Driverless Dreams: Technological Narratives and the Shape of the Automated Car

by

Erik Lee Stayton

Sc.B. Physics A.B. English, Brown University (2011)

Submitted to the Program in Comparative Media Studies/Writing
in partial fulfillment of the requirements for the degree of
Master of Science in Comparative Media Studies
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2015

Author . . . .
Program in Comparative Media Studies/Writing
May 8, 2015

Certified by . . . .
T. L. Taylor
Associate Professor of Comparative Media Studies
Thesis Supervisor

Certified by . . . .
David A. Mindell
Dibner Professor of the History of Engineering and Manufacturing
Professor of Aeronautics and Astronautics
Thesis Supervisor

Accepted by . . . .
T. L. Taylor
Department Committee Chair, Director of Graduate Studies

Driverless Dreams: Technological Narratives and the Shape of the Automated Car

by

Erik Lee Stayton

Submitted to the Program in Comparative Media Studies/Writing
on May 8, 2015, in partial fulfillment of the
requirements for the degree of
Master of Science in Comparative Media Studies

Abstract

In this work Erik Stayton examines dominant and alternative paradigms of ground vehicle automation, and concludes that current and imagined automation technology is far more hybrid than is often recognized, presenting different questions about necessary or appropriate roles for human beings.

Automated cars, popularly rendered as “driverless” or “self-driving” cars, are a major sector of technological development in artificial intelligence and present a variety of questions for design, policy, and the culture at large. This work addresses the dominant narratives and ideologies around self-driving vehicles and their historical antecedents, examining both the media’s representation of self-driving vehicles and the sources of the idea, common both among the media and many self-driving vehicle researchers, that complete vehicle autonomy is the most valuable future vision, or even the only one worth discussing and investigating. This popular story has important social stakes (including surveillance, responsibility, and access), embedded in the technologies and fields involved in visions of full automation (machine vision, mapping, algorithmic ethics), which bear investigating for the possible futures of automation that they present. However, other paradigms for automation exist, representing lenses from literature in the fields of human supervisory control and joint-cognitive systems design. These fields—compared with that of AI—provide a very different read on what automation means and where it is headed in the future, which leads to the possibility of different futures, with different stakes and trade-offs. The work examines how automation taxonomies, such as that by the NHTSA, fail to account for these possibilities. Finally, this work examines what cultural understandings need to change to make this (cyborg) picture more broadly comprehensible, and suggests potential impacts for policy and future technological development. It argues that a broader appreciation for our hybrid engagements with machines, and recognition that automation alone does not solve any social problems, can alter public opinion and policy in productive ways, away from focus on “autonomous” robots divorced from human agency, and toward system-level joint human-machine designs that address societal needs.

Thesis Supervisor: T. L. Taylor
Title: Associate Professor of Comparative Media Studies

Thesis Supervisor: David A. Mindell
Title: Dibner Professor of the History of Engineering and Manufacturing
Professor of Aeronautics and Astronautics

Acknowledgments

Any substantial work of writing has a long gestation period, and this has been no exception. A great number of people come to mind who have contributed to the existence of this thesis over the last 20 months, and I fear I must inevitably leave some out. To them I can say only *mea maxima culpa.*

No master's thesis is complete without a solid advising team, and my most heartfelt thanks go to T. L. Taylor for her careful reading and editing, and her ceaseless dedication (including periodic game nights) to ensuring we all made it through the master's process with most of our sanity intact. David Mindell has also been a continual resource throughout the writing of this thesis—and indeed is cited frequently within it—and I greatly appreciate his encouragement that I cut it down and sharpen my arguments. I look forward to continuing to work with him on future projects. Nick Montfort (who has exhorted me to, as much as possible, interrogate the technical objects themselves), William Uricchio (who has challenged my terminological precision), Heather Hendershot (who has kept my eye on science fiction), and Shannon Larkin (who has furnished continual doses of sanity) have all been indispensable in their own ways, and have given freely of both thesis and career advice. Discussions with Kate Crawford, Nick Seaver, Tarleton Gillespie, Merritt Roe Smith, Stefan Helmreich, Göde Both, and Patsy Baudoin have also shaped this work and made it what it is; as have innumerable discussions at conferences, many with people unrelated to the field. Conversations with a number of informants furnished important material, and though I cannot list all of them by name, I wish to thank them all for giving of their time. I should also recognize Wendy Hui Kyong Chun and William Keach for each, in their own way, setting me on this path years ago.

All of my classmates, in the classes of 2014, 2015 and 2016 have been great friends, supporters, and sounding boards for ideas. But I fear I must specifically thank Liam Andrew, Chelsea Barabas, Heather Craig, Suruchi Dumpawar, Sean Flynn, Desi Gonzalez, Jesse Sell, Ainsley Sutherland, and Wang Yu for discussing self-driving vehicles and technological ideology with me more times than some of them probably cared

to—and additionally Kyrie Caldwell and Lacey Lord for being available for games and gripe sessions throughout the process. On that note, my old friend Kristin Bergman, and my new friend Andrea Morales Coto, have been invaluable companions on this adventure, providing much needed levity and perspective in trying times. I also thank my forthcoming classmates at MIT HASTS for their input and encouragement, particularly Clare Kim and Alison Laurence.

My parents, Barbara Mindell and Lee Stayton, have often been the first to hear my ideas, and have provided irreplaceable feedback on this and other projects during my two years at MIT. Barbara, thank you so much for your willingness to read and edit "just one last time." I am so pleased to have been able to share this work with both of you, and both it, and I, have profited immensely from your input and your presence. Abigail Strait, whose food, hospitality, and seemingly endless supply of tea I have partaken of innumerable (at least twenty!) times during the writing of this thesis, you will always have my gratitude.

But most of all, Evan Donahue, whose keen commentary, acerbic wit, and drive to re-envision the future of humanity has spurred me on: this project exists in no small part due to his influence, and he most of all deserves thanks. As we have often said, this was supposed to be the Fourth Great and Bountiful Human Empire. Let us strive to make it so.

Contents

1 Introduction **11**

1.1 Concepts, Models, Dreams . . . 11
  1.1.1 Early Work . . . 13
  1.1.2 The Mid-Century . . . 16
  1.1.3 Research at the Turn of the Century . . . 20
  1.1.4 Automated Cars in 2015 . . . 22
1.2 Short-Term Promises . . . 24
1.3 Overview . . . 26

2 Narratives and Counternarratives **29**

2.1 The Automation of Work . . . 30
2.2 Artificial Intelligence . . . 35
  2.2.1 Logical Beginnings . . . 37
  2.2.2 Robotics, Embodiment, Probability . . . 38
  2.2.3 Learning and “Knowing” . . . 40
2.3 Aircraft and Autopilots . . . 44
2.4 NASA and High Technology . . . 45
  2.4.1 Manned Spaceflight . . . 46
  2.4.2 Researchers and “Robot Geologists” . . . 48
2.5 From the Driver’s Seat . . . 51
2.6 Conclusion . . . 52

