Tutorial 5: Privacy and Surveillance

Part 1

Q1) Despite the difference in understanding offline and online privacy, the values that privacy brings to these two spheres should remain the same. Privacy, in both offline and online environments, remains essential in one's life. Furthermore, the dignitary values that privacy brings to humans regardless of the background that person is in remain crucial.

i) Explain how online privacy and offline privacy differ from each other. (PIC: Yih Feng)

Online privacy	Offline privacy
The definition of online privacy is the **level of privacy protection an individual has while connected to the Internet.** It covers the amount of online security available for personal and financial data, communications, and preferences. The complex issue of computer privacy covers the way your personal information is used, collected, shared, and stored on your personal devices and while on the Internet. Personal information about your habits, shopping, and location can be collected from your phone, GPS, and other devices—and eventually shared with third parties. Internet and device users have the right to ask how information will be used and to review online privacy policies. Example:e-commerce -Name -Tel.no -Address -Card.no source:https://www.winston.com/en/legal-glossary/online-privacy.html	The offline privacy represents the personal information that is not captured, processed or stored through online means (through internet connection) in or on any platform. An example is the Identity Card(IC) or student id which contains information of a particular person. Example: Name, Age, Height, Weight, these are the offline data privacy which mostly has to be filled in a form. E.g Form, Questionnaire,interview Job application Bank application

ii) Comparing online and offline privacy, which type of privacy is rather challenging? Justify your reason. (PIC: T'nsam)

In my opinion, **online privacy** is more challenging than offline privacy. This is because information can be **spread more quickly through online** and it is easier to be stolen by others when compared to information stored offline. This indicates that it is more challenging to protect online privacy as there is a **high possibility to be accessed by unauthorised users**. For instance, Google uses **cookies** to **remember your preferred language**, to make the ads you see more relevant to you, to count how many visitors they receive to a page, to help you sign up for their services.

References: How Google uses cookies – Privacy & Terms – Google

Q2) In Malaysia context, briefly explain one legal act in place to protect users' privacy. (PIC: Kai Yuan)

Personal Data Protection Act 2010 (“PDPA”), which deals with personal data and focuses on regulating the processing of ‘personal data’ in commercial transactions.

The information relates directly or indirectly to a data subject, who is identified or identifiable from that information or from other information in the possession of a data user, including any sensitive personal data and expression of opinion about the data subject

In simpler terms, the aim of the PDPA is to safeguard the personal data of individuals that are collected, stored and used from being abused by the person or persons who have control over the personal data or authorises the processing of such personal data. This wide definition covers details such as name, address, contact details and your national registration identity card. It also includes sensitive personal data such as the physical or mental health condition of an individual, their political opinions and even religious beliefs.

General Principle
The General Principle prohibits a data user from processing an individual's personal data without their consent. The Personal Data Protection Regulations (Regulations) stipulate that consent must be "recorded" and "maintained" which suggests that express consent is required.

Security Principle
The PDPA imposes obligations on the data user to take steps to protect the personal data during its processing from any loss, misuse, modification, unauthorised or accidental access or disclosure, alteration or destruction.

Retention Principle
Personal data is not to be retained longer than is necessary for the fulfilment of the purpose for which it is processed. A duty is also imposed on the data user to take reasonable steps to ensure that all personal data is destroyed or permanently deleted if it is no longer required for the purpose for which it was processed.

Notice & Choice Principle
The PDPA requires a data user to inform the individual by written notice, in both the national and English languages, of certain matters including the fact that the personal data of the individual is being processed and a description of the data; the purposes for which the personal data is being collected and further processed; any information available to the data user as to the source of that personal data; the individual's right to request access to and correction of the personal data and contact particulars of the data user in the event of any inquiries or complaints; the class of third parties to whom the data is or may be disclosed; the choices and means offered to the individual to limit the processing of the data and whether it is obligatory or voluntary for the individual to supply data, and if obligatory, the consequences of not doing so.

Disclosure Principle
This Principle prohibits the disclosure, without the individual's consent, of personal data for any purpose other than that for which the data was disclosed at the time of collection, or a purpose directly related to it; and to any party other than a third party of the class notified to

the data user.

Data Integrity Principle
The data user has to take reasonable steps to ensure that the personal data is accurate, complete, not misleading and kept-up-to-date, having regard to the purpose (and any directly related purpose) for which it was collected and processed.

