Data Science Ethics

College of Science

Al-Mustaqbal University

Stage 2 , Semester 1
@ Department of Intelligent Medical Systems

**Prepared by
Dr. Samir Badrawi**

Data Privacy

The majority of this course material is based on Coursera
https://www.coursera.org/learn/data-science-ethics
"H.V. Jagadish lectures", a Professor at the University of Michigan

1. <mark>Privacy</mark>

what is privacy?. Based on Greek myth, there was a giant with "a hundred" eyes, **Argus Panoptes**. Who was ever watchful, as some of his eyes would be awake when others were asleep

Argus Panoptes

- Giant with "a hundred" eyes in Greek myth.
- Was ever watchful.

Building off this myth, *Jeremy Bentham*, in the **18th century**, designed a prison called the **Panopticon**.

The idea was that a single guard in the center of this complex could observe all the prisoners. The prisoners wouldn't know whether they were being watched, because the guard could look in any direction that he chose. Sometimes he might be looking at you, other times he might not.
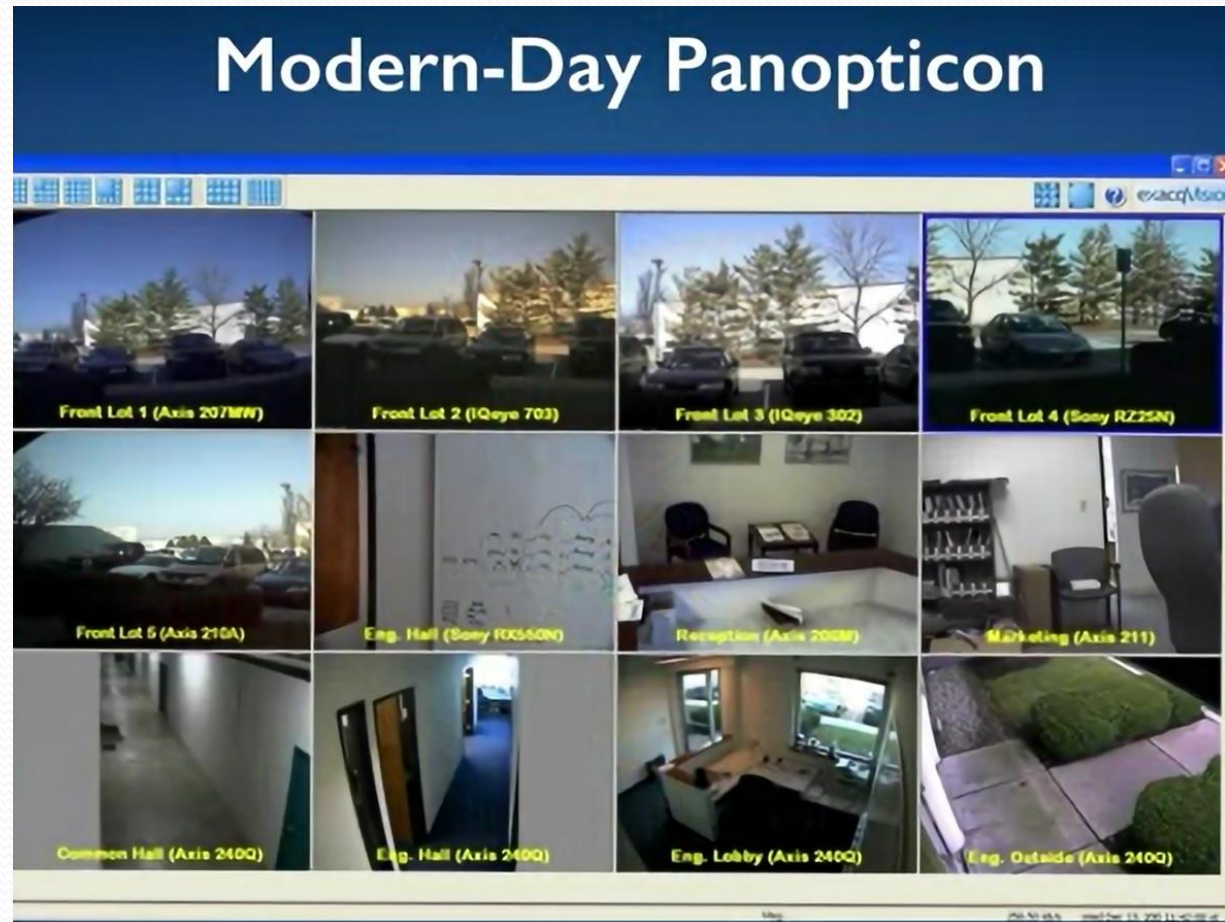
Panopticon

