

# CSC197 Lecture 1 Notes

Anatoly Zavyalov

September 16, 2020

# 1 What you will learn in CSC197

You will learn:

1. to think about different approaches to managing data and privacy
2. how to protect your own privacy online
3. to communicate and do research at a university level

# 2 Assessment Scheme

Weekly Preparation	20 points ( $2 \cdot 10$ )
Journal	18 points ( $3 \cdot 6$ )
Explainer Video	20 points
Final Paper Preparation	22 points
Final Paper	20 points

# 3 How to do well in CSC197

- Participate in class (use chat a lot!)
- Do the weekly preps
- Share your point of view
- Talk to your instructor
- Respect each other

## 4 Data Exhaust

Think about a database with all humans listed. The rows represent humans, the columns represent their features.

**Big data:** A database with MANY rows and columns.

About 25 years ago, companies decided to hang on to formerly irrelevant data, turns out lots of it can be monetized, such as:

- Cookies
- Geolocation
- Correspondents
- Musical taste
- Click rate/pattern
- Typing rate/pattern

## 5 Personally identifiable information (PII)

Some personally identifiable information can include:

- Date of Birth
- Gender
- Physical address
- Name
- Face
- Phone number
- Acquaintances
- Habits/Hobbies
- Voice
- IP address

Professor Latanya Sweeney (from Harvard), an expert on privacy showed that 87% of Americans can be uniquely identified by **date of birth**, **gender**, and **zip code**, and introduced the idea of combining different pieces of data to uniquely identify someone.

## 6 Identity, Autonomy, Privacy

**Identity** distinguishes you from others, including things like:

- Characteristics
- Appearance
- Likes/Dislikes
- Acquaintances
- Social class
- Race
- Gender
- Nationality

**Autonomy** is the ability to act on your own and make choices. It may be limited by laws, physics, customs, etc.

**Privacy** is the ability to control access to your body, image, information, thoughts, reputation, and so forth.

## 7 Seeking ‘I am’

“I think, therefore I am” (Rene Descartes) leads to the notion that there’s a tiny being experiencing it all. This leads to a very individual notion of identity, e.g. protecting individual rights to privacy, to some autonomy, etc., and often leads to the notion that privacy rights can be protected by the market.

Another point of view is that our identity is more like the symbiotic entangling of a fungus (amanita) and tree roots: “I am because we are”. This leads to collective protection of privacy rights.

## 8 Classical Privacy

Before the invention of the telegraph, the fastest that data could move was the speed of a horse, ship, or in some cases a carrier pigeon. Classical privacy back then can be modeled with concentric circles.

1. The inner circle contains information that one takes to the grave.
2. Slightly further out is information that you share with your spouse, lover, or partner.
3. Even more further out is information that you share with your family or close friends.
4. Even MORE further out is information that you share with workmates and classmates.
5. The furthest out is information that you share with strangers.

Information bleeds out slowly, and is eventually forgotten.

## 9 Networked Privacy

A social network has a very long reach, as your friends have other friends who have other friends, and so on.

Information bleeds **very quickly** and persists for a **long time**. This leads to the notion of “the right to be forgotten”.