CSI5137: ETHICS IN AI

WINTER 2020

ETHICAL AI DESIGN PROJECT

For this ethical AI design project, you will be using all of the design for human values knowledge you have gained throughout the course. The overall purpose of the ethical AI design project is for you to demonstrate your ability to use the design for human values approach to: (a) identify and analyze ethical issues in AI, and (b) identify and prototype a solution for those issues.

You must choose one (1) design topic from the list below. You will complete this project in groups of three (3) students. Your final project submission will include:

- A 10-minute (max) Video walking viewers through, and describing in detail, each of the key design elements you completed (details below). **(70% of total Project grade)**
- A 2-page (single-spaced, 12-point font) written Ethics Analysis of your key value tension. **(30% of total Project grade)**

Key design elements:

1. Value Map, including:
    a. Stakeholders; Values; Value Tensions
    b. **Your value map must be represented in a visual format of your design/choosing.** Your video must clearly include the value map and must document the brainstorming that went into creating the value map.
2. Key Value Tension:
    a. You must identify the key value tension that you chose to focus on for your design topic.
    b. You must explain why you chose to focus on that tension.
3. Ethics Analysis:
    a. You must use the course readings, and supplement the course readings with additional sources that you find on your own, to explain in detail the ethical issue(s) that are associated with the key value tension you have identified.
    b. You must describe/explain how different design decisions would result in different ethical outcomes (e.g. by emphasizing certain values, or by dissolving the tension, etc.).
4. Prototypes and Tension Resolutions:
    a. You must use all of the above work (Items 1-3) to explore two prototypes that resolve your key value tension. Your prototypes do not have to be a piece of code (though they very well could be). Your prototypes might be a process, a checklist, design principles, a user interaction/interface, a visual aid (recall the hexbug metaphor cards), a policy, etc.. You can be as creative as you like in designing your prototypes.
    b. You must show that your prototypes each represent a different approach to resolving the value tension, and describe/explain how each prototype that you designed would clearly implement a tension resolution.
    c. You must clearly explain which prototype you think is best and make a clear argument why it is the best one to pursue.
    d. Your prototype must CLEARLY resolve the key value tension and draw from the ethical analysis.

Video Marking Scheme (70% of total Project grade):

1. Overall Video Content
    a. Topic is presented clearly and ideas flow smoothly from one element to the next. Video tells a clear story about the process and results of the Ethical Robotics and AI Design Project. (15%)
    b. Visual representations of the various design elements presented throughout the video are clear and easy to understand. (10%)
2. Value Map
    a. Visual representation is clean and organized and conveys the information clearly. (5%)
    b. Students provide evidence and description of brainstorming that went into the creation of the value map. (10%)
3. Key Value Tension
    a. Description of the key value tension is clear. (10%)
    b. Explanation of why students chose that particular key value tension is clear. (5%)
4. Ethics Analysis
    a. Students clearly explain the ethical issues that they uncovered in their written ethics analysis. These ethical issues, and the details presented in the video, should provide ***reasons*** for the decisions that are made in the tension resolution and prototype. (10%)
5. Prototype and Tension Resolutions
    a. Students clearly describe/explain each of two prototypes, and corresponding tension resolution embodied by each prototype. (10%)
    b. Students clearly explain which prototype they believe is the best (i.e. most justifiable) resolution given all of the options they considered. (20%)
    c. Prototype clearly draws from the information in the Ethics Analysis. (5%)

Ethics Analysis Marking Scheme (30% of total Project Grade):

1. Grammar & Spelling:
    a. A 1 point deduction for each spelling error to a max of 5% deducted. (5%)
    b. A 1 point deduction for each major grammatical error (e.g. causes readability problems, vagueness, ambiguity) to a max of 5% deducted. Minor grammatical errors permitted. (5%)
2. Clarity:
    a. Ideas are clearly presented and flow logically from one to the next throughout. (15%)
3. Citations:
    a. All sources are clearly and consistently cited using IEEE citation style guide. (5%)
4. Key Value Tension:
    a. Students clearly introduce the key value tension on which they are focusing their ethical analysis. (10%)
5. Identification and description of ethical issue(s):
    a. Major ethical issue(s) clearly identified. (15%)
    b. Course readings used to support the description of ethical issues. (15%)
    c. Other resources identified, and clearly used, to further support the ethical analysis. (10%)
6. Design Decision Impacts
    a. Alternative resolutions/prototypes are explored and the ethical implications of each is explained. (15%)
7. Focus:
    a. Ethical issues clearly relate to the design topic and key value tension. (5%)

DESIGN TOPICS

1. Amazon Echo Smart Speaker:
    - Consider the fact that the Echo can be placed in any room/space in a home. How might you design one or more features into the Echo to make sure it is recording/sharing data appropriately in those different spaces. Some relevant ethical issues for this design topic could include, but are not limited to: privacy; explainability/understandability/interpretability; deception; anthropomorphism; trustworthiness; and metaphors.
2. Sentiment Analysis using Natural Language Processing:
    - Imagine you are asked to design an AI system that will be marketed to parents with teenage children. The system will use Sentiment Analysis to monitor kids' text messages and "alert" parents of "risky" behaviour such as aggression/bullying, drug use, and depression. How would you design the "alerts" feature? Some relevant ethical issues for this design topic could include, but are not limited to: privacy; explainability/understandability/interpretability; deception; and metaphors.
3. Automated Vehicles:
    - Current turn-by-turn applications (e.g. Google Maps, Waze) are designed to minimize the time to destination, while only allowing the user to modify the algorithm to "avoid tolls," "avoid highways," and "avoid ferries". At the same time, these systems are changing traffic patterns in cities as more and more people use them to navigate during their daily trips around town. What additional "algorithm modifiers" could be designed into the systems? Some relevant ethical issues for this design topic could include, but are not limited to: privacy; metaphors, accountability; bias; and fairness.
4. Smart Cities (e.g. Sidewalk Labs in Toronto):
    - According to the various smart city imaginings, at any point in time while walking around a public space in a smart city, citizens will be subject to an array of sensors tracking various points of data related to them, such as: location; walking direction and speed; interactions with various infrastructure like doors or traffic crossings; facial recognition; purchases; gait detection; etc.. Those data points and surveillance technologies will differ from location to location. How might we (citizens) be made aware of that data collection and surveillance technologies in those spaces? Some relevant ethical issues for this design topic could include, but are not limited to: privacy; explainability/understandability/interpretability; accountability; trust/trustworthiness; and metaphors.

Feel free to talk to me or your course TA, if you are having difficulty with any of these topics. Remember, be creative, keep an open mind, play around with your design project, and have fun!