- A prison where a single guard could observe all the prisoners.
- But prisoners would not know when they were being watched, so would "behave" all the time.
- Jeremy Bentham, 18th Cent.
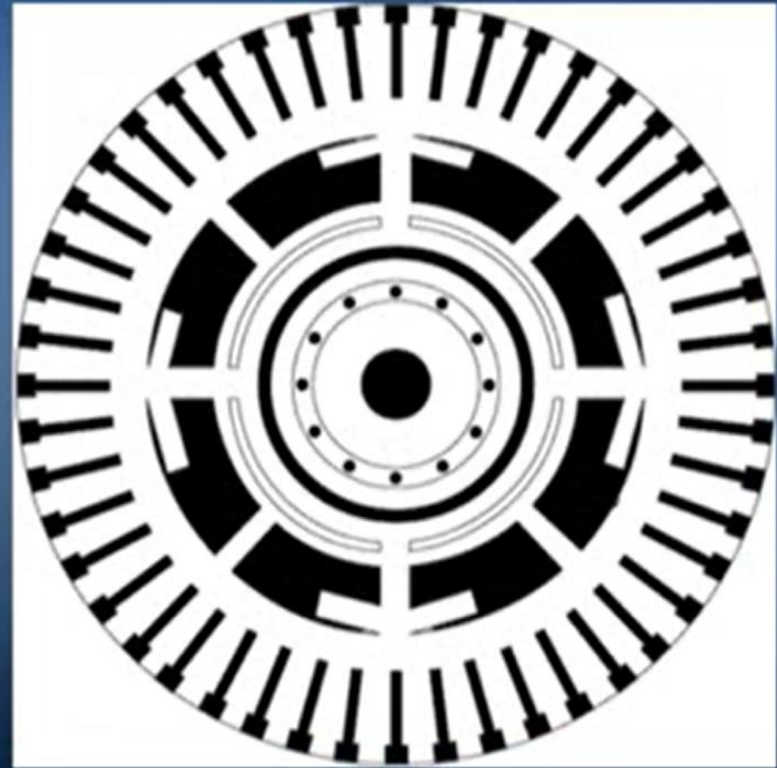
Since the prisoners wouldn't know whether they were being watched, they would "behave" all the time as if they were being watched.

This notion (always being watched) today exists. Here's a ***Modern-Day Panopticon***, in terms of a security camera or actually a bunch of security cameras in a typical house.

If you think about the fact that your interactions on the Web, your interactions online with merchants are being recorded, you end up with something that looks like this scene. And we have here,

the Facebook guard of the Panopticon.

Also, Even If Less Obvious

- OK, Facebook does not really watch everything you do.
- But how about the internet as a whole?
- Do you care if you are doing nothing wrong?

Privacy is a ***basic human need***. And this is true even for people who have "nothing to hide." Besides, privacy also has societal benefits.

A central part of our democracy is ***anonymous voting***. You need to be able to hide your vote to be able to vote freely, without having to worry about pressure from other people.

