do not know / Hard to say	1.9%	1.4%	1.3%
Total		100% (3,256)	100% (3,261)	100% (3,261)
	Mean	2.8	2.35	2.61

Table 4d(ii). Respondents' willingness to provide personal information (Cont')

How much are you willing to provide your personal information to websites or apps for the following reasons/purposes?		To get advertisements or information that are more relevant to me (like discount offers)	To get free services (like public Wi-Fi network)
	Strongly willing	4.0%	3.1%
	Somewhat willing	12.6%	12.1%
	Half-half	26.1%	43.2%
	Somewhat unwilling	29.2%	21.2%
	Strongly unwilling	26.7%	17.0%
	Do not know / Hard to say	1.3%	3.4%
Total		100% (3,256)	100% (3,261)
	Mean	2.37	2.62

Table 4d(iii). Respondents' willingness to provide personal information under certain conditions

How much are you willing or unwilling to provide your personal information to websites or apps under the following conditions?		If the website or app clearly state how they will use your data	If the website or app clearly state how they will use your data
	Strongly willing	6.1%	10.9%
	Somewhat willing	26.4%	36.3%
	Half-half	38.3%	32.9%
	Somewhat unwilling	17.7%	12.2%
	Strongly unwilling	9.2%	5.9%
	Do not know / Hard to say	2.2%	1.9%
Total		100% (3,261)	100% (3,261)
	Mean	3.02	3.35

Table 4d(ix). Respondents' perceived types of personal data collected

Do you think the following data of yours would be collected by website or app providers?		Traces of my internet activities (like my web history, search history and shopping history)	Information about me on social media (like my posts, photos and videos)	Information that I store on my electronic devices (like my mobile phone and computer)
	Definitely	52.0%	40.6%	24.5%
	Likely	37.7%	39.7%	38.4%
	About half the time	5.7%	13.9%	18.9%

	Unlikely	2.3%	2.8%	11.6%
	Definitely not	1.3%	0.8%	2.5%
	Do not know / Hard to say	0.9%	2.3%	4.0%
Total		100% (3,259)	100% (3,258)	100% (3,240)
	Mean	4.38	4.19	3.74

Table 4d(x). Respondents' perceived types of personal data collected (Cont')

Do you think the following data of yours would be collected by website or app providers?		My biometric data that have been recorded (like my fingerprints, appearance, voice)	My habits in real life (like my hobbies and places I frequent go)
	Definitely	23.2%	38.8%
	Likely	31.3%	36.7%
	About half the time	22.1%	15.6%
	Unlikely	12.0%	6.1%
	Definitely not	3.9%	1.6%
	Do not know / Hard to say	7.4%	1.1%
Total		100% (3,258)	100% (3,253)
	Mean	3.62	4.06

Table 4e(i). Respondents' perceived data truthfulness

How much do you agree with the following statements?		The information I find online is always true	I don't care if the information I find online is true or not
	Strongly agree	0.1%	0.5%
	Somewhat agree	4.4%	7.9%
	Neither agree nor disagree	43.6%	23.4%
	Somewhat disagree	25.6%	34.6%
	Strongly disagree	21.8%	33.4%
	Do not know / Hard to say	4.5%	0.3%
Total		100% (3,259)	100% (3,259)
	Mean	2.32	2.07

Table 4e(ii). Respondents' perceived data truthfulness (Cont')

How much do you agree with the following statements?		When I search the internet, I generally look for information that I find interesting	When I search the internet, I generally look for information that agrees with my view
	Strongly agree	18.5%	13.2%
	Somewhat agree	59.0%	46.7%
	Neither agree nor disagree	18.4%	30.0%
	Somewhat disagree	3.6%	7.4%

	Strongly disagree	0.3%	2.5%
	Do not know / Hard to say	0.1%	0.2%
Total		100% (3,259)	100% (3,259)
	Mean	3.92	3.61

Part 5 Respondents' Data Literacy

Table 5a. Data Thinking (*Figures of correct options for each question are highlighted in **bold**.)

	Option A	Option B	Option C	Option D	Don't know	Total
(Q16) Some of the following charts are misleading, please select the chart that you think is the most accurate (without misleading components)	12.9%	5.8%	10.4%	**50.8%**	20.1%	100% (3,174)
(Q17) When you get online, sometimes you will see the message "Do you accept to use cookies?". Which of the following descriptions about "cookies" is correct?	3.9%	**53.9%**	15.9%	1.0%	25.4%	100% (3,239)
(Q18) When you use Wi-Fi services offered in public places (such as airports or coffee shops), which of the following approaches best keep your data safe?	**61.0%**	1.3%	4.3%	3.0%	30.5%	100% (3,217)
(Q19) When you search for "dress" on Google, which of the following is the most accurate description of your first search result like below?	9.6%	**61.8%**	11.9%	1.8%	14.9%	100% (3,261)
(Q20) If the URL of a website starts with https:// (instead of http://), which of the following descriptions is correct?	6.2%	2.7%	5.0%	**31.5%**	54.6%	100% (3,257)
(Q21) When you use private browsing function of your web browser to access a website, which of the following descriptions is correct?	9.4%	**37.5%**	3.4%	4.4%	45.3%	100% (3,260)

(Q22) According to the Personal Data (Privacy) Ordinance, there are six data protection principles. Which of the following descriptions is correct regarding the data protection principles?	**66.3%**	2.8%	3.7%	1.2%	26.0%	100% (3,238)
(Q23) Many websites and apps require users to complete multi-factor authentication before they can log in. Which of the following pictures is an example of multi-factor authentication?	17.2%	4.6%	13.7%	**57.4%**	7.1%	100% (3,252)
Average correct rate	53.8% (4.3 out of 8 questions)					

Table 5b(i). Respondents' creation of data

How much do you agree with the following statements?		I often share information about me on social media	I often share community or public affairs information on social media
	Strongly agree	3.8%	4.6%
	Somewhat agree	10.0%	13.7%
	Neither agree nor disagree	37.6%	40.6%
	Somewhat disagree	26.1%	23.7%
	Strongly disagree	20.7%	15.7%
	Do not know / Hard to say	1.7%	1.7%
Total		100% (3,248)	100% (3,251)
	Mean	2.49	2.67

Table 5b(ii). Respondents' data accessing and assessing

How much do you agree with the following statements?		I often collect information from various sources, even if it is different from my position or point of view	I often collect information from various sources to verify the content's trueness
	Strongly agree	9.6%	13.7%
	Somewhat agree	40.4%	43.5%
	Neither agree nor disagree	36.6%	33.4%
	Somewhat disagree	8.0%	6.1%
	Strongly disagree	2.7%	1.6%
	Do not know / Hard to say	2.7%	1.7%
Total		100% (3,261)	100% (3,254)
	Mean	3.47	3.63

Table 5b(iii). Respondents' data management and data deletion

Would you take the following actions?		Use a firewall or other antivirus software	Use filtering software (to filter out spam, calls or advertisements)	Use a virtual private network (VPN)
	Definitely	15.2%	10.0%	7.7%
	Likely	45.4%	40.8%	13.4%
	About half the time	20.2%	28.1%	31.3%
	Unlikely	13.3%	14.3%	27.7%

	Definitely not	1.3%	1.2%	8.9%
	Do not know / Hard to say	4.6%	5.5%	11.0%
Total		100% (3,252)	100% (3,245)	100% (3,237)
	Mean	3.63	3.47	2.81

Table 5b(ix). Respondents' data management and data deletion (Cont' 1)

Would you take the following actions?		Delete cookies or browsing history in my web browser regularly	Modify the privacy settings on websites or apps when it is needed	Prefer websites or apps that are more secure (like Signal, Telegram, and Duckduckgo)
	Definitely	12.6%	7.9%	15.5%
	Likely	20.7%	22.8%	28.2%
	About half the time	27.7%	34.9%	26.5%
	Unlikely	25.6%	22.0%	19.2%
	Definitely not	4.7%	4.1%	4.0%
	Do not know / Hard to say	8.5%	8.3%	6.6%
Total		100% (3,245)	100% (3,188)	100% (3,245)
	Mean	3.12	3.09	3.34

Table 5b(x). Respondents' data management and data deletion (Cont' 2)

Would you take the following actions?		Use multi-factor authentication feature on websites or apps	Create strong passwords (with letters, numbers and symbols) on electronic devices, websites or apps	Do software updates when newer versions are available
	Definitely	11.1%	22.9%	23.2%
	Likely	35.9%	42.6%	32.4%
	About half the time	32.2%	22.2%	28.0%

	Unlikely	10.1%	9.5%	13.1%
	Definitely not	2.8%	0.5%	0.9%
	Do not know / Hard to say	7.9%	2.3%	2.3%
Total		100% (3,230)	100% (3,245)	100% (3,212)
	Mean	3.12	3.09	3.34

Table 5c(i). Respondents' usage of government services

How often would you take the following actions?		Look up public service information online via government websites (like GovHK)	Use government online services (such as making an appointment to renew my ID or passport).	Participate in government public consultations or express opinions to relevant bodies online
	Always	7.2%	14.0%	2.4%
	Often	19.0%	25.4%	5.0%
	Sometimes	32.0%	31.1%	18.2%
	Rarely	35.8%	25.0%	34.3%
	Never	5.7%	4.2%	39.5%
	Do not know / Hard to say	0.3%	0.3%	0.6%
Total		100% (3,261)	100% (3,261)	100% (3,260)
	Mean	2.86	3.20	1.96

Table 5c(ii). Respondents' online socio-political participation

How often would you take the following actions?		Participate in joint signature campaigns online	Participate in instant message or social media groups that are related to public or community affairs (includes obtaining or offering instant information, taking part in online discussions)
	Always	4.6%	2.5%
	Often	12.3%	6.2%
	Sometimes	41.5%	28.0%
	Rarely	25.9%	36.8%
	Never	14.0%	24.3%
	Do not know / Hard to say	1.8%	2.1%
Total		100% (3,192)	100% (3,257)
	Mean	2.67	2.24

Table 5c(iii). Respondents' activism in protecting data privacy

If you encounter the following situations, would you take the below corresponding actions?		When I find misinformation (like fake news) on social media, I would report the incident to the platform	When I find that the information I received on an instant message or social media group is untrue, I clarify it with others in the group
	Definitely	9.2%	10.2%
	Likely	23.8%	37.3%
	About half the time	29.0%	28.7%

	Unlikely	26.6%	17.4%
	Definitely not	7.2%	3.2%
	Do not know / Hard to say	4.1%	3.3%
Total		100% (3,261)	100% (3,261)
	Mean	3.01	3.35

Table 5c(ix). Respondents' activism in protecting data privacy (Cont')

If you encounter the following situations, would you take the below corresponding actions?		When I find misinformation (like fake news) on social media, I would report the incident to the platform	When I find that the information I received on an instant message or social media group is untrue, I clarify it with others in the group
	Definitely	14.2%	24.5%
	Likely	29.2%	31.6%
	About half the time	28.2%	22.8%
	Unlikely	21.8%	13.6%
	Definitely not	4.7%	4.1%
	Do not know / Hard to say	2.0%	3.3%
Total		100% (3,257)	100% (3,258)
	Mean	3.27	3.61

Table 5c(x). Respondents' ability to help with data literacy

		If someone needs it, would you help him/her keep his/her data safe (such as helping him/her to set strong passwords or change privacy settings)?
	Definitely	19.3%
	Likely	33.7%
	About half the time	24.4%
	Unlikely	7.9%
	Definitely not	3.2%
	Do not know / Hard to say	11.5%
Total		100% (3,261)
	Mean	3.66

List of Tables (Experiment)

Table E1. Compared Data Thinking under interventions (Figures of correct options for each question are highlighted in **bold**. Sig = Significance represented by */**/***)

	Whatsapp		Workshop		Control	
	Pre	Post	Pre	Post	Pre	Post
(Q16) Some of the following charts are misleading, please select the chart that you think is the most accurate (without misleading components)						
Option A	6.8%	8.3%	9.5%	6.0%	6.0%	8.6%
Option B	7.2%	2.9%*	3.5%	3.5%	6.9%	6.4%
Option C	6.3%	3.4%	7.0%	6.0%	6.9%	3.0%
Option D	74.9%	82.0%	70.5%	79.1%*	69.8%	72.5%
Don't know	4.8%	3.4%	9.5%	5.5%	10.3%	9.4%
(Q17) When you get online, sometimes you will see the message "Do you accept to use cookies?". Which of the following descriptions about "cookies" is correct?						
Option A	1.0%	0.5%	1.0%	1.0%	2.5%	1.3%
Option B	84.2%	89.5%	78.3%	88.1%**	72.0%	80.1%*
Option C	6.7%	3.3%	6.9%	4.5%	11.9%	4.2%**
Option D	0.5%	0.5%	1.5%	0.0%	1.3%	1.7%

Don’t know	7.7%	6.2%	12.3%	6.5%*	12.3%	12.7%
(Q18) When you use Wi-Fi services offered in public places (such as airports or coffee shops), which of the following approaches best keep your data safe?						
Option A	88.4%	93.8%	82.8%	92.1%**	82.6%	85.1%
Option B	1.0%	0.0%	1.5%	0.5%	0.0%	0.9%
Option C	1.4%	1.0%	1.5%	2.0%	1.7%	2.1%
Option D	0.5%	1.0%	0.5%	0.0%	0.9%	0.0%
Don’t know	8.7%	4.3%	13.8%	5.4%**	14.9%	11.9%
(Q19) When you search for “dress” on Google, which of the following is the most accurate description of your first search result like below?						
Option A	5.3%	5.3%	6.9%	3.0%	6%	5%
Option B	83.7%	87.1%	81.3%	86.7%	79%	82%
Option C	6.2%	5.3%	6.4%	5.9%	9%	5%
Option D	0.0%	0.0%	0.5%	0.0%	0%	0%
Don’t know	4.8%	2.4%	4.9%	4.4%	6%	8%
(Q20) If the URL of a website starts with https:// (instead of http://), which of the following descriptions is correct?						
Option A	1.0%	2.4%	3.5%	4.4%	5.2%	6.4%
Option B	1.4%	1.0%	1.0%	1.5%	0.9%	0.8%

