Conception to Reality
How data science is steering self-driving cars to the future

Swetha Chamala, Jolene Butt, Madeline Bauer and Samuel Sears

INTRODUCTION

Our generation has witnessed many technological developments and one such noteworthy advancement is autonomous vehicles. Like many technological inventions, self-driving cars were once considered science fiction. From the prototypes in the early 1900s to the present-day models, self-driving cars (SDC) have come a long way. As of today, autonomous driving has yet to overcome many technical and socioeconomic hurdles to be an everyday phenomenon.

Our research is focused on studying the involvement and importance of data science in making self-driving cars a reality. The progressive development of self-driving cars is nourished by a combination of advancements in artificial intelligence, data and data science. Artificial intelligence and machine learning are committed to overcoming these hurdles and taking autonomous driving a step closer to real-world existence. In our 'Dream project' we will demonstrate how machine learning and Big data have been the driving forces behind self-driving cars.

WHY IS THIS DATA SCIENCE?

Data Science utilizes many components and methodologies to contribute to the growth of data usage and technology. Data is collected and analyzed every day. Data is used in many ways to change the world around us; from a better way to brew coffee to self-driving golf balls and from self-driving golf balls to self-driving cars. The data collected is used to learn and adjust to the needs of their environment.

The utilization of machine learning is key to the growth of technology. Machine learning is a collaboration of the use of digital systems, processes and making use of gathered data by specifically designed algorithms. Artificial Intelligence is a growing field and with the use of Data Science the creation of self-driving cars can be accomplished.

OBJECTIVES

• Our main objective is to identify the challenges faced by SDCs with respect to data science

• Studying the impact of these challenges and understanding the outcomes if the challenges are not addressed.

• Proposing possible solutions and identifying their effects on overcoming challenges faced by SDCs.

CHALLENGES AND SOLUTIONS

We identified four challenges that limit scope of SDCs, and we are proposing corresponding solutions for the identified challenges:

Data storage:

Problem: SDC approximately generates about a petabyte of data per day. The current data storage architecture does not have the capacity to handle the massive amounts of data that will be expected to generate if each one of non-SDCs is replaced by SDCs. Large storage is essential for large data sets and flawless computing.

Solution: SDCs needs to be equipped with 1. High capacity storage which is fast, reliable and of rigorous quality for in-car data storage. 2. 5G internet connection with linear bandwidth to relay data for computing offsite for cloud-based storage and analysis.

Data acquisition:

Problem: To ensure flawless performance of SDC, neural networks need to be trained with representative data sets that cover an array of situations. The current data available is not enough to prevent such situations.

Solution: An estimate of 100 million miles of autonomous driving data is required for training algorithms for flawless driving. Corner case data sets are essential to train algorithms to prevent pedestrian fatalities and thereby ensuring safe driving.

Protecting privacy:

Problem: SDC is vulnerable to hackers who can hack into the car to take over driving and can obtain sensitive information about the driver, location and surroundings. Data ownership is unclear there by raising privacy concerns. Protecting the privacy of an individual using the SDC including location and data is of utmost importance.

Solution: 1. Laws and Policies need to be in place as to who owns the data. 2. Giving the SDC owner the ability to retain control over private data. 3. Minimizing the personal data that autonomous vehicle collects and encrypt the data it retains. 4. Anonymize the personal data generated.

Verification of Deep Neural Networks:

Problem: The major issue faced by autonomous driving is fatality cases because of corner-case behaviors. In some instances, unexpected erroneous behavior of deep neural networks used by SDC leads to safety concerns and worst-case scenario - fatalities.

Solution: 1. Training deep neural networks with data that is inclusive of wide variety of driving conditions like rain, fog, snow, extreme heat, lighting issues etc. 2. Training deep neural networks with corner case data sets that are collected from simulations and real-world driving. 3. Implementing a testing tool to automatically detect erroneous behaviors of DNN-driven vehicles that can potentially lead to fatal crashes.

DELIVERABLES

Data Monetization: Utilizing revenue from data production to fund data storage and acquisition solutions.

• Data storage: Increased development funding for cloud-based and hard drive storage.

• Data acquisition: Utilize extra revenues to bargain with companies developing 5G networks in order to meet data transfer capability requirements.

Industry Networking: Working together with other SDC industry leaders to foster societal adoption.

• Privacy protection: Create a coalition of companies in the industry and law makers to put in place laws that ensure protection of data and assurance of consumer safety.