For all of these reasons, privacy is important. We need to think about how we **manage privacy** in this complex world with **big data** and **data analytics.**

2- History of Privacy

The first known legal definition of privacy was in **1604**, when Edward Coke, then Attorney General of England, quoting the English common law, said, famously, *"the house of every one is to him as his castle and fortress"*.
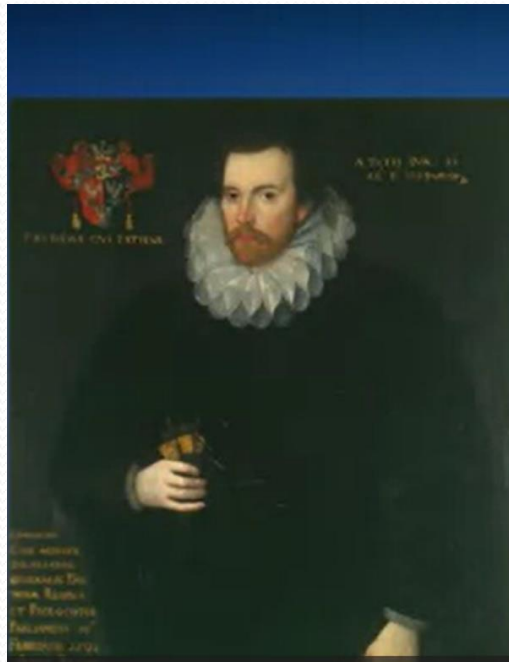
1604

- "The house of every one is to him as his castle and fortress."
- Edward Coke, Attorney General of England
- Quoting "Common Law"

In **1928**, there was a famous lawsuit at the Supreme Court in the US, and this case was about the legitimacy of *wiretaps*. The Court said that a *wiretap* wasn't an unreasonable search, that it didn't need any permission, as *wiretaps* do today.

In **2006**, Daniel Solovey, a Professor at George Washington University, defined a taxonomy of privacy, and the various ways in which privacy could be eroded.

2006

- **Taxonomy of privacy**
  - *Information Collection*
  - *Information Processing*
  - *Information Dissemination*
  - *"Invasions"*
- **Daniel Solovey, Professor at George Washington University**
- **Article in Univ. of Pennsylvania Law Review**

Information *collection*, *processing*, and *dissemination*, are greatest concern to us in the context of the data science that we are discussing here.

Invasions are things like somebody observing you from where you don't expect; bugging your apartment.

small towns vs. *Big cities*

It Takes a Village …

- In small towns, there was little privacy
    - *Everyone in town knew "everything" about everyone else*
- Big cities provide anonymity.
- Does information technology bring us back to the halcyon days in a small town?

The problem with ***big data*** is that it's universal, and it never ***forgets*** anything.

Talking about not forgetting anything, there is actually something called a ***wayback machine***, which archives pages on the web.

And the intention is to ***retain this forever***, because these public web pages are (for the most part,) ***expressions of our society and our culture,*** are things that are ***part of our heritage***, and are things that people will want to look at, that historians and sociologists, among others, would care deeply ***about in the future***.

A side effect, though, is that if you have an unflattering page written about you, it'll survive forever in an archive, even if the page itself has since been taken down.

In some countries, there is now a **right to be forgotten**. The idea is that, in terms of societal expectations of redemption, there are often laws written where a person's record is cleared after some years. For example, in the US, in most states, if you have an accident on your driving record, that gets expunged *after three years*.

For example, in the US, in most states, if you have an accident on your driving record, that gets expunged *after three years*.

And so, things that you did wrong many years ago don't continue to haunt you forever and forever.

The same idea applies to more significant harms. So, if you had a conviction for certain kinds of offenses, even if you served time in prison, in many cases, there are rules about how long that stays on your record, and after how much time it gets expunged from your official record. While this is all fine with respect to the official record.

If this information happens to be on the web, how is it ever to be removed? Well, given what we just saw on the previous slide, you really can't remove it from the web. It's there, and it's there forever

If this information happens to be on the web, how is it ever to be removed?
Well, you really can't remove it from the web. It's there, and it's there forever.