CSE 4201

Ethical Issues and Professional Practice in Computing

University of Guyana
Lecturer: Muriana McPherson
Semester 2: 2018/2019

Privacy and Cybertechnology

- Privacy concerns now affect many aspects of our day-to-day lives – from commerce to healthcare to work.
- categories include:
- consumer privacy,
- medical/healthcare privacy,
- employee/workplace privacy.

Privacy and Cybertechnology

- Are any privacy issues unique to cybertechnology?
- Privacy concerns have been exacerbated by cybertechnology in at least four ways - by the:

1. *amount* of personal information that can now be collected;
2. *speed* at which personal information can now be transferred and exchanged;
3. *duration* of time in which personal information can now be retained;
4. *kind* of personal information (such as transactional information) that can be acquired.

What is Privacy?

- Privacy is sometimes viewed in terms of something that can be *diminished* (i.e., as a repository of personal information that can be eroded gradually) or *lost* altogether.
- Privacy is sometimes also construed in terms of the metaphor of a (spatial) zone that can be *intruded upon* or *invaded*.
- Privacy is also sometimes analyzed in terms of concerns affecting the confidentiality of information, which can be *breached* or *violated*.

Classic Theories of Privacy

- Traditional (or classic) privacy theories have tended to view privacy in connection with notions such as:
- non-intrusion (into one's space),
- non-interference (with one's decisions),
- having control over/restricting access to one's personal information.

Non-intrusion Theories of Privacy

- Non-intrusion theories view privacy as either:
  - being let alone,
  - being free from government intrusion (into one’s physical space).
- This view is also sometimes referred to as *accessibility privacy* (DeCew, 1997).

Non-interference Theories of Privacy

- Non-interference theories view privacy in terms of freedom from interference in making decisions.
- This perspective emerged in the 1960s, following the *Griswold v. Connecticut* (U.S. Supreme Court) case in 1965.
- This view of privacy is also sometimes referred to as *decisional privacy*.

The Control and Limited Access Theories of Informational Privacy

- *Informational privacy* is concerned with protecting personal information in computer databases.
- Most people wish to have some *control* over their personal information.
- In some cases, “privacy zones” have been set up either to restrict or limit *access* to one’s personal data.

Three Views of Privacy

Accessibility Privacy *Non-Intrusion*	Privacy is defined in terms of one's physically "being let alone," or freedom from intrusion into one's physical space.
Decisional Privacy *Non-Interference*	Privacy is defined in terms of freedom from interference in one's choices and decisions.
Informational Privacy	Privacy is defined as control over the flow of one's personal information, including the transfer and exchange of that information.

Moor’s Comprehensive Theory of Privacy

- According to Moor:

  “an individual has privacy in a *situation* if in that particular situation the individual is *protected from intrusion*, *interference*, *and information access* by others.”

Moor’s Theory of Privacy

- For Moor, a situation can be an “activity,” a “relationship,” or the “storage and access of information” in a computer or on the Internet.
- Moor also distinguishes between “naturally private” and “normatively private” situations required for having:

a) natural privacy (in a descriptive sense);

b) a right to privacy (in a normative sense).

Descriptively Private vs. Normatively Private Situations

- Review Scenario 5-1 (in the textbook), where Tom walks into the computer lab (when no one else is around) and sees Mary in the lab.
  - In this natural/descriptively private situation, Mary’s privacy is lost but not violated.
- Review Scenario 5-2, where Tom peeps through the keyhole of Mary’s apartment door and sees Mary typing at her computer.
  - In this normatively private situation, Mary’s privacy is not only lost but is also violated.

Nissenabum's Theory of Privacy as "Contextual Integrity"

- Helen Nissenbaum's "privacy as contextual integrity" framework requires that the processes used in gathering and disseminating information are

a) "appropriate to a particular context"

b) comply with norms that govern the flow of personal information in a given context.

Nissenbaum’s Theory

- *Norms of appropriateness* determine whether a given type of personal information is either appropriate or inappropriate to divulge within a particular *context*.
- *Norms of distribution* restrict or limit the flow of information within and across *contexts*.
- When either norm is “breached,” a violation of privacy occurs.
- Conversely, the *contextual integrity* of the flow of personal information is maintained when both kinds of norms are “respected”

Can Privacy Be Preserved in the Digital Era?

