Computer Vision for Embedded Systems

Yung-Hsiang Lu
Purdue University
yunglu@purdue.edu

PURDUE UNIVERSITY

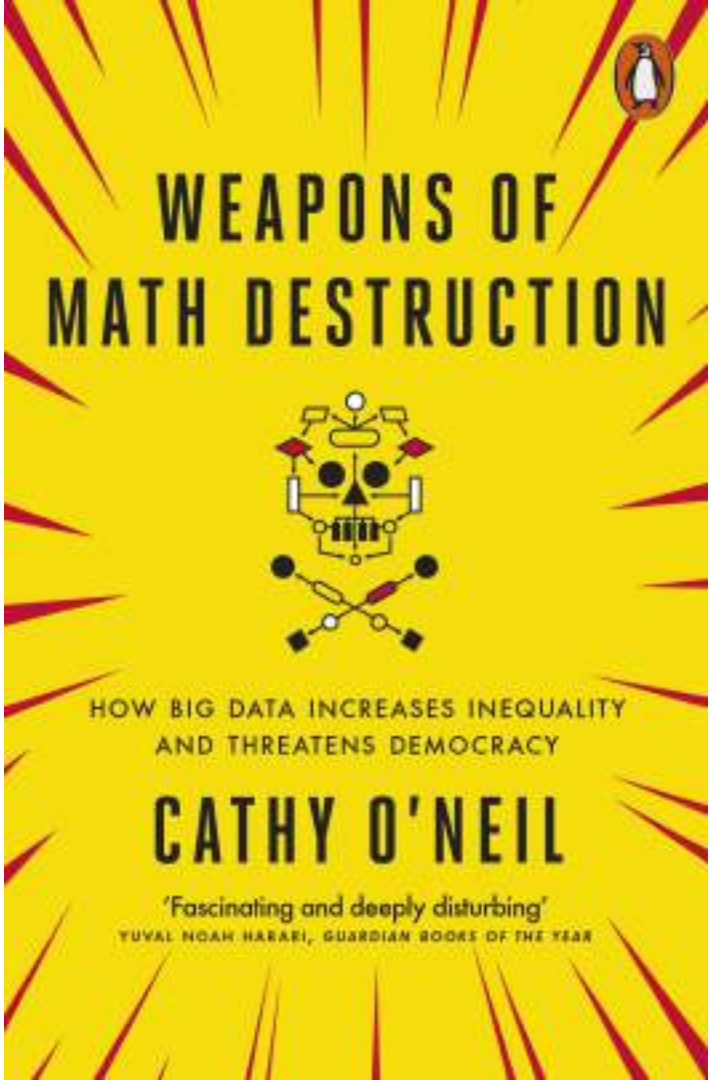
Data Bias and Privacy

Dr. Yung-Hsiang Lu has never and will never conduct any research that can identify any individual, such as face recognition.

What is Dataset Bias?

Dataset Bias: Certain elements in a dataset are disproportionately represented compared to others

If a dataset is not created randomly, it is **biased** in some way


WEAPONS OF
MATH DESTRUCTION
HOW BIG DATA INCREASES INEQUALITY
AND THREATENS DEMOCRACY
CATHY O'NEIL
'Fascinating and deeply disturbing'
YUVAL NOAH HARARI, GUARDIAN BOOKS OF THE YEAR