Access Principle
The PDPA gives the individual the right to access and correct his own data where it is inaccurate, incomplete, misleading or outdated. The PDPA provides grounds on which the data user may refuse to comply with a data access or data correction request by the individual.

Put in the remaining 4 principles as well
…..
……
…….

Or refer for more from lecture notes, chapter 7 slide 72 onwards

Q3) Do online research, discuss the following:

(i) How our privacy will be exposed when we perform online activities. (PIC: Jun Wai)

Internet Service Provider (ISP): In order to have access to the Internet, we need to register with an ISP and then we are allotted with an IP. The ISP could **track our geolocation based on the IP** given to us whenever we open a website.

Surfing the Internet: Whenever we are surfing the internet, the online companies can **trace our surfing behaviour** and **advertise you with their relevant products based on your browsing activities**. For example, if we frequently search for academic relevant topics such as Data Science, you will see advertisements about Data Science courses offered by certain parties in your social media and so on.

Search Engines: search engines are used for searching information online. At the same time, the search engine would **record all your searched items in order to know you better and gather more information about you for business purposes such as advertising**.

Cookies: cookies are referred to as the data collected on the website that you have visited and it is stored in your hard drive. For example, when we purchase a product online by providing them our name and bank details, these details could be stored as cookies but legitimate websites usually **store cookies in order for tracking the result of their advertising**. There is another kind of cookie which is the third-party cookies. These cookies are used to share data about online users with online advertising companies.

Terms and privacy policy (e.g. privacy disclaimer): Whenever we register as a user in a mobile app platform, we must **read properly on the terms and conditions** to ensure the data collected by the mobile app are safe and secure and **do not expose our personal details for third partie**s.

Online Activities That Expose Your Privacy | SafeSpace

(ii) Privacy-preserving techniques/methods that can be used for big data analytics. (No technical details required. Discuss the general concepts/approaches of the methods for preserving privacy). (PIC: Mun Jun)

Privacy Preservation in Big Data

Cryptography

- A traditional data privacy preservation method
- Plaintext is converted into cipher text using various encryption schemes.
- Various methods based on this scheme like public key cryptography, digital signatures etc.

Data Anonymization

- Process of changing data that will be used or published in a way that prevents the identification of key information
- Referred as Data de-identification
- Key pieces of confidential data are obscured in a way that maintains data privacy
- Hiding identifier attributes (attributes that uniquely identify individuals) like full name, license number, voter id etc

K-Anonymity

- A dataset is called k-anonymized if for any tuple with given attributes in the dataset there are at least k-1 other records that match those attributes.
- K-anonymity can be achieved by using suppression and generalization.
- In suppression, quasi identifiers are replaced or obscured by some constant values like 0,* etc.
- In generalization, quasi identifiers are replaced by more general values from levels up the hierarchy

L-Diversity

- L-diversity technique of data anonymization tries to bring diversity in the sensitive attribute of data.
- It ensures that each equivalence class of quasi identifiers has atleast L different values of sensitive attribute
- Improvement of K-Anonymity (to overcome to weakness of K-anonymity: sensitive value in an equivalence class lack of diversity)

T-Closeness

- An equivalence class is said to have t-closeness if the distance between the distribution of a sensitive attribute in this class and the distribution of the attribute in the whole table is no more than a threshold t.
- A table is said to have t-closeness if all equivalence classes have t-closeness.
- The main advantage of t-closeness is that it prevents attribute disclosure

Notice and Consent

- The most common privacy preservation method for web services is notice and consent.
- Every time an individual accesses a new application or service, a notice stating privacy concerns is displayed.
- The consumer needs to consent the notice before using the service.
- This method empowers an individual to ensure his privacy rights. It puts the burden of privacy preservation on the individual

Differential Privacy

- Differential Privacy is a method enabling analysts to extract useful answers from databases containing personal information while offering strong individual privacy protections.
- It aims to minimize the chances of individual identification while querying the data.

