Ethics

Claire Le Goues

Michael Hilton

institute for SOFTWARE RESEARCH | Carnegie Mellon University School of Computer Science

Administrivia

- Homework 1 grades released
  - Overall quality was high
  - Commit messages could have been improved
  - Link to specific commit, not repo in general
  - You should stage your commit properly (hint: careful when using ”git add .”)
  - We are looking for reflections to be specific (assumptions are ok), vs vague

Administrivia

- Homework 2
  - CI questions - We were not able to cover CI as early as we wanted. We will not require you to implement a CI pipeline for HW2.
  - You will need to answer:
    - What (if anything) did you do for QA during HW2? How do you feel about your QA? Was it good enough? What could have been done better?
  - In future assignments you will implement a CI pipeline.

Learning goals

- Awareness of ethical issues in software engineering
- Reflection on decision making
- Questions to ask when evaluating the ethics of software
- Starting points to dig deeper

Ethics

Activity

- Go to breakoutrooms
- Share Unethical situations with each other
- Post to slack your andewID's as well as a summary of each unethical situation

Volkswagen Scandal

VW was caught cheating on emissions for Diesel engines

What is 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**.

Why Human Flourishing?

- Universal Declaration of Human Rights: “All human beings are born free and equal in dignity and rights.”
- Declaration of Independence: “We hold these truths to be self-evident…”
- Internal Compass
- Faith

(Un)Ethical situations

EA calls its loot boxes ‘surprise mechanics,’ says they’re used ethically

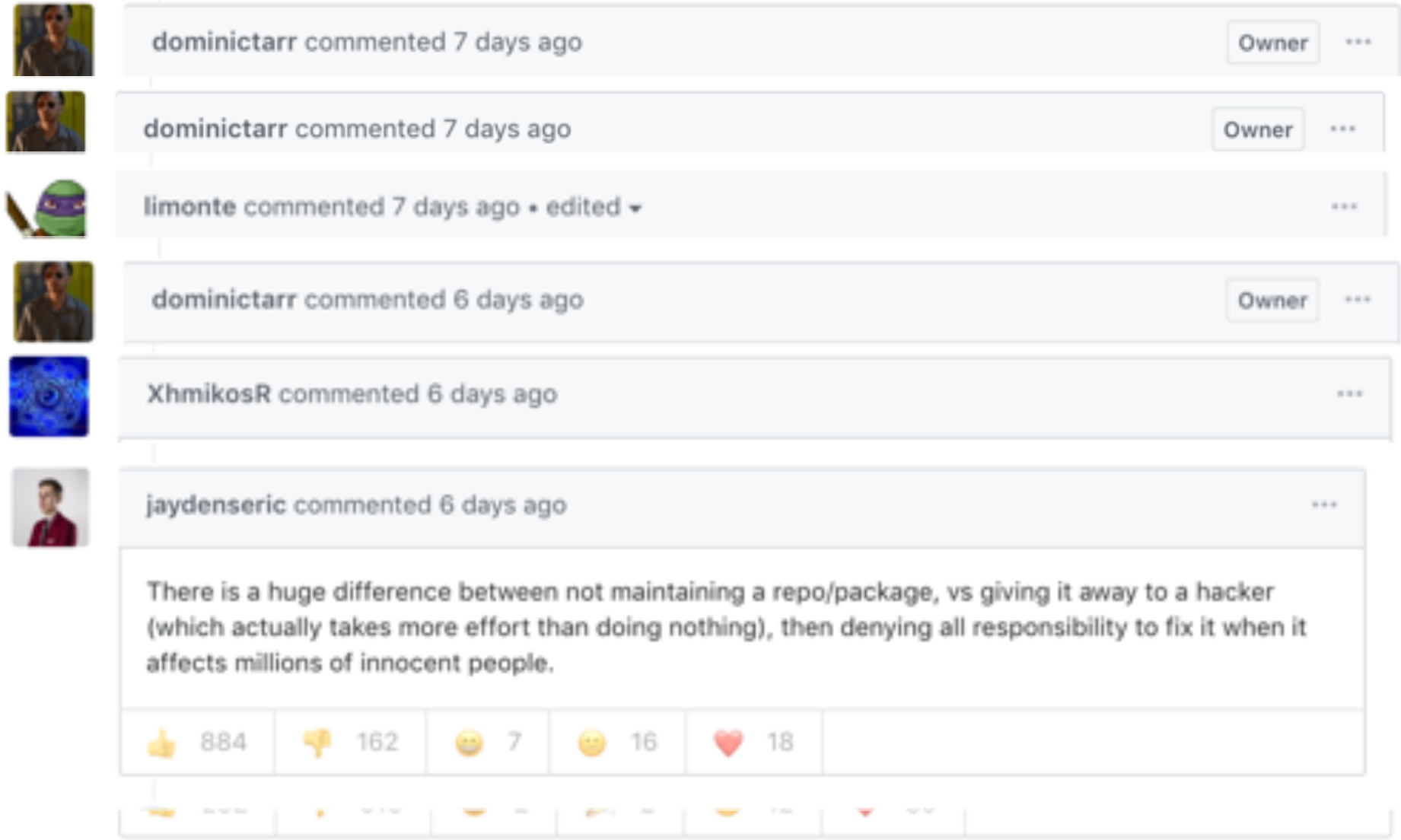
80

‘People like surprises,’ executive tells UK Parliament

By Ana Diaz | @AnaLikesPikachu | Jun 21, 2019, 9:10am EDT

SHARE

Open Source Maintainers



Domino's Would Rather Go to the Supreme Court Than Make Its Website Accessible to the Blind

Rather than developing technology to support users with disabilities, the pizza chain is taking its fight to the top

by Brenna Houck | @EaterDetroit | Jul 25, 2019, 6:00pm EDT

SHARE

Airlines

Some airlines may be using algorithms to split up families during flights

Your random airplane seat assignment might not be random at all.

