Ethics, Public Policy, and Technological Change

Rob Reich

Mehran Sahami

Head TA: Roberta Fischli

Housekeeping

- The Syllabus has been updated on Canvas and is current at the course website, cs182.Stanford.edu
- Sections begin Friday, 3-4pm. You will hear via email about your section assignment by tomorrow afternoon.
- You all get two free “late days” for assignment submissions
  - Each “late day” is a 24 hour day
  - Can use them when you like, but best not to use them at all
  - Assignments turned in late after free late days are used incur a one full grade penalty (e.g., A -> B) per 24-hour day that they are late
  - No late days may be used on the Final Reflection (last assignment), which is due **March 17th**

Today's Agenda

1. Why *You* are Taking this Class
2. Three Dimensions of Ethics
3. Professional codes of ethics
4. Disciplinary perspectives
5. The Optimization Mindset

Today's Agenda

1. Why *You* are Taking this Class
2. Three Dimensions of Ethics
3. Professional codes of ethics
4. Disciplinary perspectives
5. The Optimization Mindset

Overview of Survey Results

- Think about a moment when you saw/experienced/learned about/anticipated the impacts of a new technology that gave you pause about the costs to society of this technological change

    - What was the technology?
    - What were your concerns?
    - Do these concerns outweigh the benefits? Why?

Common Themes and Concerns

- Generative AI
  - Academic integrity
  - Job displacement
  - Environmental impact
  - Erosion of critical thinking; brain rot
- Deep Fakes/Misinformation
- Surveillance & Privacy
- Social Media and Attention Economy
- AI’s effect on creativity
- Hidden Costs and Exploitation
- Algorithmic Bias and Discrimination
  - Facial recognition in policing
  - Tenant screening systems ("technologized redlining")
  - Immigration tracking and deportation
  - Predictive policing

Sample Quotes

"I will never forget the weeks in my senior year of high-school when ChatGPT was released, it was a frenzy of my peers attempting to cheat on final projects and papers... I feared that DoorDash, ChatGPT, Amazon Fresh, and more would create a society where humans needed to do nothing but press buttons to live life."

"Two years ago, I was talking about a very niche topic with my friend... Around half an hour after we split up, I opened Instagram, and I saw a suggested post about the exact same niche topic... I don't feel happy when I'm suggested a product I was talking about - I feel creeped out."

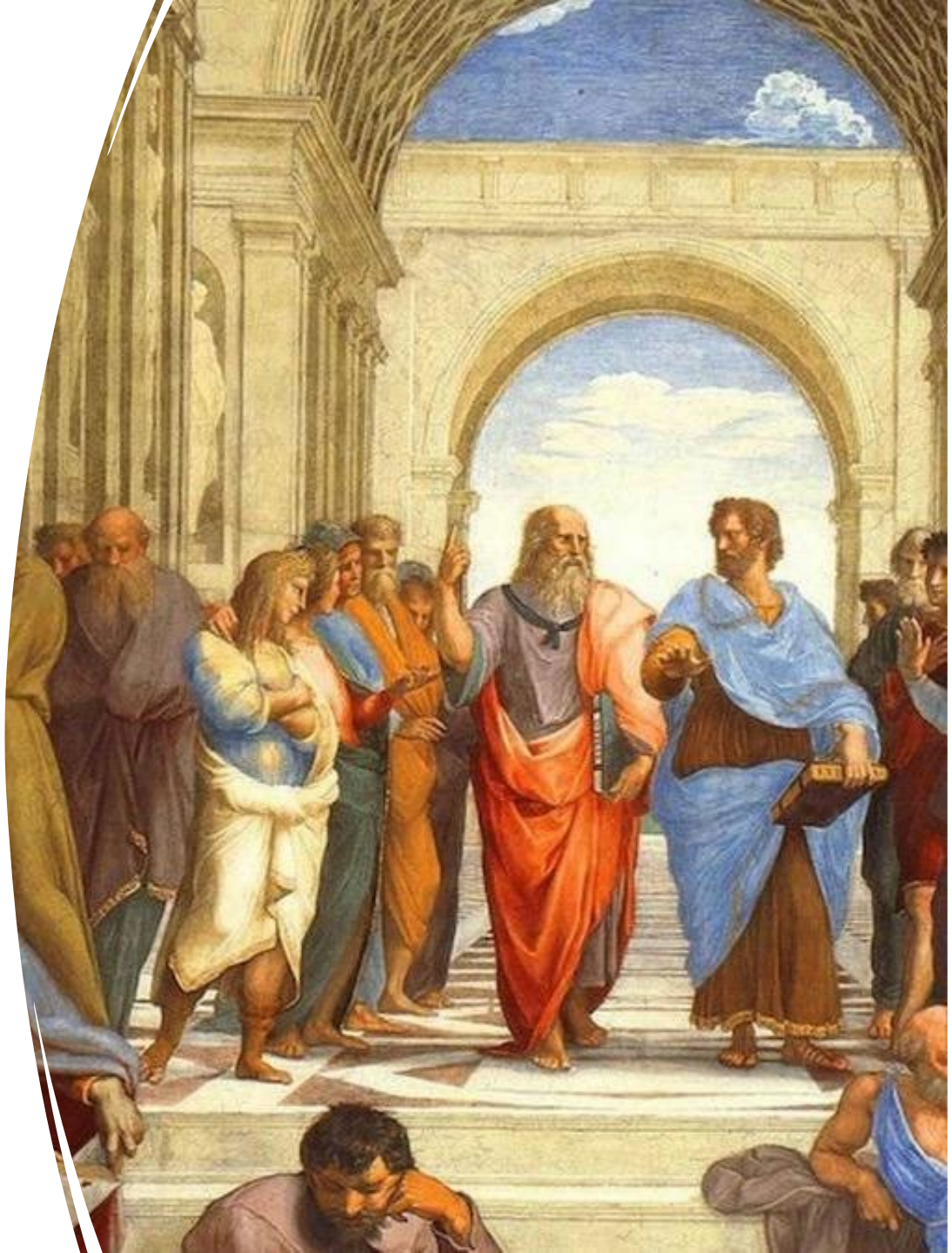
"My brother is actually a voice actor, and his job community is completely scared of their profession being wiped out... I can imagine the world losing a lot of its rich color if art becomes something only a few are creating without the use of AI."