- Scott McNealy, a former CEO of Sun Microsystems, proclaimed his now famous remark to a group of reporters: *You have zero privacy anyway. Get over it*.
- Others authors have expressed concerns about the “death of privacy.”
- But some believe that not all has yet been lost in the battle over privacy.
  - For example, some privacy advocates staunchly believe that we should be vigilant about retaining and safeguarding what little privacy we may still have.

Is Protecting Personal Privacy Still Considered an Important Goal?

- Can the current privacy debate be better understood in terms of differences that reflect *generational* attitudes?
- For many “Millennials,” who are now college-aged, privacy does not always seem to be of paramount importance.
  - Consider, for example, that many Millennials seem eager to share their personal information widely on social networking services such as Facebook.
- But for many older persons, including Baby Boomers, privacy is still highly valued.
- So the relative importance of privacy may vary considerably among the generations.

What Kind of Value is Privacy?

- Three distinct questions can be distinguished with respect to privacy as a *value*:

1. Is privacy an *intrinsic* value, or is it an *instrumental* value?
2. Is privacy *universally* valued, or is it valued mainly in Western industrialized societies (where greater importance is placed on the individual than on the broader community?)
3. Is privacy an important *social* value (as well as an individual value)?

Is Privacy an Intrinsic Value or an Instrumental Value?

- Is privacy something that is valued for its own sake?
  - In other words, is it an *intrinsic value*?
- Or, is privacy valued as a means to some further end?
  - Is it merely an *instrumental value*?
- Charles Fried (1990) privacy seems to be more than merely an instrumental value because it is *necessary* (rather than merely *contingent*) for achieving important human ends such as trust and friendship.
- Moor believes that privacy can be viewed as an expression of a "core value" – viz., security, which is essential for "human flourishing."

Privacy as a Universal Value

- Privacy has at least some importance in all societies, but it is not valued the same in all cultures.
- For example, privacy tends to be less valued in many non-Western nations, as well as in many rural societies in Western nations.
- Privacy also tends to be less valued in some democratic societies where national security and safety are considered more important than individual privacy (e.g., as in Israel).

Privacy as an Important Social Value

- Priscilla Regan notes that we tend to underestimate the importance of privacy as an important *social value* (as well as an individual value).
- Regan also believes that if we frame the privacy debate in terms of privacy as a social value (essential for democracy), as opposed to an individual good, the importance of privacy is better understood.

Cybertechology-related Techniques that Threaten Privacy

- We examine two distinct cyber-related techniques that threaten privacy:

1) *data-gathering* techniques used to collect and record personal information, often without the knowledge and consent of users.
2) *data analysis* techniques, including data mining, used to manipulate large data sets of personal information to discover patterns and generate consumer profiles (also typically without the knowledge and consent of users).

Cybertechnology Techniques Used to *Gather* Personal Data

- Personal data has been gathered at least since Roman times (census data).
- Roger Clarke uses the term *dataveillance* to capture two techniques made possible by cybertechnology:

a) surveillance (data-monitoring),

b) data-recording.

Internet Cookies as a Surveillance Technique

- Cookies technology enables Web site owners to collect data about those who access their sites.
- With cookies, information about one’s online browsing preferences can be “captured” whenever a person visits a Web site.
- The data recorded via cookies is stored on a file placed on the hard drive of the user's computer system.
- The information can then be retrieved from the user's system and resubmitted to a Web site the next time the user accesses that site.
- The exchange of data typically occurs without a user's knowledge and consent.

Can the Use of Cookies be Defended?

- Many proprietors of Web sites that use cookies maintain that they are performing a service for repeat users of their sites by customizing a user's means of information retrieval.
- For example, some point out that, because of cookies, they are able to provide a user with a list of preferences for future visits to that Web site.

Arguments Against Using Cookies

- Some privacy advocates argue that activities involving the monitoring and recording an individual's activities while visiting a Web site violates privacy.
- Some also worry that information gathered about a user via cookies can eventually be acquired by or sold to online advertising agencies.

RFID Technology as a Surveillance Technique

- RFID (Radio Frequency IDentification) consists of a *tag* (microchip) and a *reader*:
- The tag has an *electronic circuit*, which stores data, and *antenna* that broadcasts data by radio waves in response to a signal from a reader.
- The reader contains an *antenna* that receives the radio signal, and *demodulator* that transforms the analog radio into suitable data for any computer processing that will be done (Lockton and Rosenberg, 2005).