Option C	5.7%	6.3%	5.0%	3.9%	6.0%	6.4%
Option D	59.8%	72.0%**	61.4%	71.4%*	50.9%	53.4%
Don't know	32.1%	18.4%**	29.2%	18.7%*	37.1%	33.1%
(Q21) When you use private browsing function of your web browser to access a website, which of the following descriptions is correct?						
Option A	6.2%	7.2%	6.9%	5.4%	8.1%	6.0%
Option B	70.3%	77.5%	70.0%	75.7%	59.1%	64.7%
Option C	1.4%	0.5%	1.0%	0.0%	0.9%	0.9%
Option D	5.3%	4.8%	5.4%	5.4%	6.0%	3.8%
Don't know	16.7%	10.0%*	16.7%	13.4%	26.0%	24.7%
(Q22) According to the Personal Data (Privacy) Ordinance, there are six data protection principles. Which of the following descriptions is correct regarding the data protection principles?						
Option A	74.5%	90.9%**	74.9%	85.2%**	66.5%	75.3%*
Option B	1.9%	0.5%	3.0%	1.0%	1.7%	0.0%
Option C	2.9%	2.4%	2.5%	3.0%	4.2%	3.8%
Option D	1.9%	0.5%	0.5%	2.0%	0.4%	0.4%
Don't know	18.8%	5.8%**	19.2%	8.9%**	27.1%	20.4%
(Q23) Many websites and apps require users to complete multi-factor						

authentication before they can log in. Which of the following pictures is an example of multi-factor authentication?						
Option A	16.7%	8.1%**	12.8%	7.9%	16.1%	12.3%
Option B	7.2%	6.2%	6.4%	5.9%	6.8%	6.8%
Option C	3.3%	3.3%	3.9%	3.0%	6.8%	6.4%
Option D	68.9%	79.9%**	74.4%	81.8%	66.9%	71.6%
Don't know	3.8%	2.4%	2.5%	1.5%	3.4%	3.0%
Average correct rate (Out of 8 questions)	76.6%	84.4%	73.5%	83.4%	69.5%	64.2%

Table E2. Compared Data Doing under interventions

	Whatsapp		Workshop		Control	
	Pre	Post	Pre	Post	Pre	Post
Would you take the following actions?						
(Q26a) Use a firewall or other antivirus software	3.86	3.96	3.81	3.81	3.85	3.90
(Q26b) Use filtering software (to filter out spam, calls or advertisements)	3.83	3.83	3.61	3.75	3.65	3.70
(Q26c) Use a virtual private network (VPN)	3.17	3.22	3.13	3.16	3.29	3.35
(Q26d) Delete cookies or browsing history in my web browser regularly	3.20	3.34	3.16	3.27	3.18	3.23
(Q26e) Modify the privacy settings on websites or apps when it is needed	3.36	3.52	3.33	3.50	3.34	3.31
(Q26f) Prefer websites or apps that are more secure (like Signal, Telegram, and Duckduckgo)	3.99	3.98	3.88	3.84	3.87	3.93
(Q26g) Use multi-factor authentication feature on websites or apps	3.97	4.00	3.78	3.95*	3.79	3.80
(Q26h) Create strong passwords (with letters, numbers and symbols) on electronic devices, websites or apps	4.10	4.22	3.91	4.09*	4.01	4.01
(Q26i) Do software updates when newer versions are available	3.91	4.10*	3.83	3.92	3.82	3.90

Table E3. Compared acquired confidence under interventions

	Whatsapp		Workshop		Control	
	Pre	Post	Pre	Post	Pre	Post
How much do you agree with the following statements?						
(Q5a) Using a firewall or other related antivirus software keeps my information safe	3.90	4.11**	3.91	3.91	3.97	3.98
(Q5b) Using Virtual Private Network (VPN) keeps my information safe	4.04	4.12	4.03	3.87*	3.98	4.01
(Q5c) Changing privacy settings keeps my information safe	4.09	4.24	4.07	4.07	4.10	4.07
(Q5d) Enabling multi-factor authentication keeps my information safe	4.30	4.45*	4.21	4.25	4.22	4.23
(Q5e) Using websites or apps that are more secure or sensitive to privacy (like Signal, Telegram and Duckduckgo) keeps my information safe	4.18	4.19	4.13	4.04	4.13	4.15

Table E4. Compared self confidence under interventions

	Whatsapp		Workshop		Control	
	Pre	Post	Pre	Post	Pre	Post
How much do you agree with the following statements?						
(Q6a) I am confident in securing my personal privacy online	3.06	3.35**	3.12	2.98	3.02	3.08
(Q6b) I am confident in keeping my information safe online	3.10	3.33**	3.13	2.96	3.04	3.06

(Q6c) I am confident in managing who can use my personal information and data online	2.75	2.97*	2.81	2.51**	2.66	2.79
(Q6d) I am confident in managing who can see the information I share online	2.92	3.25**	3.03	2.72**	2.94	2.98
(Q6e) I am confident in determining whether online information is true or not	3.52	3.66	3.52	3.44	3.42	3.49
(Q6f) I am confident in understanding the terms and conditions offered by websites or apps	2.98	3.10	2.98	2.90	2.80	2.78
(Q6g) I am confident in understanding data analysis which I encounter online	3.21	3.46**	3.23	3.21	3.05	3.07
(Q6h) I am confident in using online public information and data to solve my problems	3.24	3.47*	3.26	3.16	3.15	3.15

Table E5. Compared perceived disempowerment under interventions

	Whatsapp		Workshop		Control	
	Pre	Post	Pre	Post	Pre	Post
How much do you agree with the following statements?						
(Q8a) I have no control over how tech companies collect and use my personal information and data	4.24	4.01**	4.21	4.25	4.41	4.31
(Q8b) It is meaningless to read terms and conditions because I have to accept them in order to use the website or app	4.43	4.19*	4.30	4.22	4.50	4.51
(Q8c) Changing privacy settings does not help to protect personal privacy	3.46	3.05**	3.39	3.31	3.59	3.55
(Q8d) Changing privacy settings takes too much time and effort	3.55	3.24**	3.53	3.65	3.63	3.67
(Q8e) Generally, even if I report inappropriate or harmful content to a	4.00	3.76**	3.87	3.88	3.98	3.89

platform, the platform will not take my report seriously						
(Q8f) Tech companies keep their users' best interests in mind when they design and operate their websites and apps	2.62	2.68	2.73	2.40**	2.59	2.58
(Q8g) Even if I think that a tech company's website or app is creating negative impacts on society, I will have no say over the issue.	3.78	3.57*	3.77	3.69	3.87	3.88
(Q8h) Being online is not safe and there will always be someone trying to deceive or harm me in some way	3.32	3.11*	3.30	3.23	3.34	3.40

Question Book

Survey on Data Literacy of Hong Kong Internet Users

The Department of Government and International Studies, Hong Kong Baptist University (HKBU) commissioned the Hong Kong Public Opinion Research Institute (PORI) to conduct a survey on Hong Kong Internet users' data literacy.

If you have questions about the survey, please email to wendyleungly@hkbu.edu.hk. If you have questions about your rights as a research participant, please contact the HKBU Research Ethics Committee: hkbu_rec@hkbu.edu.hk.

All data you provide will be kept private and confidential. Your participation is voluntary and you may skip any question or terminate your participation in this survey at any time. The result of this study will only be used for academic purposes.

◯ I understand and agree to participate in this survey.

Eligibility Confirmation

[S1] Are you a Hong Kong resident aged 18 or above (i.e., currently residing in Hong Kong)?

◯ Yes
◯ No

Digital Activities

1. In the last month, how much time **on average** do you spend online **every day**?

◯ Less than 1 hour
◯ 1 to less than 2 hour(s)
◯ 2 to less than 4 hours
◯ 4 to less than 6 hours
◯ 6 to less than 8 hours
◯ 8 hours or more
◯ I never go online
◯ Don't know / hard to say

2. In the last month**,** have you done the following online activities? Please select all that apply.

- ☐ Buy and sell online (include goods and services)
- ☐ Use online payment software (like PayMe, WeChat Pay, Alipay)
- ☐ Use instant messenger (like WhatsApp, FB Messenger, Telegram, Signal, WeChat)
- ☐ Use social media (like Facebook, Instagram, Twitter, LIHKG)
- ☐ Use search engines to search for information (like Google, Yahoo!, Bing)
- ☐ Visit news websites or websites about current affairs and politics (like MingPao, NowTV, Initium, HK01)
- ☐ Watch TV show or movie via online streaming (like Netflix, Disney+, Apple TV)
- ☐ Listen to music via online streaming (like Spotify, KKbox, Joox)
- ☐ Watch short videos online (like YouTube, Instagram, Facebook)
- ☐ Play games online
- ☐ Other: _____
- ☐ Don't know / hard to say

3. Have you ever registered for a personal account with the following social media sites or apps? Please select all that apply.
   - ☐ Facebook
   - ☐ Instagram
   - ☐ LinkedIn
   - ☐ Pinterest
   - ☐ Reddit
   - ☐ Snapchat
   - ☐ TikTok
   - ☐ Tumblr
   - ☐ Twitch
   - ☐ Twitter
   - ☐ YouTube
   - ☐ Online forums (like LIHKG, Discuss.com.hk, Baby Kingdom)
   - ☐ Other: _____
   - ☐ I do not use any social media websites or apps
   - ☐ Don't know / hard to say

4. Have you ever used the following instant messaging websites or apps? Please select all that apply.
   - ☐ Discord
   - ☐ Facebook Messenger
   - ☐ Google Hangouts
   - ☐ Instagram Direct Message
   - ☐ Line
   - ☐ Signal
   - ☐ Skype

- ☐ Telegram
- ☐ Viber
- ☐ WeChat
- ☐ WhatsApp
- ☐ Zoom
- ☐ Other: _____
- ☐ I do not use any instant messaging websites or apps
- ☐ Don’t know / hard to say

Data Technology and Confidence

5. How much do you agree or disagree with the following statements?

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Don't know / hard to say
a) Using a firewall or other related antivirus software keeps my information safe.	○	○	○	○	○	○
b) Using Virtual Private Network (VPN) keeps my information safe.	○	○	○	○	○	○
c) Changing privacy settings keeps my information safe.	○	○	○	○	○	○
d) Enabling multi-factor authentication keeps my information safe.	○	○	○	○	○	○
e) Using websites or apps that are more secure or sensitive to privacy (like Signal, Telegram and Duckduckgo) keeps my information safe.	○	○	○	○	○	○

6. How much do you agree or disagree with the following statements?

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Don't know / hard to say
a) I am confident in securing my personal privacy online.	○	○	○	○	○	○
b) I am confident in keeping my information safe online.	○	○	○	○	○	○
c) I am confident in managing who can use my personal information and data online.	○	○	○	○	○	○
d) I am confident in managing who can see the information I share online.	○	○	○	○	○	○
e) I am confident in determining whether online information is true or not.	○	○	○	○	○	○
f) I am confident in understanding the terms and conditions offered by websites or apps.	○	○	○	○	○	○
g) I am confident in understanding data analysis which I encounter online.	○	○	○	○	○	○
h) I am confident in using online public information and data to solve my problems.	○	○	○	○	○	○

Data Technology and Attitudes

7. Overall speaking, do you think data technology makes your personal life or your society better or worse?

	Much better	Somewhat better	No change	Somewhat worse	Much worse	Don't know / hard to say
a) Personal life	○	○	○	○	○	○
b) Society	○	○	○	○	○	○

8. How much do you agree or disagree with the following statements?

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Don't know / hard to say
a) I have no control over how tech companies collect and use my personal information and data.	○	○	○	○	○	○
b) It is meaningless to read terms and conditions because I have to accept them in order to use the website or app.	○	○	○	○	○	○
c) Changing privacy settings does not help to protect personal privacy.	○	○	○	○	○	○
d) Changing privacy settings takes too much time and effort.	○	○	○	○	○	○

e) Generally, even if I report inappropriate or harmful content to a platform, the platform will not take my report seriously.	○	○	○	○	○	○
f) Tech companies keep their users' best interests in mind when they design and operate their websites and apps.	○	○	○	○	○	○
g) Even if I think that a tech company's website or app is creating negative impacts on society, I will have no say over the issue.	○	○	○	○	○	○
h) Being online is not safe and there will always be someone trying to deceive or harm me in some way.	○	○	○	○	○	○

9. How much do you agree or disagree with the following statements?

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Don't know / hard to say
a) When I use a website or app, I usually accept the terms and conditions without reading or fully understanding them.	○	○	○	○	○	○
b) I welcome the government to collect my personal information and data to improve their services.	○	○	○	○	○	○

c) I welcome tech companies to collect my personal information and data to improve their services.	○	○	○	○	○	○
d) I believe that tech companies are violating my privacy by collecting my personal information and data online.	○	○	○	○	○	○
e) I believe that my personal information and data online is being collected excessively.	○	○	○	○	○	○

10. How much do you agree or disagree with the following statements?

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Don't know / hard to say
a) Internet users' personal data should be protected.	○	○	○	○	○	○
b) The government is protecting Internet users' personal data.	○	○	○	○	○	○
c) Tech companies are protecting Internet users' personal data.	○	○	○	○	○	○

11. How much do you agree or disagree with the following statements?

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Don't know / hard to say
a) Citizens should have the right to access the Internet anonymously.	○	○	○	○	○	○
b) Citizens should have the right to anonymously express their opinions online.	○	○	○	○	○	○
c) Citizens should have the right to speak freely online, even if the speech is harmful or controversial.	○	○	○	○	○	○
d) Citizens should have the right to speak freely online without being subject to criminal charges.	○	○	○	○	○	○

12. How much are you willing or unwilling to provide your personal information to websites or apps for the following reasons/purposes?

