Lecture 21: Finale

CS 5150, Spring 2026

Fall 2026 Courses: CS 6158

- CS 6158: Software Engineering in the Era of Machine Learning
- **Instructor**: Saikat Dutta
- **Goals**:
  - Study state-of-the-art research ideas in SExML
  - Hands-on exposure to Software Engineering research
    - Apply machine learning-based techniques to solve software engineering problems
    - Apply automated software engineering techniques to machine learning systems.
  - Develop and implement new research ideas
- Fall 2025 version: https://www.cs.cornell.edu/courses/cs6158/2025fa

Fall 2026 Courses: CS 5154

- CS 5154: Software Testing
- **Instructor**: Owolabi Legunsen
- **Goals**:
  - Deep dive into testing: regression testing, unit testing, mutation analysis, …
  - Design and automate the execution of high-quality software tests.
  - Generate test suites that meet coverage and other adequacy criteria.
- **Project**: Extend your 5150 project to focus on testing
- Apply for TA!

Lecture Goals

- Few notes about Ethics and Professionalism

Professionalism & Ethics

What should you do if you discover a major security vulnerability in a piece of widely-used software?

PollEv.com/cs5150sp26

Which of these development efforts would you be comfortable contributing to?

- Drug marketing campaign
- Click fraud
- Selling 0-day vulnerabilities
- Reverse engineering
- Weaponized AI
- Selling personal data
- Bitcoin mining

See: Computer Fraud and Abuse Act (CFAA).

Ethics

- Software can harm society beyond physical injury
- Personal fulfilment is important too
  - Take responsibility for your work
  - Avoid future regrets

- Compared to traditional engineering, software has ***less oversight*** and ***wider impact***
  - Amplification: One day's work can affect millions of people, consume millions of hours

Diversity

- Wider impact => more diverse user base
  - => More potential to reinforce stereotypes, inequity
- Failure to anticipate/respond to **biased systems** can lead to major societal (not to mention reputational) harm
- Need to **expand** diversity during development (shift left)
  - More diverse developer teams
  - More diverse user testing
- "*Single source of truth*" does not apply to human society
  - Disputed borders
  - Different interpretations of words/phrases/symbols
  - Different value systems

Ethics extends beyond code

- Hiring practices
  - Beware affinity bias, groupthink
- Promotions/opportunities
  - Beyond mentoring - advocate for coworkers who do good work but seem to go unnoticed
- Decision-making
  - Don’t defend decisions solely on precedent
  - Look beyond direct “bottom line” impact

Three questions to ensure human well-being

1. Does my software respect the humanity of the users?
2. Does my software amplify positive behavior, or negative behavior for users and society at large?
3. Will my software's quality impact the humanity of others?

Taken from **Human Flourishing**: According to Harvard's Human flourishing program: Human flourishing is composed of five central domains: **happiness and life satisfaction, mental and physical health, meaning and purpose, character and virtue, and close social relationships.**
https://hfh.fas.harvard.edu/

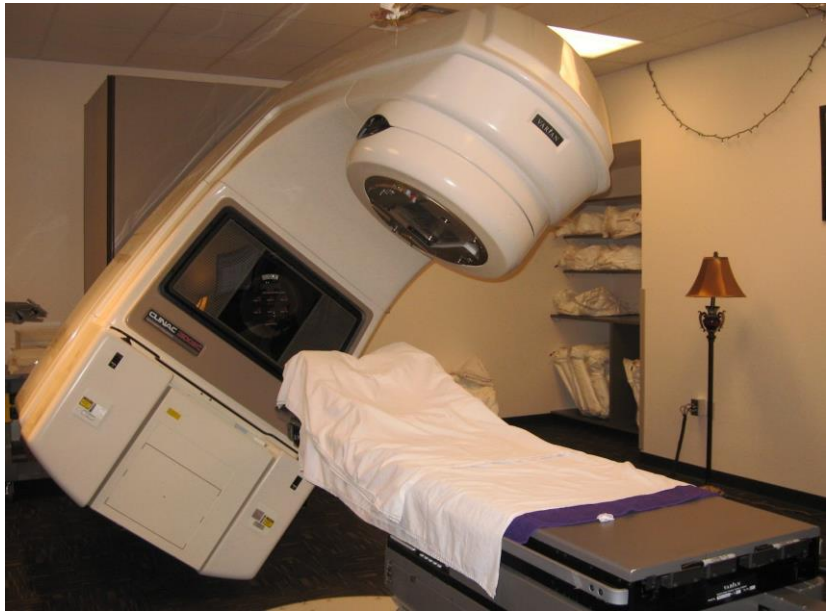
Ship or Not?

You deploy a feature that increases engagement by **20%** but **worsens** user well-being—
ship or not?

Safety, Reliability, and Harm

- Software Failures impact real people in the world (sometimes include moral factors)
- Ethical practices should be followed throughout the software lifecycle – design to operation
- Collaboration with legal teams
- "*If the development team is measured on their rate of feature development, there's a **high probability** that the **ethics of a given implementation** might not be front of mind, either at the design or at the implementation phase.*" -- Tim Mackey, head of software supply chain risk strategy at Black Duck, an application security platform

Therac-25



Goal: Radiation therapy for cancer patients

Bug (race-condition) in software lead to at least 6 deaths (overdosing)

Traced to:
- Lack of reporting bugs
- Lack of proper due diligence
- Engineers were overconfident, removed hardware locks
- Race condition of 8 seconds could lead to problems

Q: Who is responsible for errors?

Q: What steps are creators of software morally required to take to minimize the risk that they will sell flawed software with dangerous consequences?