3 Science Fictions, Designed Dreams **55**
3.1 The Stakes of Our Stories 56
  3.1.1 Data Gathering and Monitoring 58
  3.1.2 Maps and Geography 61
  3.1.3 Seeing and Sensing 64
  3.1.4 Functionalism, Utilitarianism, Ethics 68
  3.1.5 Safety and Statistical Risk 72
  3.1.6 Planning, Policy, Cities 75
3.2 Conclusion 79

4 Hybrid Controls, Hybrid Possibilities **81**
4.1 Human Supervisory Control 82
4.2 Cognitive Networks, Networked Subjectivities 88
4.3 Engineering Standards and Policy Documents 90
  4.3.1 NHTSA: Reacting to the Industry Narrative 91
  4.3.2 SAE: The Human as Engineered Component 96
4.4 Alternative Futures 100
4.5 Conclusion 105

5 Conclusion: Driving the Future **107**
5.1 Reshaping the Road: A Last Lesson in History 107
5.2 Cyborgs in Traffic 110
5.3 Automation and Social Goals 114
5.4 Lingering Questions 117

List of Figures

1-1 Houdina Radio Control's Linrrican Wonder, as pictured in *Radio News*, November 1925 [Green, 1925]. The article, in a radio enthusiast magazine, pays far more detailed attention to the operation of the vehicle than popular press accounts. 14
1-2 Image from an advertisement titled "Power Companies Build for Your New Electric Living," *Saturday Evening Post*, 1950s. Picture credit: [Novak, 2010]. 18
1-3 MIT's Urban Challenge vehicle. The grey cylindrical sensor at the top is a Velodyne LIDAR sensor. Below it sits a row of other LIDAR sensors and cameras. Picture credit: DARPA [DARPA, 2007]. 22

3-1 A simulated pedestrian detection image, showing the detected figures outlines by yellow boxes. Photo by Marjan Lazarevski [Lazarevski, 2014]. 68

Chapter 1

Introduction

It is early August, 1925, and a strange spectacle greets viewers on Broadway in New York City. An empty car lurches uncertainly down the street,[1] followed closely behind by a chase car, filled with radio equipment and several men, one driving, and one furiously operating a set of controls. The lead car features a diamond-shaped antenna sprouting from where the back seat should be, which is instead filled with a motley collection of dials and tubes, wires and batteries. A man stands on the running board of the jittering contraption, but keeps his hands well away from the controls, as if to tempt fate, as the car, unsteady in its movements, turns onto Fifth Avenue and comes, ungracefully, to a halt.

1.1 Concepts, Models, Dreams

While present work on automated vehicles by Google, Mercedes, Tesla, and others might lead one to think that driverless vehicles are a new idea, specifically enabled by recent advances in artificial intelligence, they have not merely been dreamed about, but actually built in prototype form, as far back as 1925, when Houdina Radio Control's "Linrrican Wonder" was demonstrated live on the streets of New York City. But it would be wrong to characterize a 90-year engineering journey as the continual progression toward a "modern," developed device. There is nothing fundamentally

---

[1] This description is adapted from *Time Magazine* [Science..., 1925].

natural about the way automated—or, often interchangeably "autonomous," "driverless," or "self-driving"[2]—vehicles are being envisioned, designed, and talked about. Instead, automated vehicle development has sustained several successive, overlapping paradigms, in part, though not completely, driven by the technical capabilities of and excitement about contemporary technologies. Equally important in this story are the ideologies of control and human-machine interaction that shape engineering practice and vehicle development, and alter how developments are represented and discussed in popular media. What we believe self-driving cars to be—or what we believe they soon will be—is a product of the popular, teleological understanding of technology development, which draws from existing technocratic narratives. Alternative paradigms exist, but deserve much broader recognition outside of certain engineering circles. In order to understand the way that automated vehicles have been figured popularly, the implications of that vision, and the alternatives to it, we should start by examining the varied types and dimensions of historical automotive automation research.

Self-driving cars are born from the same history as interchangeable parts, assembly lines, machine tools, and scientific management; as electrification, computerization, and networked interconnection. When asked how to make something better, safer, faster, easier, more productive, we involve machines or at least what H. M. Collins describes as "machine-like actions" [Collins, 1990, p. 42]. This is not a new predilection, but is something that has been with us since the Industrial Revolution. Furthermore, technological visions and concepts inscribe excitement about the technologies of the present moment, and this tendency has deeply affected the history of automated vehicle research. From the railroad to the Internet (passing through the telegraph, the airship, electrification, television, the automobile, the rocket), each new mode

---

[2] I indeed use these terms interchangeably at times even within this text, so I should mention the specifics of their use here. Autonomous, driverless, and self-driving vehicles are generally one and the same, though the terms "driverless" and "self-driving" are used particularly in the press when talking about these systems. These all presuppose largely self-contained operation on the part of the automated system. An "automated" vehicle however does not necessarily presuppose self-contained operation, and cars have become increasingly automated over the past 50 years (through ABS, traction control, cruise control, automatic transmissions, etc.) without, so far at least, becoming autonomous. This point, simple though it sounds, is at the center of this thesis. I use *automated* to refer to any amount or kind of automation, and the other terms interchangeably when implying more complete automation, or conventional ideas thereof.

of communication (or transportation[3]) ushers in new visions of our collective future. We are famously bad at these predictions [Riper, 2013], but they nevertheless drive the popular imaginary, which has included automated vehicles since the 1920s. Both visions of and research on self-driving technology have taken on aspects of the central information technologies of their day, and re-inscribed them as control technologies for physical devices. The history of our driverless dreams shows that the form of automated vehicles has been neither constant nor a foregone conclusion.

1.1.1 Early Work

Houdina's early experiments in radio-controlled vehicles (which began this chapter) put us into a time period in which radio was an exciting, emerging technology that was envisioned as a way to take the driver out of the vehicle, even if a human's direct commands were still necessary for vehicle operations. Aachen Motors's "Phantom Auto," operating on the same principles, toured Milwaukee a year later ['Phantom Auto' will ..., 1926], and returned as an attraction in Fredericksburg in 1932 ['Phantom Auto' to Be ..., 1932].[4] Contemporary news articles describe the machine as if it drove itself, as if the physical position of the person was what mattered, not the fact that a human was always using the radio transmitter: "Driverless, it will start its own motor, throw in its clutch, twist its steering wheel, toot its horn, and it may even 'sass' the policeman at the corner" ['Phantom Auto' will ..., 1926]. Described as "one of the most amazing products of modern science," the car proceeds "as though there were an invisible driver at the wheel" in an effect both "uncanny and mystifying" ['Phantom Auto' to Be ..., 1932]. Such work, though it seems relatively primitive in retrospect (hardly more than a glorified RC-car, though it removed the driver from a full-size vehicle), may well have helped capture the popular imagination. This compelling idea would reemerge a few years later.

---

[3] Rudolf Arnheim referred to TV, a "relative of motorcar and airplane," as a "means of cultural transportation" [Arnheim, 2006, p. 194].

[4] Radio control had at this point been public knowledge for at least two decades, since Nikola Tesla's 1898 demonstration of a radio-controlled boat, but though it was not fundamentally new, the use of the technology on public roads nevertheless shocked onlookers.