	Strongly unwilling	Somewhat unwilling	Half-half	Somewhat willing	Strongly willing	Don't know / hard to say
a) To create personal accounts (like social media accounts)	○	○	○	○	○	○

b) To get free services (like public Wi-Fi network)	○	○	○	○	○	○
c) To get personalized services (like weather service, video recommendations)	○	○	○	○	○	○
d) To get advertisements or information that are more relevant to me (like discount offers)	○	○	○	○	○	○
e) To improve the service quality of the website or app	○	○	○	○	○	○

13. How much are you willing or unwilling to provide your personal information to websites or apps under the following conditions?

	Strongly unwilling	Somewhat unwilling	Half-half	Somewhat willing	Strongly willing	Don't know / hard to say
a) If the website or app clearly state how they will use your data	○	○	○	○	○	○
b) If you can control how much personal information and data to share with the tech company	○	○	○	○	○	○

14. Do you think the following data of yours would be collected by website or app providers?

	Definitely	Likely	About half the time	Unlikely	Definitely not	Don't know / hard to say
a) Traces of my internet activities (like my web history, search history and shopping history)	○	○	○	○	○	○
b) Information about me on social media (like my posts, photos and videos)	○	○	○	○	○	○
c) Information that I store on my electronic devices (like my mobile phone and computer)	○	○	○	○	○	○
d) My biometric data that have been recorded (like my fingerprints, appearance, voice)	○	○	○	○	○	○
e) My habits in real life (like my hobbies and places I frequent go)	○	○	○	○	○	○

15. How much do you agree or disagree with the following statements?

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Don't know / hard to say
a) The information I find online is always true.	○	○	○	○	○	○
b) I don't care if the information I find online is true or not.	○	○	○	○	○	○

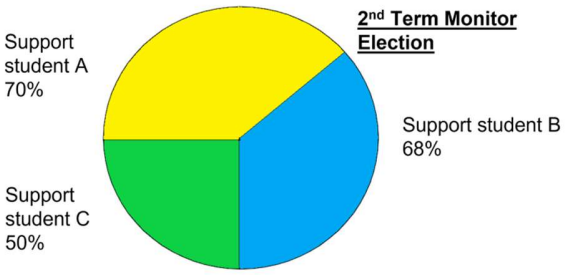
c) When I search the internet, I generally look for information that I find interesting.	○	○	○	○	○	○
d) When I search the internet, I generally look for information that agrees with my view.	○	○	○	○	○	○

Data Literacy

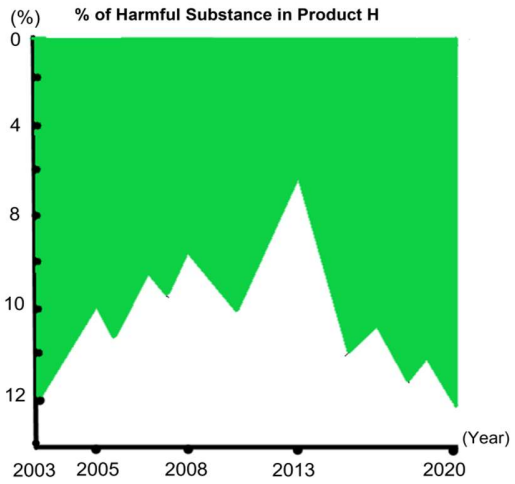
Please answer the following questions based on your existing knowledge, you **DO NOT NEED** to search for the answers.

a. Data Thinking

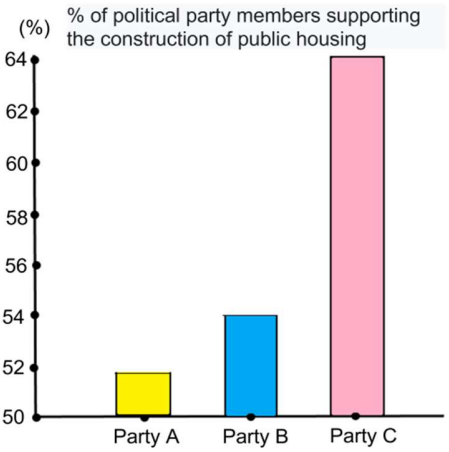
16. Some of the following charts are misleading, please select the chart that you think is the **most accurate (without misleading components)**.



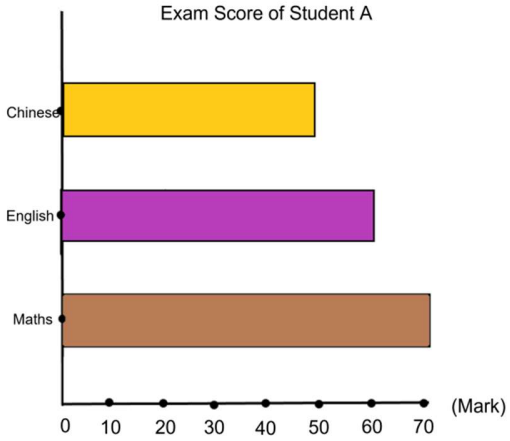
○ A.


% of Harmful Substance in Product H
(%)
0
4
8
10
12
(Year)
2003
2005
2008
2013
2020

○ B.


% of political party members supporting the construction of public housing
(%)
64
62
60
58
56
54
52
50
Party A
Party B
Party C

○ C.



◯ D.

◯ Don't know

17. When you get online, sometimes you will see the message "Do you accept to use cookies?". Which of the following descriptions about **"cookies"** is **correct**?
    - ◯ It helps you to remove advertisements
    - ◯ It temporarily stores your browsing history on the website
    - ◯ You do not have the right to clear, enable and manage cookies
    - ◯ It protects you from hackers
    - ◯ Don't know

18. When you use Wi-Fi services offered in public places (such as airports or coffee shops), which of the following approaches **best keep your data safe**?
    - ◯ Use Virtual Private Network (VPN) service before logging in the Wi-Fi
    - ◯ Conduct online activities that involve sensitive personal data such as online banking or purchasing airline tickets
    - ◯ Choose a Wi-Fi service that allows you to log in without a password
    - ◯ Turn on the auto-connect function that allows you to access Wi-Fi service faster next time
    - ◯ Don't know

19. When you search for "dress" on Google, which of the following is the **most accurate** description of your first search result like below?

- ◯ It is the best or most relevant result
- ◯ It appears as an advertisement or sponsored link
- ◯ It is the most common result found by other web users
- ◯ It is a computer virus
- ◯ Don't know

20. If the URL of a website starts with **https://** (instead of http://), which of the following descriptions is **correct**?
    - ◯ It means that only specific Internet service providers can open this website
    - ◯ It means that the current version of the website is the latest version
    - ◯ It means that the website's content has been reviewed as harmless content
    - ◯ It means that the information input to the website will be encrypted
    - ◯ Don't know

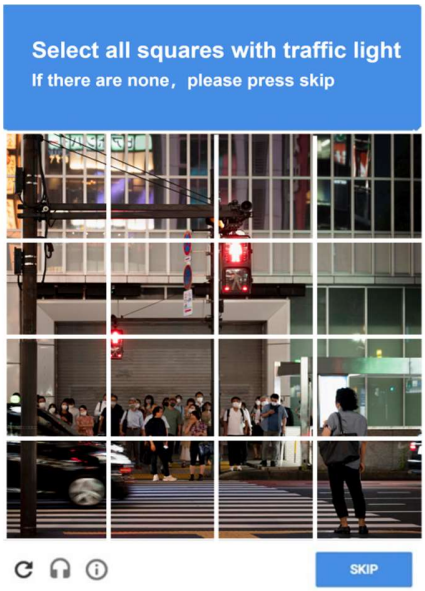
21. When you use **private browsing function** of your web browser to access a website, which of the following descriptions is **correct**?
    - ◯ This feature prevents internet service providers from viewing your online activities
    - ◯ This feature prevents web browsers from storing your information (like browsing history)
    - ◯ This feature prevents you from being hacked
    - ◯ This feature prevents your employer or school from viewing your online activities
    - ◯ Don't know

22. According to the Personal Data (Privacy) Ordinance, there are six data protection principles. Which of the following descriptions is **correct** regarding the data protection principles?
    - ◯ Data users are required to collect personal data of data subjects in a fair and lawful manner
    - ◯ Data users can freely collect personal data of data subjects without specifying their purposes
    - ◯ Data subjects have no right to request access to or correct their personal data, even if the data collected is inaccurate

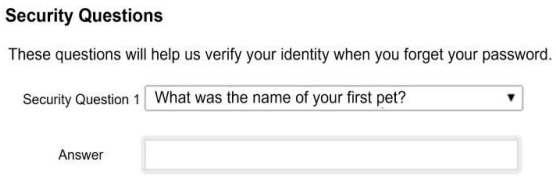
◯ Data users can freely disclose the personal data they collected without the consent of the data subjects
◯ Don't know

23. Many websites and apps require users to complete **multi-factor authentication** before they can log in. Which of the following pictures is an example of multi-factor authentication?

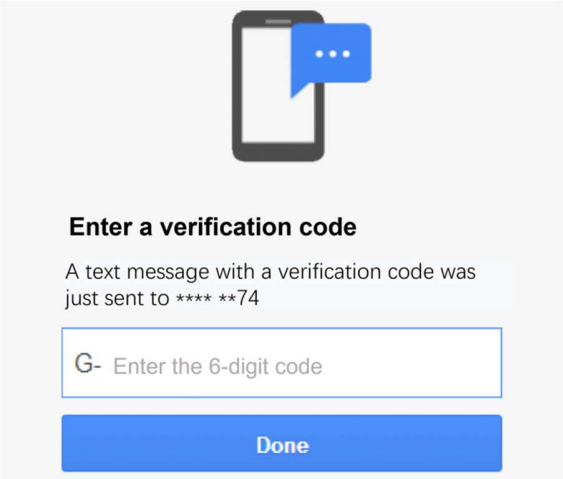
◯ A.



◯ B.



◯ C.



○ D.

○ Don't know

b. Data Doing

24. How much do you agree or disagree with the following statements?

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Don't know / hard to say
a) I often share information about me on social media.	○	○	○	○	○	○
b) I often share community or public affairs information on social media.	○	○	○	○	○	○

25. How much do you agree or disagree with the following statements?

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Don't know / hard to say
a) I often collect information from various sources, even if it is different from my position or point of view.	○	○	○	○	○	○
b) I often collect information from various sources to verify the content's trueness.	○	○	○	○	○	○

26. Would you take the following actions?

	Definitely	Likely	About half the time	Unlikely	Definitely not	Don't know / hard to say
a) Use a firewall or other antivirus software	○	○	○	○	○	○
b) Use filtering software (to filter out, like, spam, calls or advertisements)	○	○	○	○	○	○
c) Use a virtual private network (VPN)	○	○	○	○	○	○
d) Delete cookies or browsing history in my web browser regularly	○	○	○	○	○	○

e) Modify the privacy settings on websites or apps when it is needed	○	○	○	○	○	○
f) Prefer websites or apps that are more secure (like Signal, Telegram, and Duckduckgo)	○	○	○	○	○	○
g) Use multi-factor authentication feature on websites or apps	○	○	○	○	○	○
h) Create strong passwords (like passwords with letters, numbers and symbols) on electronic devices, websites or apps	○	○	○	○	○	○
i) Do software updates when newer versions are available	○	○	○	○	○	○

c. Data Participation

27. How often would you take the following actions?

	Always	Often	Sometimes	Rarely	Never	Don't know / hard to say
a) Look up public service information online via government websites (like GovHK).	○	○	○	○	○	○
b) Use government online services (such as making an appointment to renew my ID or passport).	○	○	○	○	○	○
c) Participate in government public consultations or express opinions to relevant bodies online.	○	○	○	○	○	○

28. How much do you agree or disagree with the following statements?

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Don't know / hard to say
a) As a citizen, I have a duty to try to get informed about politics and social affairs.	○	○	○	○	○	○
b) Everyone is free to not engage in political and societal activities.	○	○	○	○	○	○
c) Even as a single person, one can change something in society and politics.	○	○	○	○	○	○
d) Only as a member of a group can one change something in society and politics.	○	○	○	○	○	○

29. If you encounter the following situations, would you take the below corresponding actions?

	Definitely	Likely	About half the time	Unlikely	Definitely not	Don't know / hard to say
a) When I find misinformation (like fake news) on social media, I report to the platform.	○	○	○	○	○	○

b) When I find that the information I received on an instant message or social media group is untrue, I clarify it with others in the group.	○	○	○	○	○	○
c) When I see inappropriate or harmful content online, I report to the platform.	○	○	○	○	○	○
d) When I receive a suspected spam message on instant messenger or within a social media group, I block the sender and report to the platform.	○	○	○	○	○	○

30. How often would you:

	Always	Often	Sometimes	Rarely	Never	Don't know / hard to say
a) Participate in joint signature campaigns online	○	○	○	○	○	○
b) Participate in instant message or social media groups that are related to public affairs or community affairs (participation includes obtaining or offering instant information, taking part in online discussions)	○	○	○	○	○	○

31. If someone needs it, would you help him/her keep his/her data safe (such as helping him/her to set strong passwords or change privacy settings)?