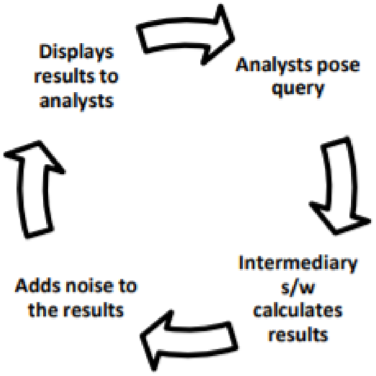

Displays results to analysts
Analysts pose query
Intermediary s/w calculates results
Adds noise to the results

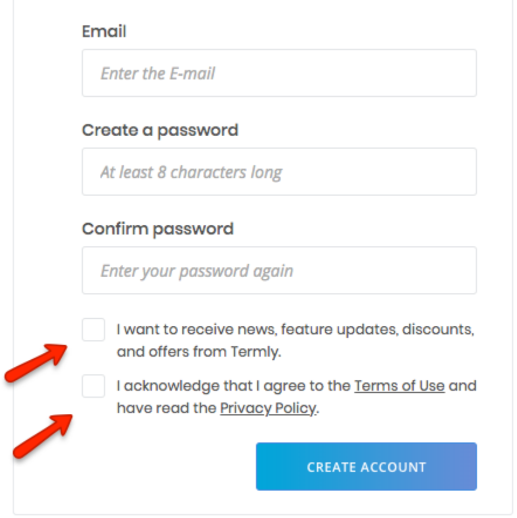
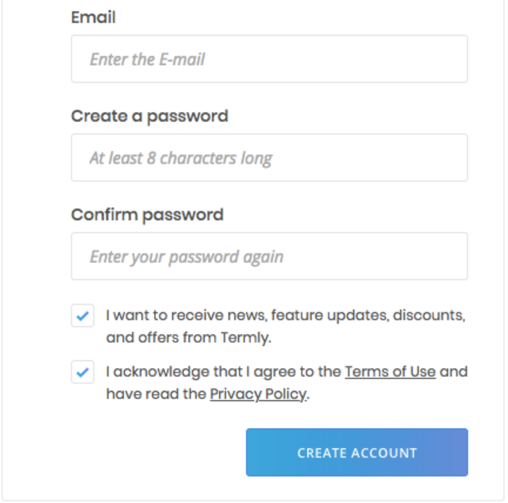
Fig. 1. Differential privacy process

Part 2

Q1) A debate has raged in the industry for many years about the use of "opt-in" or "opt-out" options during a sign up process in a web form. At the heart of this debate is user consent, i.e. user choice to reveal information. These options allow the company to get user agreement to use the person's information, such as their name, email address and so on, to then contact them, usually for marketing purposes.

i) Differentiate between Opt-In and Opt-Out. (PIC: Yee Hui)

Opt-In	Opt-Out

Opt-In means that a user will take affirmative action to offer their consent before the company uses their personal information such as their name and email address for marketing purposes. For example, the user can click the checkbox to agree their personal information to be used by the company.	Opt-Out means that the organization can share the users' information until the users explicitly forbid it. This means that the user can take action to withdraw their consent. For example, the users can uncheck a marked box if they do not agree their information to be used by the company.
Example of Opt-In:  Email Enter the E-mail Create a password At least 8 characters long Confirm password Enter your password again □ I want to receive news, feature updates, discounts, and offers from Termly. □ I acknowledge that I agree to the Terms of Use and have read the Privacy Policy. CREATE ACCOUNT In Opt-In, both checkboxes are unchecked before the users click checked.	Example of Opt-Out:  Email Enter the E-mail Create a password At least 8 characters long Confirm password Enter your password again ☑ I want to receive news, feature updates, discounts, and offers from Termly. ☑ I acknowledge that I agree to the Terms of Use and have read the Privacy Policy. CREATE ACCOUNT In Opt-Out, both checkboxes are checked before the users click unchecked.

Example of opt in

□ I would like to subscribe to ABC's Company email newsletter.

Example of Opt out

☑ I would like to subscribe to ~~company~~'s ABC's Company email newsletter.