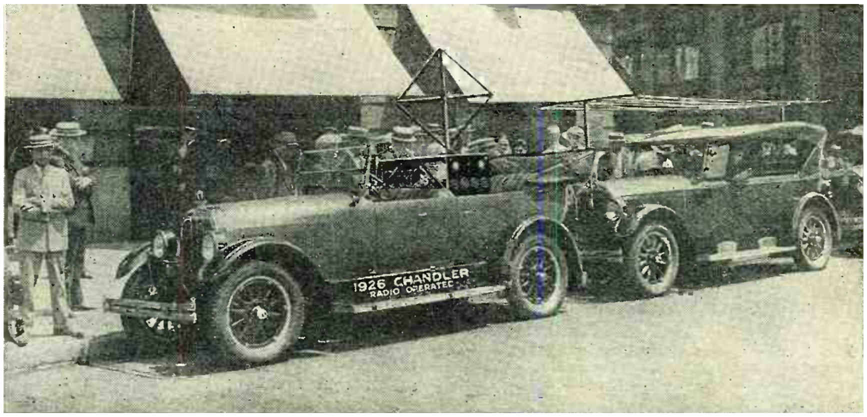

Figure 1-1: Houdina Radio Control's Linrrican Wonder, as pictured in *Radio News*, November 1925 [Green, 1925]. The article, in a radio enthusiast magazine, pays far more detailed attention to the operation of the vehicle than popular press accounts.

The 1939 New York World's Fair opened to great acclaim. Covering over a thousand acres and attracting a total of over 44 million visitors, the fair was a massive public spectacle demonstrating the technical prowess of American industry and providing a grand vision for the future of the nation. The structure of the exhibition—exemplified in its grand (and oddly-named) buildings, the Perisphere and Trylon—manifested the modernist techno-utopian ideology that drove the dramatic technological displays [Nye, 1990, p. 371]. The great success of the fair was GM's Futurama exhibit[5] which presented an idea of what the world would look like in 1960 [Wetmore, 2003, p. 3-4]. Automobiles were the focal point: the bird's-eye view presented a future city with multi-lane elevated expressways filled with largely automated vehicles. While the cars' lateral position was maintained via curved barriers, the Futurama exhibit described that distance between cars would be maintained via a "sophisticated system of radio control" [Wetmore, 2003, p. 5]. Despite the fragility of glass vacuum tubes traveling at high speed,[6] and despite the fact that the technology could work well only for sufficiently well-designed roads (high-speed highways, built

---

[5] The line for the exhibit routinely stretched to two miles in length, and 28,000 people visited the exhibit each day.

[6] Vacuum tube technology, though well-understood at the time, was subject to significant hurdles in terms of mechanical reliability. It was not until silicon transistors were invented and popularized that hardware became the relatively stable component of driverless systems, with software reliability becoming the key issue [Wetmore, 2003, p. 15].

specifically for these vehicles, with attendant infrastructural costs), such a system was reasonably coherent as a vision of what 1960 might be, from the perspective of attendees. This aspirational future was clean, modernist, and efficient; the rationalization of the factory[7] had spread to the city and countryside, where working hours were reduced through electrification, and even driving could become leisure, the mere inhabitation of a device operating via radio control.

Radio was not the only revolutionary medium to be included in portrayals of the self-driving car of the future. The cover of the April 1936 edition of the magazine *Modern Mechanix* [Modern Mechanix, 1936] betrays a different vision of the autonomous vehicle: the so-called "electric eye automobile" which would steer itself via a control loop, using an array of photocells on the car to track a light beam projected from the car and reflected by mirrors in the road surface.[8] As the article, titled "Light Beams Steer Super Racing Cars," describes:

> With speeds, such as recently attained by the famous Sir Malcom Campbell, already approaching the point where human reflexes are too slow to insure safe control of the car, science has turned to the photo electric cell for a possible solution. A proposed driverless car involves the use of multiple electric eyes as the heart of its steering mechanism. A powerful beam of light directed at a large lens on the front of the car is concentrated on steel mirrors set at an angle in the trackbed. The reflections are "caught" by the electric eyes which convey the electrical impulses to a mechanical-electrical brain which keeps the speeding car on its course [Light Beams ..., 1936, p. 71].

The cover pictures an intrepid "driver" using a camera to take a "movie recording" of the car's performance, while he leaves the "driving responsibilities to the mechanical and electrical brain" [Light Beams ..., 1936, p. 71]. Responsible for initially putting

---

[7] Public interest in this topic was demonstrated through popular factory tours of major Detroit automotive plants, and working GM and Ford assembly lines at the earlier Century of Progress Exhibition in Chicago (1933-1934) [Nye, 2013, p. 63-64].

[8] Intriguingly, this simple application has become one of the classic robotics teaching experiments. I remember building similar robots in high school robotics club.

the car into motion, the driver can then step back from the task of driving, allowing the electrical controls of mechanical linkages within the car—meticulously, if only artistically, diagrammed within the magazine article—to convey him or her safely and speedily down the track. The photocells within the vehicle's sensing mechanism would operate relays controlling the steering linkage, closing the loop between sensing and acting in the manner of a "teleological" [Rosenblueth et al., 1943], self-governing mechanism with corrective feedback: the simple idea of feedback control opens a vast array of new potentialities for the future. But how the car's governor could be modified to allow it to race beside other cars is not mentioned, which is telling since the complexities of the environment are what make automated driving so difficult.

1.1.2 The Mid-Century

By 1960, driverless vehicles had not been brought to market, but interest in the technology continued. Ford's vice president of engineering and research, Andrew Kucher, was referenced in the *Chicago Daily Tribune*, April 25, 1959, in a speech he gave at Northwestern University, taking seriously the idea of autonomous cars. The article, titled "In 50 Years: Cars Flying Like Missiles!" asks readers: "Can you imagine flying automobiles directed by automatic guidance systems?" [In 50 Years ..., 1959]. "Arthur Radebaugh's syndicated Sunday comic 'Closer Than We Think' was also a likely inspiration" for contemporary images of advanced car technologies (*The Jetsons*), as in 1958 it depicted hovering cars floating on an air cushion in an "already proved" concept publicized by the same Andrew Kucher [Novak, 2012]. Newspapers from April of that year also describe a 3-foot-long model of Kucher's "Glideair" that was demonstrated to reporters in Detroit [Novak, 2011]. The car may not be driverless, but the people depicted riding in the "flying carpet car" in Radebaugh's illustration do not seem to be paying much attention to where it is going, and the road surface is bordered with ridges that visually suggest the idea that the car will keep itself within the demarcated space.

One particularly salient public image from the 1950s comes from the *Saturday Evening Post*, which ran an advertisement for America's Electric Light and Power

Companies that depicted self-driving cars coasting through an idyllic landscape, guided by electricity [Weber, 2014]. The roadways are clean and clear, stretching off confidently into the distance. The landscape, tamed and controlled, yet retains its natural beauty: nowhere to be seen are smokestacks and the power plants generating the guiding force of these vehicles. The rural, even pastoral, landscape is largely unbesmirched, only cut across by the rationalization of a futuristic interstate highway system.[9] Within the finned, bubble canopied auto sits a nuclear family—father, mother, daughter, and son—enjoying family time. Perfectly coiffed, snappily dressed, and totally at ease, the mother and daughter play dominoes, while the son admires a model of a futuristic delta-wing jet aircraft and the father looks on. Their faces glow with expressions of domestic contentment. The father, in the "driver's seat," is turned away from the wheel, facing backward. The caption reads:

> One day your car may speed along an electric super-highway, its speed and steering automatically controlled by electronic devices embedded in the road. Travel will be more enjoyable. Highways will be made safe—by electricity! No traffic jams ... no collisions ... no driver fatigue [Weber, 2014].