"I worry that I've become so reliant on Chat-GPT/Claude/etc that I no longer have the capacity for as much critical thought as I did when entering Stanford, which is a depressing realization."

Today's Agenda

1. Why *You* are Taking this Class
2. **Disciplinary Perspectives: philosophy**
   - **Three Dimensions of Ethics**
3. Professional codes of ethics
4. Disciplinary perspectives
5. The Optimization Mindset

What role can philosophy play in thinking about questions and problems concerning technology and science?



Raphael, “The School of Athens”
Source: Wikimedia Commons

What's Distinctive about Ethical Inquiry?

- Normative questions and conclusions; not about explanation
- Prescriptive, not descriptive
- Not about understanding the world as it is, but as it ought to be

Three levels of ethics

Personal Ethics

Professional Ethics

Social and Political Ethics

Three levels of ethics

Personal Ethics

What does it mean to be a good person?
What is your moral compass?

Professional Ethics

Social and Political Ethics

ORTUNE RANKINGS MAGAZINE NEWSLETTERS VIDEO PODCASTS CONFERENCES COVID-19 SEARCH SIGN IN Subscri

The Ugly Unethical Underside of Silicon Valley

Too many startups are taking "fake it till you make it" too far.

BY ERIN GRIFFITH

December 28, 2016 3:30 AM PST

Source: *Fortune*, Erin Griffith, 2016

The New York Times

The Epic Rise and Fall of Elizabeth Holmes

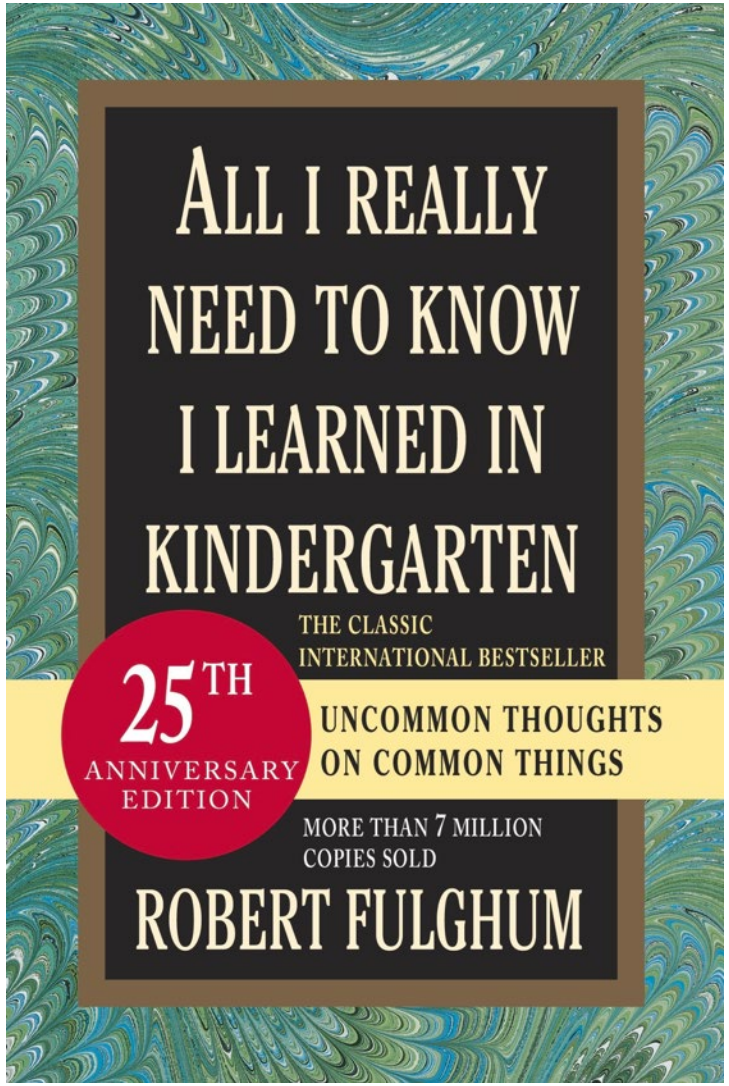
In Silicon Valley's world of make-believe, the philosophy of "fake it until you make it" finally gets its comeuppance.

