Personal AI/Data Code of Ethics

|  |  |  |
| --- | --- | --- |
| **Principle** | **What I Will Do.** | **What I Will Not Do.** |
| Bias in training datasets for machine learning and AI | I will always try to understand the context of the dataset. I will start by asking: are there any predispositions or stereotypes being unintentionally demonstrated by the benchmark? Then, I will look at data from an open perspective. I will constantly consider the viewpoints of others. I will also avoid any problematic value judgments that could cause the model to be harmful because some bias is more harmful than others. | I will NOT fall into any of the following traps: framing, probability, formalism, ripple effect, and solutionism. I will ensure that what my group works on does not entrench inequities or marginalize minority groups. I will always question the values latent within models because bias is inescapable. |
| Achieving algorithmic fairness | Reducing discrimination and bias can be accomplished by implementing ethically aware human oversight in planning, production, and post by testing for potentially discriminatory behavior. I can also accomplish this through digital insight tools that focus on pre-/in-/post-processing bias in algorithms. I will employ human discretion together with technology to achieve algorithmic fairness, even if forbidding or excluding data is useless. | I will NOT release an algorithm before ensuring that it is fair through an in-depth exploration into its inherent biases or potentially discriminatory outcomes. I will NOT allow coworkers to neglect their duty to reduce biases in their work throughout every step of the planning and production processes. |
| Data privacy and transparency with data usage | I will always ask for consent before collecting data: I will document informed consent and exclude individuals who opt out. I will further ensure the informed consent of the individuals contained within the dataset to eliminate any potential privacy breach concerns when creating datasets. Lastly, I will employ a variety of oversight mechanisms. This will help provide ethical, legal, and professional accountability. | I will NOT take information from individuals without their informed consent. I will NOT engage in any project that neglects to collect consent for data collection. I will NOT engage in any project that lacks transparency to its end users on the purposes for the data being collected. |
| Eliminating differential impacts of biased technology | Since technology is not neutral in its purpose, worldview, or impact, I must be cognizant of potential actions that will have unjust or harmful effects on particular groups of people. I will actively seek out these differential impacts and combat them by ensuring that everyone in my group takes active steps to reduce these biases. | I will not allow any interaction, latent, selective, or response bias in my models. If these biases are not properly acknowledged, then they will lead to differential or disparate discrimination. I will not allow my coworkers to get away with these biases. |
| Data governance and dataset documentation | I will ensure that all datasets used for projects are accurately documented and analyzed before they are used. By using Denton’s guide for dataset documentation, I can be sure that the implications of the elements within the dataset are fully understood before they are used. In addition, I will accurately document any changes to datasets through external programs so that processes can be accurately traced. | I will NOT allow data to be used freely without properly understanding where it is coming from and the context it is being used for. I will NOT permit coworkers to make changes to processes without properly documenting their work, especially with shared documents and processes. |
| Autonomy and freedom with AI | I will ensure that the use of artificial intelligence still upholds the liberty, command, independence, integrity, agency, and virtue of individuals. Data privacy is directly linked to autonomy, so a line must be drawn for AI systems to honor our secrecy and anonymity. | I will NOT allow AI models to compromise the autonomy and freedom of its users and constituents. Doing so will allow models to do more harm than good. |
| Dependency on AI | Since many jobs are now reliant on AI to bring the information needed to do work, there are many ethical implications. I will ensure that this dependency on AI does not presuppose the needs for human employees. | I will NOT allow companies to lay off workers in favor of robotic processes. Human intervention is still required for AI/ML to be entirely effective. |
| AI as a moral actor | Despite appearing as “mindless machines”, AI is still a moral actor. I can help make sure AI improves personalization, but there are also associated with potential moral hazards. In the back of my mind, I will always ask: Is AI approaching the natural human limit that humans must responsibly make ethical decisions based on our moral compasses? | I will NOT allow new technologies to have unpredictable, uncontrollable, or irreversible negative effects on human flourishing. I will NOT allow technology to become too powerful for us to use it ethically or responsibly, as this can have negative impacts for future generations. |
| Protection of youth and minors | I will actively search for potentially harmful or dangerous content that may be considered invasive or manipulative business strategies towards minors | I will NOT engage in any project that takes advantage of children, regardless of the potential rewards. |
| Creating safe and effective systems | I will ensure that every model and process has a purpose to which is effectively satisfied. I will make sure that this is accomplished safely, without compromising the integrity of the data and without leading to any potential data breaches. | I will NOT create systems that are potentially harmful to end-users. I will NOT create systems that are useless or entirely ineffective. |