The limitless possibilities of electricity, measured and controlled by America's Electric Light and Power Companies, reproduced a conservative present into an equally conservative future.

Wartime and postwar developments in electronics, the same ones that inspired new science fiction visions of the future, drove research in automated vehicle systems sensing and control. In 1953, General Motors and RCA completed a scale model of an automated highway system that would allow the two companies to experiment with developing steering and distance control systems [Wetmore, 2003, p. 6]. General Motors's 1950s Firebird concept vehicles marketed themselves on this research, despite the fact that Firebird II "had no automated capabilities whatsoever" [Wetmore, 2003, p. 7]. GM's promotional films suggested the vehicle could be controlled electronically from traffic control towers—like those in use in aviation—placed along major high-

---

[9] Recall that the Federal Aid Highway Act of 1956 began the creation of the interstate system as we have it today.

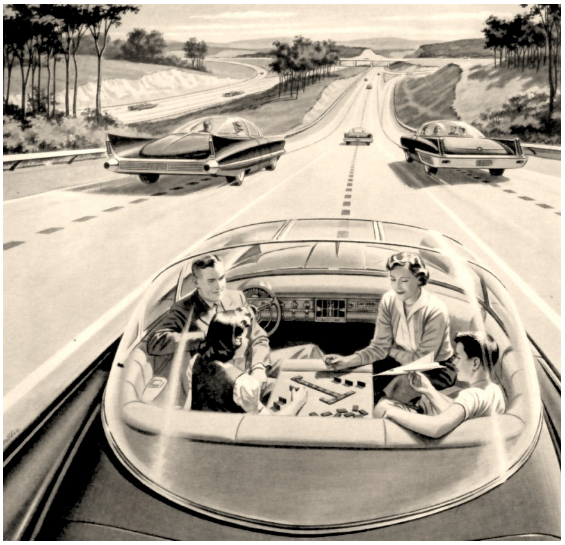

Figure 1-2: Image from an advertisement titled "Power Companies Build for Your New Electric Living," *Saturday Evening Post*, 1950s. Picture credit: [Novak, 2010].

ways: the car was "under the direction of an 'electronic brain' on a dream highway of the future" [Wetmore, 2003, p. 7]. This would have been the realization of the New York World's Fair exhibit from two decades earlier. Building to make this optimistic vision of GM's engineering ability a reality, engineers led by Joseph Bidwell and Lawrence Hafstad within GM Research installed "pick-up coils" on a 1958 Chevrolet: through a feedback system like that of the "electric eye automobile," these coils would detect a wire embedded in the road and adjust the vehicle's steering [Wetmore, 2003, p. 7]. Consistent with the image presented in the *Saturday Evening Post*, and a significant advance from Futurama's mechanical "half-pipes," alternating current would be the control mechanism of the future. A GM press release announced:

> An automatically guided automobile cruised along a one-mile check road at General Motors Technical Center today, steered by an electric cable beneath the concrete surface. It was the first demonstration of its kind with a full-size passenger car, indicating the possibility of a built-in guidance system for tomorrow's highways.... The car rolled along the two-lane check road and negotiated the banked turn-around loops at either end without the driver's hands on the steering wheel [Wetmore, 2003, p. 7].

Also in the 1950s, RCA's Vladimir Zworykin, a lead inventor of television technology, was working on an intelligent road system of his own. His concept, inspired by "railroad block signals," used circuits embedded in the road to magnetically sense vehicle speed and location, placing sensing and coordination capabilities outside of the vehicle: Zworykin's centralized planning model would send instructions to individual cars, and a 1/40th scale demonstration system was built for the 1960 Highway Research Board meeting in Washington D.C [Wetmore, 2003, p. 9].

The 1964 New York World's Fair hosted a second Futurama exhibit, in which GM presented an automated highway system much like Zworykin's system next to multitudinous nuclear-powered concepts. The exhibit was thus described:

> A revolutionary 'Autoline' expands the capacity of a three-lane expressway: Electronically, a control-tower operator steers, brakes and sets the speed of each car in an automatic lane; groups of cars move at equal intervals as a group [Wetmore, 2003, p. 9].[10]

Advances in electronics, and the expansion of early control theory into cybernetics through the work of Norbert Wiener and his contemporaries, seemed poised to make revolutionary applications of electrical control possible. Concern about human reaction times persisted. In a lecture given in 1960, Wiener made the case for automated control systems: "by the time we are able to react to our senses and stop the car which we are driving, it may already have run head on into a wall," he wrote, and the answer to this problem was the lightning speed of cybernetic control [Wiener, 1960]. But real developments in automated driving did not materialize, and research languished, out of the public eye, throughout the late 1960s and 1970s [Wetmore, 2003, p. 10]. Instead of autonomous vehicles—in a trend that arguably began with *Safety Last* [O'Connell and Myers, 1966] and *Unsafe at Any Speed* [Nader, 1965]—the public pushed for an increasing array of safety systems designed to make vehicles more forgiving of human errors—starting with rudimentary mechanical restraints such as

---

[10] This idea may seem quaint, but actually represents an important operations paradigm that is the focus of the last two chapters of this thesis. Public proposals to this effect are rare, but the concept remains important.

seat belts, and continuing with airbags, ABS, and other more complex automation systems.

1.1.3 Research at the Turn of the Century

Not all vehicle research focused on cars themselves, and active efforts to design and develop intelligent vehicle-highway systems (IVHS) occurred in the United States and elsewhere during the 1980s and 1990s [Wetmore, 2003, p. 11-12]. These efforts focused primarily on "nearer term" technologies, such as electronic tolls and driver information systems,[11] located outside the vehicle. In contrast to earlier radio or track-based electrical control, which similarly involved infrastructural changes, the "modern" era of self-driving vehicle research begins with computer-vision based approaches. While computer vision techniques had been experimented with for vehicle guidance since 1969, it was not until the 1980s that microelectronics became powerful enough to process images in near real time and compact enough to be placed on the vehicle itself [Dickmanns, 2007, p. v] [Dickmanns et al., 1994]. In the 1980s, Ernst Dickmanns's lab at Bundeswehr University, in Munich, Germany was active in some of the earliest of this self-driving car research and development. Their early vehicle, a specially equipped 5-ton Mercedes-Benz van—computerized controls could be used to perform all necessary driver inputs—was fitted with cameras, other sensors, and an image-processing system to close the control loop and drive the vehicle based on visual information [Dickmanns et al., 1994]. This early experiment, VaMoRs,[12] paralleled in the United States with a Carnegie Mellon-developed vehicle based on a Chevrolet panel van [Bogost, 2014], was successfully tested on roads without traffic at up to 60 miles per hour, and was soon followed by new projects produced as part of the EUREKA PROMETHEUS[13] project [EUREKA, 2014].