Volkswagen Scandal

How Volkswagen's 'Defeat Devices' Worked

By GUILBERT GATES, JACK EWING, KARL RUSSELL and DEREK WATKINS UPDATED March 16, 2017

Volkswagen admitted that 11 million of its vehicles were equipped with software that was used to cheat on emissions tests. This is how the technology works and what it now means for vehicle owners. RELATED ARTICLE

"The software sensed when the car was being tested and then activated equipment that reduced emissions, United States officials said. But the software turned the equipment down during regular driving, increasing emissions far above legal limits, most likely to save fuel or to improve the car's torque and acceleration. "

*" … The researchers found that when tested on the road, some cars emitted almost **40 times** the permitted levels of nitrogen oxides… "*

*VW eventually paid **$30 Billion** in fines …*

https://www.nytimes.com/interactive/2015/business/international/vw-diesel-emissions-scandal-explained.html

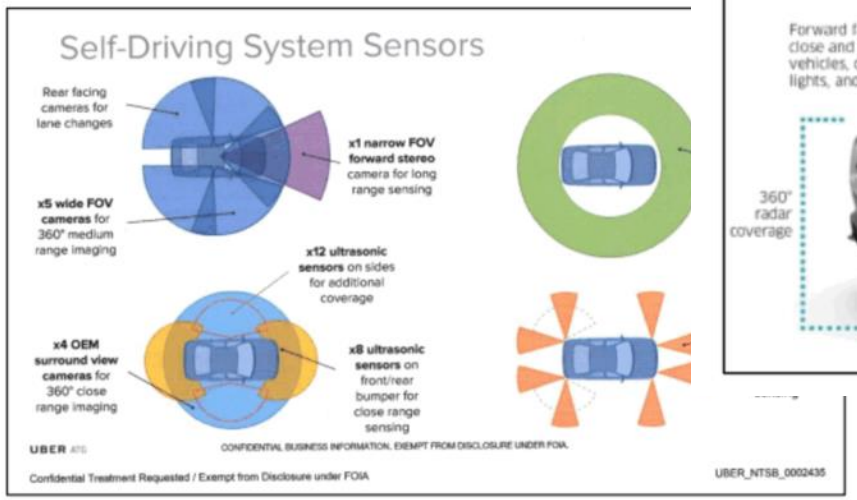
Uber self-driving car involved in fatal crash couldn't detect jaywalkers

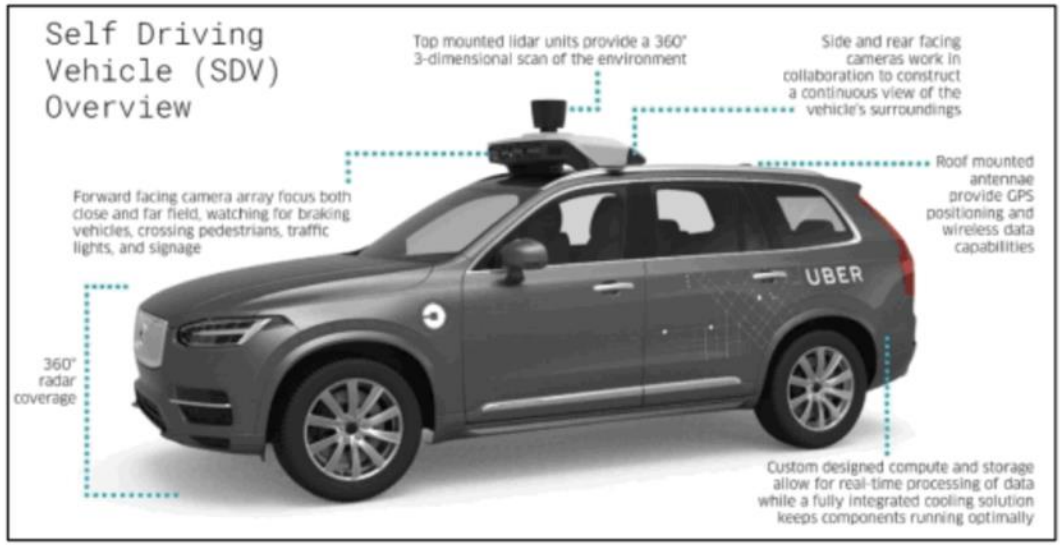
The system had several serious software flaws, the NTSB said.

Steve Dent, @stevetdent
11.06.19 in Transportation

25
Comments

1131
Shares





https://www.npr.org/2019/11/07/777438412/feds-say-self-driving-uber-suv-did-not-recognize-jaywalking-pedestrian-in-fatal-

Privacy and Data Ethics

- ”Anonymous” Data isn’t anonymous
  - Netflix Prize Dataset (2007) case: Researchers link this dataset with **public data sources, individual are re-identified**
- Even aggregated data can reveal patterns (Strav fitness tracking data)
- Internal access is a privacy risk—not just externa leaks (Internal tool at **Uber** allowed employees t track riders in real time)
- **Cambridge Analytica** (collected 87 M FB user da – used for political targeting in 2016 elections)

Algorithmic Bias/Fairness

- Technology can amplify existing biases
- Computers themselves have no inherent moral framework.
- Software can only reflect the biases of its creators.

How AI systems amplify bias

Image recognition systems that use biased machine learning data sets will inadvertently magnify that bias. Researchers are examining ways to reduce the effects.

COOKING

ROLE	VALUE
AGENT	WOMAN
FOOD	PASTA
HEAT	STOVE
TOOL	SPATULA
PLACE	KITCHEN

COOKING

ROLE	VALUE
AGENT	WOMAN
FOOD	FRUIT
HEAT	—
TOOL	KNIFE
PLACE	KITCHEN

COOKING

ROLE	VALUE
AGENT	WOMAN
FOOD	MEAT
HEAT	GRILL
TOOL	TONGS
PLACE	OUTSIDE

COOKING

ROLE	VALUE
AGENT	WOMAN
FOOD	VEGETABLES
HEAT	STOVE
TOOL	TONGS
PLACE	KITCHEN

COOKING

ROLE	VALUE
AGENT	MAN
FOOD	—
HEAT	STOVE
TOOL	SPATULA
PLACE	KITCHEN

In this example of gender bias, adapted from a report published by researchers from the University of Virginia and the University of Washington, a visual semantic role labeling system has learned to identify a person cooking as female, even when the image is male.

IMAGES (FROM LEFT): ALEXEY STIOP/ADOBE STOCK, JACOB LUND/ADOBE STOCK, BERND JÜRGENS/ADOBE STOCK, GRANGER WOOTZ/GETTY IMAGES

Currently, the AI portrait generator has been trained mostly on portraits of people of European ethnicity. We're planning to expand our dataset and fix this in the future. At the time of conceptualizing this AI, authors were not certain it would turn out to work at all. This is close to state of the art in AI at the moment.

Sorry for the bias in the meanwhile. Have fun!