By Aditi Shrikant | aditi@vox.com | Nov 27, 2018, 6:10pm EST

SHARE

Passengers boarding a Boeing aircraft of the low cost airline carrier Ryanair in Thessaloniki Macedonia Airport, Greece. | Nicolas Economou/NurPhoto/Getty Images

TC

Login

Startups
Apps
Gadgets
Videos
Audio
Extra Crunch
Newsletters
Events
Advertise

—

Crunchbase
More

Search

Facebook privacy
Transportation
Enterprise
Def Con 2019

Lime halts scooter service in Switzerland after possible software glitch throws users off mid-ride

Ingrid Lunden @ingridlunden / 9:51 am EST • January 12, 2019

Comment

xing.com search for “Brand Strategist”

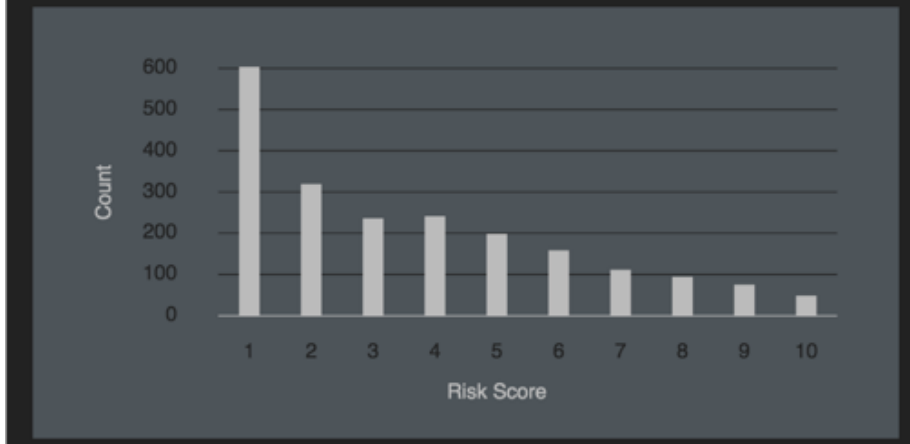
Search query	Work experience	Education experience	Profile views	Candidate	Xing ranking
Brand Strategist	146	57	12992	male	1
Brand Strategist	327	0	4715	female	2
Brand Strategist	502	74	6978	male	3
Brand Strategist	444	56	1504	female	4
Brand Strategist	139	25	63	male	5
Brand Strategist	110	65	3479	female	6
Brand Strategist	12	73	846	male	7
Brand Strategist	99	41	3019	male	8
Brand Strategist	42	51	1359	female	9
Brand Strategist	220	102	17186	female	10

Lahoti, Preethi, Krishna P. Gummadi, and Gerhard Weikum. “iFair: Learning Individually Fair Data Representations for Algorithmic Decision Making.” 2019 IEEE 35th International Conference on Data Engineering (ICDE) (2019

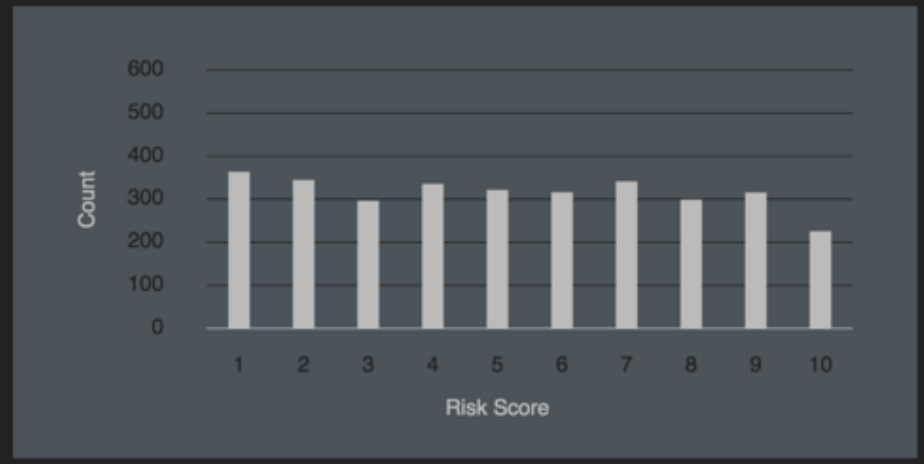
Twitter cropping photos

Twitter cropping photos

White Defendants' Risk Scores



Black Defendants' Risk Scores



Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Algorithmic Bias

Algorithms affect:

Where we go to school

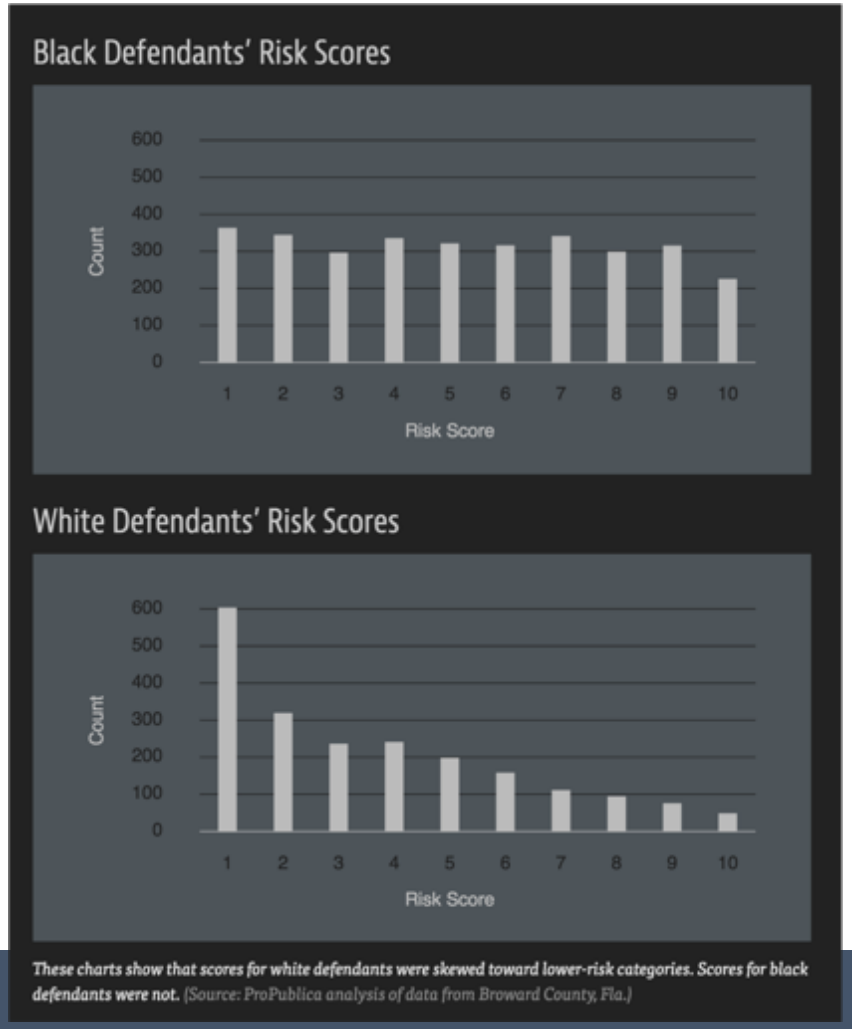
Access to money

Access to health care

Receiving parole

Possibility of Bail

Risk Scores



These charts show that scores for white defendants were skewed toward lower-risk categories. Scores for black defendants were not. *(Source: ProPublica analysis of data from Broward County, Fla.)*