□ I would like to unsubscribe to ABC's Company email newsletter

ii) Identify which option is preferred by privacy advocates. (PIC: Pei Xuan)

Privacy advocate --- a person/group who publicly supports or recommends the privacy right protection / promote the policy to protect privacy right

Privacy advocates will prefer **Opt-In**. Opt-In means that consumer must explicitly give permission before the organization can share info. The company will give the permission for the users whether they allow the company to use their personal information for business purposes. If the user don't want to let the company use their information, they can deny the permission. Therefore, consumer can freely choose according to their opinion. In conclusion, opt-in can **draw the attention of users to the existing privacy policy setting and ask them to take an action** such as checking the checkbox to note consent in order to ensure that the **user is aware of the right to voluntarily provide the information** and **well aware who is collecting it and to serve what objective.**

Q2) Why is information about individuals so important to organizations? Give examples of the uses of personal information by **private** and **public organizations**. (PIC: Xin Yi)

Personal information for the business such as customer details. For example, the organization can use the customer information to do the analysis for their daily business such as the grocery company, the company can be based on the customer purchasing behaviour to understand how the decision to buy was made and how the customer hunted for the product. All of this

information helps their companies and business managers to know the reasons behind the purchase or rejection of a product or service by the customer.
Besides, the company can have some promotions based on the customer purchasing behaviour in order to produce more products and increase the revenue for the company. Not only this, the company also can promote their new product to the customer by selling the new product combined with the best selling product as a package to be sold together. Therefore, the information about the individual for the organization is important because it can be considered a valuable asset for their company. Moreover, in the healthcare industry like the hospital or clinic, the patient's medical record is important for the doctor to keep track of the history in the system. The medical record needs to be kept accurate and updated in order to ensure that the medical treatment is administered to the correct patient.

Examples of private organization: Grocery company (e.g. Wal-mart, Lotus, AEON….)
Usage of personal data from such private organization: customer purchasing behaviour, date of birth
Objective:
1. Companies can base on the customer purchasing behaviour to do the business analysis.
2. Companies will have the promotion to the customer/member when it is the birthdate month such as giving the cash voucher, discount and other.

Examples of public organization: Government hospital (Hospital Kuala Lumpur)
Usage of personal data from such public organization: Patient's medical record
Objective: To keep track the history report for the patient such as treatment record, allergic

Q3) XYZ company is planning to provide a new online service to users. The online service will provide a user sign-up page for users to register for the online service provided by XYZ company. After the sign-up page, users will be required to input their data such as full name, age, email address, etc. Discuss THREE (3) considerations when collecting data from users for the online service to avoid legal issues. (PIC: Jun Xian)

Considerations:
Consent: user data should not be shared or processed without the users' consent.
E.g consider adopting either opt-in or opt-out policy. Opt-in: requires the user to explicitly agree with the terms and conditions whereas Opt-out: requires the user to explicitly disagree with the terms and condition because by default the t&c is that the the user agrees.
Security: collected data should be kept secure.

- the data processor must make sufficient guarantees in respect of the technical and organizational security measures governing the processing to be carried out
  E.g. via encryption method to prevent eavesdropping, or data stealing

Disclosure: users should be informed about who is collecting their data.

- The data you have previously collected for a purpose other than for which it was initially collected, will need the data provider's consent again in order to release their

data to a third party that is in a different class.
E.g. publish privacy disclaimer and require the user to read through the disclaimer by ensuring the user scroll through the disclaimer until the end or tick the check box that states the user has read the disclaimer before continuing the process.

Q4) Identify the current weaknesses of Malaysia's Personal Data Protection Act (PDPA). (PIC: Kah Wei)

Personal Data Protection Act 2010 (PDPA) only applies to personal data handling during the commercial transaction and **does not affect the data stored outside Malaysia** and it for example in American Social Media like Facebook, Twitter are not restricted under this PDPA.-->**subject to territory/geographical region/countries basis--**

This Act also left untouched for about a decade while the technology and size of data are rapidly growing and **it makes the businesses and also consumers do not know who to reach out** in the event of major data breach. →

The act should be **reviewed from time to time** or even 5 years a time because there will be new technology that is emerging and technology is reviewed from time to time and this act could not catch up with the pace of the technology.

There is **not enough public awareness of this act**. For example, in just a case of data breach or data leakage, the consumer or business **would not know who to reach out** because they are not really aware of this act.

No early precautions were imposed based on the act to prevent this event of data breach to occur. For example, there should be some rules and regulations or policies that should be followed by the respective parties to prevent the event from going to occur. It could be **too late for the business to be aware of this act after the incidents have happened** and it could cause permanent money loss to the business.