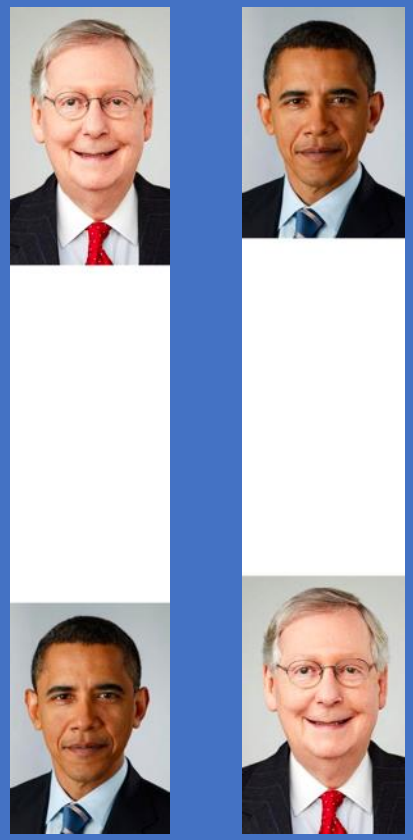
MMitchell
@mmitchell_ai

I really love the active discussion abt the role of ethics in AI, spurred by Google Gemini's text-to-image launch & its relative lack of white representation. As one of the most experienced AI ethics people in the world (>4 years! ha), let me help explain what's going on a bit.

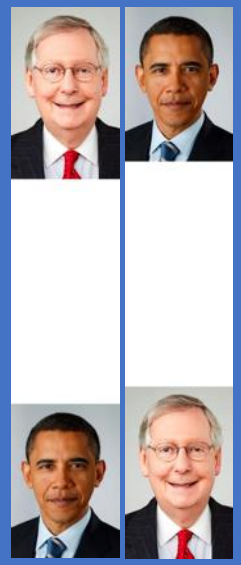
4:08 PM · Feb 25, 2024 · **182.3K** Views

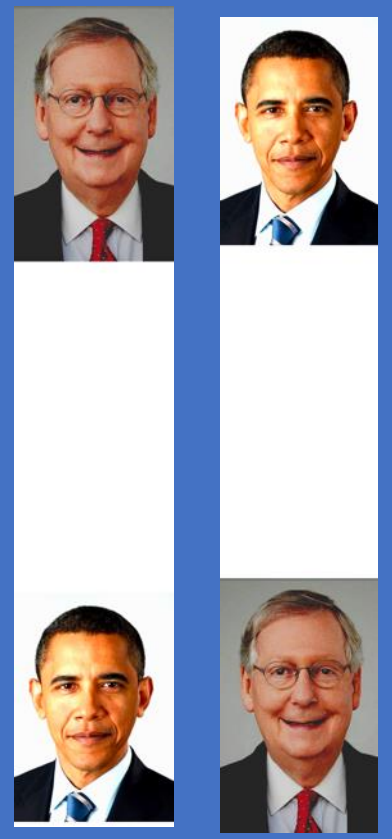
36 231 795 731

Twitter cropping photos



Twitter cropping photos





Security and Responsible Disclosure

- Security is often an afterthought
- Security failures can lead to **privacy violations** (Anthem medical data breach 2015), **financial loss** (Bitcoin losses), **physical harm**
- *Q: Once you know about the security vulnerability, do you protect users, protect the company, or protect yourself?*

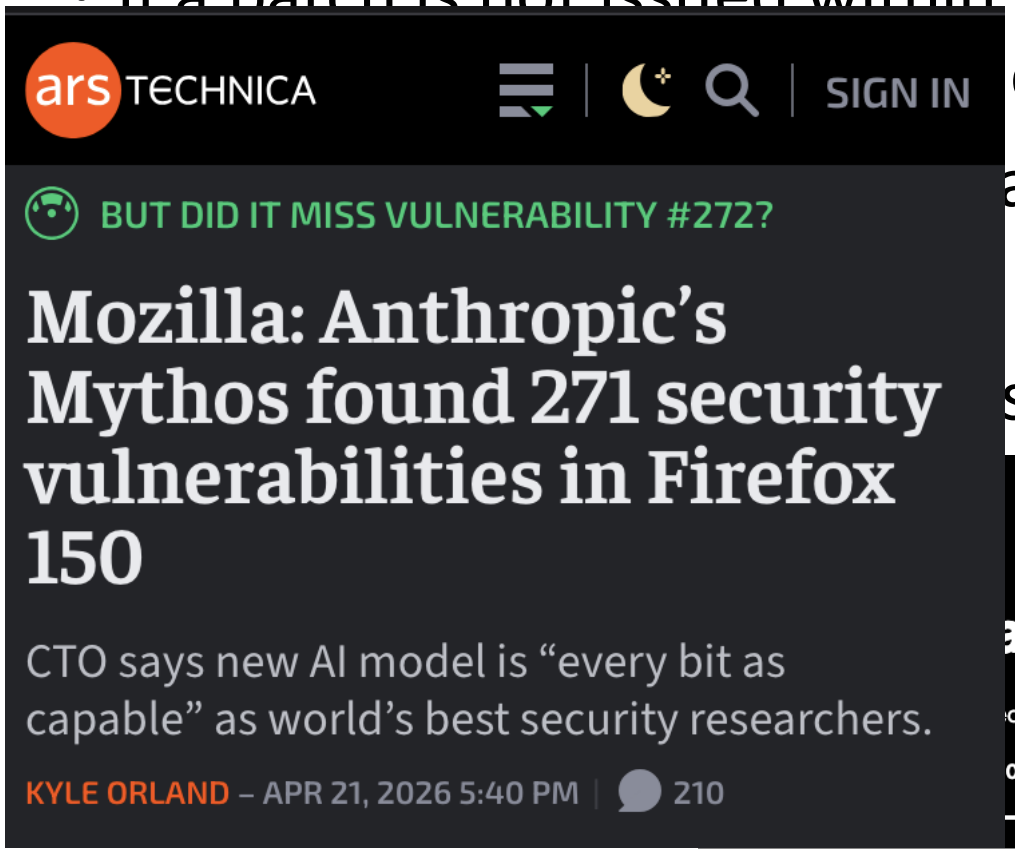
Responsible disclosure (or CVD)