◯ Definitely
◯ Likely
◯ About half the time

◯ Unlikely
◯ Definitely not
◯ Don't know / hard to say

Basic Demographics

32. Gender:
◯ Male
◯ Female
◯ Other

33. Age:
◯ ____ year-old
◯ 18 – 29
◯ 30 – 39
◯ 40 – 49
◯ 50 – 59
◯ 60 – 69
◯ 70 or above

34. Education attainment (including current study):
◯ Primary or below
◯ Lower secondary (S1-S3)
◯ Upper secondary (S4-S7 / DSE / YiJin)
◯ Tertiary: non-degree course (including diploma / certificate / sub-degree course)
◯ Tertiary: bachelor degree course
◯ Tertiary: postgraduate school or above

35. Current work status:
◯ Employer
◯ Permanent Work (No renewal required)
◯ Contract work (One year or more)
◯ Short-term contract work (Less than one year)
◯ Freelancer / Part-Time work
◯ Retired / Pensioner
◯ Housekeeper
◯ Full-time student
◯ Unemployed
◯ Other

36. Monthly household income:

◯ No income
◯ Below $10,000
◯ $10,001 to $20,000
◯ $20,001 to $30,000
◯ $30,001 to $50,000
◯ $50,001 to $70,000
◯ $70,001 to $100,000
◯ Over $100,000
◯ Don't know

37. Which of the following best describes your political inclination?

- ○ Localist
- ○ Pro-democracy camp
- ○ Centrist
- ○ Pro-establishment camp
- ○ Other
- ○ No political inclination / politically neutral / do not belong to any camp
- ○ Don't know / hard to say

Lastly, our team will conduct a follow-up study in the coming few months to explore the mechanisms leading to enhanced data literacy, cash incentives will be offered. If you are interested in participating in our follow-up study, please indicate your interest below, and we will contact you soon.

◯ I'm interested in participating in the follow-up study.

End of Questionnaire

Thank you for your participation.

If you are one of the first 1,000 respondents who fully complete the survey, we will pay you $20 in cash as a token of appreciation. The arrangement for collecting the incentives will be sent in another email (if you are still our panel member at that time). To avoid missing the email, please add panel@pori.hk to your contacts.

WhatsApp Short Posts

（1A）什麼是多重要素驗證？ 為什麼我們需要它？

大部分網站與應用程式都需要用戶設立個人賬戶，使用密碼是保護用戶個人資料與數據的有效方法。然而，對於一些涉及個人敏感資料的網站與應用程式，如網上銀行與大學賬戶，僅使用密碼進行保護則未必足夠，因為密碼有機會被竊取、猜測或遭黑客入侵，用戶亦難以察覺有人正在訪問自己的賬戶。多重要素驗證的出現，為用戶的賬戶增加了額外的安全性，使用其他因素，如帶有多重要素驗證應用程序的智能電話，去驗證用戶的身份並批准身份驗證請求，可防止用戶以外的任何人登錄用戶的賬戶，即使他們知道用戶的密碼。

資料來源： “DUO TWO-FACTOR AUTHENTICATION (2FA)”，香港浸會大學資訊科技處網站，
(https://ito.hkbu.edu.hk/services/it-security/2fa.html)

（1B）What Is Multi-Factor Authentication? Why Do We Need It?

Most websites and apps require their users to create a personal account, and passwords effectively protect users' personal information and data. Yet, password protection alone may not be sufficient for some websites and applications involving sensitive personal information, such as online banking and university accounts, as passwords may be stolen, guessed, or hacked. It is also difficult for users to know that someone is accessing their account.

Source: “DUO TWO-FACTOR AUTHENTICATION (2FA)”, Office of Information Technology, Hong Kong Baptist University (https://ito.hkbu.edu.hk/services/it-security/2fa.html)

（2A）認識《個人資料（私隱）條例》：安裝閉路電視或網絡攝錄機合法嗎？

現時有不少地方都會安裝閉路電視或攝錄機，這個行為有否違反《個人資料（私隱）條例》取決於其有否涉及「收集」個人資料。只作實時監察，沒有拍攝功能的閉路電視與有開啟攝錄功能，但未能清楚拍攝容貌的閉路電視都不涉及「收集」個人資料。然而，若有僱主安裝能拍攝與清楚記錄容貌的閉路電視，並會儲存影片以監察僱員的工作情況，他就有機

會因為過度或不合法公平「收集」個人資料而違例。近年出現了不少與攝錄機有關的爭議性新聞，如的士司機在網上公開乘客哺乳的照片；及政府為了加強監察而在公眾地方安裝攝錄機，你認為這些行為有沒有違反《個人資料（私隱）條例》呢？

資料來源：「從真實個案輕鬆認識《個人資料(私隱)條例》」講座資料，香港個人資料私隱專員公署網站 (https://www.pcpd.org.hk/misc/dpoc/files/case_sharing27112107.pdf)

（2B）Introduce Personal Data (Privacy) Ordinance: Is Installing CCTV or Video Cameras Legal?

CCTV and video cameras are everywhere. Whether their installations have convicted the Personal Data (Privacy) Ordinance depends on whether it involves the "collection" of personal data. CCTV, which is only for real-time monitoring without filming, and CCTV with recording function but unable to capture appearances do not involve "collecting" personal data. However, if an employer installed CCTV that recorded and captured appearance, storing videos to monitor employees' work, he would risk violating the law by excessive or unlawful "collection" of personal data. Recently, there has been much controversial news related to video cameras, such as a taxi driver posting photos of a breastfeeding woman on the Internet; and the government installing cameras in public places to strengthen surveillance. Do you think these actions violate the Personal Data (Privacy) Ordinance?

Source：「從真實個案輕鬆認識《個人資料(私隱)條例》」Talk Slides, the Office of the Privacy Commissioner for Personal Data's Website (https://www.pcpd.org.hk/misc/dpoc/files/case_sharing27112107.pdf)

（3A）Business Insider：全球超過 5 億 Facebook 用戶的個人數據被洩露

2021 年 4 月，Business Insider 發布了一篇報導，稱 5.33 億 Facebook 用戶的個人數據遭外洩。這些數據被發佈在一個低級別的駭客論壇；數據包括用戶的電話號碼、全名、位置與電子郵件地址。研究人員表示，黑客可以利用這些數據冒充他人並進行欺詐。這並不是 Facebook 第一次發生用戶數據外洩事件。許多大公司，如 LinkedIn、雅虎和 Adobe，亦曾發生過類似的用戶數據外洩事件。用戶使用如 Facebook 等知名科技公司的服務是信任他們能夠保護用戶資料，Facebook 應該小心對待這些數據，用戶個人資料外洩因此屬嚴重失信行為，應慎重處理。你認為在這件事誰應該承擔最大的責任呢？你認為應該做些什麼來防止同樣問題再發生呢？

資料來源： “533 million Facebook users' phone numbers and personal data have been leaked online”, Business Insider (https://www.businessinsider.com/stolen-data-of-533-million-facebook-users-leaked-online-2021-4)
「FB 全球 5 億個資外洩》住家地址、信用卡號都可能外流…使用者如何應對？」，商周 (https://www.businessweekly.com.tw/business/blog/3006048)

（3B）Business Insider: Personal Data of Over 500 Million Facebook Users was Leaked

In April 2021, Business Insider released an article stating that 533 million Facebook users' data was leaked online. The personal data was posted in a low-level hacking forum; it includes users' phone numbers, full names, locations, and email addresses. Security researchers said hackers could use the data to impersonate people and commit fraud. Facebook has experienced more than one user data breach. Like LinkedIn, Yahoo, and Adobe, many large corporations have experienced typical data breaches. Users signing up to reputable tech companies like Facebook trust them with their data, and Facebook is supposed to treat the data with the utmost care. Therefore, users' personal information leaked is a significant breach of trust and should be carefully handled. Who do you think should bear the most responsibility? What do you think should be done to prevent the same issue?

Source: “533 million Facebook users' phone numbers and personal data have been leaked online”, Business Insider (https://www.businessinsider.com/stolen-data-of-533-million-facebook-users-leaked-online-2021-4)
「FB 全球 5 億個資外洩》住家地址、信用卡號都可能外流…使用者如何應對？」，商周 (https://www.businessweekly.com.tw/business/blog/3006048)

（4A）認識《個人資料（私隱）條例》: 我可以拒絕接受直銷訊息嗎？

《個人資料（私隱）條例》在二零一二年針對使用個人資料作直接促銷的新規定進行了主要修訂。條例規定資料使用者必須向資料當事人提供清楚易懂的訂明資訊及回應途徑，讓資料當事人自願選擇同意或表示「不反對」個人資料被用作直銷用途。因此，如商戶沒有依從客戶的拒收直銷訊息要求，繼續對客戶進行直銷，商戶就有機會被定罪及罰款。即使

客戶曾同意接受直銷條款，他仍有權隨時要求商戶停止使用或提供個人信息給第三方進行直銷，他亦可以要求更改收到的產品或服務信息類型及提供哪種資訊予商戶進行直接營銷。

資料來源：「從真實個案輕鬆認識《個人資料(私隱)條例》」講座資料，香港個人資料私隱專員公署網站 (https://www.pcpd.org.hk/misc/dpoc/files/case_sharing27112107.pdf)

（4B）Introduce Personal Data (Privacy) Ordinance： Can I Opt-out of Direct Marketing?

The Personal Data (Privacy) Ordinance underwent significant amendments in 2012 to introduce the direct marketing provisions. The Ordinance stipulates that data users must provide their data subjects with clear and understandable prescribed information and response channels to allow data subjects to voluntarily consent to or express "no objection" to using their data for direct marketing. Therefore, if a company does not comply with the customer's request to opt out of direct marketing and continues to promote information to the customer, the company may risk being convicted and fined. Even if the customer has once agreed with the terms of direct marketing, he still has the right to request the seller to stop using or offering personal information to third parties for direct marketing. The customer can also ask to change the type of product or service information received and change which type of personal data can be used by the company for direct marketing.

Source：「從真實個案輕鬆認識《個人資料(私隱)條例》」Talk Slides, the Office of the Privacy Commissioner for Personal Data's Website (https://www.pcpd.org.hk/misc/dpoc/files/case_sharing27112107.pdf)

（5A）向網絡欺凌說「不！」

網絡欺凌泛指欺凌者在社交平台作出令欺凌對象困擾的行為，當中包括：發送騷擾、羞辱及恐嚇的訊息給欺凌對象；進行網絡公審以取笑或批評欺凌對象；及在社交平台上發佈欺凌對象的個人資料（俗稱「起底」）。雖然欺凌者可能只因一時氣憤，貪玩或無聊而作出網絡欺凌行為，但這不但會對欺凌對象造成極大傷害，還可能構成刑事罪行。若你在網上發現欺凌他人的行為時，請不要以任何方式（如轉發）助長欺凌行為，你可以刪除該訊息及封鎖欺凌者的賬戶，並向有關平台作出舉報。如你不幸成為網絡欺凌的對象，請避免與欺凌者爭議，向有關平台反映及要求刪除欺凌訊息，並暫時離開該社交平台，有需要時亦可向個人資料私隱專員公署等相關機構尋求協助。

資料來源：數碼新世代， 向網絡欺凌 說「不」 (2022 年 2 月) 單張，香港個人資料私隱專員公署網站，
(https://www.pcpd.org.hk//tc_chi/resources_centre/publications/files/children_privacy_cyberbullying.pdf)

（5B）Say “No!” to Cyberbullying

Cyberbullying refers to the behaviors that the bullies (or haters) do online to harass their targets. They include attacking the targets via humiliating and intimidating messages, ridiculing and criticizing the targets as Internet judges, and disclosing the targets' personal information online (doxxing). While the bullies may engage in cyberbullying out of anger, playfulness, or boredom, it does not only create significant harm to the targets but may also constitute criminal offenses. If you spot any cyberbullying behaviors, please do not promote them in any way (such as forwarding). You can delete the messages, block the bullies' accounts, and report them to the relevant platform. If you become the target of cyberbullying, please avoid disputes with the bullies, write to the platform and request the deletion of the bullying messages, and leave the online platform temporarily. If necessary, you can also seek help from the Office of the Privacy Commissioner for Personal Data or other relevant organizations.

Source: “New Digital Era, Say "No" to Cyberbullying (February 2022)” Leaflet, the Office of the Privacy Commissioner for Personal Data’s Website
(https://www.pcpd.org.hk//english/resources_centre/publications/files/children_privacy_cyberbullying.pdf)

（6A）使用公共 Wi-Fi 的六大貼士

很多公眾場所都會提供免費的 Wi-Fi 上網設施。用戶使用 Wi-Fi 服務時，須注意潛在的保安風險，如將流動裝置連接至黑客裝設或被黑客入侵的 Wi-Fi 網絡，導致個人資料外泄。以下提供六個使用公共 Wi-Fi 的保安貼士：（1）所有軟件，尤其是安全防護軟件及瀏覽器，須經常保持更新；（2）小心核實所選用網絡的名稱，避免連結可疑的 Wi-Fi 網絡；（3）連接前採取適當的加密連接，如虛擬私有網絡（VPN）；（4）避免使用電子服務處理個人敏感資料；（5）避免連接可疑網頁，亦不要按下任何連結；及（6）停用流動裝置的自動連接功能，在不使用 Wi-Fi 時，亦應關閉流動裝置的 Wi-Fi 連接。

資料來源：「使用公共 Wi-Fi 的保安貼士」，政府資訊科技總監辦公室資訊安全網 (https://www.infosec.gov.hk/tc/best-practices/person/tips-on-using-public-wi-fi)

（6B）Six Tips on Using Public Wi-Fi

Free Wi-Fi facilities are available in many public areas. Users should pay attention to potential security risks when using Wi-Fi services. For example, mobile devices may connect to the Wi-Fi network set up or compromised by hackers, leading to personal information leakage. Here are six tips on using public Wi-Fi: (1) All software, particularly security applications and browsers, should constantly be updated. (2) Verify the selected network's name. Avoid using dubious Wi-Fi networks. (3) Use appropriate encrypted connections like Virtual Private Network (VPN) before connection. (4) Avoid handling sensitive personal information via e-services. (5) Avoid accessing suspicious web pages and not clicking any links. (6) Disable the auto-connection function on mobile devices. Turn off the Wi-Fi connection of your mobile devices when not in use.