RFID Technology (Continued)

- RFID transponders in the form of "smart labels" make it much easier to track inventory and protect goods from theft or imitation.
- RFID technology also poses a significant threat to individual privacy.
- Critics worry about the accumulation of RFID transaction data by RFID owners and how that data will be used in the future.
- Privacy advocates note that RFID technology has been included in chips embedded in humans, which enables them to be tracked.

Cybertechnology and Government Surveillance

- As of 2005, cell phone companies are required by the FCC to install a GPS (Global Positioning System) locator chip in all new cell phones.
- This technology, which assists 911 operators, enables the location of a cell phone user to be tracked within 100 meters.
- Privacy advocates worry that this information can also be used by the government to spy on individuals.

Analyzing Personal Data: Big Data, Data Mining, and Web Mining

- John Stuart Ward and Adam Barker (2013) note that while the term "big data" has become "ubiquitous," it also has no precise or "unified single" meaning.
- They also point out that the definitions of big data put forth thus far are not only "diverse," but are often "contradictory" as well.

Big Data

- Initially, one might assume that the concept of *big data* simply refers to the size or scale of the data being analyzed.
- For example, Danah Boyd and Kate Crawford (2012) suggest that big data can be understood mainly in terms of its "capacity to search, aggregate and cross-reference *large data sets*."

Big Data (Continued)

- Definitions of big data that focus on capturing the large size of the data sets involved often view big data primarily in terms of its *volume*.
- Other definitions include factors affecting the "three V's":
  - *variety*,
  - *velocity*,
  - *veracity*.
- Whereas "velocity" captures the speed ("fast data in/out") involved in the process, "veracity, refers to the notion of trust in the (big) data analysis that needs to be established for business decision making.

Big Data

- The concept of big data is a far more complex phenomenon than merely the size, or volume, of the data involved.
- As Deborah Poskanzer (2015) points out, in the case of big data, “more isn’t just more—more is different.”
- She further suggests that big data can be better understood as a “new mode of knowledge production.”

Big Data (Continued)

- Regardless of which expression we use to describe this phenomenon – big data, data mining, or Knowledge Discovery Database – serious privacy concerns have been generated by it.
- Some believe that these kinds of concerns justify the need for a new legal category of privacy, which some call "group privacy.“
- Many, if not most, of the kinds of privacy concerns currently associated with big data had already been introduced by the use of various *data mining* techniques, beginning in the 1990s.

Data Mining

- Data-mining activities can generate new and sometimes non-obvious classifications or categories.
- Individuals whose data is mined could become identified with or linked to certain newly created groups that they might never have imagined to exist.
- Current privacy laws offer individuals little-to-no protection for how personal information that is acquired through data-mining activities is subsequently used.
- Yet, important decisions can be made about individuals based on the patterns found in the personal data that has been "mined."
- Some uses of data-mining technology raise special concerns for personal privacy.

Why is mining personal data controversial?

- Unlike personal data that resides in explicit records in databases, information acquired about persons via data mining is often derived from implicit patterns in the data.
- The patterns can suggest "new" facts, relationships, or associations about that person, such as that person's membership in a newly "discovered" category or group.
- Much personal data collected and used in data-mining applications is generally considered to be information that is neither confidential nor intimate.
- So, there is a tendency to presume that personal information generated by or acquired via data mining techniques must by default be *public* data.

Data Mining (Continued)

- Review Scenario 5-6 (in the text) involving Lee, a (hypothetical) junior executive, who:
- applies for a mortgage at XYZ Bank;
- has an impeccable credit history.
- A data-mining algorithm “discovers” that:

I. Lee belongs to a group of individuals likely to start their own business;

II. people who start business in this field are also likely to declare bankruptcy within the first year;

- Lee is denied the mortgage loan based on the profile revealed by the data-mining algorithms, despite his credit score.

Data Mining (Continued)

- Although the preceding scenario (involving Lee) is merely hypothetical, an actual case (that was similar to this) occurred in 2008.
- In that incident, a person had two credit cards revoked and had the limit on a third credit card reduced because of certain associations that the company made with respect to *where* this person:
  - shopped,
  - lived,
  - did his banking (Stuckey 2009).