- AKA "coordinated vulnerability disclosure"
- Coordinate timing of announcement with vendor
  - Give them time to patch products, prepare press response
  - Upper bound on timing to hasten vendor action (typ. 90 days)
- For open-source projects, look for security policy (SECURITY.md)
  - Contact Vulnerability Management Team or owner
  - Do not post details to public mailing lists, chat rooms
- May be assigned placeholder CVE to coordinate efforts without disclosing details

https://github.blog/2022-02-09-coordinated-vulnerability-disclosure-cvd-open-source-projects/

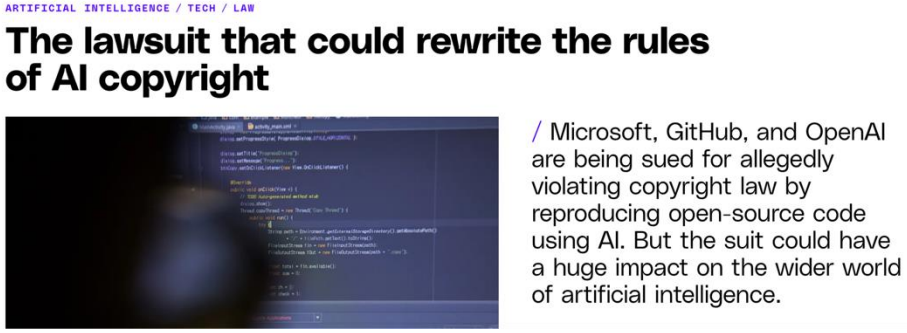
Google Project Zero disclosure policy

- Project Zero quietly notifies vendor with security leak information
- If a patch is not issued within 90 days, the vulnerability details are
  cly (otherwise 30 days after patch)
  adline applies
  soft, and others) to prioritize patching!



Intellectual Property & Open Source

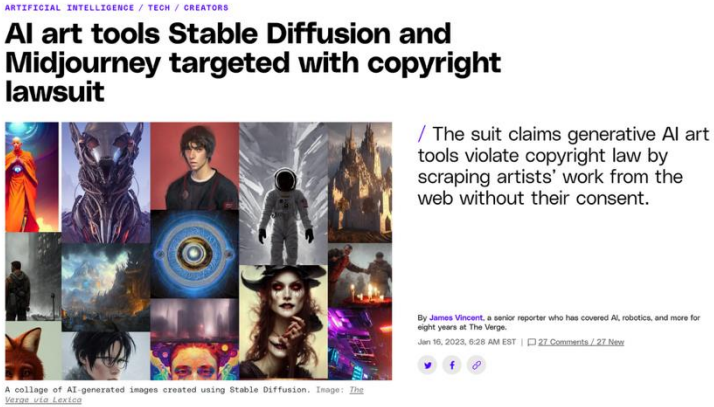
- Was the data used to train these LLMs obtained illegally?
- Who owns the IP associated with LLM outputs?
- Should sensitive information be provided as inputs to LLMs?



ARTIFICIAL INTELLIGENCE / TECH / LAW

The lawsuit that could rewrite the rules of AI copyright

/ Microsoft, GitHub, and OpenAI are being sued for allegedly violating copyright law by reproducing open-source code using AI. But the suit could have a huge impact on the wider world of artificial intelligence.



ARTIFICIAL INTELLIGENCE / TECH / CREATORS

AI art tools Stable Diffusion and Midjourney targeted with copyright lawsuit

/ The suit claims generative AI art tools violate copyright law by scraping artists' work from the web without their consent.

By James Vincent, a senior reporter who has covered AI, robotics, and more for eight years at The Verge.

Jan 16, 2023, 6:28 AM EST | 27 Comments / 27 New

A collage of AI-generated images created using Stable Diffusion. Image: The Verge via Lexica

Whoops, Samsung workers accidentally leaked trade secrets via ChatGPT

ChatGPT doesn't keep secrets.

By Cecily Mauran on April 6, 2023

Workplace Ethics & Professional Pressure

- Deadlines, metrics, and management shape decisions
- Ethical issues often look like: “Just ship it” , “We’ll fix it later”, “Don’t raise this externally”
- **Theranos case**: Employees knew blood-testing tech was unreliable
- Organizational ethics:
  - What should employers be able to track about employees? Keystrokes, Screenshots, Location? (Org. Culture)
  - **Typical Values:** Honesty, Fairness, Accountability, Respect
  - **Do your values align with organizational values?**

https://lpsonline.sas.upenn.edu/features/ethics-modern-workplace-lessons-organizational-culture-and-collaboration

Activity: Discuss (5 mins)

- You are a senior engineer at a startup building **SafeRoute**, an AI-powered navigation app for urban commuters. The app is about to launch publicly.
- SafeRoute recommends routes based on:
  - Shortest time
  - Traffic conditions
  - "Safety score" (based on historical crime data + user reports)
- ***Your manager says: "This feature is our differentiator. We need to launch now—we'll improve it later."***
- You discover: Bias in "Safety Score": Neighborhoods with higher minority populations are consistently labeled "unsafe" even when actual incidents are comparable (Cause: biased data).
- **Q1**: Would you: Ship, Ship with warning, Delay, Remove Safety Feature?
- **Q2**: Is a biased system better than no system?
- **Q3**: If your manager insists on shipping, what would you do?

ACM Code of Ethics and professional practice

1. PUBLIC – Software engineers shall act consistently with the public interest.
2. CLIENT AND EMPLOYER – Software engineers shall act in a manner that is in the best interests of their client and employer consistent with the public interest.
3. PRODUCT – Software engineers shall ensure that their products and related modifications meet the highest professional standards possible.
4. JUDGMENT – Software engineers shall maintain integrity and independence in their professional judgment.
5. MANAGEMENT – Software engineering managers and leaders shall subscribe to and promote an ethical approach to the management of software development and maintenance.
6. PROFESSION – Software engineers shall advance the integrity and reputation of the profession consistent with the public interest.
7. COLLEAGUES – Software engineers shall be fair to and supportive of their colleagues.
8. SELF – Software engineers shall participate in lifelong learning regarding the practice of their profession and shall promote an ethical approach to the practice of the profession.

https://ethics.acm.org/code-of-ethics/software-engineering-code

ACM Code of Ethics

As an ACM member I will ....

Contribute to society and human well-being.

Avoid harm to others.

Be honest and trustworthy.

Be fair and take action not to discriminate.

