



# Using the Knowledge

Toni-Ann McDaniel  
Joseph LoGatto  
Karan Nahar  
Daima Sheikh

# Using Tech/AI Ethics: Why are we doing this?

01

---

## **METRICS FOR SUCCESS**

Ethics measure how well a tool works in human terms

02

## **ENHANCE TECHNICAL DESIGN**

Ethics guide the design process to ensure the tool is aimed at the right target

03

## **BUILD USER TRUST**

Users are more likely to trust and interact with a company that respects their autonomy and privacy

04

---

## **SHAPE COMPANY CULTURE**

Ethical practices increase work satisfaction and productivity

05

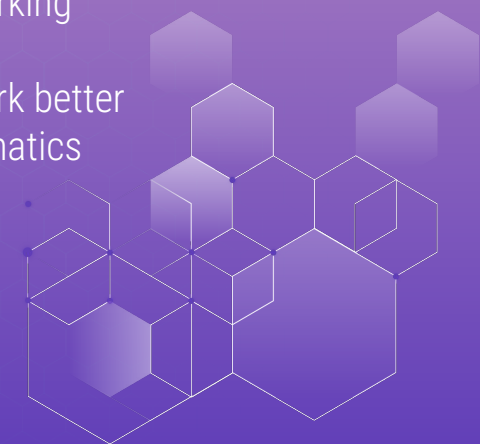
---

## **RISK MITIGATION**

Ethics help companies comply with regulations and avoid legal issues



# METRICS FOR SUCCESS

- Provide human language/metrics to determine success of AI tools
  - Allows people in computer science the language and the metrics to determine whether their tool is actually working
  - Makes the machine work better apart from just mathematics
- 

# EXAMPLE: CARDISIO CASE

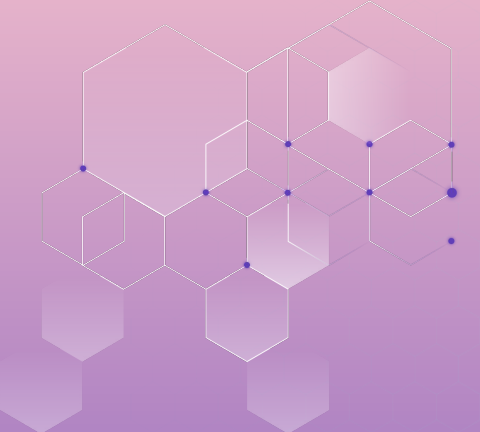
## TECHNICAL

- Great numerical success
- Identified everyone who was going to have heart issues in the future
- Very sensitive with high detection rate; technical accomplishment


## REALITY

- Flagged many false positives
- Caused unnecessary anxiety
- Ultimately negatively impacted patient autonomy

- Patients would live with anxiety fearing a health issue that may never happen due to low probability
- Ethics audit allowed us to see the tool wasn't working
  - Not just talking about numerical results; talking about how people live in the real world
- Tool needed to be evaluated in terms of patient autonomy (ethics) instead of purely numerical correctness (mathematics)




# ENHANCING TECHNICAL DESIGN

- Understanding ethics empowers designing for success on the right target
  - Not how accurate/precise the AI is, but where does the AI aim? Not how close to the target, but which target?
  - What's more important than how precise or how accurate your data is with respect to your target is **WHAT** your target is
- 



# BUILD USER TRUST

- Users sensing that their fundamental rights are protected are more likely to trust and interact with the device
  - Users are more likely to interact with a company when they feel like the company is respecting their autonomy
- 

