Memo

To: Professor Enos & Classmates

From: Brittany Chiang

Date: February 18, 2018

Re: Mandating ethics settings for self-driving cars

Topic Overview

With the arrival of self-driving cars, we are on the brink of a transportation revolution. Advocates of this new technology claim that autonomous vehicles (AVs) will be safer, cheaper, and more convenient. Instead of spending so much time at the wheel, people will be able to dedicate more time to things they care about. However, the rise in popularity of these self-driving cars comes with some complex issues. One problem that still remains open is the regulation of programming ethics into these cars. Unlike human drivers, the behavior of artificial intelligence (AI) is programmed before any accident even happens. At this time, there is no official regulation for how ethics are being programmed. This must change before autonomous vehicles are released to the public, otherwise, we may be putting innocent lives in jeopardy without knowing it.

Purpose

As a potential solution to this issue, I am proposing for all self-driving car technology to adopt a mandated ethics setting. This ethics setting will simply be a set of laws enforced by all governments. Under these laws, self-driving cars will be required to comply with the predetermined ethical settings. Similar to how there are Federal Motor Vehicle Safety Standards, there should be federal self-driving car ethics standards. I will also stress that this ethics setting should be standardized internationally. Obviously, the United States is not the only nation that is involved with developing these new technologies. If we are all to live harmoniously with artificial intelligence behind the wheel, self-driving cars should be required

to meet consistent, safe standards. With this document, I will convince professionals who have the power to regulate the self-driving car industry to construct and mandate a standardized ethics setting for all autonomous vehicles.

Genre

To persuade the federal and international government to implement these regulations, I will be writing a proposal. A proposal will allow me to clearly present the problem, the background of that problem, and finally present my solution as a call to action for my audience. Enough research has been done on the topic of self-driving cars that I will have ample sources to confidently support my case. It is time that professionals came up with a sensible, solid solution to programming ethics into self-driving technology. Supported by scholarly sources, this proposal will convince the discourse community that this course of action is the best solution.

Motive

Up until recent years, the U.S. Department of Transportation has been virtually silent about autonomous vehicles. In 2016, the National Highway Traffic Safety Administration (NHTSA) released a 15-point checklist for self-driving cars. This safety checklist outlined best practices for companies developing self-driving cars. It set a range of goals for manufacturers around how the car perceives objects and responds, how it records and shares data, the human-machine interface, ethical considerations (i.e. the "trolley problem"), how manufacturers program the cars to follow traffic laws, and so on. However, the key here is that the NHTSA only *recommended* this checklist. Meeting all requirements on this checklist is not mandatory for all self-driving cars. This leaves room for autonomous vehicle manufacturers to potentially cut corners or only meet several of the safety guidelines that the NHTSA has outlined. Not only does this pose a threat to all passengers of those AVs, but also to the other vehicles on the road.

Audience

The professional audience for this proposal will be the U.S. Department of Transportation – more specifically, the National Highway Traffic Safety Administration (NHTSA). The NHTSA is responsible for keeping people safe on America's roadways. More specifically, the NHTSA is accountable for writing and enforcing Federal Motor Vehicle Safety Standards as well as licensing vehicle manufacturers and importers. As such an influential organization within the

federal government, the NHTSA is the perfect audience because they have the ability to actually create and enforce the proposed mandated ethics setting as they see fit.

Although the NHTSA will be the main audience of this paper, I also want to address traffic safety administrations on an international scale. To this extent, my second audience will be the International Transport Forum (ITF), which is an inter-governmental organization that aims to advance the global transport policy agenda. Even though the ITF does not have any legal power, it still holds a great amount of influence in the way many countries around the world are regulating their traffic safety.

Tentative Outline

I. Abstract (250 words)
    a. Short summary of purpose of proposal

II. Background (500 words)
    a. Explain how AVs will transform transportation
    b. Pros and Cons, emphasizes benefits

III. The Problem (500 words)
    a. Lose-lose scenarios
    b. Trolley problem
    c. Describe scenarios to make the problem more understandable
    d. Explain why the problem needs a solution

IV. Current regulation (200 words)
    a. Explain what currently exists in the form of regulating AVs

V. Mandated Ethics Settings (500 words)
    a. Explain the solution and its many benefits
    b. Explain why the solution is the best course of action compared to other options

Annotated Bibliography

1. **Belay, Nick. "Robot Ethics and Self-Driving Cars: How Ethical Determinations in Software Will Require a New Legal Framework." *Journal of the Legal Profession*, vol. 40, no. 1, 2015, pp. 119–130.**

   This article aims to survey the various approaches to the legal and ethical aspects of self-driving cars and offer the best strategy going forward to meet these considerations without deterring innovation in the market. I'll use it to present several different solutions to the major ethical dilemmas of self-driving technology. It will support my argument that a mandated ethics setting is the best way to go.