Honor property rights including copyrights and patent.

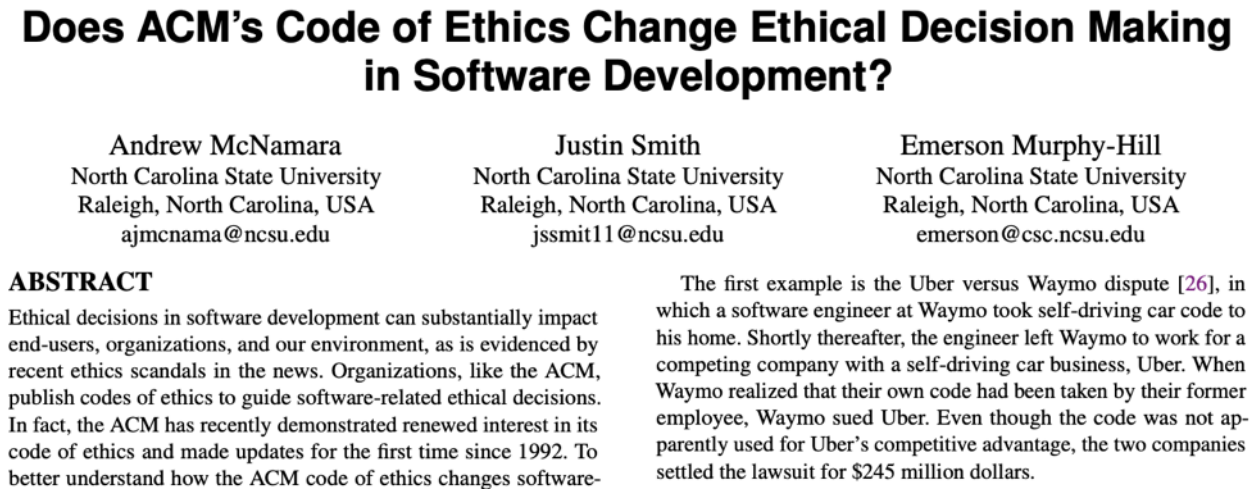
Give proper credit for intellectual property.

Respect the privacy of others.

Honor confidentiality.

Code of Ethics

Research shows that the code of ethics does not appear to affect the decisions made by software developers.



Does ACM's Code of Ethics Change Ethical Decision Making in Software Development?

Andrew McNamara
North Carolina State University
Raleigh, North Carolina, USA
ajmcnama@ncsu.edu

Justin Smith
North Carolina State University
Raleigh, North Carolina, USA
jssmit11@ncsu.edu

Emerson Murphy-Hill
North Carolina State University
Raleigh, North Carolina, USA
emerson@csc.ncsu.edu

ABSTRACT

Ethical decisions in software development can substantially impact end-users, organizations, and our environment, as is evidenced by recent ethics scandals in the news. Organizations, like the ACM, publish codes of ethics to guide software-related ethical decisions. In fact, the ACM has recently demonstrated renewed interest in its code of ethics and made updates for the first time since 1992. To better understand how the ACM code of ethics changes software-

The first example is the Uber versus Waymo dispute [26], in which a software engineer at Waymo took self-driving car code to his home. Shortly thereafter, the engineer left Waymo to work for a competing company with a self-driving car business, Uber. When Waymo realized that their own code had been taken by their former employee, Waymo sued Uber. Even though the code was not apparently used for Uber's competitive advantage, the two companies settled the lawsuit for $245 million dollars.

How do we enforce these codes?

Challenge

- How do we apply ethics to a field (Software Engineering) that is changes so often?
- The incidents might predate the actual laws …
- To handle this uncertainty about the future, let’s focus on three questions we can ask to remind ourselves to focus on promoting human flourishing.

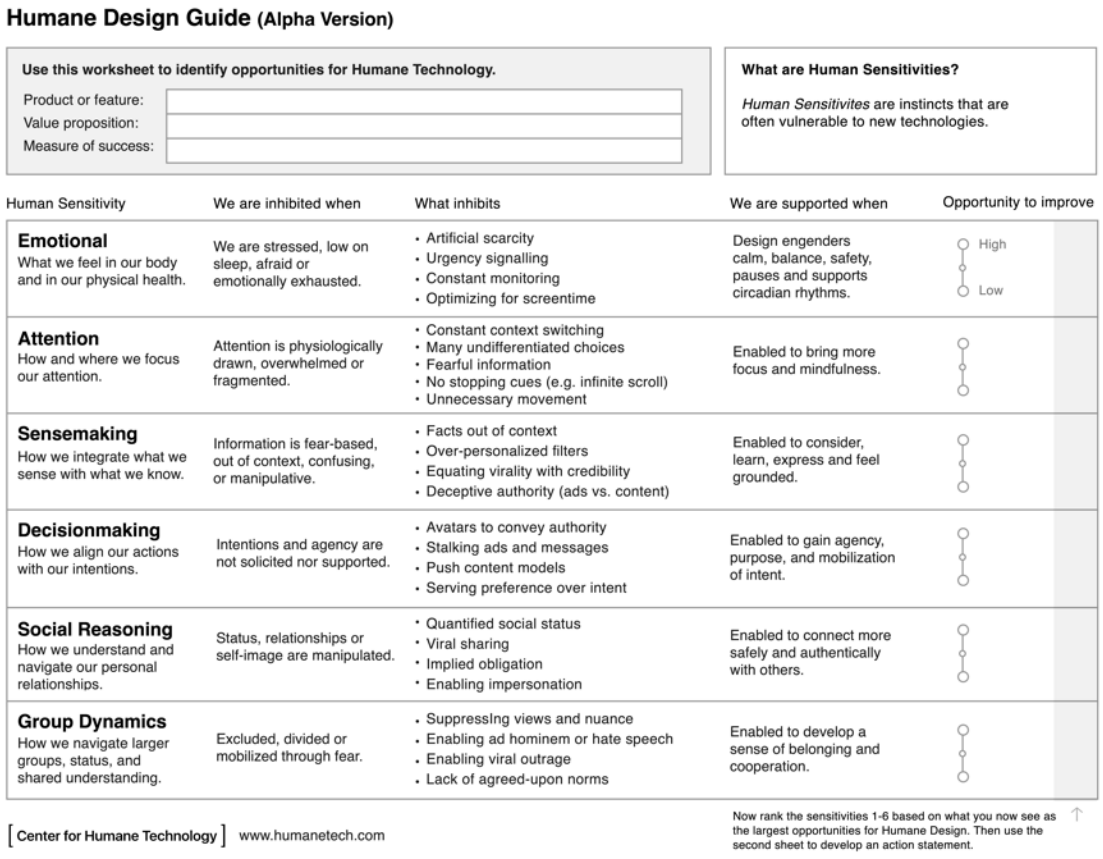
Three questions to ensure human well-being

1. Does my software respect the humanity of the users?
2. Does my software amplify positive behavior, or negative behavior for users and society at large?
3. Will my software's quality impact the humanity of others?