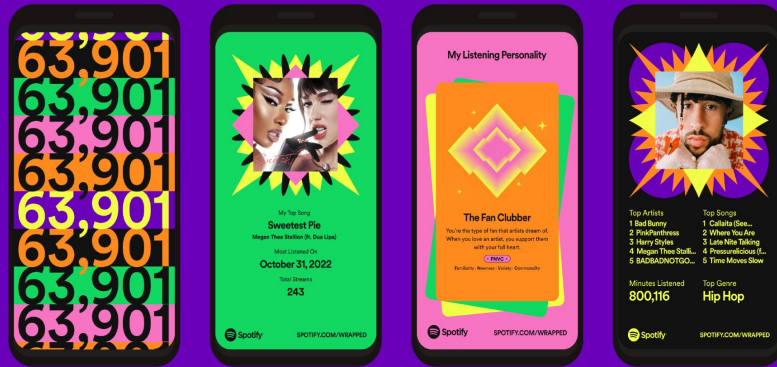
# EXAMPLE: APPLE VS ANDROID

- General consensus amongst users is apple has better privacy
  - Privacy is seen as user right
- Apple ecosystem is built to ensure strong security and privacy controls
- Android's open model allows for greater customization, but leaves users more vulnerable to security breaches and invasive data collection practices





# EXAMPLE: SPOTIFY DATA COLLECTION




#SPOTIFYWRAPPED

- Spotify's extensive user data collection leads to privacy concerns
- Users feel uncomfortable with the lack of transparency and control over their personal information
- Data being collected is being used for personalized recommendations and user experience
  - Popular feature, spotify wrapped, demonstrates how much data they collect throughout the year
- Users concerned with potential misuse of data





# SHAPE COMPANY CULTURE

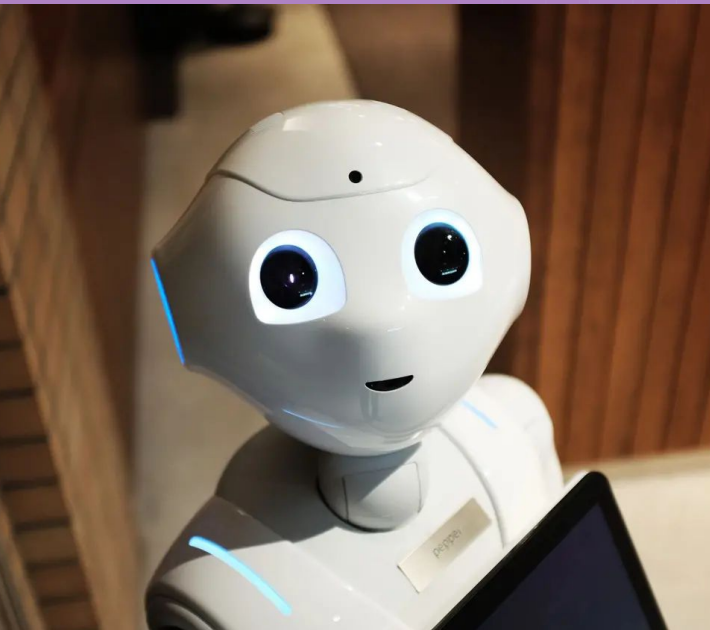
- Work satisfaction increases with knowledge of outcomes
  - People are more productive at work if they have a sense of/believe in what they are doing
  - Ethics encourages people to be engaged in and enjoy their own work
- 



# RISK MITIGATION

- Innovation is surpassing legal regulation
  - Incorporation of ethics provides justification for decisions about the design of AI models
    - Also clarifies accountability
  - Being involved in ethics makes you directly involved in the fundamental decisions in what direction the company goes in when developing products
- 

# EXAMPLE: EU AI ACT



- Act was supposed to be publicized about 9 months before it was
  - Just as they were about to release the legislation, ChatGPT, Dall-E, and other LLMs started coming out in droves
- Had to rework the legislation to include these LLMs
  - As they reworked the legislation, more advancements were being made with AI
- Government is having a hard time catching up with technical innovation


# EXAMPLE: FACIAL RECOGNITION

- Facial recognition in public surveillance in London was seen as a tool for enhancing public safety
- The technology has discriminatory biases and potentially violates citizens' privacy
  - Led to criticisms in autonomy
- Ethics audit was necessary to address autonomy concerns





# EVALUATIONS (AUDITS)

- Required by law
    - Governments are mandating this ethical action due to the inability of the government to keep up with tech innovation
  - Ethics plays a defensive role in ensuring that bad things don't happen and when they do, there is accountability
- 