Source: *The New York Times*

Star Technologist Who Crossed Google Sentenced to 18 Months in Prison

Anthony Levandowski, a onetime star Silicon Valley engineer of self-driving cars, had pleaded guilty to stealing trade secrets.

Source: *The New York Times*

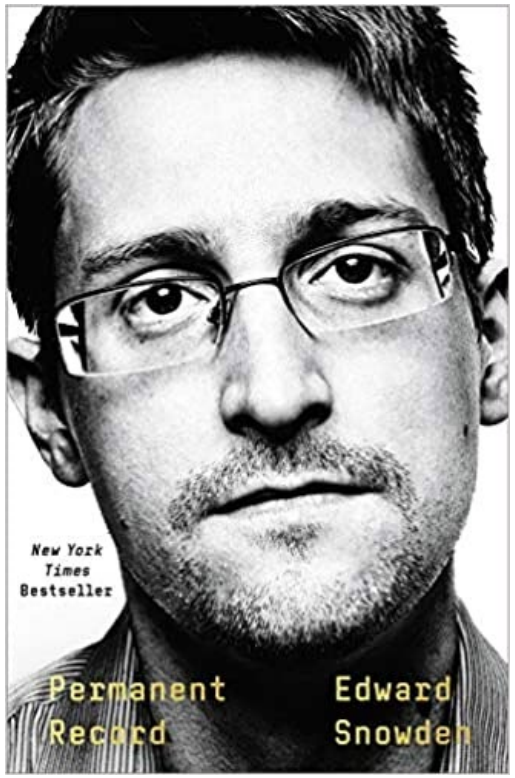


Important, but not especially interesting.

No good argument for cheating, lying, stealing.

Must design society to assume moral imperfection.

Ethics of Whistleblowing



Theranos Whistleblower Shook the Company—and His Family

Tyler Shultz says he wanted to shield reputation of former Secretary of State George Shultz, a Theranos director and his grandfather; $400,000 in legal fees

By John Carreyrou

After working at Theranos Inc. for eight months, Tyler Shultz decided he had seen enough. On April 11, 2014, he emailed company founder Elizabeth Holmes to complain that Theranos had doctored research and ignored failed quality-control checks.

Source: The Wall Street Journal, November 18th, 2016

Frances Haugen

Image Source: Wikimedia Commons, CC BY SA 4.0, Creator: Stephan Rohl

Three levels of ethics

Personal Ethics	**What does it mean to be a good person? What is your moral compass?**
Professional Ethics	**What are the standards of behavior that govern my professional role? What regulates responsible conduct as a member of a profession?**

Social and Political Ethics

Compare Professional Ethics in Biomedical Research vs Computer Science Research

BIOMEDICAL RESEARCH

- Hippocratic oath: do no harm
- Professional licensure requirements
- Institutional Review Board (IRB) with ethics review
- Strong regulation by the Food and Drug Administration (FDA)
- Well-developed institutional footprint in universities and hospitals (bioethics as a scholarly field, ethics committees at hospitals)

COMPUTER SCIENCE RESEARCH

- ACM Code of Ethics
- No general licensing requirements
- No IRB for data science (except if research is done on human subjects or human generated data, as is often the case in HCI)
- No single regulatory body
- AI ethics is a growing body of work

Screenshot taken from faception.com home page, January 8, 2025

Professional Norms Governing Use of ML?

21st Century Scientific Racism?

Hashemi and Hall *J Big Data* (2020) 7:2
https://doi.org/10.1186/s40537-019-0282-4

Journal of Big Data

RESEARCH Open Access

Criminal tendency detection from facial images and the gender bias effect

Mahdi Hashemi[1*] and Margeret Hall[2]

*Correspondence:
mhashem2@gmu.edu
[1] Department of Information Sciences and Technology, George Mason University, 4400 University Dr, Fairfax, VA 22030, USA
Full list of author information is available at the end of the article