Taken from **Human Flourishing**: According to Harvard's Human flourishing program: Human flourishing is composed of five central domains: **happiness and life satisfaction, mental and physical health, meaning and purpose, character and virtue, and close social relationships.**
https://hfh.fas.harvard.edu/

1.Does my software respect the **humanity** of the **users**?

Humane Design Guide



Humane Design Guide (Alpha Version)

Use this worksheet to identify opportunities for Humane Technology.

Product or feature:
Value proposition:
Measure of success:

What are Human Sensitivities?

Human Sensitivites are instincts that are often vulnerable to new technologies.

Human Sensitivity	We are inhibited when	What inhibits	We are supported when	Opportunity to improve
Emotional What we feel in our body and in our physical health.	We are stressed, low on sleep, afraid or emotionally exhausted.	• Artificial scarcity • Urgency signalling • Constant monitoring • Optimizing for screentime	Design engenders calm, balance, safety, pauses and supports circadian rhythms.	High / Low
Attention How and where we focus our attention.	Attention is physiologically drawn, overwhelmed or fragmented.	• Constant context switching • Many undifferentiated choices • Fearful information • No stopping cues (e.g. infinite scroll) • Unnecessary movement	Enabled to bring more focus and mindfulness.	
Sensemaking How we integrate what we sense with what we know.	Information is fear-based, out of context, confusing, or manipulative.	• Facts out of context • Over-personalized filters • Equating virality with credibility • Deceptive authority (ads vs. content)	Enabled to consider, learn, express and feel grounded.	
Decisionmaking How we align our actions with our intentions.	Intentions and agency are not solicited nor supported.	• Avatars to convey authority • Stalking ads and messages • Push content models • Serving preference over intent	Enabled to gain agency, purpose, and mobilization of intent.	
Social Reasoning How we understand and navigate our personal relationships.	Status, relationships or self-image are manipulated.	• Quantified social status • Viral sharing • Implied obligation • Enabling impersonation	Enabled to connect more safely and authentically with others.	
Group Dynamics How we navigate larger groups, status, and shared understanding.	Excluded, divided or mobilized through fear.	• Suppressing views and nuance • Enabling ad hominem or hate speech • Enabling viral outrage • Lack of agreed-upon norms	Enabled to develop a sense of belonging and cooperation.	

[Center for Humane Technology] www.humanetech.com

Now rank the sensitivities 1-6 based on what you now see as the largest opportunities for Humane Design. Then use the second sheet to develop an action statement.

Humane Design Guide

Provides a template for considering a piece of software, and asking questions to help us arrive at a “humane design”

Consider 6 human sensitivities: Emotional, Attention, Sense making, Decision making, Social Reasoning, and Group Dynamics

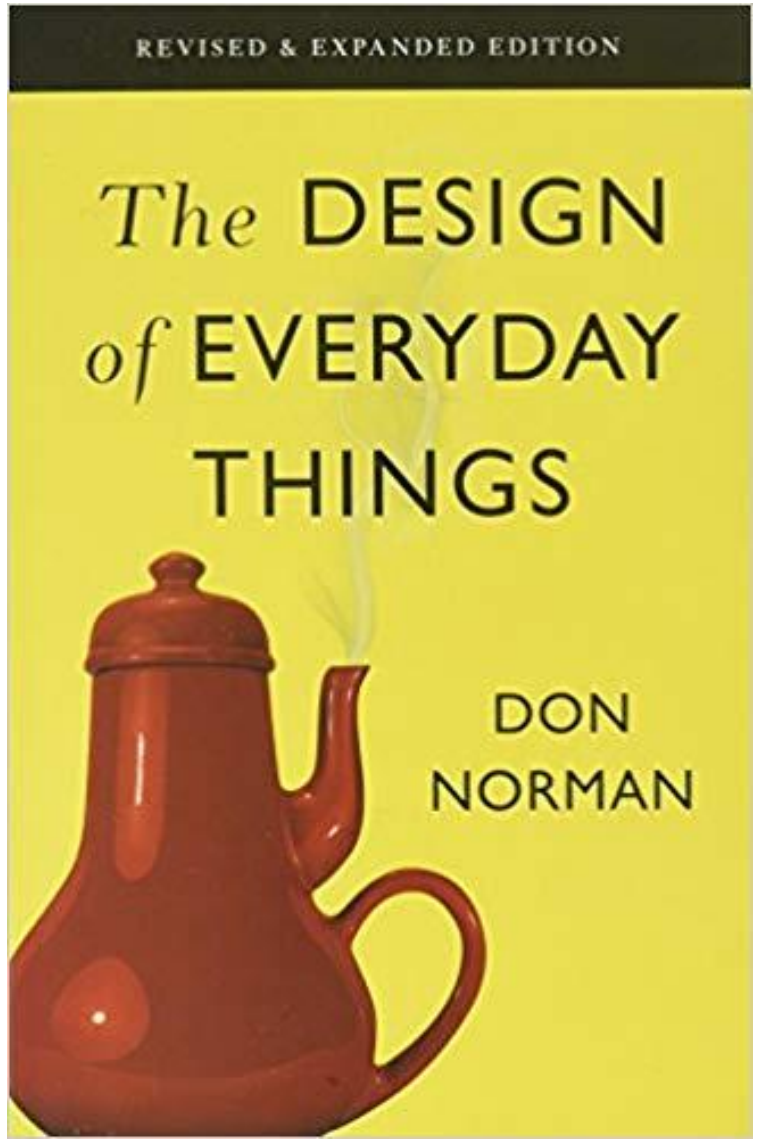
Human Sensitivity	We are inhibited when	What inhibits	We are supported when	Opportunity to improve
Attention How and where we focus our attention.	Attention is physiologically drawn, overwhelmed or fragmented.	• Constant context switching • Many undifferentiated choices • Fearful information • No stopping cues (e.g. infinite scroll) • Unnecessary movement	Enabled to bring more focus and mindfulness.	