Source: "Tips on Using Public Wi-Fi", InfoSec, Office of the Government Chief Information Officer (https://www.infosec.gov.hk/en/best-practices/person/tips-on-using-public-wi-fi)

（7A）保護隱私，你我有責

我們在保護個人私隱的同時，也不要忽視保護別人私隱的重要性。首先，我們應尊重別人的私隱權。在得到他人的同意前，我們不應隨便翻看及公開他人的個人資料，也不應隨便使用他人的電子設備與個人賬戶。此外，我們亦應該幫助身邊有需要的人保護他們的個人私隱，例如協助他人：（1）理解網站或應用程式提供的條款和條件；（2）根據需要更改隱私設置；（3）設置安全度高的密碼（如包含大小階字母，數字與符號）；（4）刪除網絡痕跡，如瀏覽記錄與 Cookies（指為了辨別使用者，而儲存在用戶端，包含個人資料的小型文字檔案）；及（5）根據需要安裝安全防護軟件，過濾器軟件與私隱保密度較高的應用程式。

資料來源：「尊重他人私隱，我做得到 (2021 年 7 月)」單張，香港個人資料私隱專員公署網站，
(https://www.pcpd.org.hk//english/resources_centre/publications/files/children_privacy_respect.pdf)

（7B）Protect Privacy, We Can Do It!

While protecting personal privacy, we should not ignore the importance of protecting the privacy of others. First, We should respect others' right to privacy. Before obtaining their consent, we should not casually browse and disclose other people's personal information, nor should we casually use other people's electronic devices and personal accounts. In addition, we should also help those in need to protect their privacy, such as assisting others to (1) understand the terms and conditions provided by websites or applications; (2) change the privacy settings as needed; (3) set a highly secure password (ones containing letters, numbers, and symbols); (4) delete network traces, such as browsing records and Cookies (referring to small text files stored on the user side, containing personal data for identification purpose); and (5) install anti-virus software, filtering software and applications with high privacy protection as needed.

Source: "Respect Others' Privacy, I Can Do It (July 2021)" Leaflet, the Office of the Privacy Commissioner for Personal Data's Website (https://www.pcpd.org.hk//english/resources_centre/publications/files/children_privacy_respect.pdf)

（8A）虛擬私人網路（VPN）：你需要知道的

不少網民都會考慮訂閱 VPN，但到底什麼是 VPN 呢？VPN 是虛擬私人網路的縮寫，旨在幫助用戶在上網時保障私隱和安全。它能隱藏用戶的網路協定位址（IP）並加密他們的互聯網流量，防止用戶的身份和私人資料被科技公司、互聯網服務提供商和網絡犯罪分子跟踪和記錄，而且通過重新定位用戶連接的位置，用戶可以訪問受地區限制和政府審查限制的在線內容。

資料來源： "What Is A VPN And How Does It Work?", Forbes, (https://www.forbes.com/advisor/business/software/what-is-a-vpn-and-how-does-it-work/)

（8B）Virtual Private Network (VPN): What You Should Know

Many online users are considering subscribing to VPN, but what is VPN? VPN is short for virtual private networks; it is a privacy protection tool that helps people stay private and secure online. It hides users' Internet Protocol (IP) addresses. It encrypts their internet traffic, preventing users'

identity and personal information from being tracked and recorded by tech companies, internet service providers, and cybercriminals. And by re-allocating the location of users' connections, they can access online content restricted under regional restrictions and government censorship.

Source: "What Is A VPN And How Does It Work?", Forbes, (https://www.forbes.com/advisor/business/software/what-is-a-vpn-and-how-does-it-work/)

（9A）謹防網上詐騙！

網絡世界危機處處，騙徒行騙手法層出不窮，隨著網上詐騙的風險不斷增加，網民應小心提防。以下是三個常見的網上騙局例子。（1）騙徒會開設投資股票社交媒體群組，訛稱提供專家內幕消息，吸引網民加入，群組表面讓股民交流心得，實際以「唱高散貨」詐騙。網民應小心查證消息來源，如有疑慮應馬上退出與舉報群組。（2）釣魚是常見的詐騙形式，騙徒會佯稱銀行發電郵，誘騙網民打開連結與登入，盜取他們的戶口資料。網民應核實可疑電郵的地址與連結，切勿隨便「更新」個人資料。（3）騙徒亦會在交友軟件以「美人計」哄騙目標在非法網站投資，簽署投資授權書，高價購買不需要的產品與服務及傳送裸照從而勒索。網民在向「網友」提供個人信息或金錢前，應謹慎考慮。

資料來源：「網上詐騙」，錢家有道，消委會網站 (https://www.ifec.org.hk/web/tc/moneyessentials/scams/scam-websites.page)

（9B）Beware of Online Scams!

The Internet can be dangerous when scammers have endless tricks to fool online users. With the ever-present risk of online scams, netizens should be careful. Below are three examples of common online scams. (1) Scammers set up social media groups related to investments, falsely claiming to provide experts' insider information to attract netizens to join. While allowing shareholders to exchange their experiences on the surface, it is a scam to manipulate stock prices. Netizens should be careful about information sources; leave and report the group immediately if in doubt. (2) Phishing is a typical online fraud. Scammers pretend to send emails from banks to trick netizens into opening links and logging in, stealing their account information. Netizens should verify the addresses and links of suspicious emails and not "update" their personal information causally. (3)

Scammers, via “honey trap”, deceive their targets in dating software to invest in illegal websites, sign investment authorization letters, buy unwanted products and services at high prices, and send nude photos for blackmail. Before offering personal information or money to “online friends”, netizens should think twice.

Source: “Beware of online scams”, The Chin Family, The Investor and Financial Education Council Website (https://www.ifec.org.hk/web/en/moneyessentials/scams/scam-websites.page)

(10A) 數碼鴻溝：以網上教學與抗疫措施為事例

科技資源、知識與能力不足是造成社會數碼鴻溝(digital dive)擴大的主要成因，長者與低收入人士等弱勢往往最容易被忽略。當數碼設備與科技滲透生活，數碼能力較貧乏的人如長者除了較難追趕科技潮流，亦更容易陷入資料洩露與網絡騙案等危機。全面數碼化的病毒風險通知、疫苗注射預約與消費券系統都顯示著數碼化進程的必然性，市民若不跟隨，就有機會失去自身的自由、權利，甚至個人身份。疫情讓網上教學成為主流，進一步加深了基層家庭學童的數碼鴻溝，他們除了沒有適合的工具如熒幕設備，數碼能力亦因為缺乏家長支援而難以發展，如何把提升數碼能力融入主流教育，讓所有學生能公平地享受數碼化教學，成為值得深思的重要議題。

資料來源：「解構香港數碼鴻溝 資源、知識不足都是成因 羅陸慧英教授：你冇得選擇上唔上網，只可選擇存唔存在」，明周，https://www.mpweekly.com/culture/zoom-%E4%B8%8A%E7%B6%B2%E8%AA%B2-%E5%85%AC%E5%85%B1%E6%9C%8D%E5%8B%99%E6%95%B8%E7%A2%BC%E5%8C%96-195047

(10B) The Digital Divide: Examples of Online Teaching and Anti-epidemic Measures

Insufficient technological resources, knowledge, and ability are the main reasons for the widening digital divide in society. The disadvantaged, such as the elderly and low-income groups, are often the most easily overlooked. When digital devices and technology have permeated our daily lives, people who are less digitally capable, like the elderly, are more difficult to keep up with technological trends and are more likely to fall into digital crises such as data leakage and internet fraud. The fully digitized system of virus risk notification, vaccination appointments, and consumption voucher scheme all illustrate the inevitability of the digitization process. If citizens

do not follow, they may lose their freedom, rights, and even personal identity. The epidemic necessitates online teaching, further deepening the digital divide among students from grassroots families. Apart from not having suitable tools such as screen devices, their digital skills are also difficult to develop due to the lack of parental support. How to integrate digital skills into mainstream education, so that all students can enjoy online teaching equitably, has become an important issue worth pondering.

Source:「解構香港數碼鴻溝 資源、知識不足都是成因 羅陸慧英教授：你冇得選擇上唔上網，只可選擇存唔存在」, Ming Pao Weekly, https://www.mpweekly.com/culture/zoom-%E4%B8%8A%E7%B6%B2%E8%AA%B2-%E5%85%AC%E5%85%B1%E6%9C%8D%E5%8B%99%E6%95%B8%E7%A2%BC%E5%8C%96-195047

（11A）假新聞與事實查核

相信不少人都曾被假新聞誤導，事實查核能幫助人們破解假新聞，避免以訛傳訛的情況出現。事實查核可分為 4 個 W。第一個與第二個 W（When & Where）的重點是時間與地點。善用以圖搜圖的功能，尋找圖片或影片的真正出處，從而還原情境脈絡，避免被錯置或修改過的資訊矇騙。第三個與第四個 W（Who & Why）的重點是來源與原因。透過評估上傳者或創作者的立場、信用與動機等來分析新聞的可信度，識別意在攻擊、誤導、圖利與宣傳的信息，尤其需留意包含著極端正面或負面消息的帖文，它們會較易出現虛假成份。若覺得辨別假新聞太困難，也可以訪問坊間的事實核查平台，如事實查核實驗室(Factcheck Lab)與香港浸會大學事實查核中心(BU Fact Check)。

資料來源：【社群內容打假術】Google 教你深度事實查核，人人都是數位打假王，新聞實驗室網站，https://newslab.pts.org.tw/news/81

（11B）Fake News and Fact-Checking

It is not uncommon for people to be misled by fake news. Fact-checking helps people to identify fake news, avoiding the circulation of such misinformation. Fact-checking can be divided into 4 Ws. The first and second W (When & Where) focus on time and space. People should make good use of the image search function to seek the real source of the picture or video to restore the context of the situation. Avoid being deceived by misplaced or modified information. The third and fourth W (Who & Why) focus on source and reason. People should analyze the credibility of news pieces

by evaluating the uploader or creator's stance, credit, and motivation and identify information intended to attack, mislead, make a profit, and promote. Especially those news that contains exceptionally positive or negative messages, which are more likely to be false. If you find it too difficult to identify fake news, you can visit fact-checking platforms such as Factcheck Lab and BU Fact Check.

Source:「【社群內容打假術】Google 教你深度事實查核，人人都是數位打假王」，新聞實驗室網站，https://newslab.pts.org.tw/news/81

（12A）網絡公民參與：以香港為例

新媒體的崛起創造了公民參與的新方式，網絡和社交媒體在公民意識的形成和社會事務的參與均扮演著重要角色。網絡公民參與的形式眾多，除了獲取社會事務資訊，還包括評論、討論、加入群組、聯署和眾籌等。然而，社交媒體的信息並不完整，討論傾向一面倒，反對意見亦可能會被屏蔽。人們若不加注意，就有可能陷入偏聽偏信，無法把握事實全貌。此外，皆因人人都可以在網上匿名發言，社交媒體內容的質量和可信度一直成疑。面對海量信息，分辨和處理信息的能力就變得至關重要。學校教育讓學生建立完整的知識體系，並培養他們的批判思維，讓他們能以不同學科的知識和分析方法來解讀信息。除了學校教育，媒體素養教育亦能培養人們解讀信息和產生資訊的能力。近年，社交媒體的地區群組成為了人們了解自己社區的常用平台。與鄰居的互動、社區活動的參與和針對社會事務的討論都有助培養了人們的社會和公民意識。

資料來源:「【網絡公民參與】教育脫節　市民難辨海量信息」，香港 01，https://www.hk01.com/sns/article/418223；「【網絡公民參與】媒體素養教育　校園不是唯一擔責」，香港 01，https://www.hk01.com/sns/article/418302

（12B）Online Civic Engagement: The Case of Hong Kong

The rise of new media creates new forms of civic participation. The Internet and social media play an increasingly important role in forming civic awareness and involvement in social affairs. There are various forms of online civic participation. On top of obtaining social affairs information, it includes commenting, discussing, joining groups, co-signing, and crowdfunding. However, the information on social media is often incomplete, discussions tend to be one-sided, and counter-opinions may be blocked. If people are unaware, they may fall into partiality and fail to grasp the whole picture. Moreover, the quality and credibility of social media content are questionable as

everyone can post online anonymously. Facing massive information, the ability to distinguish and process information becomes crucial. School education allows students to establish complete knowledge systems and nourish their critical thinking skills, allowing them to interpret data with different subjects' knowledge and analytical methods. In addition to school education, media literacy education cultivates people's ability to interpret the information they receive and their ability to generate information. In recent years, community-based social media groups have become common platforms for people to understand their community better. People's interactions with neighbors, community activities participation, and discussion regarding social affairs cultivate their social and civic awareness.