Abstract

Explosive performance and memory space growth in computing machines, along with recent specialization of deep learning models have radically boosted the role of images in semantic pattern recognition. In the same way that a textual post on social media reveals individual characteristics of its author, facial images may manifest some personality traits. This work is the first milestone in our attempt to infer personality traits from facial images. With this ultimate goal in mind, here we explore a new level of image understanding, inferring criminal tendency from facial images via deep learning. In particular, two deep learning models, including a standard feedforward neural network (SNN) and a convolutional neural network (CNN) are applied to discriminate criminal and non-criminal facial images. Confusion matrix and training and test accuracies are reported for both models, using tenfold cross-validation on a set of 10,000 facial images. The CNN was more consistent than the SNN in learning to reach its best test accuracy, which was 8% higher than the SNN's test accuracy. Next, to explore the classifier's hypothetical bias due to gender, we controlled for gender by applying only male facial images. No meaningful discrepancies in classification accuracies or learning consistencies were observed, suggesting little to no gender bias in the classifier. Finally, dissecting and visualizing convolutional layers in CNN showed that the shape of the face, eyebrows, top of the eye, pupils, nostrils, and lips are taken advantage of by CNN in order to classify the two sets of images.

Keywords: Image classification, Facial images, Convolutional neural network, Deep learning, Machine learning, Personality traits

CrowdTrack: A Benchmark for Difficult Multiple Pedestrian Tracking in Real Scenarios

Teng Fu
Fudan University
tfu23@m.fudan.edu.cn

Yuwen Chen
Fudan University
ywchen23@m.fudan.edu.cn

Zhuofan Chen
Fudan University
zfchen23@m.fudan.edu.cn

Mengyang Zhao
Fudan University
myzhao@fudan.edu.cn

Bin Li*
Fudan University
libin@fudan.edu.cn

Xiangyang Xue*
Fudan University
xyxue@fudan.edu.cn

Abstract

Multi-object tracking is a classic field in computer vision. Among them, pedestrian tracking has extremely high application value and has become the most popular research category. Existing methods mainly use motion or appearance information for tracking, which is often difficult in complex scenarios. For the motion information, mutual occlusions between objects often prevent updating of the motion state; for the appearance information, non-robust results are often obtained due to reasons such as only partial visibility of the object or blurred images. Although learning how to perform tracking in these situations from the annotated data is the simplest solution, the existing MOT dataset fails to satisfy this solution. Existing methods mainly have two drawbacks: relatively simple scene composition and non-realistic scenarios. Although some of the video sequences in existing dataset do not have the above-mentioned drawbacks, the number is far from adequate for research purposes. To this end, we propose a difficult large-scale dataset for multi-pedestrian tracking, shot mainly from the first-person view and all from real-life complex scenarios. We name it "CrowdTrack" because there are numerous objects in most of the sequences. Our dataset consists of 33 videos, containing a total of

Arxiv, July 2025

Use Cases for CrowdTrack?

- “Turnkey authoritarianism via AI”
- “Scalable authoritarianism”
  - One of the things that makes authoritarianism expensive is the overhead that comes from building out and running a large-scale police state. One of the things AI does is make it much, much cheaper to do large-scale surveillance. Datasets like CrowdTrack are a symptom of the way AI is making it cheaper and easier to do surveillance that the dictators of the 20th century would have fantasized about but always been unable to fully fund. “Our dataset can be used for tasks like visual grounding, captioning, and appearance feature extraction,” the researchers write.
- Jack Clark, Co-Founder Anthropic (in his Import AI newsletter)

Three levels of ethics

Personal Ethics	**What does it mean to be a good person? What is your moral compass?**
Professional Ethics	**What are the standards of behavior that govern my professional role? What regulates responsible conduct as a member of a profession?**
Social and Political Ethics	**How to create rules and laws that facilitate well-being, cooperation, and human rights? Governance, law, regulation, and policy.**

Third Lens: Social/Political Ethics

Identifying and refereeing value tradeoffs and the fair or legitimate process for making collective decisions

- liberty vs. equality
- convenience vs. privacy
- privacy vs. security/safety
- efficiency vs. human autonomy

And a bonus Q: **who gets to decide?**

→ These are questions of governance

Philosopher
ISAIAH BERLIN
1909-1997

“Total liberty for the wolves is certain death to the lambs.”

Today's Agenda

1. Why *You* are Taking this Class
2. Three Dimensions of Ethics
3. **Professional codes of ethics**
4. Disciplinary perspectives
5. The Optimization Mindset

ACM Code of Ethics

- First created in the early 1990's and adopted by ACM in 1992
- Updated in 2018
  - Catalyst was recent negative spotlight shined on the tech sector
- Breaking the guidelines of the Code of Ethics can result in immediate ACM membership termination
  - Who here are members of ACM?
  - Who here even knew what ACM was before this class?
  - Does membership termination carry any threat to you?

ACM Code of Ethics – Preamble

Computing professionals' actions change the world. **To act responsibly, they should reflect upon the wider impacts of their work, consistently supporting the public good.** The ACM Code of Ethics and Professional Conduct ("the Code") expresses the **conscience of the profession**.