Identify Opportunities to improve

User Centred Design

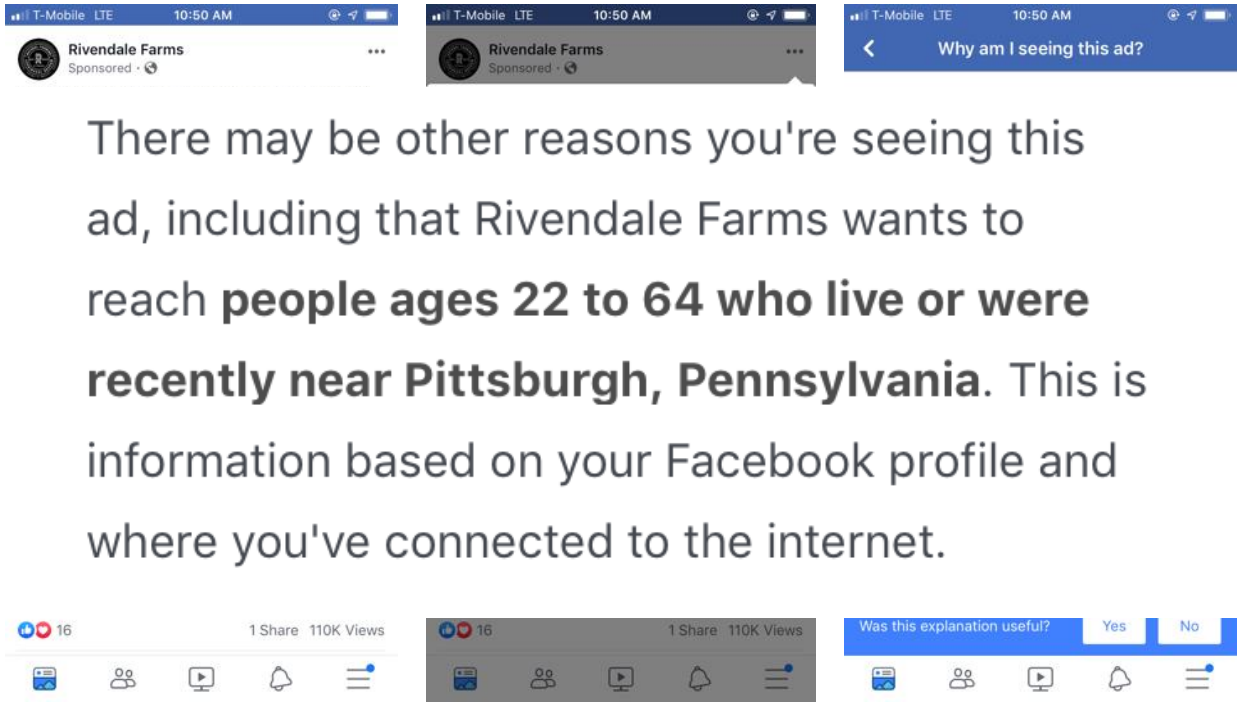
User-centered design tries to optimize the product around how users can, want, or need to use the product, rather than forcing the users to change their behavior to accommodate the product.

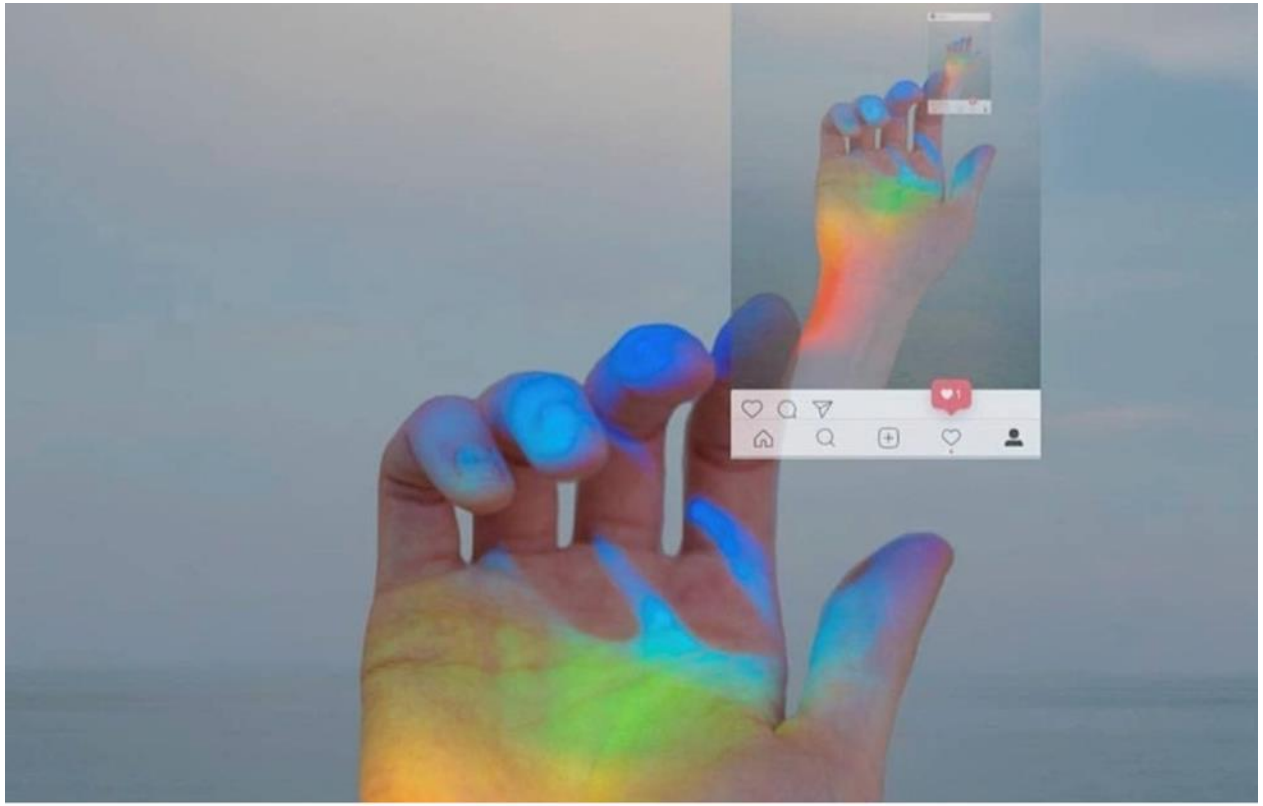
-Wikipedia



2.Does my software amplify positive or negative behavior for users and society at large?