The largest driverless vehicle research project of the time, funded by EU member states to the tune of almost 750 million Euros, PROMETHEUS ran from 1987 to

---

[11] Such as 511 systems [Costello, 2007].

[12] Versuchsfahrzeug für autonome Mobilität und Rechnersehen, or "test vehicle for autonomous mobility and computer vision"

[13] PROgraMme for a European Traffic of Highest Efficiency and Unprecedented Safety

1995 and involved the VaMP[14] vehicle from Dickmanns's research lab along with its Daimler-Benz twin, VITA-II.[15] Like its forebear VITA, the VITA-II system, scaled down to fit in a passenger car (an S-class SEL 500), used digitized analog video signals to detect lanes and other vehicles [Ulmer, 1994b, p. 2]. Additional sensors detected brake pressure, temperature, steering angle, acceleration in lateral and longitudinal directions, and yaw [Ulmer, 1994b, p. 2]. The camera systems utilized two-camera setups with stabilized two-axis rotation to allow them to follow objects of interest in the scene [Schiehlen and Dickmanns, 1994]. And the technology largely worked, with the PROMETHEUS project coming to a successful conclusion in 1994 and 1995 with 1000 kilometers of largely autonomous operation in normal traffic conditions on Paris motorways, as well as a finale drive from Munich to Copenhagen [Dickmann et al., 2014]. Meanwhile, vision-based research also occurred actively in the United States; Congress's passing of the Intermodal Surface Transportation Efficiency Act in 1991 allowed for $650 million in funding for driverless vehicle research over the following six years [Novak, 2013].[16] The act spurred a consortium of nine organizations to work on meeting the 1997 deadline for demonstrations. In contrast to IVHS, which was seen as fundamentally limited by the human driver, this automated driving effort promised computationally-driven improvements in safety and environmental impact, as Rodney Slater of the Federal Highway Administration described in a 1993 announcement [Wetmore, 2003, p. 30]. Notably, a Congressional report described the goal as "hands-off, feet-off" driving [Novak, 2013], language which returns in the National Highway Traffic Safety Administration's recent *Preliminary Statement of Policy Concerning Automated Vehicles* and proves similarly unhelpful.

Further autonomous vehicle research has progressed from these vision-based beginnings. The DARPA Grand Challenge in 2004 and 2005 (evolving from the earlier DARPA Strategic Computing project), and the DARPA Urban Challenge in 2007,

---

[14] VaMoRs Passenger Car

[15] VIsion Technology Application. The original VITA was a D811 Mercedes van, chosen in order to store both computing equipment, power supplies, and environmental control for those electronics [Ulmer, 1994a, p. 37].

[16] Novak suggests that the biggest problem with the legislation was that it failed to define what a "fully automated highway system" would be, such that it could be achieved in the time allotted [Novak, 2013]. This is an important consideration for future work.

Figure 1-3: MIT's Urban Challenge vehicle. The grey cylindrical sensor at the top is a Velodyne LIDAR sensor. Below it sits a row of other LIDAR sensors and cameras. Picture credit: DARPA [DARPA, 2007].

spurred significant interest, and brought together industry groups and universities to try to extend the capabilities of fully autonomous vehicles.[17] While the Grand Challenge technologies bear little resemblance to current navigation approaches designed to deal with the complexities of real-world road environments,[18] the Urban Challenge vehicles, both via their sensor setups and algorithmic approaches to navigation, can be interpreted as another step along the path to the modern self-driving car, as instantiated within Google's design approach.

1.1.4 Automated Cars in 2015

Approaches today have a variety of forms. While Google is making headlines with its self-driving vehicle research, and is defining the look of the autonomous car with its

---

[17] The military should not be discounted as a contributor to and funder of this research, and military visions of automated vehicles—as a way to remove human soldiers and transport drivers from harm's way—are sources of inspiration that deserve fuller exploration elsewhere.

[18] These vehicles basically followed a predetermined GPS path, with gross navigation skills to avoid obstacles at smaller scales. However, the desert environment, though hostile in many ways, is not nearly so complex as an urban or suburban road. The navigational approaches applied there are fundamentally different than those necessary for an every-day self-driving vehicle. John Leonard, discussion with the author, December 3, 2014.

roof mounted LIDAR[19] scanner, other companies are developing vehicles with slightly different sensors and techniques. Google supplements cameras and LIDAR with detailed pre-mapping and highly accurate differential GPS [Guizzo, 2011]. Mercedes, so far, focuses more on cameras and radar, without using expensive LIDAR systems—and some of these features are already seen in the new S-class luxury sedan, though these vehicles have not yet reached the level of capability of Mercedes's test vehicle [Dickmann et al., 2014]. But these sensing choices come with their own trade-offs in terms of operations, safety, and how safety can be determined and proven. Tesla Motors is also working on an iterated approach to vehicle autonomy, focusing publicly on software updates to existing cars that add new automated features. Volvo is both involved in a public test of vehicle autonomy in Sweden [Volvo Car..., 2014], and is moving forward with automated systems in its production cars, building from years of research in Advanced Driver Assistance Systems (ADAS) and Advanced Vehicle Control Systems (AVCS). It is important to recognize that a spectrum of approaches are being developed today, from iterated improvements on current designs to radical approaches that attempt to achieve full self-driving at once. Automated[20]—or even "self-driving"—vehicles are not just one, uniform type of device. And, critically, there is no clear line to be drawn between their extremes.

One would be excused, however, for thinking differently. Portrayals of self-driving technology are often lacking in nuance. While there is a fair amount of popular argument about whether these vehicles will be beneficial (or how beneficial they will be) along various axes of interest, there is relatively little discussion of *what* these vehicles will actually be: the implicit assumption is that everyone is headed for the same technological goals, the same ultimate type of device, approached gradually or all-at-once.

At the Consumer Electronics Show in Las Vegas, in January 2015, visions of the driverless car stood front and center. Prototypes and concepts at CES worked glitz

---

[19] LIght Detection and Ranging

[20] I use "automated" precisely here. One of the main failures of the "self-driving" term is that it appears to make a clear distinction: either something can drive itself or it cannot. There are levels even here: Under what circumstances can it be self-driving? Is complete self-driving capability necessary before the term applies? These questions return again and again throughout this thesis.

and glamour into the dream of a driverless future. Mercedes-Benz's F015 autonomous car concept forgoes the traditional seating arrangement in favor of rotating chairs that turn the center of the car into a lounge or meeting space [Riofrio, 2015]. The car's side windows are relatively minimal, its doors taken up primarily by touch screens, set up for an immersive digital experience. But beneath this shiny media exterior sits a relatively unglamorous purpose: its publicity photos show serious-looking, young, white businesspeople in uniform grey work clothing [Simpson, 2015]. The environment is high tech and sharply clinical. Far from an exuberant depiction of the promise of media technology in the automobile, this future is so serious as to be dull: a homogeneous work space bleeding out into other parts of life. Compared to America's Electric Light and Power Companies' ad from the 1950s, where the car was a space for idyllic, family life, what we think we will use these prospective vehicles for has changed to the point where productivity is the stated or unstated goal. Our assumptions about use have changed (with seemingly little investigation into why), but they continue to be the background of both optimistic and pessimistic viewpoints.