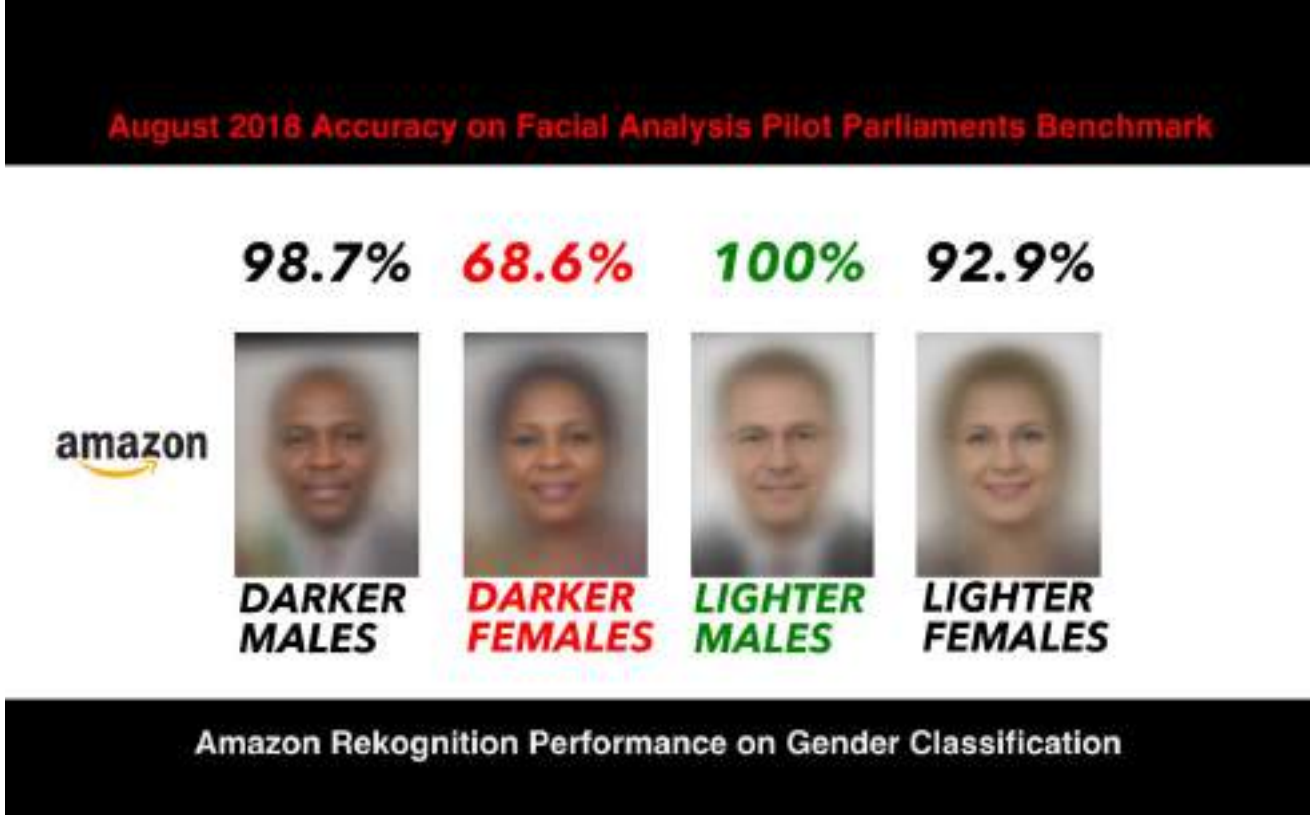
How are these images biased?

https://www.thehappycatsite.com/white-cat/
https://www.dailypaws.com/living-with-pets/pet-compatibility/white-cat-breeds
https://www.catwatchnewsletter.com/health/disease/the-odds-of-deafness-in-white-cats/

How are these images biased?

Depending on the purpose of the dataset.
Dataset of "cat" or "white cat"?