Source: 「【網絡公民參與】教育脫節　市民難辨海量信息」，HK01，https://www.hk01.com/sns/article/418223；「【網絡公民參與】媒體素養教育　校園不是唯一擔責」，HK01，https://www.hk01.com/sns/article/418302

（13A）數據可視化錯誤：三個例子

數據可視化旨在通過視覺手段簡潔地傳達數據的模式、趨勢和相關性。然而，糟糕的數據可視化可能會誤導並導致錯誤解釋。以下是三個常見的錯誤。 (1) 尺度具誤導性的軸會形成有重大差異和變化的錯誤印象，標準化軸的尺度從零開始，能有助避免小尺度造成的"間隙錯覺"。 (2) 圓餅圖不適合用作標準比較，例如月度或年度比較，這基於大多數圓餅圖只顯示總數的百分比，不提供絕對值，百分比的變化不一定表示絕對值的變化。 (3) 無法突顯差異作比較的數據可視化是無效的，添加標籤來提供額外的上下文和信息是一種方法，根據邏輯比例調整軸也能有助於改進圖表。

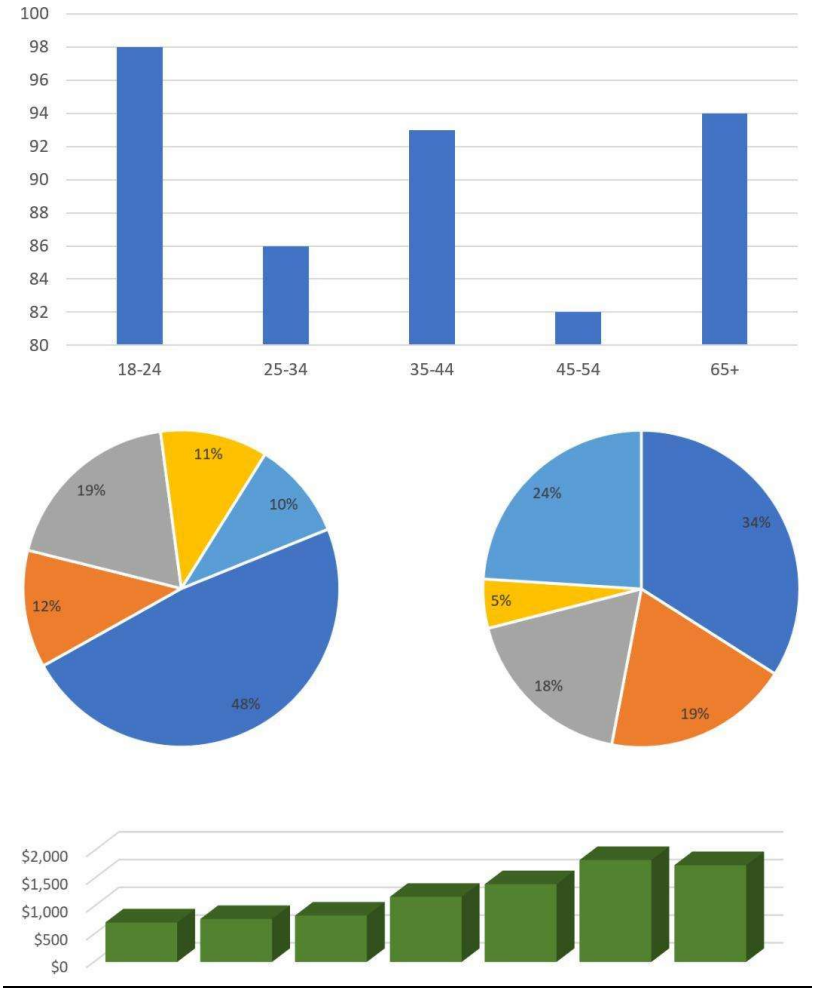
資料來源："Stop Making These Data Visualization Mistakes in 2019", Liquid's Website, https://www.liquidint.com/blog/marketing/data-visualization-mistakes-to-avoid

（13B）Data Visualization Mistakes: Three Examples

Data visualization aims to concisely communicate patterns, trends and correlations of data through visual means. However, bad data visualization can be misleading and lead to wrong interpretation. Here are three common mistakes spotted. (1) Axis with misleading scales creates the wrong impression of significant difference and change. Standardization of the axis to start at zero helps

to avoid the “gap illusion” caused by small scales. (2) Pie charts are not fit for standard comparisons like monthly or yearly comparison as most pie charts only show the percent of the total, without absolute values, a change in percentage does not necessarily indicates a change in absolute value. (3) Data visualization that fails to emphasize clear differences for comparison is ineffective. While adding labels to offer additional context and information is a way out, adjusting the axis with logical measurement scales also works to improve the charts.

Source: “Stop Making These Data Visualization Mistakes in 2019”, Liquid’s Website, https://www.liquidint.com/blog/marketing/data-visualization-mistakes-to-avoid



（14A）香港智慧城市發展：開放數據

政府在 2018 年把開放數據納入「香港智慧城市藍圖」發展計劃，在「資料一線通」網站開放政府各部門的數據，供公眾瀏覽及使用。香港互聯網協會於 2021 年進行市民對開放數據評價和需求的調查，發現幾乎所有參加者都曾查閱過公共數據，最常用的數據為交通資訊及天氣資料。參加者一般從公共服務機構網站得到所需資訊，使用「資料一線通」的比例並不高。針對研究結果，香港互聯網協會認為政府應進一步整合各政府網站的數據發佈方式，並持續完善搜索功能；亦認為可以參考台灣與新加坡的開放數據措施，設立平台讓市民反映意見，以實際需求導向開放數據。開放數據的發展其實並不單只把數據上傳到網上，而是應提高市民的數據素養，加強大眾對開放數據的認知，學習如何分析和運用有關數據。

資料來源：「【新聞稿】「香港開放數據指數」民意調查 顯示市民對搜尋數據過程滿意程度較低 香港互聯網協會建議政府基於市民需求開放數據」，香港開放數據指數網站，https://zh.opendata.isoc.hk/updates/press-release-hong-kong-open-data-index-survey-finds-people-less-satisfied-with-the-process-of-finding-public-data-internet-society-hong-kong-recommends-a-demand-driven-approach-for-the-government-to-release-data

(14B) The Development of Smart City in Hong Kong: Open Data

In 2018, the government incorporated Open Data into the "Smart City Blueprint for Hong Kong", releasing open data of various government departments on the "DATA.GOV.HK" website for the public to browse and use. In 2021, Internet Society Hong Kong (ISOC HK) conducted a survey on people's view and demand on open data. They found that almost all respondents have viewed public data, the most-used data are transportation and weather. Respondents generally get the information they need from the websites of public institutions, and the proportion of using the "DATA.GOV.HK" is not high. In response to survey results, ISOC HK believed that the government should reconcile the data publication standards of various government websites and continue to improve the search function. They also believed that the government can refer to open data measures of Taiwan and Singapore, opening up channels for citizens to voice their opinions, and maintaining data availability with actual demand. The development of open data is not just about uploading data online but should improve citizens' data literacy, enhance their public awareness of open data, and learn how to analyze and use relevant data.

Source: "【Press Release】'Hong Kong Open Data Index' survey finds people less satisfied with the process of finding public data — Internet Society Hong Kong recommends a demand-driven approach for the government to release data", Hong Kong Open Data Index website, https://opendata.isoc.hk/updates/press-release-hong-kong-open-data-index-survey-finds-people-

less-satisfied-with-the-process-of-finding-public-data-internet-society-hong-kong-recommends-a-demand-driven-approach-for-the-government-to-release-data

（15A）辨別假新聞：以俄烏戰爭消息作為事例

俄羅斯與烏克蘭 2022 年 2 月爆發戰爭，網上關於烏克蘭平民死亡的討論一直持續不斷。香港浸會大學事實查核中心近期針對一些備受爭議的俄烏戰爭消息作出事實查核，以下將節錄其中一個報告。有 Facebook 用戶於 3 月發佈一則名為「戲唔係咁演」的影片，影片裡一名男子正躺在黑色的裹屍袋裡吸煙。影片的語境，如烏克蘭國旗，驚恐表情圖片與「專業啲啦」字幕，都意指烏克蘭使用演員扮演俄烏戰爭的死者。浸大事實查核團隊將影片的畫面截圖，以圖搜圖後發現影片其實出自 Instagram 用戶「vasya_ivanov」於 2021 年 3 月 25 日發布的一則限時動態，並收錄在其名為「Никогда Никогда」的動態精選集，發布時間遠早於俄烏戰爭。以「Никогда Никогда」為線索，團隊發現影片為音樂影片《Хаски — Никогда-нибудь (Official Music Video)》的幕後製作花絮，與俄烏戰爭無關。事實查核的技巧很容易掌握，只要多注意多查證，任何人也能夠準確識別假新聞。

資料來源：「【錯誤】影片顯示，俄烏戰爭中烏克蘭「死難者」在裹屍袋中吸煙？」，事實查核，浸大事實查核網站，https://factcheck.hkbu.edu.hk/home/2022/05/10/body-bag_smoking/

（15B）Identifying Fake News: A Case of the Russian-Ukrainian War

The war between Russia and Ukraine broke out in February 2022, and online discussions about the deaths of Ukrainian civilians have continued. HKBU Fact Check has recently released a few fact-checking reports about controversial news of the Russian-Ukrainian war. The following is an excerpt from one of the reports. In March, a Facebook user posted a video, "Poor Performance, " showing a man smoking while lying in a black body bag. The video’s context, such as the Ukrainian flag, the frightened expression image, and the "be professional" subtitle, all imply Ukraine's use of actors to play the war's dead. The HKBU Fact Check team took screenshots of the video and searched for pictures. They found that the video was a “story” posted by Instagram user “vasya_ivanov” on March 25, 2021, and included in his story album, titled “Никогда Никогда”. The story was published much before the Russian-Ukrainian war. Using "Никогда Никогда" as a clue, the team found that the video was a behind-the-scenes of the music video "Хаски — Никогда-нибудь (Official Music Video)", which had nothing to do with the Russian-

Ukrainian war. Fact-checking skills are easy to master. As long as people pay attention and verify, everyone can accurately identify fake news.

Source: “[False] Video shows Ukrainian 'victim' smoking in body bags during the Russian-Ukrainian war?” Fact Check, BU Fact Check Website, https://factcheck.hkbu.edu.hk/home/2022/05/10/body-bag_smoking/

（16A）香港智慧城市發展：簡介

智慧城市，指以創新科技提升城市的管理成效與改善市民的生活質素，增強城市的可持續發展，競爭力，效率及安全。香港政府於 2017 年公佈首份《香港智慧城市藍圖》，藍圖涵蓋六大範疇：出行，生活，環境，市民，政府，經濟，並提出多項措施。不少措施已融入市民生活，如電子支付系統「轉數快」，讓跨銀行轉賬更方便，同時簡化商戶的收款程序；及交通運輸流動應用程式「香港出行易」，讓市民知道不同交通方式的路線，時間與費用，選擇最合適的出行安排。在 2020 年 12 月，政府推出了《香港智慧城市藍圖 2.0》，持續改善智慧城市的策略，並添加了善用創新科技應對疫情的新章節，如針對抗疫而推出的感染風險通知流動應用「安心出行」與家居檢疫系統「居安抗疫」。然而，智慧城市發

展一直面對著不少挑戰，如創新科技產業的人才短缺；數據收集引發的私隱安全隱患；及社會公眾對於智慧城市策略的不認識與低接受度。

資料來源：香港智慧城市藍圖網站，https://www.smartcity.gov.hk/tc.html

(16B) The Development of Smart City in Hong Kong: Introduction

“Smart city” refers to utilizing innovation and technology to enhance the city’s management effectiveness and improve citizens’ quality of life, aiming to enhance the city's sustainable development, competitiveness, efficiency, and safety. The Hong Kong government released the first "Smart City Blueprint for Hong Kong" in 2017. The “blueprint” covers six areas: mobility, living, environment, people, government, and economy, and has proposed various measures. Many measures have already been integrated into people's daily life. For example, the electronic payment system "FPS" makes inter-bank transfers easy and simplifies the merchants’ transaction procedures. The transportation mobile app “HKeMobility” allows citizens to investigate the route, time and cost of different modes of transportation and choose the most suitable travel option. In December 2020, the government launched the "Smart City Blueprint for Hong Kong 2.0" to continuously improve smart city strategies. They have added a new chapter on using innovation and technologies to combat COVID-19, such as the exposure notification mobile app “LeaveHomeSafe” and the home quarantine system “StayHomeSafe.” However, the development of smart cities has been facing plenty of challenges, such as the shortage of talents in the innovation and technology industries, data collection privacy and security issues, and the public's lack of awareness and low acceptance of the smart city strategy.