Therac-25

Bug in software lead to at least 6 deaths

Traced to:
Lack of reporting bugs
Lack of proper due diligence
Engineers were overconfident

The New York Times

SUBSCRIBE NOW LOG IN

BUSINESS DAY

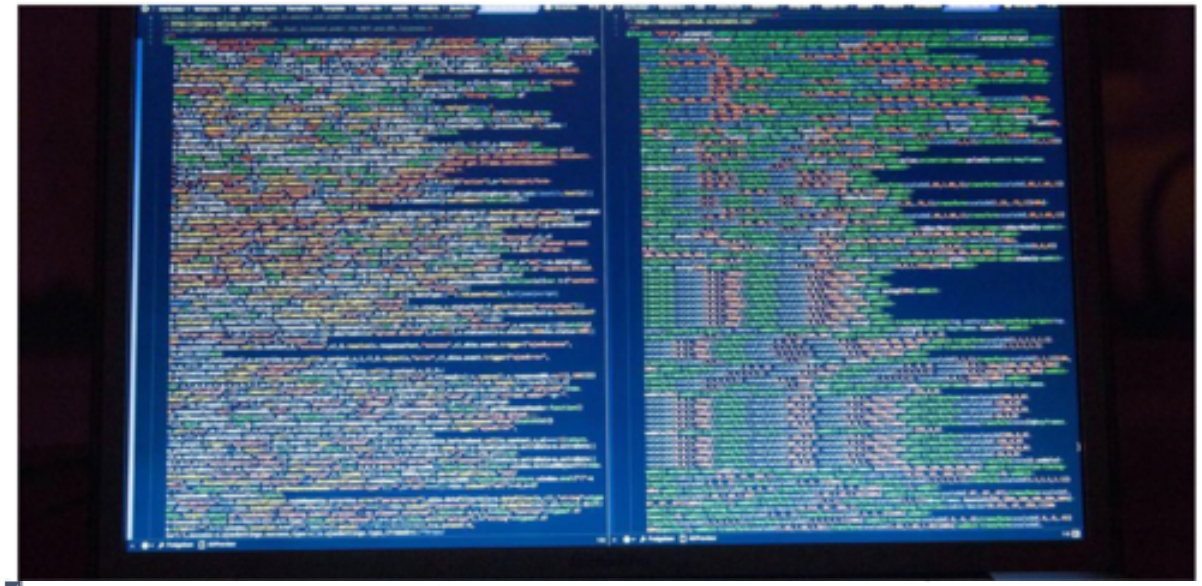
4,331 views | Oct 17, 2018, 06:13pm

We Need To Work Harder To Make Software Engineering More Ethical

Jessica Baron Contributor

Consumer Tech

I write about the ethics of science and technology.



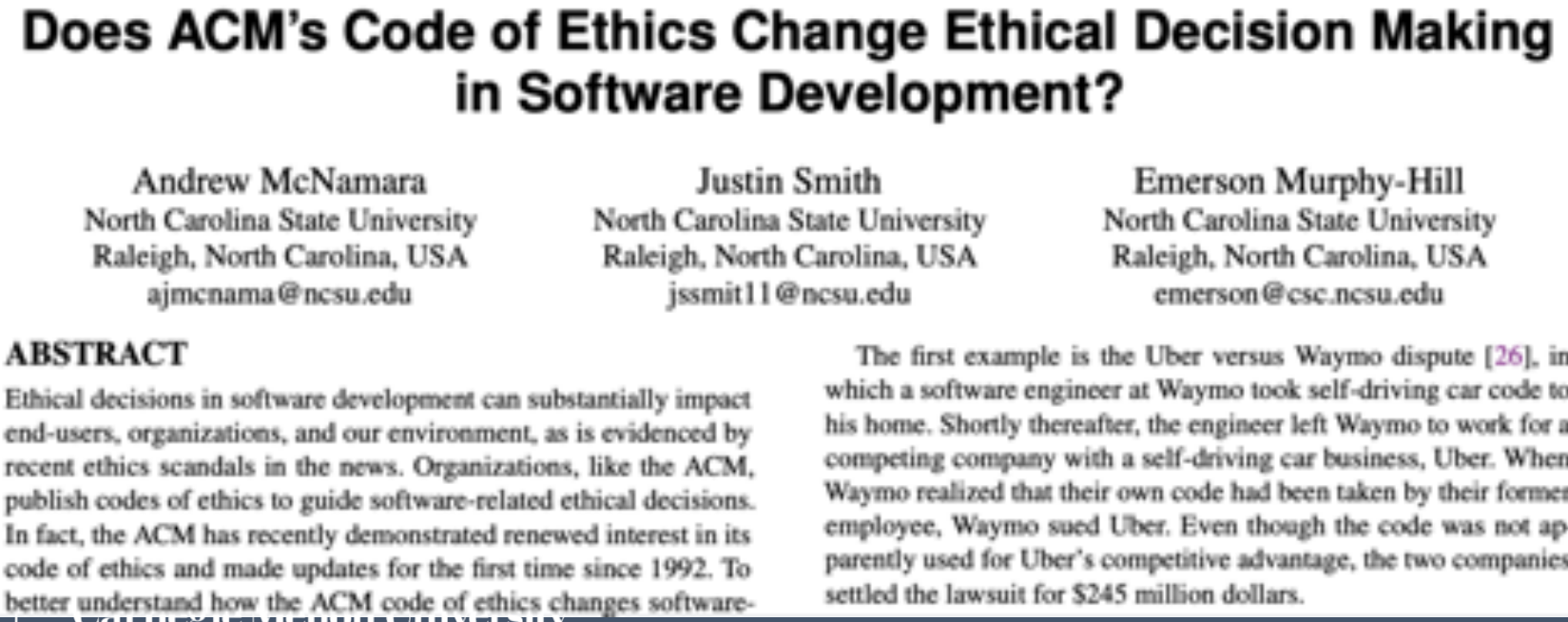
patch the software, but you can't patch a person if you, you know, damage someone's reputation." Sam Hodgson for The New York Times

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.

Challenge:

How do we apply ethics to a field (Software Engineering) that is changes so often?

Remember the Dominos case? The ADA law was written before the first website (1990)

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 promote human flourishing

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?