• Verification of Deep Neural Networks: Combine training data for all companies in the industry in order to increase the coverage of corner cases and verification of correct situation handling by SDCs.
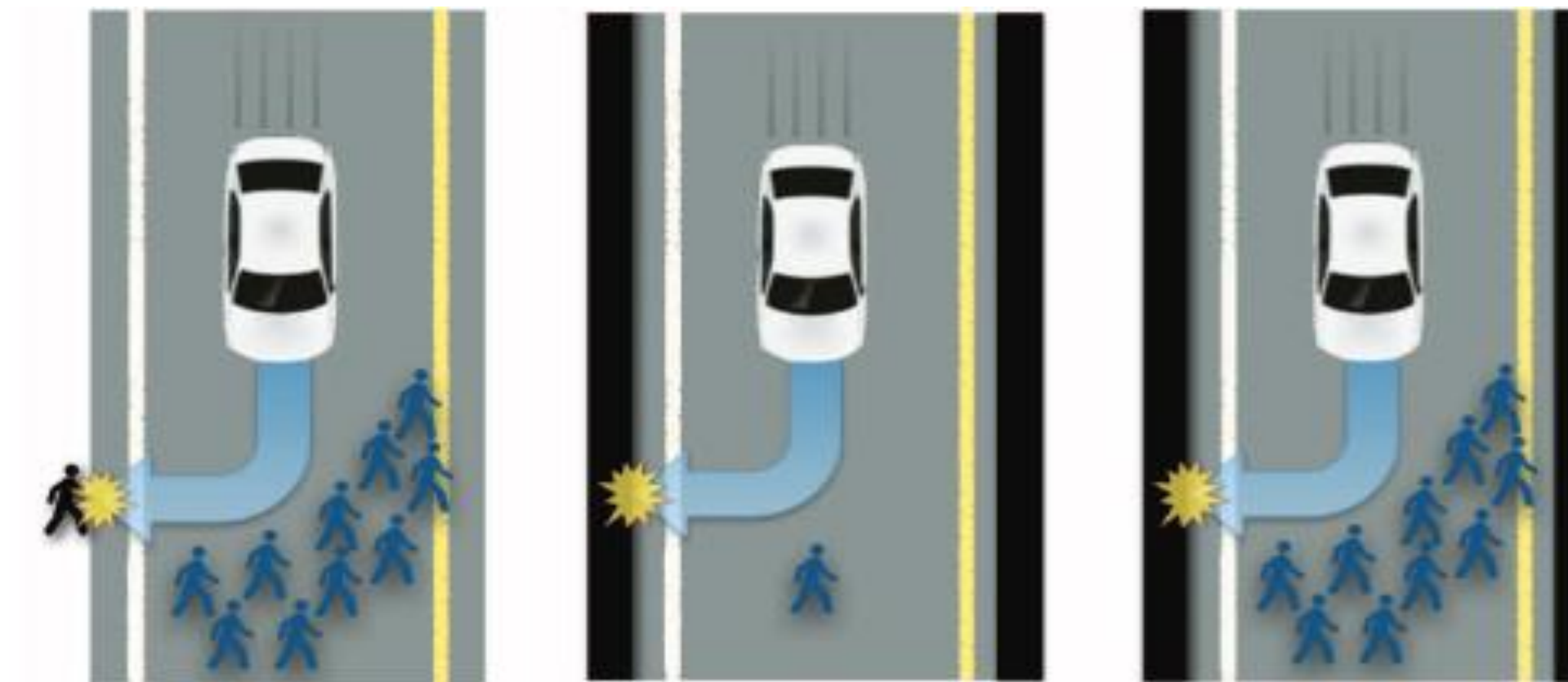

LITERATURE CITED

Boeglin, J. (2015). The Costs of self-driving cars: Reconciling freedom and privacy with tort liability in autonomous vehicle regulation. *The Yale Journal of Law and Technology*, Vol. 17

Collingwood, L. (2017). Privacy implications and liability issues of autonomous vehicles. *Information & Communications Technology Law, 26*(1), 32-45. Retrieved from http://dx.doi.org/10.1080/13600834.2017.1269871

LLANASAS, R. (2019). 5G's Importance for Self-Driving Cars. Machine Design. *91*(4), 38. Retrieved from http://search.ebscohost.com.ezproxy.bellevue.edu/login.aspx?direct=true&db=f6h&AN=135831515&site=eds-live

Phillips, J. (2018). The automotive industry needs and deserves better tools. *T&M Supplement*, 34-36

Tian, Y. (2018). DeepTest: Automated Testing of Deep-Neural-Network-driven Autonomous Cars, International Conference on Software Engineering , https://arxiv.org/pdf/1708.08559.pdf

Zawadzki, J. (2018, June 1). *The State of Self-Driving Cars for Everybody*. Retrieved from Towards Data Science: https://towardsdatascience.com/the-state-of-self-driving-cars-for-everybody-29446c1c2e2c

CONCLUSION

While the concept of self-driving cars has been around for many years, there have been several challenges that are hindering the process from becoming our reality. Some of the challenges that have been discussed include data storage and management, training data acquisition, protecting a user's privacy and being able to verify the work of deep neural networks. Research and technology continue to improve in hopes of making this concept become reality sooner than most people think.