The Code is designed to inspire and guide the ethical conduct of all computing professionals, including current and aspiring practitioners, instructors, **students**, influencers, and anyone who uses computing technology in an impactful way.

General Principles

1. Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing
2. Avoid harm ("harm" means negative consequences)
3. Be honest and trustworthy
4. Be fair and take action not to discriminate
5. Respect the work required to produce new ideas, inventions, creative works, and computing artifacts
6. Respect privacy
7. Honor confidentiality

Anything you would add, delete, or change?

Professional Responsibilities

1. Strive to achieve high quality in both the processes and products of professional work
2. Maintain high standards of professional competence, conduct, and ethical practice
3. Know and respect existing rules pertaining to professional work
4. Accept and provide appropriate professional review
5. Give comprehensive and thorough evaluations of computer systems and their impacts, including analysis of possible risks
6. Perform work only in areas of competence
7. Foster public awareness and understanding of computing, related technologies, and their consequences
8. Access computing and communication resources only when authorized or when compelled by the public good
9. Design and implement systems that are robustly and usably secure

Does this change your behavior?

Professional Licensing

- Many professions are licensed
  - Lawyers are admitted to the bar association
  - Doctors have to pass board exams
- Licensing bodies maintain ethical standards for profession
  - Violating ethical standard results in losing license
  - You can’t practice if your license is revoked
- Software Engineering is not a licensed profession
  - Although there is licensing of software engineers in Texas – only required if you sell your software directly, not if you work for a software company

Professional Licensing

- Should Software Engineering be a licensed profession?

- If so, what should the requirements be?
  - Do you need an accredited degree in CS (or related discipline)?
  - If so, then these people wouldn't qualify…

Professional Licensing

- Notion of licensing of Software Engineers was explored by an ACM panel in 1999

- A sample of their findings:
  - Licensing of software engineers provides no guarantee of software quality
  - Obtaining a professional license requires demonstrating mastery of a specified field of knowledge, usually requiring a test
  - Difficult to specify widely accepted body of knowledge defining competency in Software Engineering (which is rapidly changing)

- Panel could not reach consensus on licensing software engineers
  - Majority were opposed

Legal Liability

- Professionals are often legally liable for their work product
  - Medical and legal malpractice

- Should software professionals be liable for their work product
  - Should you be able to sue Microsoft for time lost if your word processor crashes? Or negative result of miscalculation in Excel?
  - Generally, you give up those rights when you agree to the End User Licensing Agreement (EULA)… that you don't read

- Excerpt from the Windows 11 EULA:

  Except for any repair, replacement, or refund that Microsoft, or the device manufacturer or installer, may provide, **you may not under this limited warranty, under any other part of this agreement, or under any theory, recover any damages or other remedy**, including lost profits or direct, consequential, special, indirect, or incidental damages.

Legal Liability

- So, should a doctor be able to have you agree to a EULA to perform a standard operation, so you are not entitled to any damages as a result of negative outcomes?

- Why is that acceptable for software?

Today's Agenda

1. Why *You* are Taking this Class
2. Three Dimensions of Ethics
3. Professional codes of ethics
4. Disciplinary perspectives: social science
5. The Optimization Mindset

Disciplinary Perspectives: the social scientist

A Social Scientist's View of the ACM Code

- It is a code.

  <u>What</u> is a code?

- It applies to computing professionals.

  <u>When</u> is that triggered?

- It is produced and governed by the ACM.

  <u>Where</u> does ACM come from?

- A nonprofit association that says the *public good* matters.

  <u>How</u> do we determine the public good?

- It threatens removal from the ACM for non-compliance.

  <u>Why</u> comply?

The Governance Spectrum

Free market

The Policy Menu

Free market

Ban

The Policy Menu

Free market

Self-regulation

Ban

The Policy Menu

Free market

Self-regulation

Ban

Self-regulation

=

Oxymoron?

The Policy Menu

Free market

Self-regulation

Private litigation

Ban

Radio Host Sues OpenAI for Libel After ChatGPT Accuses Him of Crime

The Georgia-based broadcaster claims defamation after ChatGPT allegedly said he embezzled funds.