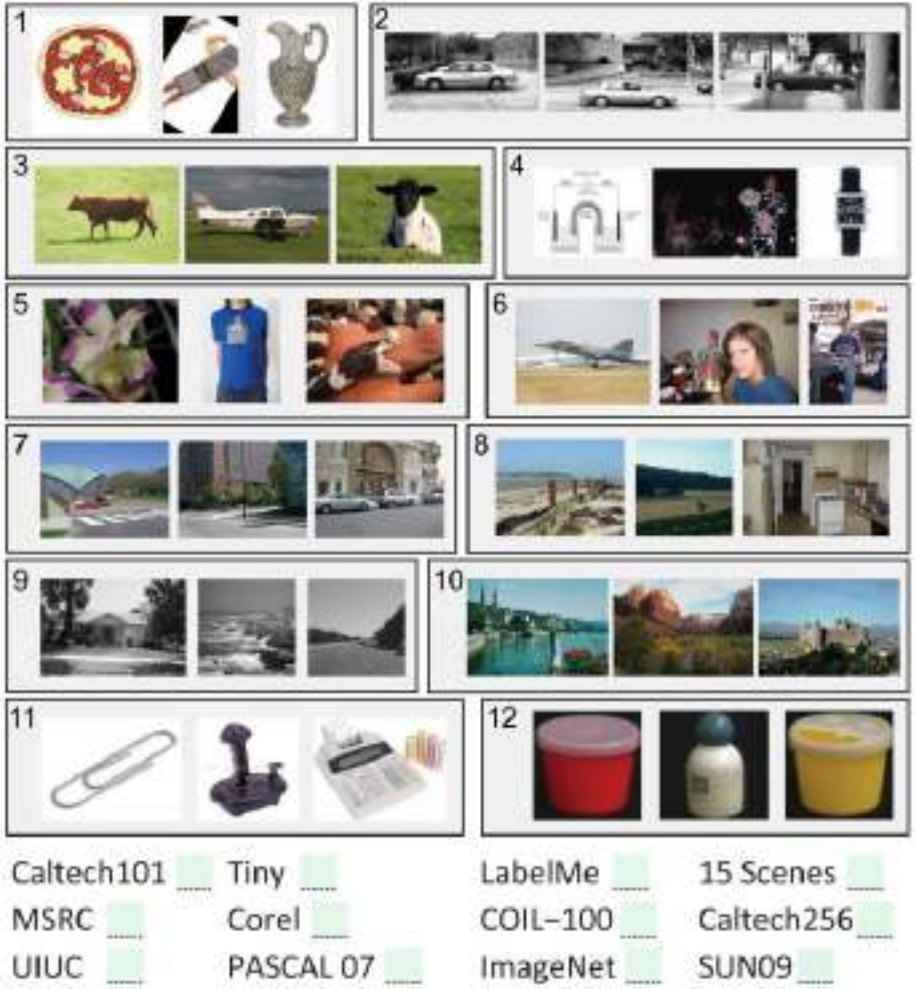
https://www.thehappycatsite.com/white-cat/
https://www.dailypaws.com/living-with-pets/pet-compatibility/white-cat-breeds
https://www.catwatchnewsletter.com/health/disease/the-odds-of-deafness-in-white-cats/



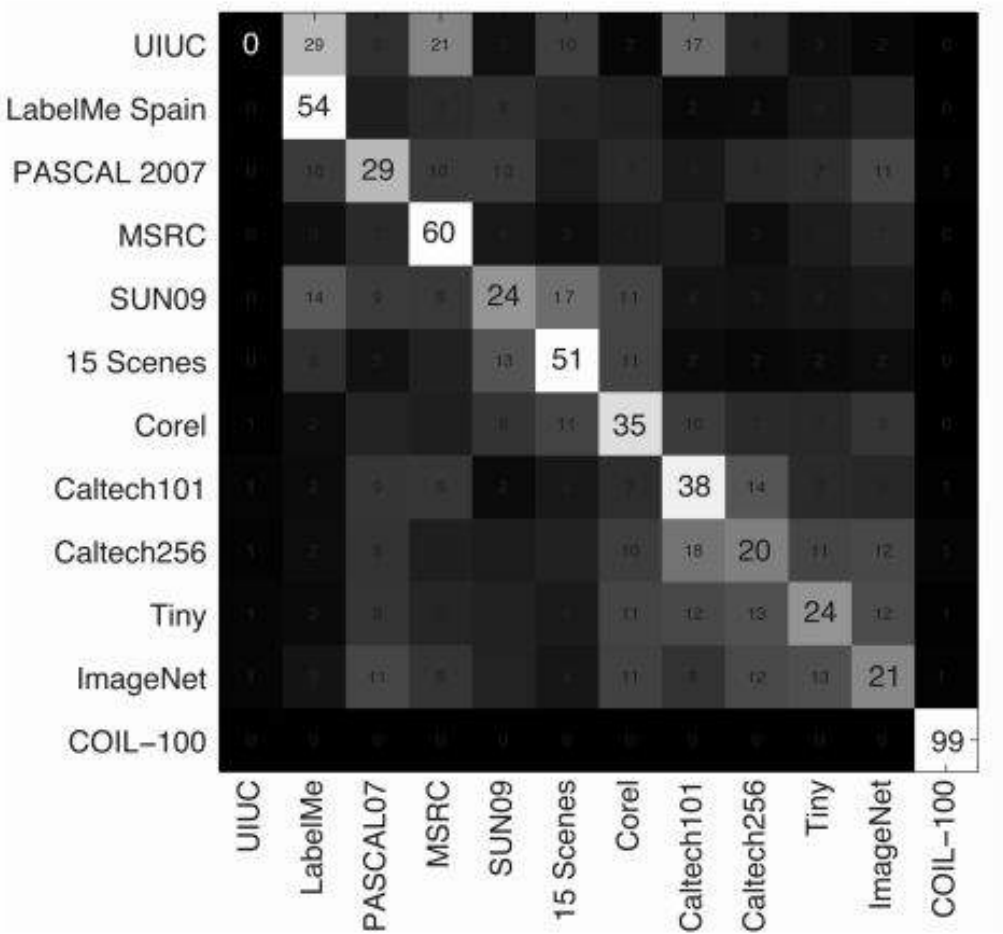
https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced

Unbiased Look at Dataset Bias

Antonio Torralba
Massachusetts Institute of Technology
torralba@csail.mit.edu
Alexei A. Efros
Carnegie Mellon University
efros@cs.cmu.edu



created by a SVM
(support vector machine)





Is it "style" or "bias"? Distinguish desired style from unintended bias.

https://www.biography.com/musician/ludwig-van-beethoven
http://www.audio-music.info/htm/m/Mahler_Gustav.htm
https://www.hifinews.com/content/johannes-brahms-symphonies

Different Types of Bias

- Selection Bias: Where and how the samples are selected (e.g. Internet)
- Capture Bias: How is the data captured? (e.g. focused at center)
- Negative Set Bias: Lack of negative examples

Crowdsourcing Detection of Sampling Biases in Image Datasets (The Web Conference 2020)

Xiao Hu, Haobo Wang, Anirudh Vegesana, Somesh Dube, Kaiwen Yu, Gore Kao, Shuo-Han Chen, Yung-Hsiang Lu, George K. Thiruvathukal, Ming Yin

1. show groups of photos and ask people to describe the photos
2. merge the descriptions (natural language processing)
3. ask whether the descriptions are intrinsic to the objects in the photos
4. If the descriptions appear often and is not intrinsic, the dataset has unintended bias.

Sample Images of Input Dataset

Biases Detected by the Crowd

1: The plane is facing right. [KB]	6: 1 airplane is in the air. [U]
2: The photos are orientated landscape. [AB]	7: Zero people are visible. [AB]
3: There is only 1 airplane in each image. [AB]	8: The location of the airplane is ground. [KB]
4: The photos are taken during the day. [AB]	9: The type of the airplane is commercial. [KB]
5. The main color of the airplanes is white. [KB]	10: The airplanes have wings on both sides and back. [US]

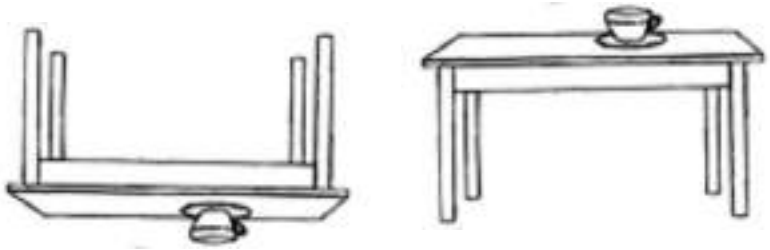
Known bias (KB), additional bias (AB), unbiased similarity (US) or unrelated (U)

Theory of Learning

Theory of learning

- No learning algorithm is universally better for all possible problems.
- Anything that can be learned must have patterns (or rules).
- Learning means the ability to handle unseen situations correctly.
- Designing an "unbiased" machine learning model is impossible
- It is necessary to know the specific applications and purposes before deciding whether bias is acceptable or not.
- Unintended biases can have surprising consequences

- Q: Is it possible to "learn" by simply analyzing pixels, without knowledge about the underlying physical properties (such as gravity, materials …)?



https://www.researchgate.net/figure/Cup-on-table-Prompt-Where-is-the-cup_fig1_41555072