Data Mining (Continued)

- In that (2008) case, a data-mining algorithm used by the bank “discovered” that this person (whose credit cards were revoked):
  - purchased goods at a store where typical patrons who also purchased items there defaulted on their credited card payments;
  - lived in an area that had a high rate of home foreclosures, even though he made his mortgage payments on time.

Web Mining: Data Mining on the Web

- Traditionally, most data mining was done in large “data warehouses” (i.e., off-line).
- Data mining is now also used by commercial Web sites to analyze data about Internet users, which can then be sold to third parties.
- This process is sometimes referred to as “Web mining.”
- Examine the “Facebook Beacon” incident (Scenario 5.7) as an example of Web mining.

Public vs. Non-Public Personal Information

- *Non-Public Personal Information* (or *NPI*) refers to sensitive information such as in one's financial and medical records.
- NPI currently enjoys some legal protection.
- Many privacy analysts are now concerned about a different kind of personal information called *Public Personal Information* (or *PPI*).
- PPI is non-confidential and non-intimate in character, and is generally not legally protected.

Public Personal Information (or PPI)

- In the past, it was assumed that there was no need to protect the kind of information we now call PPI, because it was viewed as simply public information.
- Nissenbaum (2004) believes that our assumptions about not needing to protect PPI are no longer tenable because of what she views as a misleading assumption:

> *There is a realm of public information about persons to which no privacy norms apply*.

Search Engines and Personal Information

- Search engines can be used to:

i. acquire personal information about individuals.

ii. reveal to search facilities data about which Web sites you have visited, as illustrated in the controversial incident (Scenario 5.10) in which describes how Google users' search requests were subpoenaed by the U.S. Government.

Privacy Law: European Union (EU) General Data Protection Regulation (GDPR)

- Privacy policies will have to be written in a clear, straightforward language.
- The user will need to give an affirmative consent before his/her data can be used by a business. Silence is no consent
- Businesses will need to clearly inform the user about transfer of data outside of EU
- Businesses will be able to collect and process data only for a well-defined purpose. They will have to inform the user about new purposes for processing

GDPR

- Businesses use algorithms to make decisions about the user based on his/her personal data (e.g. when applying for a loan); the user is often unaware about this. Businesses will have to inform the user whether the decision is automated and give him/her a possibility to contest it
- Businesses will have to inform users without delay in case of harmful data breach
- The user will be able to move his/her data, for instance to another competing service if desired.
- The user will have the right to access and get a copy of his/her data, a business has on him/her
- Users will have a clearly defined "right to be forgotten" (right to erasure), with clear safeguards

Right to be Forgotten (RTBF)

- A relatively recent challenge for the GDPR is whether users should have a right to have certain kinds of personal information about them deleted, or at least "delinked" from search-engine indexes.
- This "right" is controversial because it:
  - has international implications, given the flow of information across the porous boundaries of cyberspace.
  - applies to online personal information that is shown to be either inaccurate or no longer "relevant."

Arguments Opposing RTBF

- Search engine companies generally make two different kinds of claims, arguing that they:

1) do not control *content* and on the Internet (and thus cannot be held responsible for the relevance, or accuracy, of the content on sites to which they provide links);
2) cannot be expected to respond to all of the links requested by users (even if the information being linked to is either inaccurate or no longer relevant, because doing so would be too *impractical*, if not impossible).

- Being required to comply with RTBF is:

a) tantamount to "Internet censorship" (because it violates "freedom of expression");
b) harmful to the general public (because it interferes with a citizen's "right to know").

Arguments Defending RTBF

- Arguments supporting RTBF generally fall into two broad categories.
- Supporters claim that this privacy principle is needed to:

1) Prevent innocent people from being harmed;
2) Protect people whose personal identity evolves over time.

Establishing "Appropriate" Criteria for RTBF

- While the European Court of Justice ruled in favor of RTBF (May 2014), it did not provide precise criteria for search engine companies to comply with the new privacy principle.
- Google has since established an advisory council to come up with appropriate criteria.
- Arguably, two important factors need to be taken into consideration:

1. the nature of the *personal information* itself;
2. the *context(s)* in which this information flows.