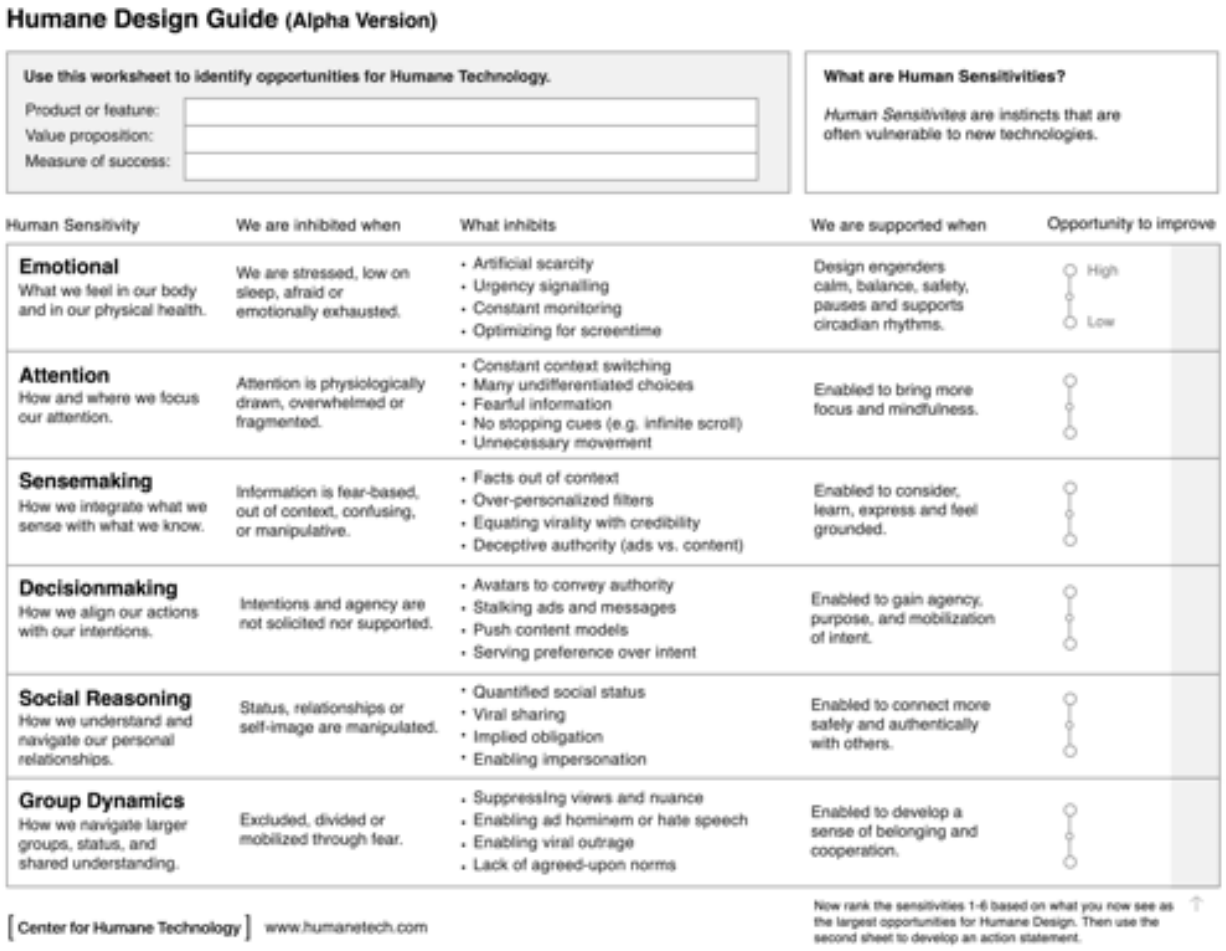
1.Does my software respect the **humanity** of the **users**?

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

Lets consider some tools and processes…

Humane Design Guide

http://humanetech.com



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
http://humanetech.com

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

Humane Design Guide

http://humanetech.com

After analysis step, develop plan of action:

1. In what ways does your product/feature currently engage Human Sensitivities?

2. How might your product/feature support or elevate human sensitivities?

3. Action Statement

GenderMag
https://gendermag.org

Abby Jones[1]

You can edit anything in blue print

- 28 years old
- Employed as an Accountant
- Lives in Cardiff, Wales

Abby has always liked music. When she is on her way to work in the morning, she listens to music that spans a wide variety of styles. But when she arrives at work, she turns it off, and begins her day by scanning all her emails first to get an overall picture before answering any of them. (This extra pass takes time but seems worth it.) Some nights she exercises or stretches, and sometimes she likes to play computer puzzle games like Sudoku

Background and skills

Abby works as an accountant. She is comfortable with the technologies she uses regularly, but she just moved to this employer 1 week ago, and their software systems are new to her.

Abby says she's a "numbers person", but she has never taken any computer programming or IT systems classes. She likes Math and knows how to think with numbers. She writes and edits spreadsheet formulas in her work.

In her free time, she also enjoys working with numbers and logic. She especially likes working out puzzles and puzzle games, either on paper or on the computer

Motivations and Attitudes

- ***Motivations:*** Abby uses technologies to accomplish her tasks. She learns new technologies if and when she needs to, but prefers to use methods she is already familiar and comfortable with, to keep her focus on the tasks she cares about.
- ***Computer Self-Efficacy:*** Abby has low confidence about doing unfamiliar computing tasks. If problems arise with her technology, she often blames herself for these problems. This affects whether and how she will persevere with a task if technology problems have arisen.
- ***Attitude toward Risk:*** Abby's life is a little complicated and she rarely has spare time. So she is risk averse about using unfamiliar technologies that might need her to spend extra time on them, even if the new features might be relevant. She instead performs tasks using familiar features, because they're more predictable about what she will get from them and how much time they will take.