Source: HKSMART CITY BLUEPRINT PORTAL, https://www.smartcity.gov.hk/index.html

Workshop Moderator Guide

第一部分

（1）你有沒有使用網上商店的習慣？

（2）你覺得 XY 網上超市（虛構）的私隱政策是否合理？
- 收集個人資料用途
- 所收集的個人資料類別
- 向第三方披露的資料

（3）假設你選擇支持此網上商店，你會否採取其他措施保護個人資料？
（其他措施可包括提供假資料，使用 VPN）

（4）根據 XY 網上商店所提供的私隱政策，你作為客戶會否擔心個人資料外洩？

（5）你有沒有信心政府的條例可以保護個人私隱？
- 個人私隱條例

第二部分

（1）在數碼社會中，我們有機會遭遇以下的風險：
    - 網絡欺凌與性騷擾；網絡詐騙；網絡安全隱患（資料外洩，黑客入侵）
- 你有沒有遇過或聽過身邊的人遇上這些風險？面對這些風險，你會如何應對呢？
- 你有沒有曾經幫助親友或身邊有需要的人應對這些風險？你會做些什麼？

(2) 事例：疫情下，網上教學成為主流，進一步加深了基層家庭學童的數碼鴻溝
    - 數碼鴻溝：指一些科技資源、知識與能力不足的弱勢社群，如長者與低收入人士，可能會在數碼化的進程中被忽略，因而失去部分個人自由、權利，甚至身份
- 你認為基層家庭學童有機會面對什麼難題？
- 你認為可以如何解決社會的數碼鴻溝，達致數碼共融？

（3）你認為我們該如何共同締造安全健康的線上環境？

- 社會層面持份者（政府，商業）
    - 法律與政策
    - 教育與培訓
    - 商業
- 個人層面持份者（家長，個人）
    - 家人與照顧者
    - 互聯網使用者

Pre-workshop Material

事前資料

主題 (一)
私隱政策

香港浸會大學

認識《個人資料（私隱）條例》

《個人資料（私隱）條例》於一九九五年通過，並在一九九六年十二月正式生效，是亞洲區內最早全面保障個人資料私隱的法例之一，適用於公私營機構（包括政府），屬科技中立及原則性的法例。條例的六大原則，規定資料使用者：

（1）收集資料的目的要與其職能活動有關，須告知資料當事人收集資料的目的，收集資料應有實際需要，不可過量；
（2）須確保資料準確無誤，保留時間不應超過實際所需；
（3）使用資料只限於收集時述明的目的，除非得到資料當事人的同意；
（4）須保障資料不會未經授權被查閱、處理、刪除、喪失或使用；
（5）須公開其處理資料的政策和方式，並交代其資料的類別和用途。
（6）而資料當事人亦有權要求查閱其個人資料；若發現資料不準確，有權要求更正。

XY 網上超市

收集個人資料的目的及用途

你毋須提供任何個人資料以瀏覽或使用本網頁及移動應用程式。當你參加活動或登記使用我們的服務或網上內容，我們會收集你的個人資料使我們能夠為你提供服務。你可拒絕向我們提供個人資料，但在此情況下，我們可能無法為你提供服務。閣下提供個人資料，即代表閣下同意我們按本私隱政策聲明使用閣下的個人資料。

XY 網上超市從你所收集的個人資料可能會用於以下用途（包括但不限於）：

- 識別你的身分及你開設的任何賬戶
- 為你提供服務
- 核實身分及／或作信貸審查
- 確定及核實你享用商品和服務的折扣及促銷的資格
- 為你處理與服務有關的付款指示或追收欠款
- 直接促銷我們的 服務（詳情見「直接促銷」部份）
- 直接促銷本公司業務夥伴的產品及服務（詳情見「直接促銷」部份）
- 進行市場研究、統計分析及行為分析
- 讓你選擇參與我們為你提供的服務的互動功能，包括識別你的朋友，並與他們溝通和分享你的購物體驗
- 就我們的服務或你感興趣的商品或服務向你推薦
- 處理你的投訴及賬戶查詢，對本公司或任何一方之索償及/或訴訟
- 任何與收集個人資料的原來目的直接有關之其他用途

所收集的個人資料類別

1. 你的個人資料及聯絡資料，如姓名、性別、出生日期、身分證號碼、電話號碼、社交媒體連結、電 郵地址、住址、郵寄地址及/或寄發賬單的地址；
2. 你的商業資料，如公司的名稱及職銜；
3. 你的賬戶資料，如信用卡賬戶號碼；
4. 你與商戶的付款交易詳情；
5. 你的家庭收入及個人興趣；及
6. 你的電腦或移動裝置IP地址、即時位置資料、瀏覽器設定、瀏覽紀錄及／或其他互聯網記錄的資料；及
7. 你的電話簿中包含的電話號碼和電子郵件地址(當你使用我們為你提供的服務的互動功)。當你接受我們的服務我們會通知你。當你提供數據給我們，你確認已經從電話簿的聯繫人取得同意。

個人資料之保密、披露及保安

本公司收集及持有關於你的個人資料均會被保密處理；但若需要按法例的規定或要求而履行法律責任，或為你提供服務，或執行收集個人資料的原來目的或直接有關之目的，本公司有可能會向下述人士披露該等資料（不論其身處香港或香港境外）：

- **具管轄權的法院、執法機關、或其他政府法定或監管部門、機構或組織**
- **本公司之聯繫公司、合作夥伴、參與服務銷售及市場推廣或行政，或提供貨品／服務的賣家或承辦商、代理或其他服務供應商**

我們有可能會公開共享非個人資料，例如包括但不限於與銷售交易、用戶流量、物流及倉庫績效有關之匿名數據和匯總數據，並有可能會與我們的合作夥伴共享，包括但不限於與現有及潛在商業合作夥伴、提供貨品／服務的賣家或承辦商、初創企業家及學界等。

事前資料

主題（二）
探討青年在線

香港浸會大學

背景

- 青年出生在互聯網年代，可說是在網絡世界長大
- 透過網絡，他們培養興趣，學習知識，認識朋友，消閒娛樂及分享創作
- 雖然網絡世界為他們帶來歡樂和益處，但也對他們構成不少風險與危機
- 近年，新冠疫情讓線上學習成為常態，隔離亦讓青年的上網時間增加
- 本主題旨在探討青年在線的有關議題，並探索社會各界如何提供支援，讓青年能安全健康地使用網絡

青年在網絡世界的現況

- 根據救助兒童會近期發表的報告（2022），受訪青年們：
  - 約七成會每週在網上學習新事物
  - 近四成曾參與創作，分享個人想法與發表作品
  - 超過八成會每週通過互聯網（如社交媒體與即時通訊軟件）進行社交活動
    - 約三成會每月結交新朋友，有些（13%）曾與網友親身見面
  - 絕大部分（93%）受訪青少年在社交媒體或遊戲網站上有賬戶

資料來源：「香港兒童在線」，救助兒童會網站，https://savethechildren.org.hk/wp-content/uploads/2022/05/Hong-Kong-Kids-Online-Report-Chinese-Final.pdf?utm_source=Meal&utm_medium=report&utm_campaign=kids-online

青年在網絡世界的現況 - 網絡遊戲

- 網絡遊戲在青年間十分流行，不少青年會每天玩網絡遊戲
  - 他們認為網絡遊戲為他們的心理、學習和社交都帶來正面影響
    - 除了消閒娛樂，遊戲能舒緩他們的學習與生活壓力
    - 在遊戲裡的成功（如勝利或升級）讓他們得到很大的成就感
  - 很多受歡迎的遊戲都能同時讓多人參與
    - 玩家組成團隊，透過互動與合作進行遊戲
    - 玩家與不相識的人組隊的情況十分普遍
    - 遊戲中的交流幫助青年學習團隊合作，鍛煉語言表達與溝通技巧

青年在網絡世界的現況

互聯網幫助青年：

- 提升知識，技術與創造力
- 開拓視野與世界觀
- 發展自尊，進行自我實現
- 認識朋友，發展社交圈子

- **然而，不少青年都發現要在現實與網上生活取得平衡並不容易**

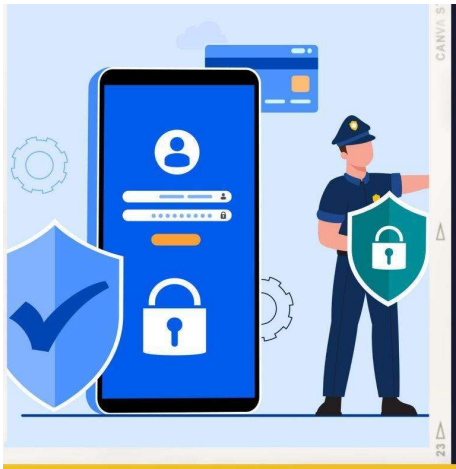
  例如他們會被家人投訴花太多時間上網，不少人更因為限制上網時間的問題與家人發生衝突

資料來源：「香港兒童在線」，救助兒童會網站，https://savethechildren.org.hk/wp-content/uploads/2022/05/Hong-Kong-Kids-Online-Report-Chinese-Final.pdf?utm_source=Meal&utm_medium=report&utm_campaign=kids-online

數碼鴻溝

- 在數碼設備與科技漸漸滲透生活的同時，社會的數碼鴻溝(digital divide)亦逐漸擴大
  - 科技資源、知識與能力不足的弱勢社群，如長者與低收入人士，很容易在數碼化的進程被忽略，他們有機會因此而失去自由、權利，甚至個人身份
- 疫情讓網上教學成為主流，亦進一步加深了基層家庭學童的數碼鴻溝
  - 不具備適當設備與穩定網絡的學童在進行網上學習時困難重重
  - 他們的數碼能力亦有機會因為缺乏適當支援而難以發展
- 如何把提升數碼能力融入主流教育，讓所有青年能公平地享受數碼化教學，是值得深思的重要議題

資料來源：「解構香港數碼鴻溝 資源、知識不足都是成因 羅陸慧英教授：你有得選擇上唔上網，只可選擇存唔存在」，明周，https://www.mpweekly.com/culture/zoom-%E4%B8%8A%E7%B6%B2%E8%AA%B2-%E5%85%AC%E5%85%B1%E6%9C%8D%E5%8B%99%E6%95%B8%E7%A2%BC%E5%8C%96-195047



網絡安全

青年對網絡安全的概念並不陌生，具備一定程度的網絡安全知識，但他們仍然面對著一些安全隱患

- 例如，不少青年都有經營社交媒體，部分社交媒體賬號會有用家清晰展示樣貌的照片，全名，電郵，年齡，出生日期，電話號碼，學校名稱，甚至是住址等個人資料
- 2021年4月，Business Insider的報導稱 5.33 億 Facebook 用戶的個人數據遭外洩。

這些數據被發佈在一個低級別的駭客論壇；數據包括用戶的電話號碼、全名、位置與電子郵件地址。研究人員表示，黑客可以利用這些數據冒充他人並進行欺詐

這並不是Facebook第一次發生用戶數據外洩事件。許多大公司，如LinkedIn、雅虎和Adobe，亦曾發生過類似事件

網絡與真實身份的連結，讓社交媒體用家在網絡的風險增加，甚至有機會被有心人利用

資料來源：“533 million Facebook users' phone numbers and personal data have been leaked online”, Business Insider (https://www.businessinsider.com/stolen-data-of-533-million-facebook-users-leaked-online-2021-4)
「FB全球5億個資外洩》住家地址、信用卡號都可能外流…使用者如何應對？」，商周 (https://www.businessweekly.com.tw/business/blog/3006048)

保障私隱和安全的方法

不少青年都從他們長年累月使用互聯網的經驗中，總結出保障個人私隱和網絡安全的方法，例如：

- 根據需要安裝安全防護軟件，過濾器軟件與使用私隱保密度較高的應用程式，並經常保持更新。
- 刪除網絡痕跡，如瀏覽記錄與Cookies（指為了辨別使用者，而儲存在用戶端，包含個人資料的小型文字檔案）
- 根據個人實際需要，更改網站或應用程式的隱私設置
- 根據需要使用虛擬私人網絡 （VPN）
  - VPN能隱藏用戶的網路協定位址（IP）並加密他們的互聯網流量，防止用戶的身份和私人資料被跟踪和記錄
- 通過重新定位用戶連接的位置，用戶可以訪問受地區限制和政府審查限制的在線內容
- 設置安全度高的密碼（如包含大小階字母，數字與符號）
- 啟用多重要素驗證功能
- 近年，部分涉及個人敏感資料的網站與應用程式（如網上銀行與大學），都會以多重要素驗證為用戶的賬戶增加額外安全性
- 除了輸入密碼，用戶還需使用其他因素，如帶有多重要素驗證應用程序的智能電話，去驗證用戶的身份並批准身份驗證請求

資料來源：“DUO TWO-FACTOR AUTHENTICATION (2FA)”，香港浸會大學資訊科技處網站，(https://ito.hkbu.edu.hk/services/it-security/2fa.html)；「尊重他人私隱，我做得到 (2021年7月)」單張，香港個人資料私隱專員公署網站，(https://www.pcpd.org.hk//english/resources_centre/publications/files/children_privacy_respect.pdf)；“What Is A VPN And How Does It Work?”, Forbes, (https://www.forbes.com/advisor/business/software/what-is-a-vpn-and-how-does-it-work/)

網絡資訊

網絡資訊常見現象：信息不完整，質量參差，可信度成疑，立場單一，相反意見被屏蔽

- 若不加注意，人們就有可能陷入偏聽偏信，難以掌握資訊的真確性
- 如何分辨，處理與查核網絡資訊成為了青年需要把握的必要技能

技能一：事實查核

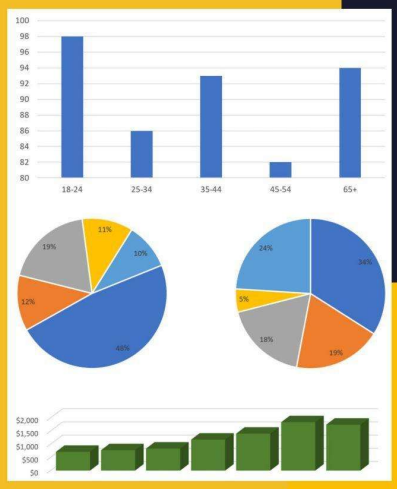
- 事實查核(Fact-checking)能幫助人們破解假新聞，避免以訛傳訛的情況出現
  - 事實查核可分為4個W
    - 第一個與第二個W（When & Where）的重點是時間與地點。善用以圖搜圖的功能，尋找圖片或影片的真正出處，從而還原情境脈絡，避免被錯置或修改過的資訊矇騙
    - 第三個與第四個W（Who & Why）的重點是來源與原因。透過評估上傳者或創作者的立場、信用與動機等來分析新聞的可信度，識別意在攻擊、誤導、圖利與宣傳的信息
  - 包含著極端正面或負面消息的帖文會較易出現虛假成份。
  - 若有需要，也可以訪問坊間的事實核查平台，如事實查核實驗室(Factcheck Lab)與香港浸會大學事實查核中心(BU Fact Check)。