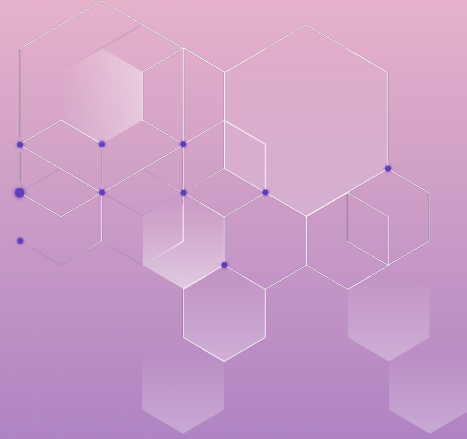
# ETHICS AUDIT

- Ethics audit casts technology in ethical/humanist terms as opposed to scientific, economic, or legal terms
- Evaluates how well the technology serves fundamental human rights
- AI/tech ethics audits are composed of three parts:
  - Social/technical scenario
  - Evaluation
  - Report/recommend

Report/Recommend

Evaluate

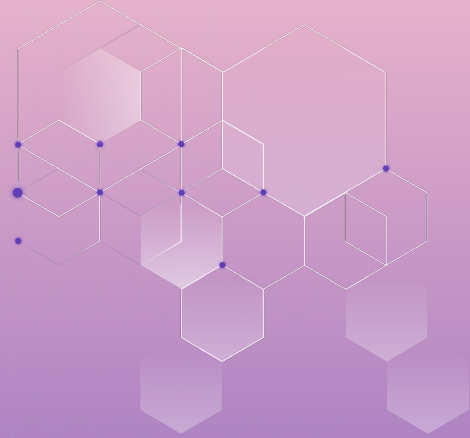
Social/technical scenario





# SOCIAL/TECHNICAL SCENARIO

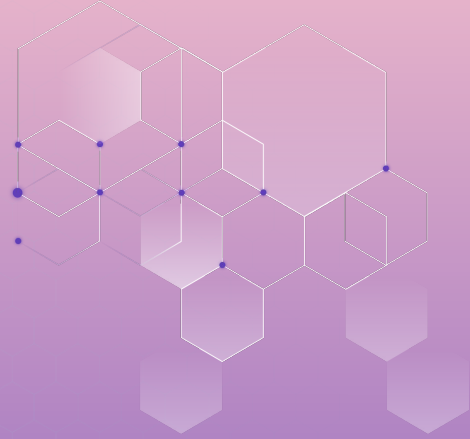
- Gather all stakeholders in a meeting (everyone who is involved in the tool/technology being evaluated)
  - People involved and know about the machine from each angle are gathered in this meeting
- Asking questions: What is it we are talking about? What does the AI/technology do? Who does it do it for? What do you need to know to know what we are talking about?
- Explain to everyone what we see in the machine and ask everyone else if there are any questions or ask how our vision of the machine may differ





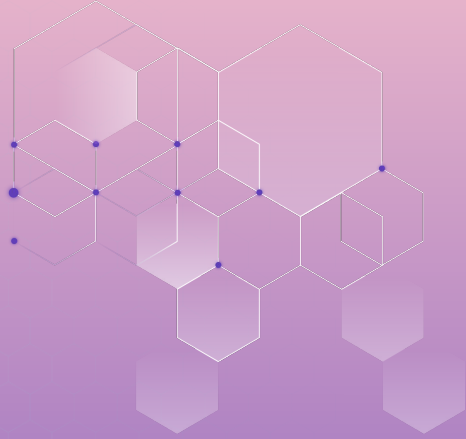
# SOCIAL/TECHNICAL SCENARIO

- A different explanation may be given to each group of people
- Each group participates by explaining their perspective on how they view the machine
- All of these questions are repeated over and over until everybody understands all aspects of the machine
  - A process of translation from group to group
- Some members of the group take responsibility for writing down what was said
  - Drafted in a google doc
- All members asynchronously read the document, correcting what they have said if necessary