1.2 Short-Term Promises

One popular vision of this technology is fundamentally optimistic: it puts forward the idea that self-driving vehicles will be commercially available in as little as five years, and that in just two decades the majority of cars on the road will be fully autonomous. Even respected business-information and consulting bodies have bought into this dream.[21] In these vehicles, the users would step in, select a destination, and would then be free to read, sleep, watch a movie, answer emails, or otherwise occupy themselves without needing to pay any attention to the operation of the vehicle. While this vision has its benefits, it makes many people nervous about ceding their driving agency to a computer system, especially on such a potentially short timescale [Lytton, 2014]. Google's Chris Urmson is still claiming that self-driving cars will be

---

[21] For example, IHS Automotive predicts 54 million such vehicles by 2035, which is not as extreme, but still a sizable fraction of road vehicles [Self-Driving Cars Moving..., 2014].

available in 5 years (a claim first made in August, 2014, a personal goal based on when his eldest son can get his license [Gomes, 2014b]), and expects these vehicles not to have steering wheels or the possibility of human control. That these vehicles would operate on public roadways is not stated explicitly—and seems optimistic given the speed of regulation alone—but is implied in these kinds of statements. Elon Musk's electric car company, Tesla Motors, is working on an "autopilot"[22] mode for their cars, intended to be released publicly by the summer of 2015. Though Tesla is clear that human involvement and judgment will be required at first, Musk suggests that once sufficiently capable vehicles come to market, human drivers may be outlawed [King, 2015]. In this vision, the technology is ready, or almost ready, and law and policy need to adapt. "Robots can already outdrive humans," the optimistic view says, "now everyone needs to get out of the way" [Fisher, 2013].

However, another common viewpoint is that major problems remain to be solved, and that fully automated vehicle technology is not yet ready for commercialization, whatever Urmson and Google say. Questions of safety, legality, and insurance cast doubts on the optimistic timescale. Moral and ethical quandaries—driverless car "trolley problems"—question how autonomous machines will respond to unusual situations and whether people will accept those responses. Would American freedom-rhetoric allow the outlawing of human drivers [Badger, 2015]? Perhaps not. Would we accept a computerized device that kills "3,300 Americans" per year, but replaces the "roughly 33,000 lives a year that perish on U.S. roads as a result of human error" [McFarland, 2015]? This is a difficult empirical question about the propensity of people to want to have a recognizable site for blame, and the tolerance of human beings for computer error. Among others, Bill Gurley, an Uber investor interviewed because of his connection to companies engaged in driverless vehicle research, believes that given the catastrophic potential for errors, the reliability requirement for these systems—"four nines," as he describes it—renders the technology "a long way

---

[22] This name explicitly invokes aircraft automation, and repeats the name of Chrysler's early "Auto Pilot" cruise-control system, from the late 1950s [Kessler, 2015]. Note that automated driving will not be available in residential neighborhoods because Tesla does not yet believe it is safe enough [Davies, 2015c].

off" [Shontell, 2015]. John Leonard, notable as one of the researchers who worked on MIT's entrant to the DARPA Urban Challenge in 2007, has gone on the record multiple times about his doubts, and has become a very public face for driverless car skepticism [Ramsey, 2015] [Gomes, 2014a].

There is much more to the story than is apparent in this simplistic dual-narrative. A few articles about automated vehicles remember the history, from the 1939 World's Fair [Petersen, 2014] to the Firebird concept cars [Walsh, 2013] to Dickmanns's work in the 1990s [Peseri, 2013] and DARPA's investments in automated driving—the latter largely stripped of its military valences [Look, no hands..., 2013]. But these are generally subsumed by an imagined historical through-line, and both popular visions of this technology focus on the future, and do relatively little to investigate the past or present: how we got to where we are now, or even what level of autonomy we currently see in our vehicles.[23] The idea that "if you are asked to take control, we've failed" has become deeply ingrained in our development culture, more than just as a risk-mitigation strategy but as a guiding ideology. But as we will see, this view does not hold together given historical examples.

1.3 Overview

This thesis investigates dominant narratives of self-driving vehicles and their gaps, and assesses the importance of alternative paradigms or narratives of automation. In chapter 2, I continue examining the sources of the idea, common both among the media and among self-driving vehicle researchers, that complete vehicle autonomy is the only future vision worth discussing and investigating. And I describe historical automation precedents that bear on automated cars. In chapter 3, I address design fictions and the issues presented by technologies involved in automation (and the fields with which visions of full automation are intertwined: computer vision, urban planning, machine ethics, and others). Taking the popular narrative at face value

---

[23] Though it should be said that John Leonard, for example, has much more nuanced views on the subject than get represented in these narratives.

presents the possibility of certain kinds of futures—with important disadvantages—while foreclosing on others in ways that must be carefully examined. In chapter 4, I describe an alternative paradigm for thinking about automation, and make the case for why it is an acceptable—and in many ways, more apt—alternative for considering what automation is and does. As I describe, supervisory control or joint-cognitive systems research presents a very different read on what automation means, and allows for the possibility of different futures with different stakes and trade-offs. Finally, in the conclusion, I briefly examine the historical re-shaping of roadways, and argue that adding autonomy to cars does not, alone, solve any social problems. Successfully addressing social issues requires careful design at a whole-system level, involving people, vehicles, infrastructure, and a broader appreciation for our hybrid engagements with machines. This appreciation allows us to shift the fundamental questions we ask about technology in a way that is, hopefully, productive for research and policy-making.

I originally set out to examine critically the cultural and social implications of self-driving vehicle technology. In my naïveté, I conceived of driverless cars as just-another-AI-technology. But as I researched, I discovered that not only did self-driving cars lack a fixed social meaning—their uses and contexts still up for grabs—but the technology itself is unspecified in profound and important ways. In an apt lesson for an early researcher, here was a thing that had not reached the stage of becoming, in Latour's parlance, a "black box," despite the media's treatment of it as such [Latour, 1988]. Below the surface, controversy bubbles. Turning up other research paths—and indeed, numerous historical counterexamples—made it painfully clear that the story I wanted was not the story I could tell. Each social vision is contingent, and will ultimately have to come into contact not only with physical reality, but the social realms of law, technology, and public acceptance. My research materials have therefore consisted not only of public media portrayals of automated vehicle efforts—comprising hundreds of news articles accessed between March 2014 and April 2015—but also other public documentation including patents; research papers in robotics, computer science, law, and ethics; and books on computers, automation systems, and their human interrelationships. I attended three information-technology or urban-planning related conferences at MIT to examine the presentation of automated vehicles within technical communities. I have also conducted several one-hour interviews with automated vehicle and transport researchers to attempt to gain internal perspective on their concerns and motivations. Rather than projecting from visions to cultural effects, I have had to dig deeper: putting these visions in the context of historical fields and precedents, exposing their gaps. And I hope this project has become a deep look into the ideologies and rhetorics of the self-driving car, one which hints at what society might be like once these vehicles are common, but more so investigates what different models for these vehicles exist, and what their public face might be if different people wrote, and different ideas guided, their stories.