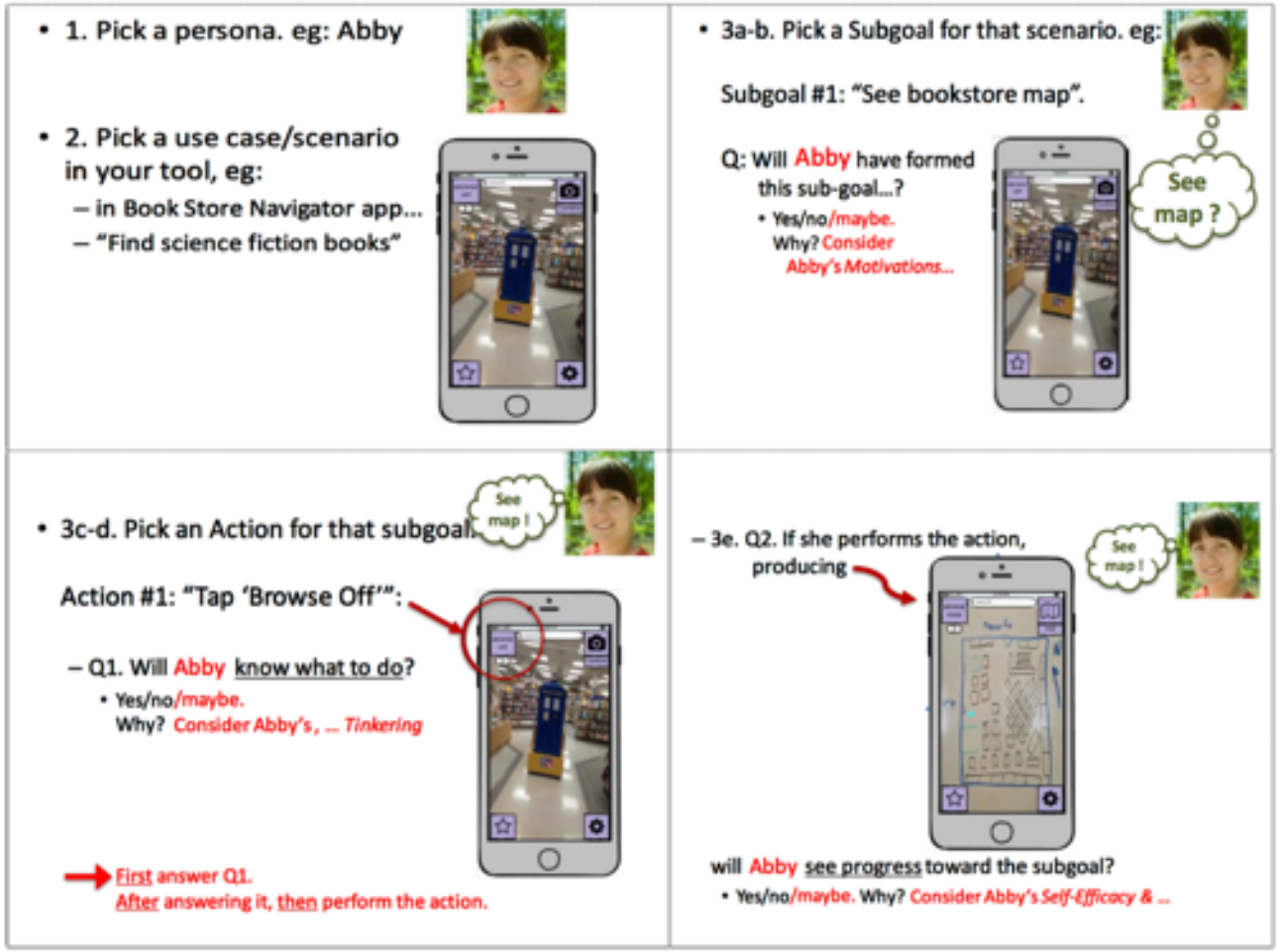
How Abby Works with Information and Learns:

- ***Information Processing Style:*** Abby tends towards a *comprehensive information processing style* when she needs to more information. So, instead of acting upon the first option that seems promising, she gathers information comprehensively to try to form a complete understanding of the problem before trying to solve it. Thus, her style is "burst-y"; first she reads a lot, then she acts on it in a batch of activity.
- ***Learning: by Process vs. by Tinkering:*** When learning new technology, Abby leans toward process-oriented learning, e.g., tutorials, step-by-step processes, wizards, online how-to videos, etc. She doesn't particularly like learning by tinkering with software (i.e., just trying out new features or commands to see what they do), but when she does tinker, it has positive effects on her understanding of the software.

[1] Abby represents users with motivations/attitudes and information/learning styles similar to hers. For data on females and males similar to and different from Abby, see http://eusesconsortium.org/gender/gender.php