By Jason Nelson

Jun 8, 2023

3 min read

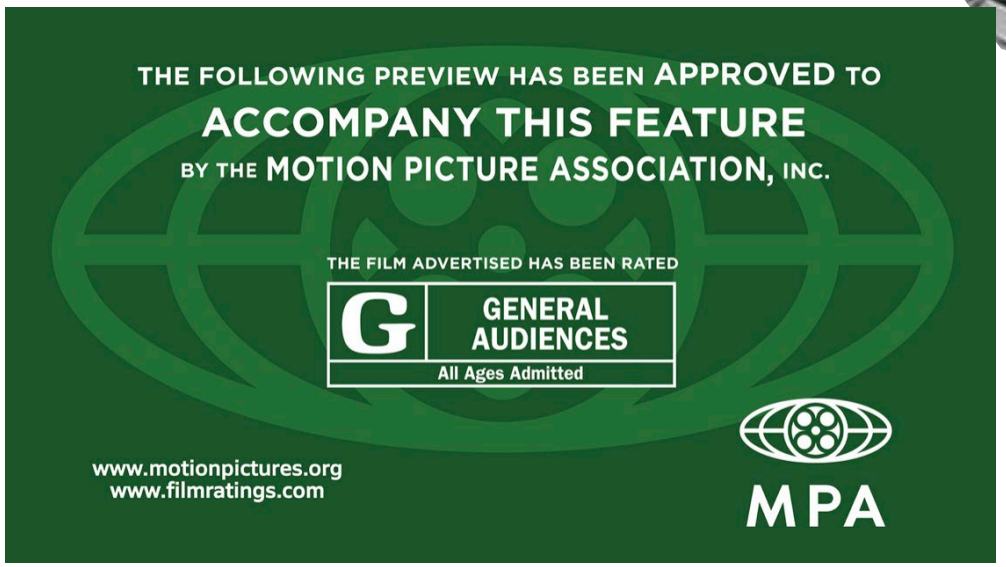
The Policy Menu

Free market

Self-regulation
Private litigation
Disclosure

Source: Imatag

Ban



THE RATINGS

G	All ages admitted	PARENTAL GUIDANCE SUGGESTED PG	Some material may not be suitable for children	PARENTS STRONGLY CAUTIONED PG-13	Some material may be inappropriate for children under 13	RESTRICTED R	Under 17 requires accompanying parent or adult guardian	ADULTS ONLY NC-17	No One 17 and Under Admitted

The Policy Menu

Free market

ethics.acm.org ACM Code of Ethics and Professional Conduct

Ban

The Policy Menu

Free market

Self-regulation
Private litigation
Disclosure
Licensing

Ban

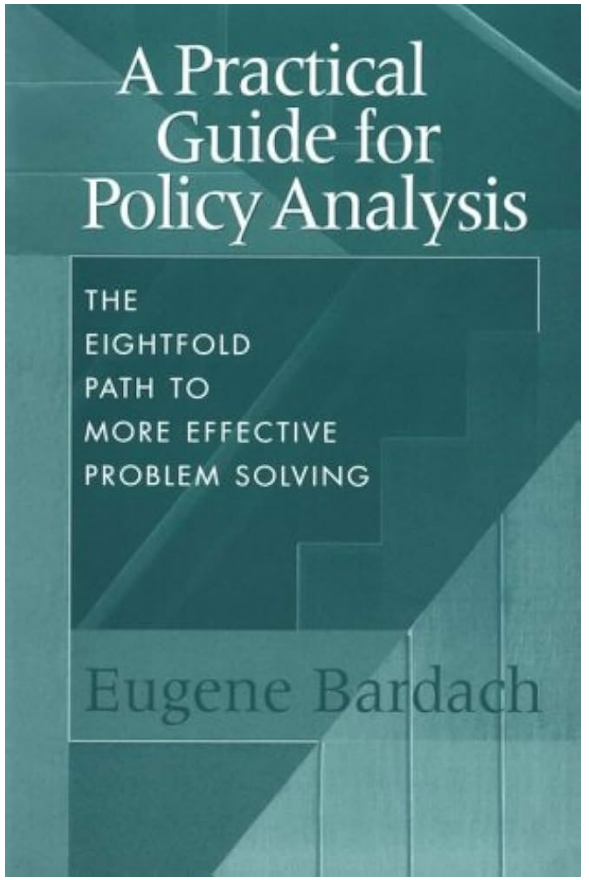
Source: ABC

The Policy Menu

Free market

Self-regulation
Private litigation
Disclosure
Licensing
Tax
Auditing
Premarket approval

Ban



1. Problem
2. Evidence
3. Alternatives
4. Criteria
5. Outcomes
6. Tradeoffs
7. Decide
8. Story

Evaluation

1. If we are going to weigh the benefits and costs of new technologies, we need to know what those benefits and costs actually are. This is a *measurement* challenge.

1. Once we assess the benefits and costs of new technologies, we need to make a judgment about what outcomes we want to achieve, what costs we are willing to accept, and how we will make a legitimate choice that is accepted by all. This is a *governance* challenge.

Social Science

- Methods for thinking systematically about the impacts of new technologies on human behaviors, interactions, and institutions

- Offers a framework for thinking about how new technologies are currently and perhaps should be governed

Example: Fake News

Source: The Verge

"Personally, I think the idea that fake news on Facebook, which is a very small amount of the content, influenced the election in any way – I think is a pretty crazy idea. Voters make decisions based on their lived experience."