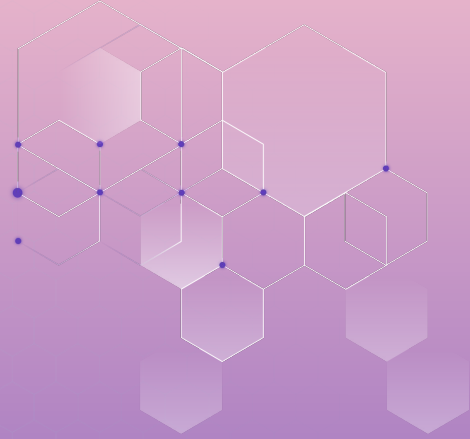
# EVALUATE

- The actual audit
  - Takes place in a meeting/zoom
- If the whole group of evaluators is large, group is broken down into smaller groups.
  - If not, a few people are chosen to take the lead
- Typically in these groups there is an ethicist, one computer scientist, and any other related professionals
  - Such as a doctor if it is a medical tool
- Social/technical scenario is available to refer to



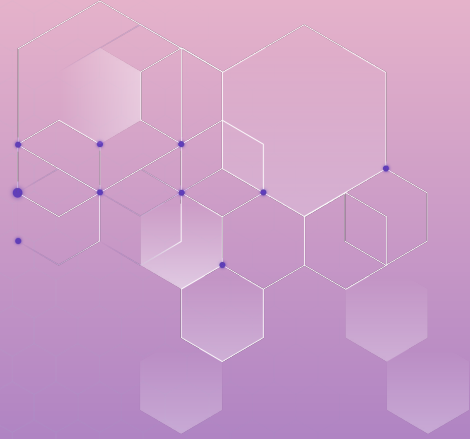
# EVALUATE

- Use ethical principles (autonomy, privacy, fairness, etc.) to analyze the tool
- Different approaches for evaluation: top-down and bottom-up
- Bottom-up: starts at the user experience of the machine and go to the principles
- Top-down: starts with the principle, and then look at the machine and the user experience of the machine
- Draft in a Google doc
  - Pose all of the issues that come up and results of discussion for each ethics principle
  - This is the heart of the paper; must be written in an organized way



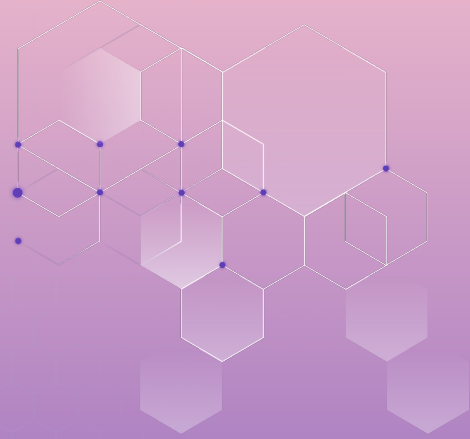
# EVALUATION EXAMPLE

- Ethics audit Professor Brusseau partook in about the tech ethics of a music program
- Top-down: At what point does the music being played no longer belong to the original composer? When does it become a new creation?
- Can this player take credit for this music? Or is it repackaging music someone else used?
  - Dignity question
- Bottom up: New way of experiencing music which gives a hearing impaired person a new outlook on music, etc.



# REPORT/RECOMMEND

- The Google doc drafted in the evaluation and social/technical scenario sections are converted into academic narrative
  - When done, there is a paper that is ready to be published
- Report is not action-oriented
  - An ethics evaluation may just say we better understand a machine and how it works on a human level
- Major areas of ethics being addressed:
  - Autonomy
  - Privacy
  - Fairness and Explainability



# EXAMPLE CASE (MAYA.AI)

- Trying to make a bot (James bot) to answer ethics questions 24/7
  - For students to use when professor Brusseau is not available
- Ethics audit:
  - Gather everyone involved (students, computer scientist(s), professor Brusseau, etc.)
  - Talk about the LLM, where the data comes from, why it is used, etc.
  - Evaluate and understand what it does for autonomy
- Evaluation makes the machine work better.
  - How much performance we are losing versus how much privacy and autonomy we are gaining





**THANK YOU!**