事實查核：事例

俄羅斯與烏克蘭於2022年2月爆發戰爭，網上關於烏克蘭平民死亡的討論一直持續不斷。香港浸會大學事實查核中心曾針對一些備受爭議的俄烏戰爭消息作出事實查核

- 有Facebook用戶於3月發佈一則名為「戲唔係咁演」的影片，影片裡一名男子正躺在黑色的裹屍袋裡吸煙
- 影片的語境，如烏克蘭國旗，驚恐表情圖片與「專業啲啦」字幕，都意指烏克蘭使用演員扮演俄烏戰爭的死者
- 浸大事實查核團隊將影片的畫面截圖，以圖搜圖後發現影片其實出自Instagram用戶「vasya_ivanov」於2021年3月25日發布的一則限時動態，並收錄在其名為「Никогда Никогда」的動態精選集，發布時間遠早於俄烏戰爭
- 以「Никогда Никогда」為線索，團隊發現影片為音樂影片《Хаски — Никогда-нибудь (Official Music Video)》的幕後製作花絮，與俄烏戰爭無關

資料來源：「【錯誤】影片顯示，俄烏戰爭中烏克蘭「死難者」在裹屍袋中吸煙？」，事實查核，浸大事實查核網站，https://factcheck.hkbu.edu.hk/home/2022/05/10/body-bag_smoking/



技能二：數據解讀

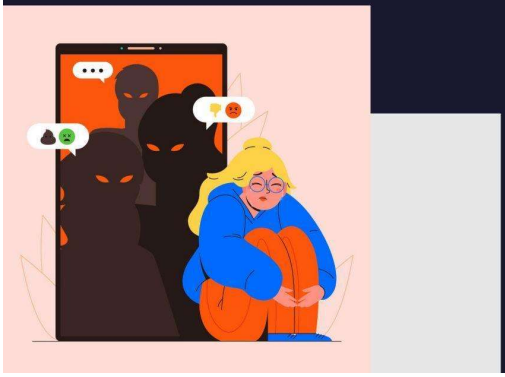
- 除了事實查核，青年也需要具備理解數據的基本能力
  - 數據可視化指以圖表、圖形、地圖等視覺形式呈現數據，可視化能讓數據模式與趨勢的呈現變得更容易，從而幫助人們了解數據所表達的含義。
  - 糟糕的數據可視化傳達具誤導性的數據模式、趨勢和相關性，可能會導致錯誤解讀。

- **(1) 尺度具誤導性的軸會形成有重大差異和變化的錯誤印象，標準化軸的尺度從零開始，能有助避免小尺度造成的“間隙錯覺”。**
- **(2) 圓餅圖不適合用作標準比較，例如月度或年度比較，這基於大多數圓餅圖只顯示總數的百分比，不提供絕對值，百分比的變化不一定表示絕對值的變化。**
- **(3) 無法突顯差異作比較的數據可視化是無效的，添加標籤來提供額外的上下文和信息是一種方法，根據邏輯比例調整軸也能有助於改進圖表。**

資料來源：“Stop Making These Data Visualization Mistakes in 2019”, Liquid's Website, https://www.liquidint.com/blog/marketing/data-visualization-mistakes-to-avoid

網絡欺凌

網絡欺凌泛指欺凌者在網絡（如社交媒體，即時通訊或遊戲等平台）作出令欺凌對象困擾的行為



包括

- 發送騷擾、羞辱、批評及恐嚇的訊息或圖像給欺凌對象；進行網絡公審以取笑或批評欺凌對象；及在社交平台上散播欺凌對象的個人資料（俗稱「起底」）
- 雖然欺凌者可能只因一時氣憤，貪玩或無聊而作出網絡欺凌行為，但這不但會對欺凌對象的身心靈健康造成極大影響，還可能構成刑事罪行
- 不少在現實世界發生的欺凌都會延伸到網絡世界，即使青少年逗留在安全的環境（如家裡），網絡會讓欺凌無時無刻，無處不在地持續進行

資料來源：數碼新世代，向網絡欺凌 說「不」 (2022年2月) 單張，香港個人資料私隱專員公署網站，(https://www.pcpd.org.hk//tc_chi/resources_centre/publications/files/children_privacy_cyberbullying.pdf)

網絡性騷擾

除了網絡欺凌，網絡性騷擾亦是青年在網絡有機會遇到的問題

- 網絡常見的性騷擾形式包括收到令人反感的性訊息，遭要求提供裸體自拍照及作出性誘惑。
- 網絡閃現指通過通訊軟件強行傳遞裸體性器官照片給他人的行為
  - 網絡閃現會讓受害者感到不安，羞辱，冒犯與威脅
  - 跟據報告（Save the Children, 2022），近兩成的受訪青年曾非自願地收到含有裸體及性圖片的網上訊息
  - 亦有部分青年曾向朋輩發送自己的裸體或性器官照片
    - 自行製作色情資訊的風險很大，這些照片有機會留存在網絡世界，並被有心人利用

報告亦指接近四成的受訪青年曾在網上接觸過色情廣告

- 這些廣告讓青年在心智未成熟時就非自願地接觸到色情資訊
- 過早接觸這些資訊會讓他們感到尷尬與不安，甚至對心理造成影響

網絡詐騙

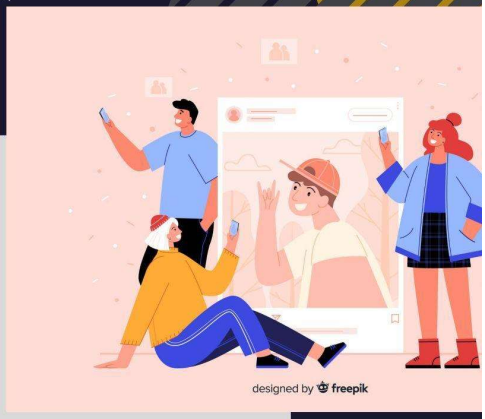
網絡世界危機處處，騙徒行騙手法層出不窮，大部分網民都曾遇上或聽說過網絡詐騙的例子：

- 騙徒盜取賬戶冒認事主親友，以不同的理由（如購買點數卡或借錢）欺騙事主向其匯款。
- 騙徒在購物網站販賣虛假商品，在事主完成付款後，騙徒就會消失或向事主發送名不副實的產品。
- 騙徒佯稱機構（如銀行或科技公司）向事主發送訊息或電郵，誘騙事主打開連結，登入賬戶，更改密碼或提供個人驗證碼，藉機盜取他們的戶口資料。
- 騙徒在交友軟件以「美人計」哄騙事主在非法網站投資，高價購買不需要的產品與服務及傳送裸照從而勒索。

資料來源：「網上詐騙」，錢家有道，消委會網站
(https://www.ifec.org.hk/web/tc/moneyessentials/scams/scam-websites.page)

網絡交友

- 網絡交友讓青年能夠接觸生活圈子外志同道合的人，這種聯繫對不少人而言普遍且必要
- 然而在網絡認識的人不一定表裡如一，他們部分動機不純，會利用青年進行剝削
  - 如帶著歪念的「怪叔叔」藉機接近未成年少女，誘惑她們提供裸照，出售校服或發生性行為）
- 感到孤獨與被父母或照顧者忽略的青年，會比朋輩較傾向在網上交友，尋求陪伴與接納，他們會更容易沉迷互聯網，亦有較大機會在網絡被利用或受到傷害



資料來源：「香港兒童在線」，救助兒童會網站，https://savethechildren.org.hk/wp-content/uploads/2022/05/Hong-Kong-Kids-Online-Report-Chinese-Final.pdf?utm_source=Meal&utm_medium=report&utm_campaign=kids-online

網絡公民參與

- 新媒體的崛起創造了公民參與的新方式，網絡和社交媒體在青年的公民意識形成和社會事務參與均扮演著重要角色
  - 網絡公民參與的形式眾多，除了獲取社會事務資訊，還包括評論、討論、加入群組、聯署和眾籌等
  - 近年，社交媒體的地區群組與文化關注組成為了青年了解社區與文化的常用平台，與網友互動、參與活動和討論社會事務都有助培養了社會和公民意識

資料來源:「【網絡公民參與】教育脫節　市民難辨海量信息」，香港01，https://www.hk01.com/sns/article/418223；「【網絡公民參與】媒體素養教育　校園不是唯一擔責」，香港01，https://www.hk01.com/sns/article/418302

Workshop Discussion Material

第一部分

(1) 個人私隱條例

1. 收集資料的目的要與其職能活動有關,須告知資料當事人收集資料的目的,收集資料應有實際需要,不可過量;

2. 須確保資料準確無誤,保留時間不應超過實際所需;

3. 使用資料只限於收集時述明的目的,除非得到資料當事人的同意;

4. 須保障資料不會未經授權被查閱、處理、刪除、喪失或使用;

5. 須公開其處理資料的政策和方式,並交代其資料的類別和用途。

6. 而資料當事人亦有權要求查閱其個人資料;若發現資料不準確,有權要求更正。

(2) 所收集的個人資料類別

1. 姓名、性別、出生日期、身分證號碼、電話號碼、社交媒體連結、電郵地址、郵寄地址

2. 商業資料,如公司的名稱及職銜

3. 賬戶資料,如信用卡賬戶號碼

4. 付款交易詳情

5. 家庭收入及個人興趣

6. 電腦或移動裝置 IP 地址、即時位置資料、瀏覽器設定、瀏覽紀錄

7. 電話簿中包含的電話號碼和電子郵件地址

(3) 個人資料之保密、披露及保安

我們有可能會公開共享非個人資料,例如包括但不限於與銷售交易、用戶流量、物流及倉庫績效有關之匿名數據和匯總數據,並有可能會與我們的合作夥伴共享,包括但不限於與現有及潛在商業合作夥伴、提供貨品／服務的賣家或承辦商、初創企業家及學界等。

本公司有可能會向下述人士披露該等資料(不論其身處香港或香港境外):

• 具管轄權的法院、執法機關、或其他政府法定或監管部門、機構或組織

• 本公司之聯繫公司、合作夥伴、參與服務銷售及市場推廣或行政

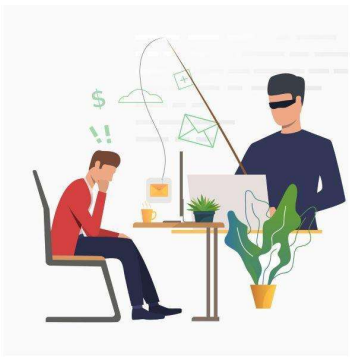
(4) 討論問題

你會否使用 XY 網上超市的服務?

第二部分

（1）在數碼社會中，我們有機會遭遇以下的風險：

網絡欺凌與性騷擾



網絡詐騙

網絡安全隱患
（資料外洩，黑客入侵）

（2）疫情下，網上教學成為主流，進一步加深了基層家庭學童的數碼鴻溝

（3）我們該如何共同締造安全健康的線上環境？

Focus Group Question Guide

-首先，請各位簡單介紹自己（希望我們如何稱呼你）。

-你與資訊安全的關係是什麼？（職業與資訊安全有關或修讀與資訊安全有關的科目）

工作坊

-你為什麼會參加工作坊？你參與工作坊的體驗是怎樣的？

-你認為工作坊的討論內容有趣嗎？對你有沒有用？

-你覺得工作坊有沒有提供足夠空間讓你與其他人交流你的睇法？

-你覺得工作坊有沒有提升你對數據素養的認識或興趣或關注？

-你覺得事前參考資料的資訊性足夠嗎？

-你認為參加工作坊有沒有令你接收到一些與你觀點不同的看法？當中有沒有令你印象深刻的看法？

-你認為參加工作坊有沒有改變你對數據科技的看法？

-你認為參加工作坊有沒有改變你使用數據科技的習慣？

WhatsApp Group

-你為什麼會加入 WhatsApp 群組？你參與 WhatsApp 群組的體驗是怎樣的？

-你認為群組裡的資訊有趣嗎？對你有沒有用？

-你覺得加入群組有沒有提升你對數據素養的認識或興趣或關注？

-你認為加入群組有沒有令你接收到一些與你觀點不同的看法？當中有沒有令你印象深刻的看法？

-你認為加入群組有沒有改變你對數據科技的看法？

-你認為加入群組有沒有改變你使用數據科技的習慣？

Follow Up

-你會如何形容你日常觀察到的網上環境？你理想的網上環境是怎樣的？你會做什麼去達致你理想的網上環境？

-社會的數碼化進程進一步加快，你覺得有什麼好處？你有沒有觀察到一些問題？

-有人認為公眾服務在數碼化的同時都需要提供適當的線下選擇方便有需要的人士，你覺得香港的線下選擇多嗎？

-有人認為：在數碼社會，保護個人私隱是不可能的，你有什麼看法？