There are many benefits of implementing self-driving cars, including an increase in safety, decrease in accidents worldwide, better traffic flow and efficiency. With continuous improvements in data science, the implementation of self-driving cars is right around the corner. Self-driving cars will create a safer, easier and more efficient lifestyle for all.
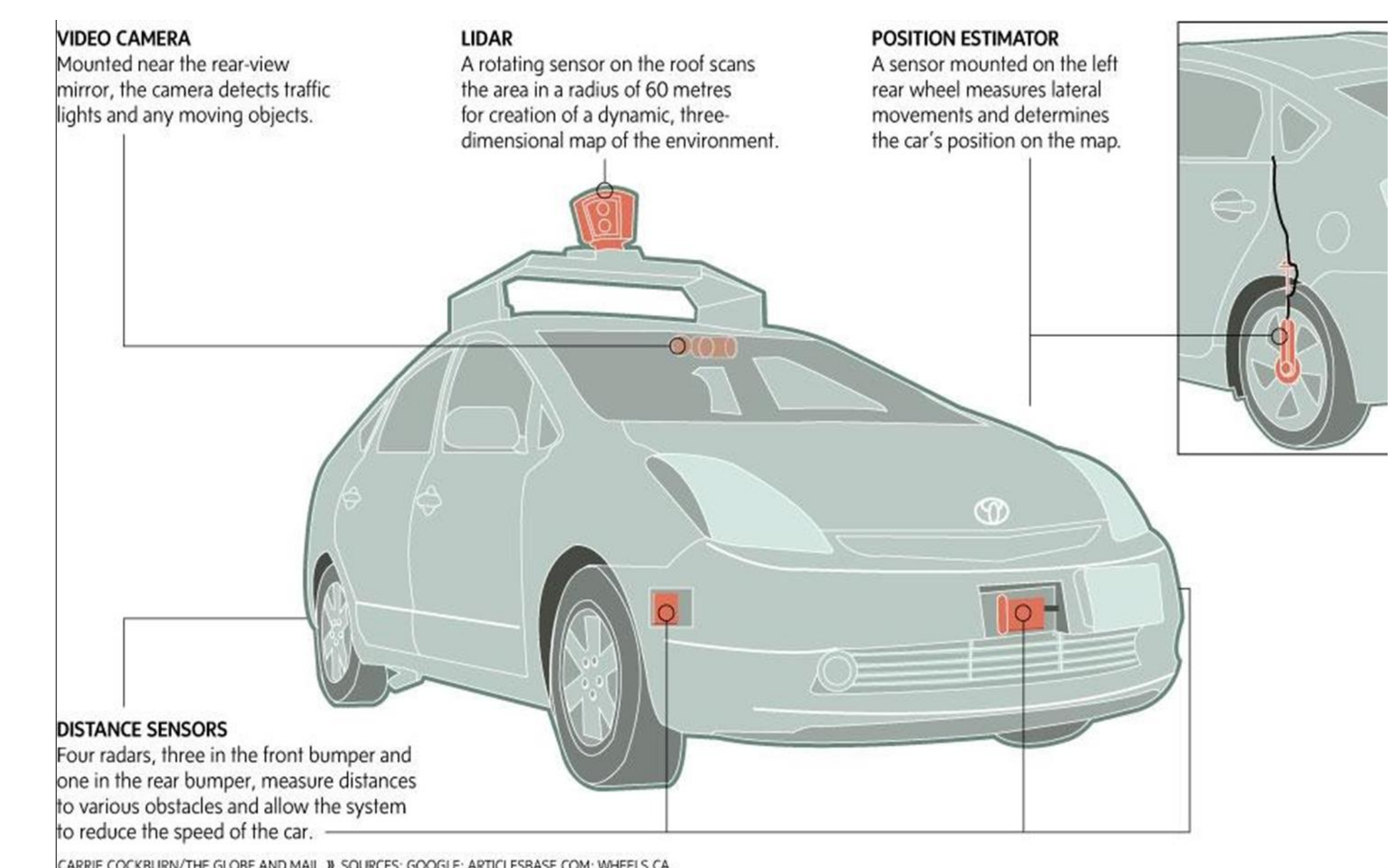
Bagloee, S.A., Tavana, M., Asadi, M. et al., 2016

ACKNOWLEDGEMENTS