GenderMag

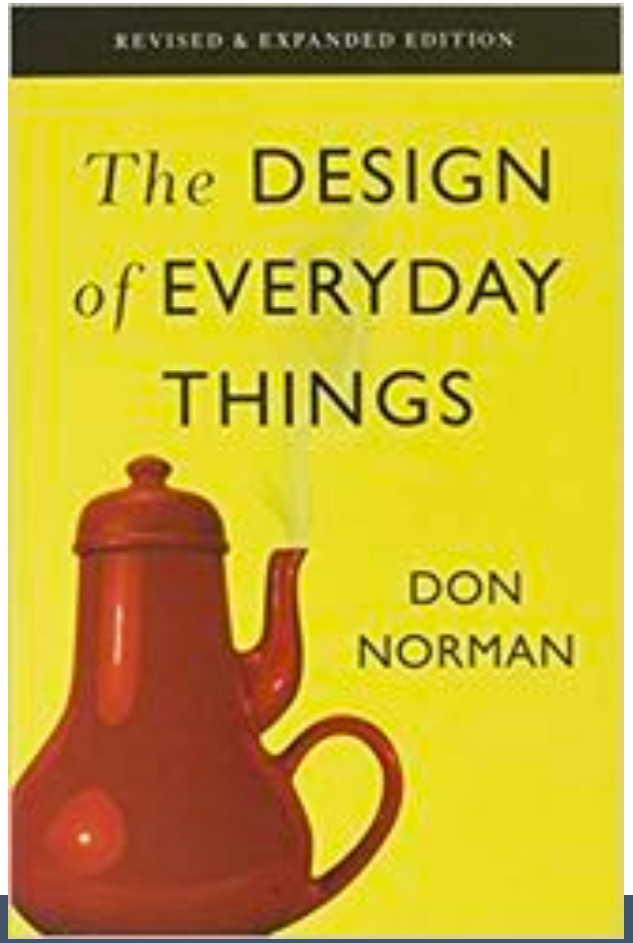
https://gendermag.org



User Centered 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



Agile

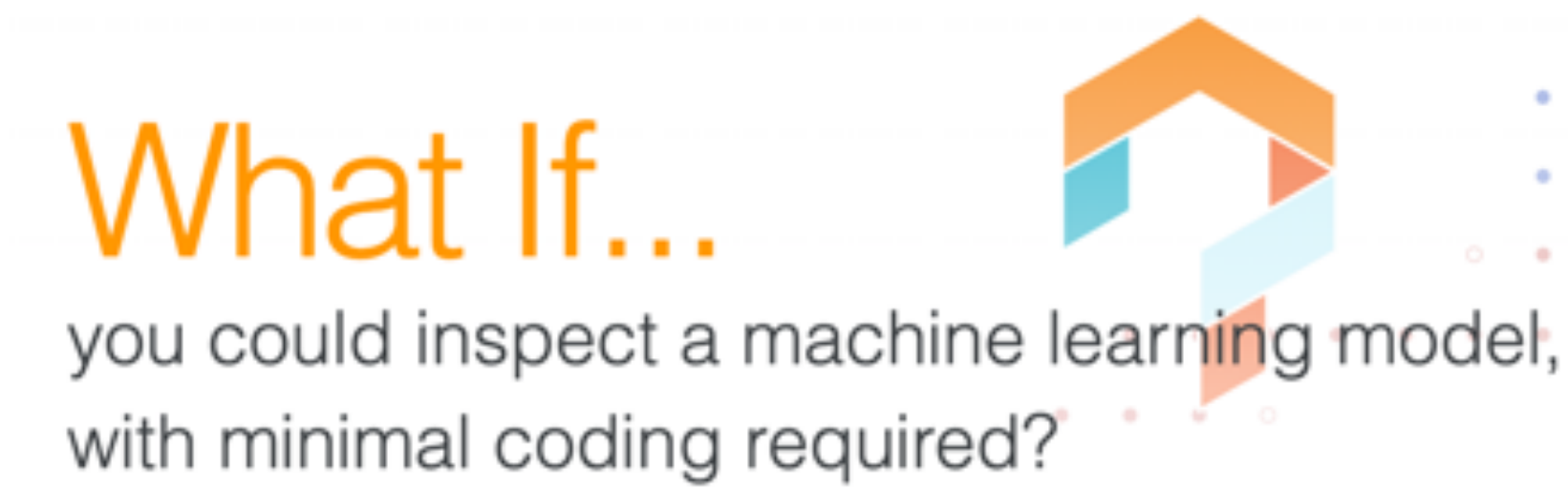
User Cer

Agile cus

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

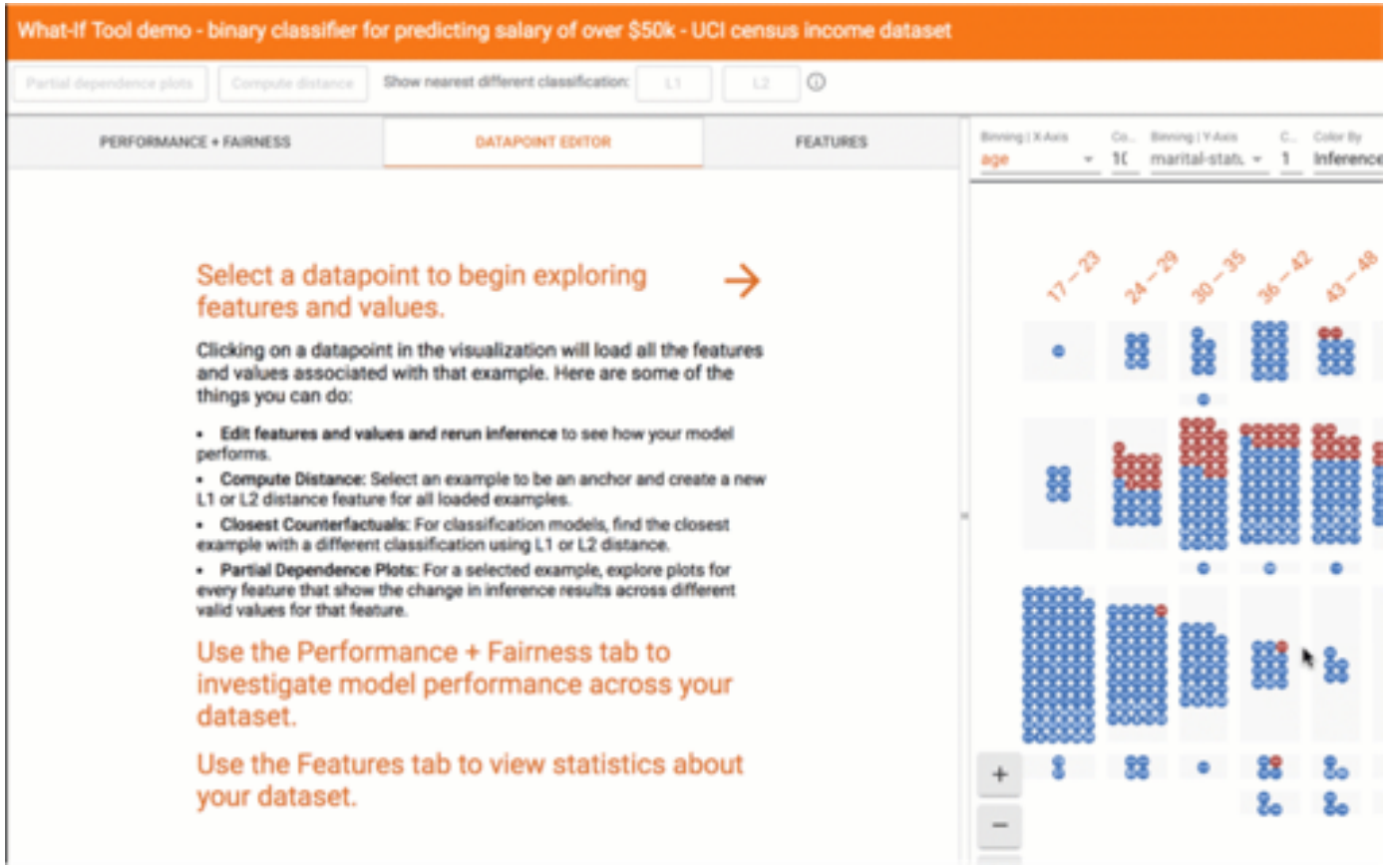
What if...