-- Mark Zuckerberg,

November 10, 2016

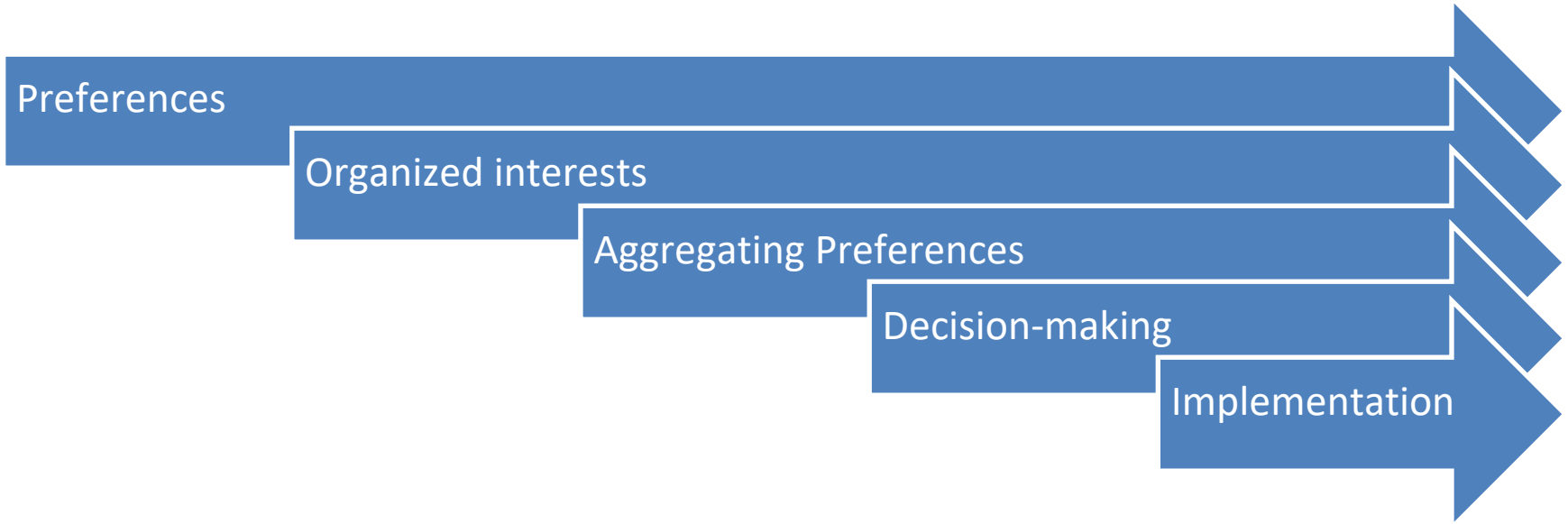
Example: Fake News

Before we can consider policy responses, there are a series of questions we might want to answer:

- How common is fake news?
- Are people able to identify fake news?
- Does exposure to fake news change people's underlying beliefs/views?
- Does fake news crowd out other kinds of information?
- Are people more or less likely to discount news that comes from a well-established media source?
- How much does the source matter as compared to the content of the news?
- What strategies might counteract the effect of fake news?

Breaking Governance Down

What does it mean to govern new technologies? It requires thinking about:



Who are the different players? Who are the winners? Who might lose? How do divergent preferences get reconciled in our institutions?

Today's Agenda

1. Why *You* are Taking this Class
2. Three Dimensions of Ethics
3. Professional codes of ethics
4. Disciplinary perspectives
5. **The Optimization Mindset**

The Optimization Mindset

The Engineer's Error? Misunderstanding Optimization

It's better to get something done efficiently rather than inefficiently.

It's better still to optimize.

"Everything is an optimization problem"

Conclusion: efficiency/maximization/optimization are good, even intrinsically good.

“Everything is an optimization problem.”

WIRED

ERICA KLARREICH SCIENCE 01.30.2013 09:30 AM

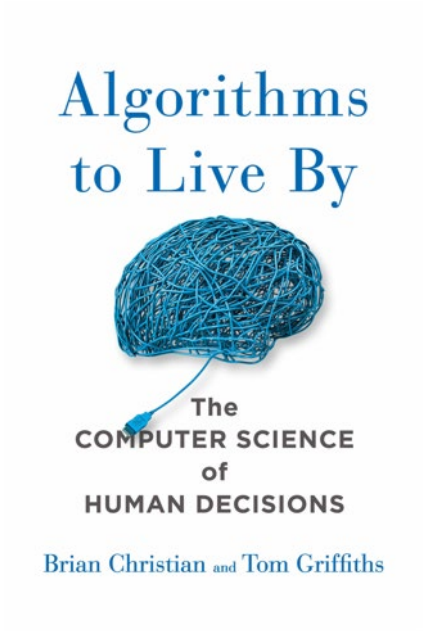
Computer Scientists Find New Shortcuts for Infamous Traveling Salesman Problem