Explain “why” to customers





@dovneon

What Instagram removing likes may mean for influencers and our self-esteem

SCIENCE & TECH - FEATURE

The decision could have a positive impact on the way people use the platform, but harm those trying to use it professionally

Anil Dash on how to prevent abuse

- You should have real humans dedicated to monitoring and responding to your community.
- You should have community policies about what is and isn't acceptable behavior.
- Your site should have accountable identities.
- You should have the technology to easily identify and stop bad behaviors.
- You should make a budget that supports having a good community, or you should find another line of work.

http://anildash.com/2011/07/20/if_your_websites_full_of_assholes_its_your_fault-2/

3. Will my software’s quality impact the humanity of others?

Engineering ethics

Ethics applies and is formalized in many professional fields: medical, legal, business, and engineering.

The first codes of engineering ethics were formally adopted by American engineering societies in 1912-1914. In 1946 the National Society of Professional Engineers **(NSPE)** adopted their first formal Canons of Ethics.

“Engineers shall hold paramount the safety, health, and welfare of the public.”

“Engineers shall perform services only in the areas of their competence”

“Engineers shall at all times strive to serve the public interest”

…

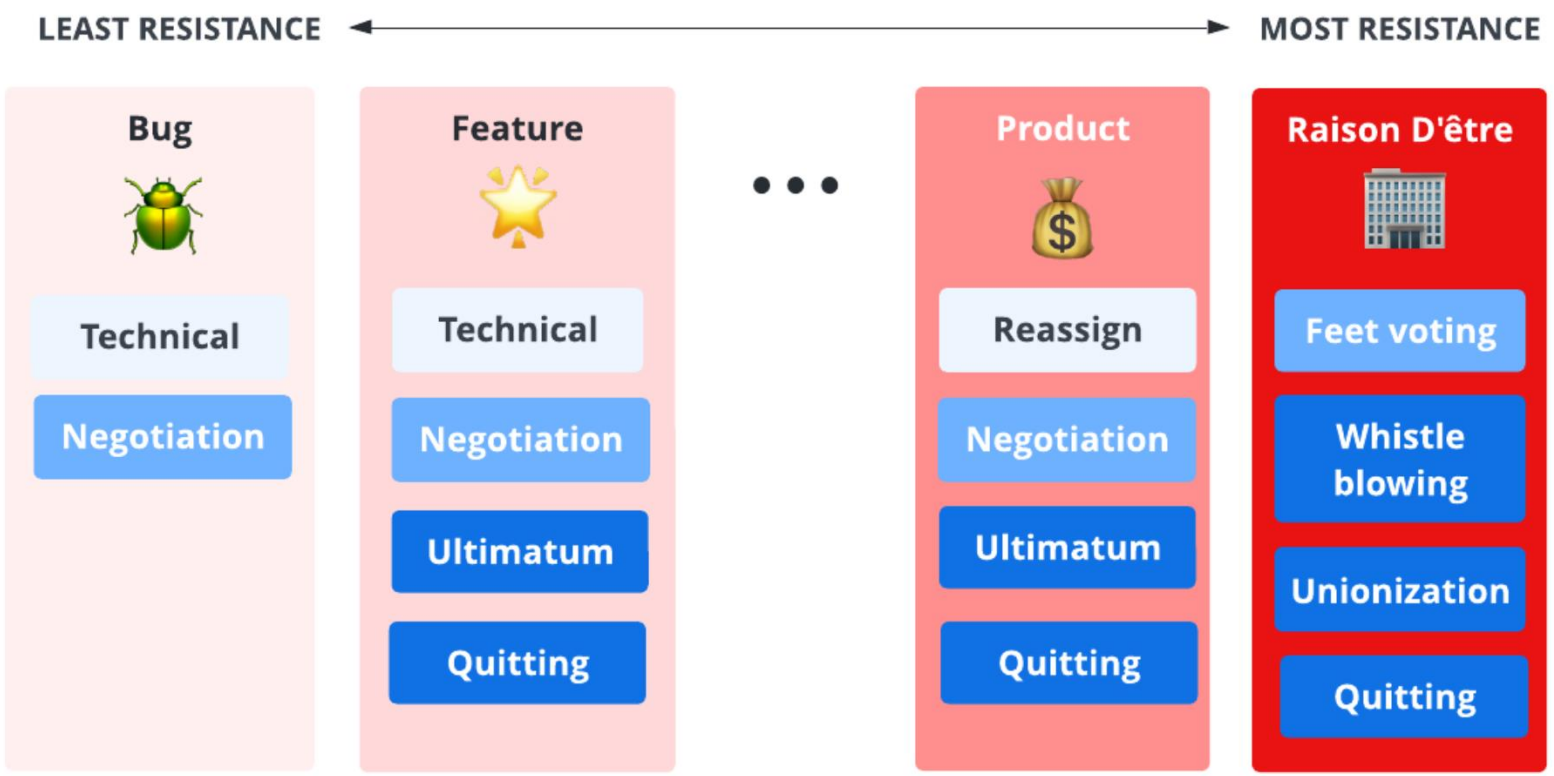
https://www.nspe.org/career-growth/ethics/nspe-code-ethics-engineers

Professional Engineers

Professional Engineers typically do a 4 year degree and take rigorous exams to obtain an engineering license

What {is / could be} the role of **professional engineers** in software?

Different scope of concerns addressed differently



What can we do for Ethics in software?

Most traditional emphasis of “engineering ethics”

What can we learn from other professions?

Should software have ”Professional Engineers”?

How do we define “safety critical systems”?

How much testing is enough? How can we convince others to do that much testing?

These questions are the **start** of the **conversation**, but as technology evolves, we must be **vigilant** to ensure we are promoting human flourishing