Fundamentally, I am a scientist at heart,[24] and I seek to approach my work in that light. What idea, or hypothesis, does this thesis attempt to disprove? While I say many things that have significant backing, though they may not themselves be provable[25] in any real sense, my core case is that there is more than one story to tell about road vehicle automation. One can argue about approaches to automation and autonomy, based on historical examples, empirical data, or ethnographic findings, but the idea that the conventional story about automation (that it is clearly and naturally headed toward completeness and totality) is right, true, and the only story to tell is clearly without basis. Ultimately, I argue that the narrative of self-driving cars, as generally conceived, is far narrower than history suggests it should be. A broader view of this narrative shows a different side of automated vehicles that is more hybrid, more complex, and presents very different questions for policymakers, engineers, designers, and the public. Stories have power, and it behooves me, on finding the stories we tell to be so deeply one-sided, to try to rebalance the narrative.

---

[24] I mean this not necessarily in the methods of what is considered modern, experimental science, which does not fit easily within the realms of historical critique, but in terms of ethos, and some notion of falsifiability.

[25] I hope they are convincing, or at least suggestive.

Chapter 2

Narratives and Counternarratives

The popular narratives about self-driving vehicles have not appeared out of nowhere. But what are their roots? From what histories, fields, and ideas about technology do these stories come? Journalistic accounts of technological change are of course impacted by a wide variety of practices and perspectives, from the market and readership needs of news organizations to a pervasive culture of commodity scientism [Smith, 1983]. But a number of specific histories and fields influence the way driverless cars are figured in the press as autonomous machines: that the only image of the technology that bears investigation is that of fully self-driving vehicles, an imagined technological peak. Factory automation history, artificial intelligence and robotics, and automation experience in air and space each impact the way automated vehicles are conceptualized, but the dominant narratives of automation ignore the finer points of these histories. A fuller exploration of them is necessary to answer the question: Why does vehicle automation appear, at least on the surface, in the way that it does? And what relationship does this bear to actual history?

Human technological progress since antiquity has involved continual re-negotiations of human labor and the roles of animals and mechanisms in the labor process. Due to a confluence of factors—the miniaturization of computing technology, new advances in machine learning and artificial intelligence algorithms, a gradual increase in battery capacities, faster wireless networks—the horizons for everyday automation are broader now than ever before. But though media focus is on the future, our past is deeply involved in its presentation. News articles fret about what will happen when no one knows how to drive manually any more [Ross, 2014a], a classic fear of "de-skilling" that is implicated in so many other implementations of computers. Coexistence with autonomous or automated systems is sometimes presented as a fundamentally new situation, as if human beings had never before had to work and live with and next to automated systems—at the same time, robotic cars are sometimes situated as the next step for robots after the factory, their final emergence into the real world having conquered the factory floor.[1] Automation already has a deep history in the industrial sector, presenting new benefits and dangers, and requiring new roles for human laborers. Current debates and fears about de-skilling, human jobs, and the role and value of human labor return us to questions that have plagued factories, and labor's relationship to machinery, since the early Industrial Revolution.

2.1 The Automation of Work

The story of early 1800s textile mills is a familiar one [Smith and Anderson, 2014]: skilled artisans made obsolete by the lower cost and higher productive capacity of mechanical labor. By analogy, new types of skilled labor (taxi and limousine drivers, or even bus drivers) are now under threat, and their fate should be no different than that of the workers who came before. But just what the processes of standardization, mechanization, and automation (or "automatization" as the cyberneticists referred to it) have done to the factory, and to laborers in it, is not clearly understood among many who write about autonomous cars. This elided history—which is substituted for by a non-existent person-less factory fabricated by the collective imagination—is relevant, perhaps more than ever, to the future of transportation.

A search for the beginnings of industrial automation takes us to the middle of the 18th century: Vaucanson's mechanical loom dates to 1741, and formed the basis of later developments in weaving by Joseph Marie Jacquard [Diebold, 1959, p. 9].

---

[1] See for example "Robot Vehicles" in *RobotWorx* [Robot Vehicles, 2014], which describes automated cars as having the sensors industrial robots have had for many years.

But the first example of "complete" industrial automation originating in the United States does not come until Oliver Evans's work in the 1780s on automated grist mills [Smith, 2014, p. 5]. Through a series of elevators and descenders, horizontal screws, spreaders and rakes, his mill moved grain from raw agricultural commodity to finished product: sifted flour. And ideally, all parts of the process would occur without human intervention.[2] Evans's innovation was to place these devices in succession so as to allow continuous production, and the elimination of many slow human jobs that degraded the quality of the product by tracking dirt and contaminants around inside the mill [Evans, 1848, p. 203].

It took some time for the automation found in the Evans mill to spread across other industries, but Evans's contemporaries were not uninterested in increasing efficiency and output.[3] Manufacturing itself was a site of public debate, pitted against the "inherent virtue" of agricultural pursuits. Tench Coxe, a political economist, wrote in 1810 that "new machines and power sources allowed even greater productivity with less labor, further underscoring the connection between technology and republican virtue" [Martello, 2010, p. 217]. To Coxe's romantic view, these machines worked "as if they were animated beings, endowed with all the talents of their inventors, laboring with organs that never tire, and subject to no expense of food, or bed, or raiment, or dwelling" [Coxe, 1814, p. xxv]. Though we may have lost this romanticism, we haven't lost this perceived animism. Automated machines, like self-driving cars, continue to excite, impress, and cause fear due to this transformative, if mechanistic, aliveness.

But these romantic words did not represent the whole reality of industrial machine labor. Human labor of maintenance and supervision is implicit in these manufacturing machines—even the Evans Mill—but it is rendered invisible by the rhetoric that the machines themselves require no bed or board. At the same time, Coxe's use of the

---

[2] Evans cited a reduction in labor and expense of "fully one-half," which is not the same as its complete elimination. Behind his concession to manual switching of machines on and off lies a great volume of labor that is excluded from the traditional narrative: tending and configuring the machines during their operation; examining machines for wear, degradation or failure; fixing the machines when they break down [Evans, 1848].

[3] Paul Revere, one of America's early industrialists, applied alterations in manufacturing techniques to transition himself from an artisan worker to manager and overseer of others over his long metalworking career [Martello, 2010, p. 187].

word "endowed" should focus our attention on exactly which "talents" of the inventors have been automatized, and the human labor necessarily involved in that conferring of capabilities.

Driven by the plight of those displaced—like Detroit autoworkers who faced unemployment with the rise of Japanese (popularly read as "highly automated") automotive might[4]—the history of automation that gets mobilized is one of a teleological progression toward complete automation of all sectors of work. Actual historical circumstances are less important, popularly, than the general perception that automated systems, such as the robotic arms used for painting and assembly, are reducing the overall pool of available jobs.[5] Modern automotive manufacturing is not alone in presenting conflicting ideas of what automation can do. Military manufacturing and assembly line history is rife with contradictions. While certain competencies were transferred from the skilled worker to the technical apparatus, human oversight and operation were still integral to the production of weaponry using the new technology of the "American system" of interchangeable parts. And it was not clear until after the fact that more mechanization was necessarily better.[6] The same exchange of competencies characterized Ford's assembly line, which began true mass production in America [Hounshell, 1984, p. 217]. Fixtures and gauges were designed to allow for use by unskilled machine tenders. But while the gauge simplifies the assurance of quality, it does not automate it: it simply changes the effort from a more complex judgment of quality and measurement to a simpler one.[7] Ford also attracted a wide variety of well-educated skilled mechanics to his automobile plants [Hounshell, 1984, p. 223]. Like Evans, Blanchard, Hall, and others before them [Smith, 1977], these

---

[4] It is worth mentioning here that the core of *kaizen* "lean" manufacturing is not automation but increased trust between management and labor, assisted by continual communications and high job security [Nye, 2013, p. 198–199].