https://pair-code.github.io/what-if-tool/



What if…

https://pair-code.github.io/what-if-tool/



Dog vs Wolf

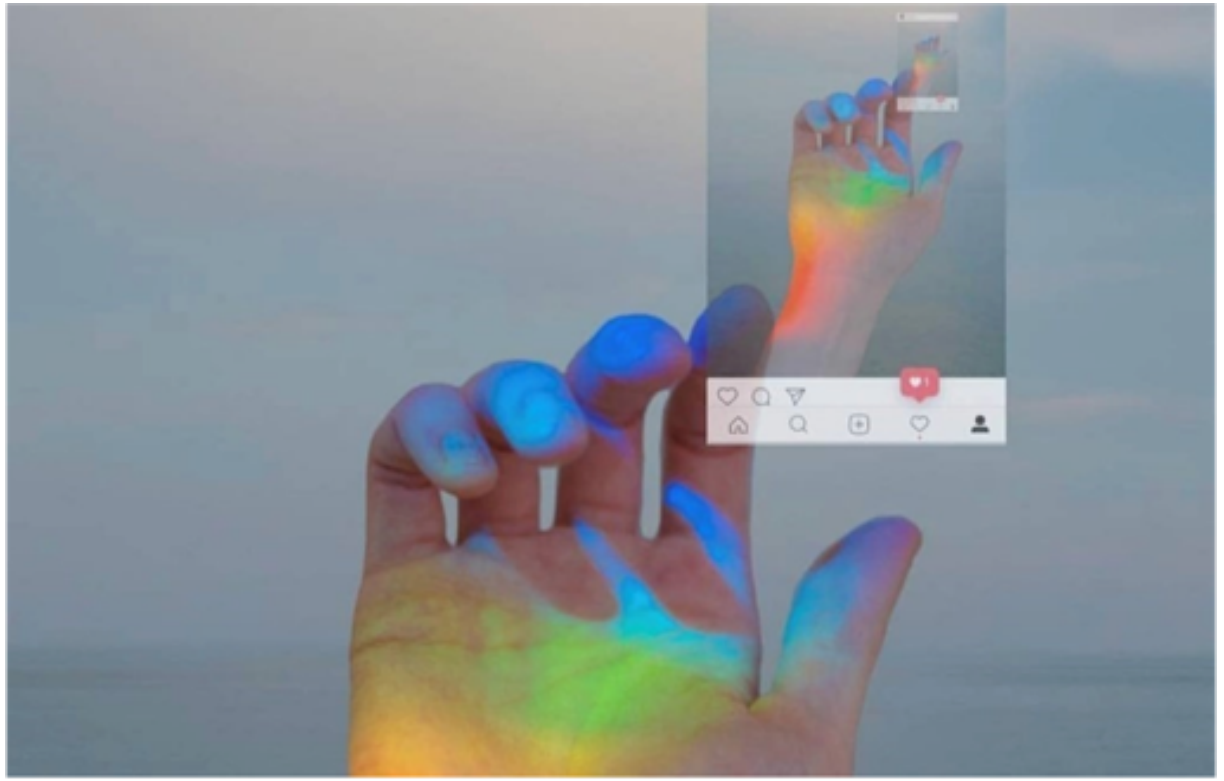
(a) Husky classified as wolf

(b) Explanation

Figure 11: Raw data and explanation of a bad model's prediction in the "Husky vs Wolf" task.

	Before	After
Trusted the bad model	10 out of 27	3 out of 27
Snow as a potential feature	12 out of 27	25 out of 27

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

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

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.

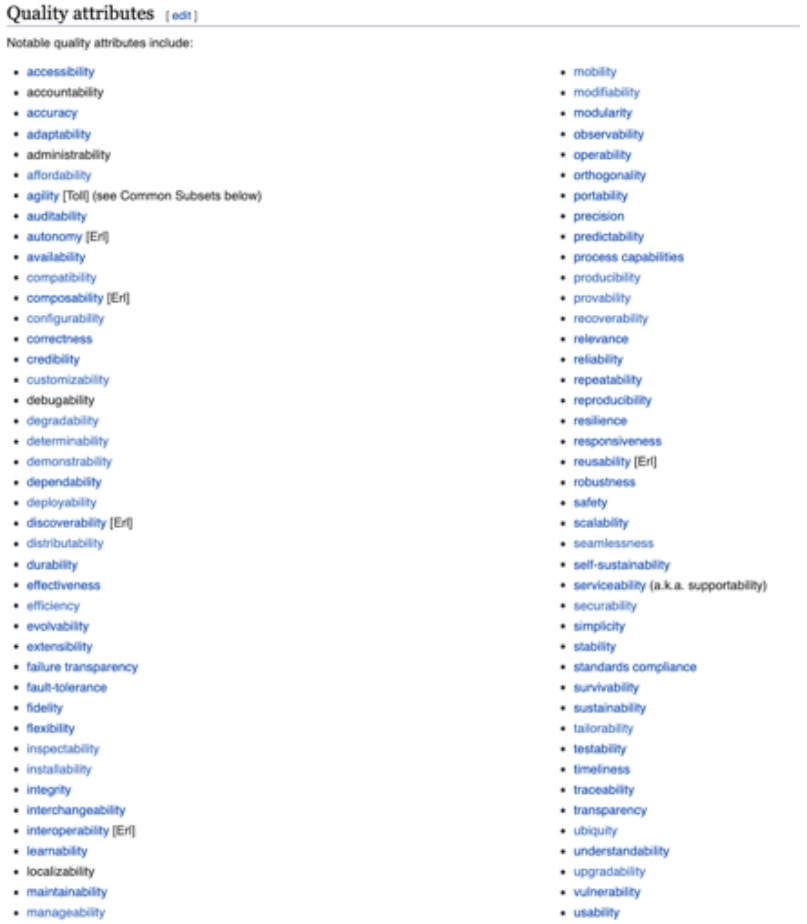
Your site should have accountable identities.

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.

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

Quality has long been considered



Quality attributes [edit]

Notable quality attributes include:

- accessibility
- accountability
- accuracy
- adaptability
- administrability
- affordability
- agility [Toll] (see Common Subsets below)
- auditability
- autonomy [Erl]
- availability
- compatibility
- composability [Erl]
- configurability
- correctness
- credibility
- customizability
- debugability
- degradability
- determinability
- demonstrability
- dependability
- deployability
- discoverability [Erl]
- distributability
- durability
- effectiveness
- efficiency
- evolvability
- extensibility
- failure transparency
- fault-tolerance
- fidelity
- flexibility
- inspectability
- installability
- integrity
- interchangeability
- interoperability [Erl]
- learnability
- localizability
- maintainability
- manageability
- mobility
- modifiability
- modularity
- observability
- operability
- orthogonality
- portability
- precision
- predictability
- process capabilities
- producibility
- provability
- recoverability
- relevance
- reliability
- repeatability
- reproducibility
- resilience
- responsiveness
- reusability [Erl]
- robustness
- safety
- scalability
- seamlessness
- self-sustainability
- serviceability (a.k.a. supportability)
- securability
- simplicity
- stability
- standards compliance
- survivability
- sustainability
- tailorability
- testability
- timeliness
- traceability
- transparency
- ubiquity
- understandability
- upgradability
- vulnerability
- usability

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.