Bias Is Common

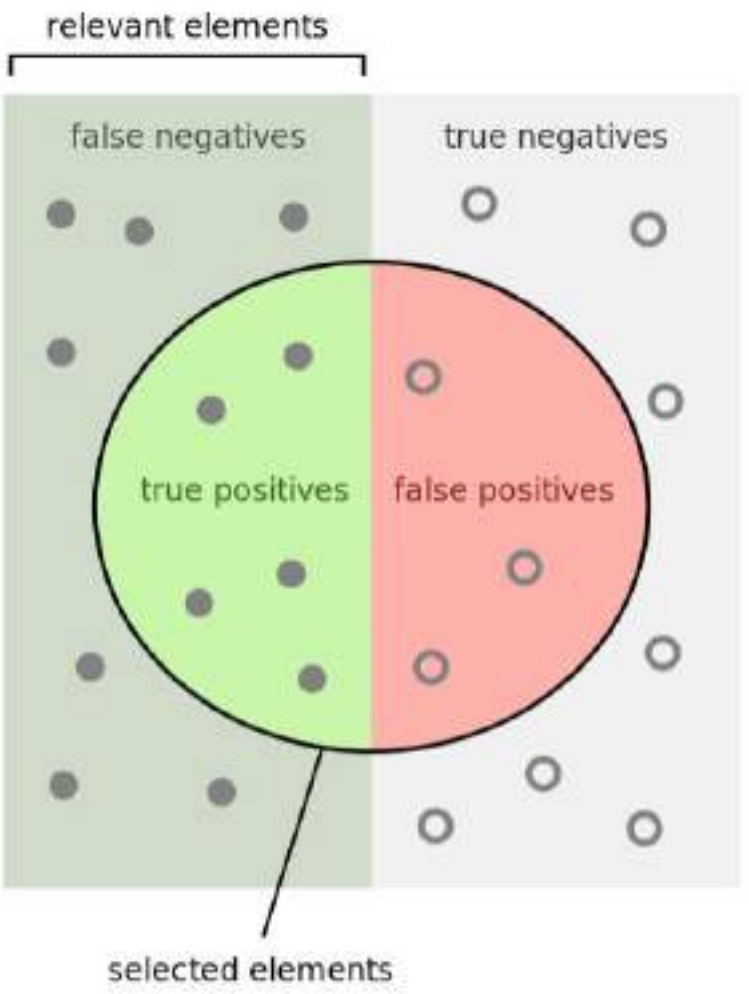
- Purdue has snow in winter. You must enjoy skiing.
- You are born in XXX. You must love pizza, or rice, or steak, or corn …
- You like to play basketball. You must be very tall.

Quantify Biases

Quantify Success

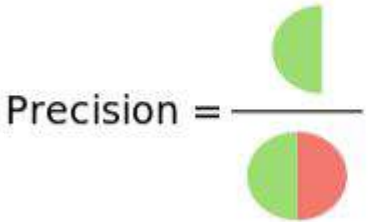
- A recognition method is "99% accurate". Is it good?
- Is the data "balanced"?
- Balanced data: each type has about the same rate of occurrences, or proportional to the occurrences of the population.

https://en.wikipedia.org/wiki/Precision_and_recall

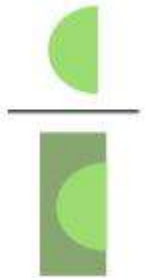


		Detection	
		True	False
Actual	True	True Positive	False Negative
	False	False Positive	True Negative

How many selected items are relevant?



How many relevant items are selected?



$$\text{precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}$$

$$\text{recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}$$

$$F_1 = \frac{2}{\text{recall}^{-1} + \text{precision}^{-1}} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} = \frac{\text{tp}}{\text{tp} + \frac{1}{2}(\text{fp} + \text{fn})}.$$

Why does this matter?

Can computer vision successful detect rare events?