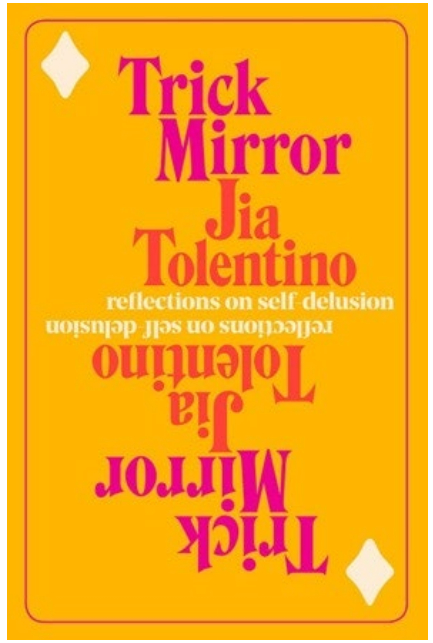
The traveling salesman problem asks: Given a collection of cities connected by highways, what is the shortest route that visits every city and returns to t
starting place? The answer has practical applications to processes such as drilling holes in circuit boards, scheduling tasks on a computer and ordering
features of a genome. Now, a long-sought advance in the traveling salesman problem is breathing new life into the search for improved approximate solu

The shortest traveling salesman route going through all 13,509 cities in the United States with a population of at least 500 (as of 1998). ILLUSTRATION: COURTESY OF DAVID APPLEGATE, ROBERT BIXBY, VASEK CHVATAL AND WILLIAM COOK

Source: Wired Magazine, January 30, 2013; Flickr CC BY 2.0

Optimization becomes an orientation to life





Algorithms to Live By by Brian Christian and Tom Griffiths; Lifehacker.com May 19, 2021; Trick Mirror by Jia Tolentino

What's the Value of Efficiency?

Efficiency/maximization/optimization are
derivative or **second-order** values.

They are instrumentally good.
Not intrinsically or fundamentally good.

Optimization is an instrument to some end or goal. We must independently assess the worthiness of that end or goal.

BLOGS BOOKS THE NEW ATLANTIS ABOUT SUBSCRIBE

The Architecture of Evil

On the delusion that technical work is morally neutral

Roger Forsgren

AGRICULTURAL HISTORY

VOLUME IV APRIL, 1930 NUMBER 2

ECONOMIC EFFICIENCY AND COMPETITIVE ADVANTAGES OF SLAVERY UNDER THE PLANTATION SYSTEM

BY L. C. GRAY[1]

Sources The New Atlantis Summer 2012; Agricultural History by LC Gray, April 1930)

Three Problems of the Optimization Mindset

1. The Problem of Bad Ends/Goals/Objectives
2. The Problem of Finding Measurable Proxies for Good Goals
3. The Problem of Multiple Valuable Goals

Problem 1: means over ends

The Problem of Bad Ends/Goals/Objectives

- Doing something efficiently is only a means to accomplish an end (or an objective).
- From a moral standpoint, what's important is the goodness or badness of end, not the means to accomplish it.
- Obsessing about efficiency can lead to optimizing for bad ends. And thereby making the world worse, not better.
- Also, many technologies are dual purpose: have good and bad use cases

Problem 2: Proxies and Measuring What's Meaningful

2. The Problem of Finding Measurable Proxies for Good Goals

What should an engineer optimize?

Need something *computationally tractable* and *representationally adequate* – a quantifiable measure of what we are trying to accomplish.

But goals like "happiness" or "safety" or "connecting people" are not easy to measure.

Facebook's Mission Statement

Facebook VP on Measuring Progress

So we connect more people

That can be bad if they make it negative. Maybe it costs a life by exposing someone to bullies. Maybe someone dies in a terrorist attack coordinated on our tools.

And still we connect people.

The ugly truth is that we believe in connecting people so deeply that **anything that allows us to connect more people more often is *de facto* good.** It is perhaps the only area where the metrics do tell the true story as far as we are concerned.

Source: BuzzFeed News March 2018

Problem 3: Multiple Valuable Goals

3. The Problem of Multiple Valuable Goals

Optimizing for one (good) end can threaten other valuable goals.

soylent

SHOP ABOUT BLOG EMERGENCY

STORE LOCATOR

The Soylent Story

From Silicon Valley Powdered Meals to International Complete Nutrition Platform.

“I started wondering why something as simple and important as food was still so inefficient, given how streamlined and optimized other modern things are.”

Rob Rhinehart, Co-Founder

What's Wrong With Soylent? (P.S. It's not the taste)

About a week and a half ago, I began drinking Soylent every day. I can't recommend that you do the same. For a purported breakthrough with such grand plans for reshaping the food industry, I found Soylent to be a punishingly boring, joyless product. From the plain white packaging to the purposefully bland, barely sweet flavor to the motel-carpet beige hue of the drink itself, everything about Soylent screams function, not fun. It may offer complete nourishment, but only at the expense of the aesthetic and emotional pleasures many of us crave in food.

Source: Farhad Manjoo, NY Times, May 28, 2014

The Multiple Purposes of Food

Bodily nutrition

Social connection

Gustatory Pleasure

Cultural identity

Optimize for one alone, and you have a problem.