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

- Citigroup Center, Designed by Structural engineer William LeMessurier
- Followed calculations required by building codes
- Civil Engineering student Diane Hartley realized there was a problem
- Tests showed that winds needed to bring it down would happen every 55 years

Professional Ethics

Professional ethics encompass the personal, and corporate standards of behavior expected by professionals.

First three “professions”

-Divinity,

-Law

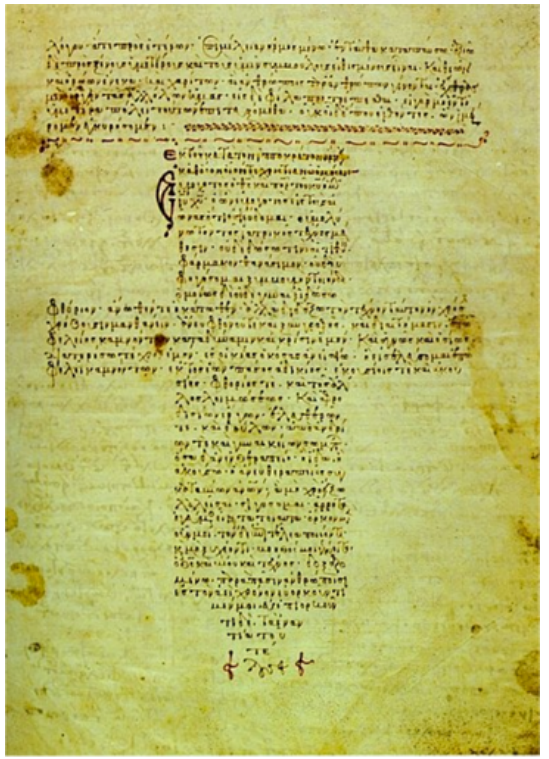
-Medicine

Medicine - Intrinsic

Hippocratic Oath

~450BC

“Do no Harm”



Law -Extrinsic

Bar regulates behavior

Oath to follow rules

Malpractice

Legal Malpractice

Not every mistake is legal malpractice. For malpractice to exist:

Attorney must handle a case inappropriately

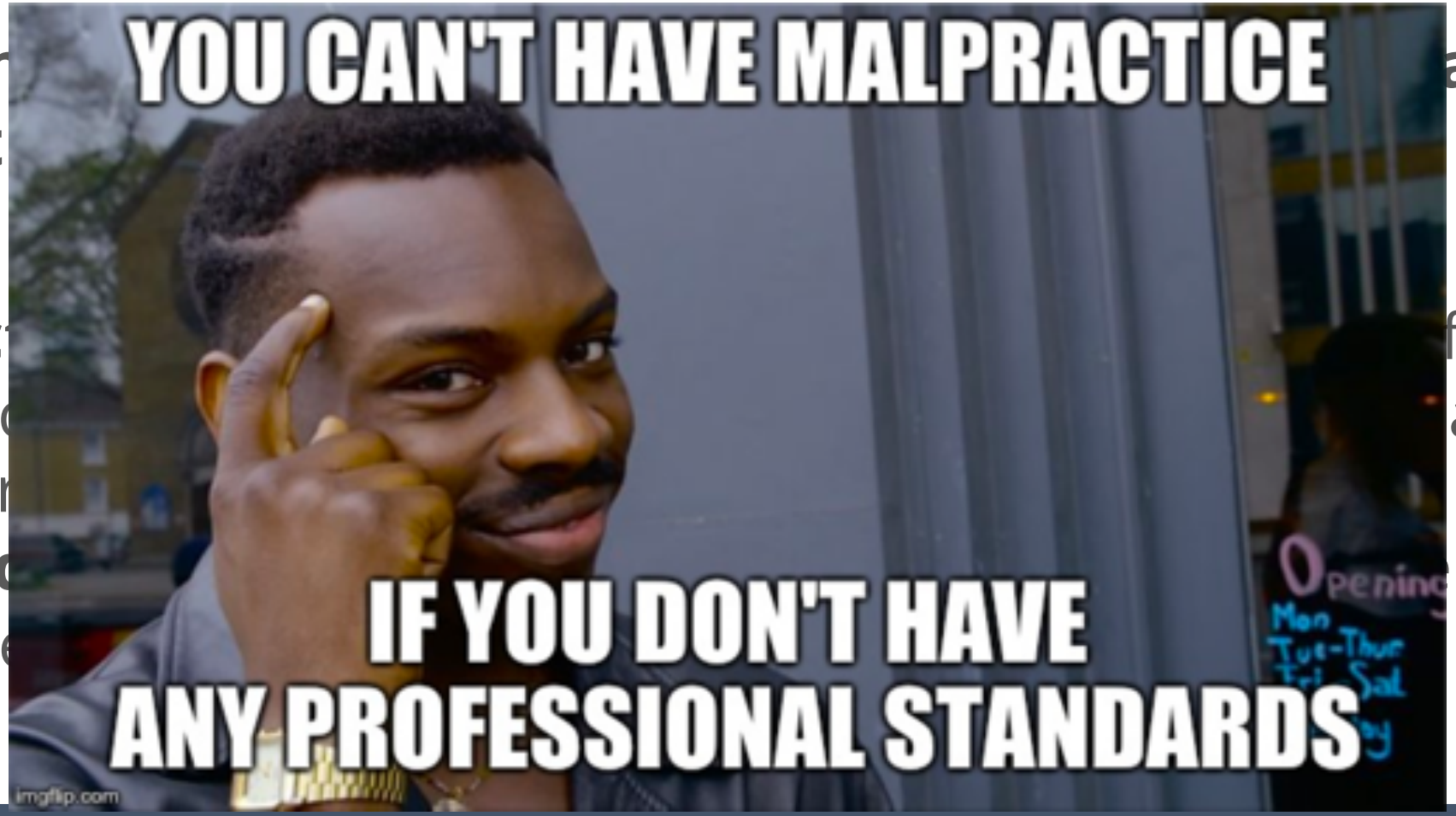
due to negligence or with intent to harm

And cause damages to a client

Malpractice vs. Negligence

Negligen **ably prudent**

Malprac fessional negligen a doctor, lawyer or **standard** "), subseque



DISCUSSION: WHAT SHOULD WE DO GOING FORWARD?

Bioengineering Ethics:

- Respect for Autonomy
- Beneficence
- Nonmaleficence
- Justice

Will software quality impact human flourishing?

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

Three questions to promote human flourishing

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?