- illness in medical images
- intruders
- crimes
- natural disasters
- …

		Detection	
		True	False
Actual	True	0	0.01%
	False	0	99.99%

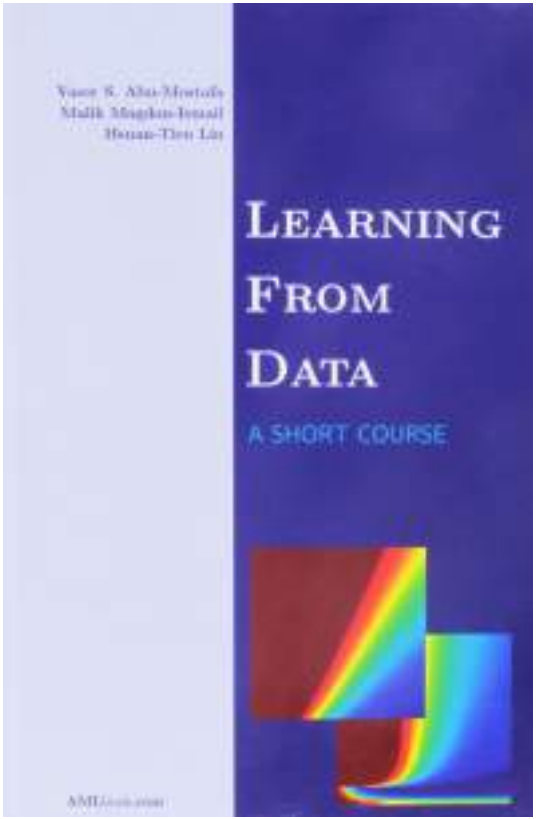
What is the precision? recall?

If a type of illness occurs to only 0.01% people, always predicting negative is "pretty accurate".

Required accuracy depends on the next step

Consider a face recognition system that is imperfect. What should the system do if it is unsure whether to accept a person?

- Customer discount ⇒ accept (happy customers)
- Secure area ⇒ reject (protect the area)



Whether to Curve an Illness?

If the detection of an illness is not perfect, what should the patients and doctors do?

- Get better diagnosis
- Evaluate trade-offs of risk and rewards
  - high fatality?
  - dangerous procedure?

		Detection	
		True	False
Actual	True	True Positive	False Negative
	False	False Positive	True Negative

Privacy

Privacy

Visual Privacy is the relationship between collection and dissemination of visual information, the **expectation** of privacy, and the **legal** issues surrounding them. These days digital cameras are ubiquitous. They are one of the most common sensors found in electronic devices, ranging from smartphones to tablets, and laptops to surveillance cams. However, privacy and trust implications surrounding it limit its ability to seamlessly blend into computing environment. In particular, large-scale camera networks have created increasing interest in understanding the **advantages** and **disadvantages** of such deployments. (Wikipedia)

Privacy is a complex topic

involving laws, emotion, expectations, social norms ...

Smart Cameras and the Right to Privacy

Designs that take account of existing norms of privacy minimize the likelihood of triggering an outcry for legal change to prohibit or restrict use of smart camera systems.

By William H. Widen

Proceedings of the IEEE, 2008-10, Vol.96 (10), p.1688-1697

	Observer's Vantage Point			
		Subject's Property	Observer's Property	Public Property
Subject's Location	Subject's Property	1.Subject Protected/ **Property** Regime	2.Subject **Unprotected**	3.Subject **Unprotected**
	Observer's Property	4.Subject Protected/ **Property** Regime ?	5.Subject Protected/ **Contract** Regime	6.Subject **Unprotected** ?
	Public Property	7.Subject Protected/ **Property** Regime ?	8.Subject **Unprotected**	9.Subject **Unprotected**

Expectation of Privacy



https://www.forbes.com/sites/davidbressan/2020/10/14/eruptions-of-old-faithful-geyser-in-yellowstone-could-become-less-frequent-or-completely-cease-in-a-warmer-future/?sh=3cde7cfe73b8
https://livejapan.com/en/in-tokyo/in-pref-tokyo/in-shibuya/article-a0002128/
https://www.youtube.com/watch?v=lDFDMwpWWXQ&ab_channel=PurdueSports

Before you worry about "privacy", understand

- Why are cameras deployed?
- What do they intend to accomplish?
- Who can access the data?
- Do you actually have the right or expectation of privacy?
- Can you protect your privacy?
- Do you "voluntarily" give away information about your age, diet preference, location, network of friends …?

Privacy, Surveillance, and Law by Richard A. Posner 2008

- People give personal information for services or privileges, such as healthcare, mortgage, driving, international travel …
- In the past, information was rarely shared. Now, sharing is common.
- People want to protect personal information for different reasons. One reason is the perceived potential harm in the future.
- Privacy can be a trade-off of public safety and personal preference.

Boston Marathon Bombing (2013/04/15)

Surveillance cameras captured suspect walking

Wikipedia
https://www.chicagotribune.com/sports/breaking/ct-boston-marathon-bombing-key-moments-20180413-story.html

Do surveillance cameras reduce crimes?