2. **Bonnefon, Jean-François, et al. "The Social Dilemma of Autonomous Vehicles." *Science (New York, N.Y.)*, vol. 352, no. 6293, 2016, p. 1573., doi:10.1126/science.aaf2654.**

   This article presents a study in which they found that participants in six Amazon Mechanical Turk studies approved of utilitarian AVs (that is, AVs that sacrifice their passengers for the greater good) and would like others to buy them, but they would themselves prefer to ride in AVs that protect their passengers at all costs. The study participants disapprove of enforcing utilitarian regulations for AVs and would be less willing to buy such an AV. Accordingly, regulating for utilitarian algorithms may paradoxically increase casualties by postponing the adoption of a safer technology. I will use the article to give some background information. Since it also presents evidence that works against my argument, I will refute its claims and convince the audience that my solution is more sensible.

3. **Fagnant, Daniel J., and Kara Kockelman. "Preparing a Nation for Autonomous Vehicles: Opportunities, Barriers and Policy Recommendations." *Transportation Research Part A*, vol. 77, 2015, pp. 167–181., doi:10.1016/j.tra.2015.04.003.**

   This article describes several barriers to implementation and mass-market penetration that autonomous vehicles still face. They describe how licensing and testing standards in the U.S. are being developed at the state level, rather than nationally, which may lead to inconsistencies across states.

4. **Gogoll, Jan, and Julian F. Muller. "Autonomous Cars: In Favor of a Mandatory Ethics Setting.(Report)." *Science and Engineering Ethics*, vol. 23, no. 3, 2017, p. 681., doi:10.1007/s11948-016-9806-x.**

   This article weighs the pros and cons of a mandatory ethics setting (MES) versus a personal ethics setting (PES) for self-driving cars. While the consensus view seems to be that people would not be willing to use an automated car that might sacrifice themselves in a dilemma situation, the authors defend that a mandated ethics setting would be nevertheless in their best interest. The reason is, simply put, that a PES regime would most likely result in a prisoner's dilemma. Although this scholarly article is more ethics based, it still directly supports my argument for this proposal.

5. **Goodall, Noah J. "Can You Program Ethics into a Self-Driving Car?" *Spectrum, IEEE*, vol. 53, no. 6, 2016, pp. 28–58., doi:10.1109/MSPEC.2016.7473149.**

   This was the article I used for the initial inquiry. In this article, the author effectively explains the multifaceted issue of programming ethics into self-driving cars. By presenting the reader with multiple hypothetical situations, the audience gets a better understanding of what's at stake. I will use this article to introduce my topic and provide the audience with background information.

6. **Hanna, Mina J., and Shawn C. Kimmel. "Current US Federal Policy Framework for Self-Driving Vehicles: Opportunities and Challenges." *Computer*, vol. 50, no. 12, 2017, pp. 32–40., doi:10.1109/MC.2017.4451211.**

   This article surveys the current federal policies and activities impacting technology developers, with special emphasis on privacy, cybersecurity, safety regulation, energy and environment, and ethical issues. This article will help me target my audience of engineers who build self-driving technology. It elaborates on the different challenges to development and deployment these engineers face, and will help me make a convincing argument.

7. **Kornhauser, Al. "Transportation Engineering and SmartDrivingCars A Perspective on the New Federal Automated Vehicles Policy." *Ite Journal-Institute Of Transportation Engineers*, vol. 86, no. 11, 2016, pp. 11–13.**

   This article presents a perspective on the NHTSA's AV policy guidelines. It combs through the different sections of the guidelines and points out the flaws in the policies. The author

stresses that the NHTSA has a very important mission of keeping human passengers and drivers safe.

8. **Nyholm, Sven, and Jilles Smids. "The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem?"** ***Ethical Theory and Moral Practice*****, vol. 19, no. 5, 2016, pp. 1275–1289., doi:10.1007/s10677-016-9745-2.**

   This article examines the differences between human reaction and the accident algorithms for self-driving cars. By presenting the trolley problem, they isolate and identify a number of basic issues and complexities that arise within the ethics of the programming of self-driving cars. This article will help me give the audience context for the problem and convince them that implementing mandatory regulations is the best course of action for all involved.