- It is usually difficult to have evidence of causal relationships. Many factors can affect crime rates, such as the economy, policies.
- Possible factors:
  + better lighting, evidence for prosecution, faster response

…

- more crimes are reported?

? people feel safer and give criminals more opportunities

...

Panel: Privacy Protection in Online Multimedia

Yung-Hsiang Lu
Purdue University
yunglu@purdue.edu

Andrea Cavallaro
Queen Mary University of London
a.cavallaro@qmul.ac.uk

Catherine Crump
University of California Berkeley
ccrump@law.berkeley.edu

Gerald Friedland
University of California Berkeley
fractor@icsi.berkeley.edu

Keith Winstein
Stanford University
keithw@cs.stanford.edu

ACM International Conference on Multimedia 2017

Billions of Cameras

- phones, dashcams, street corners, stores, gopro …
- social networks for sharing images and videos
- What are the right of the people captured by these cameras? In USA? No simple and clear answer.
- In general, people have the right to capture images and videos in ***public locations***.
- Some legal restrictions are applicable only if the people capturing the data represents governments.

US Constitution

Fourth Amendment: The right of the people to be secure in their persons, houses, papers, and effects, ***against unreasonable searches*** and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.

US Constitution

First Amendment: ***Congress*** shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.

Interpretation: obstacle to government attempts to regulate video recording

It does not prevent individuals (not representing government) requesting others to stop disinformation.

Bill of Rights

The Bill of Rights is the first 10 Amendments to the Constitution. It spells out Americans' ***rights in relation to their government***. It guarantees civil rights and liberties to the individual—like freedom of speech, press, and religion. It sets rules for due process of law and reserves all powers not delegated to the Federal Government to the people or the States. And it specifies that "the enumeration in the Constitution, of certain rights, shall not be construed to deny or disparage others retained by the people."

https://www.archives.gov/founding-docs/bill-of-rights/what-does-it-say

Questions Raised by the Paper

- Do people care about privacy? When? Where? Differences in ages, cultures, education, locations, time of day?
- Who has the rights to acquire the data (in particular, in public locations)? What are the restrictions of the rights?
- Who has the rights to view the data?
- Who has the rights to keep the data? How long can the data be kept?
- What are the social, economic, legal, and commercial values (or barriers) to protect privacy?

- How to protect against unauthorized access to data?
- What analyses can be performed on the data?
- What are the rights of the people that appear in the data?
- What technologies can protect privacy? Are the technologies ready? What are the costs of using these technologies?
- Can money be made by protecting (or violating) privacy? How? Who can benefit? Who loses?
- Should users set their own privacy rules? Or privacy should be protected by law?

Protect Your Own Privacy

- In general, it is legal to check candidates' social networks before hiring decisions
- Hiring managers should ***not*** check candidates' social networks to prevent discrimination liability
- Outsource to a third party (neither the employer nor the candidate) to ensure only valid information is considered
- To protect yourself: either have no social network presence (difficult) or ensure the information is accurate and presentable. Hide "unprofessional information" in private locations with limited accesses (and prevent sharing).

How to protect your privacy?

- Know what you are talking about
- Think carefully before posting on social media
- Restrict who can see your data
- Read and select privacy policies
- Remove metadata (such as time and location)
- Wait for several days before you post
- Encrypt your data
- Tell your friends to respect your privacy
- Wear a hat or a face cover when you are in public places
- Inform your credit card companies not to share data
- Do not sign up for "free" stuff (it is not free)

(Tom Cruise in Jack Reacher)

https://www.highlandernews.org/

http://extension.msstate.edu/publications/be-hero-wear-mask

Getting Started

The steps below provide a guided procedure from beginning to end of the IRB process. Reference these steps to navigate the most common steps of an IRB protocol application.

Before Submission

- Does the IRB need to review my study?
- Does your "human subjects research" fit the criteria for exemption from IRB review?
- If your study does not fit the criteria for exemption, complete a non-exempt application form and associated consent/assent documents
- Train all personnel who interact with or analyze data from human subjects


CITI PROGRAM
Courses
Organizations
Individuals
About
Support FAQ Contact Us
Register
Log In
Responsible Conduct of Research (RCR)
RCR covers core norms, principles, regulations, and rules governing the practice of research.
ORGANIZATIONS
LEARN MORE
LEARNERS
EXPLORE COURSES