[5] A report produced for the Society of Motor Manufacturers and Traders in the UK suggests an alternative: automated cars could be expected to create 320,000 new jobs in the United Kingdom by 2030, only 25,000 of which would be in automotive manufacturing directly [Tovey, 2015], which has everything to do with new labor involved in their development, production, testing, and use.

[6] For example, Harper's Ferry, where armorers resisted the mechanization of their craft, remained "competitive" with costs at the more highly automated Springfield Armory through the mid 1830s [Smith, 1977, p. 324].

[7] As Donald Norman writes of this fundamental principle: "the world remembers things for us, just by being there" [Norman, 1993, p. 147].

mechanics applied their skills to design machines, and simplify and standardize work processes. The individual judgment of the assembly line laborer was displaced into standardized tools and fixtures, built into these technologies by the labor of skilled machinists and designers. But the new labor generated by automation has often been made structurally invisible.

Also around the turn of the century, Taylorism in factories created "new managerial functions" performed by "new classes of people with new titles and more clearly specified responsibilities" [Aitken, 1960, p. 120]. A focus on the people—who are they? where are they? what are they doing?—shows that one of the fundamental and enduring characteristics of Taylor's system, the expansion of management roles and the further division of labor, is not about automation but about new and altered types of human work: industry continued the removal of strategic decision-making from the workers most physically engaged in production, installing it instead within formal organizational structures and the managers that constituted them.

Machines "coupled to the outer world through the mechanical equivalents of sense organs" allowed the the Rockford Ordinance Plant to operate in 1953 with a "largely automatic process of turning steel bars into 155 millimetre shells" [Wiener, 1953]. Around the same time, in Ford Motor Company's Brooks Park engine plants, near Cleveland, forty-two automatic machines "linked together by transfer devices that automatically move the blocks through the complete process, perform 530 precision cutting and drilling operations" [Diebold, 1959, p. 9]: through 1,545 feet of assembly line, no human touched the parts. Such is the contemporary vision of the automated factory. In truth an operator stood by each machine, ensuring its continued operation. This human labor, menial as it may be, is not often recognized.[8] The worker no longer makes choices about how to bore a part—choices now built into the industrial equipment she oversees—but chooses to turn the machine on and off.

Patterns of contingent change repeat for numerical control (NC): adoption of

---

[8] One worker described his experience: "I don't do nothing but press those two buttons . . . Sometimes I use my thumbs, sometimes I use my wrists and sometimes I lay my whole arm across" [Diebold, 1959, p. 10]. And yet, despite the meniality of his labor, it is still integral to the process of production.

robots to replace assembly line jobs such as spray painting and welding was gradual, with only about 6,000 robots in use in American factories by the mid 1970s [Nye, 2013, p. 159]. Industrial robots had a way of generating large contingents of skilled human laborers who still needed to be paid for their services: to tend them, and to repair them when the broke down.[9] The development of NC machines proceeded with a specific interest in eliminating skilled workers, but the jobs that disappeared were largely unskilled or semi-skilled laborers [Nye, 2013, p. 164]. And while Norbert Weiner prophesied in 1950 the end of "deadly uninteresting" jobs within 20 years, such changes have still not totally come to pass [Nye, 2013, p. 161]. To compound the problem, new industries of skilled workers—record-and-playback machine designers, and NC machine programmers—sprang up to furnish factories with their tools. This historical thread should focus our attention on what is added, rather than removed, by autonomous vehicles: more complex computer systems may make the driving task simpler for a given level of safety, but make the system engineering task, and the tasks of maintenance and repair, more and more complicated.

As control is further constituted within management, the roles of management may be rendered more and more menial themselves: the Air Force's Integrated Computer-Aided Manufacturing program (ICAM) attempted to automate management functions, "to try to reduce the enormous indirect costs that have resulted from the effort to reduce labor costs and remove power and judgment from the shop floor," costs that have continued to dog new rationalization strategies [Noble, 1984, p. 330]. Automation can be used both to routinize work—for the manual laborer—and to eliminate the routine in favor of the creative—for managers and newly generated classes of creative workers.[10] This is analogous to the current process of self-driving vehicle design: automated systems substitute the mechanical process of controlling the vehicle inputs with the new task of the intellectual supervision of the driving system. This new task

---

[9] These early experiments did not increase profits because of the volume of highly skilled labor needed to keep the machines operating [Nye, 2013, p. 162].

[10] ICAM, like the mythical ouroboros, sought to offer automation as the "solution to the problems generated by automation," providing automated scheduling functions, inventory control, and design tools to "provide better management control" and "free management from excessive routine duties to do creative work" [Noble, 1984, p. 330].

is not necessarily simpler or easier, though it may be intended to be, but can be seen as representing a management, rather than labor, role in vehicle operations.

Self-driving vehicles make sense today because of a general climate that believes in the possibility of automating knowledge work, and their development feeds back into the perception that other complex tasks will soon yield themselves to automation. Bill Gates, speaking at the American Enterprise Institute, suggested that a large portion of the workforce will find itself displaced by robots in the next 20 years, including accountants and other white-collar jobs [Reed, 2014].[11] His comments play directly into contemporary fears about automation, and support only the narrative that job-loss due to automation is inevitable, and workers (or drivers) should just get out of the way. To this view, the last two hundred years of innovation in automation is unidirectional and largely undifferentiated: from steam power to the assembly line, from Taylorism to roboticization, each was yet another nail in the coffin of the human worker. But while it is undeniable that automation has changed the character of human labor, this perceived uniformity in automation processes is a figment of the collective imagination. Control and rationalization of the work processes of the individual—in a way mechanizing her—also created new classes of worker and expanded the role of human managers in the labor process. Automation may look very different depending on where in the hierarchy a person happens to fall, but the historical lesson is that human involvement remains, though altered in space, time, and kind. As John Diebold pointed out in 1953, there will be “no worker-less factories as a result of automation” [Diebold, 1953, p. 63-64] precisely because human beings will be needed to construct, to repair, to manage, and to oversee.

2.2 Artificial Intelligence

Artificial intelligence has its own history, intertwined with that of the automation of work, that feeds into driverless car narratives. Though many popular portrayals

---

[11] Unfortunately, Gate’s suggested remedies (low-to-no taxes and decreasing minimum wages) are painfully biased toward corporate profits.

of AI care little about the actual history of the field,[12] advances in AI are seen to cross-pollinate: culturally accepted narratives of AI are used as models of progress and evidence of the technological inevitability of self-driving vehicles.

Intelligent machines are not a new idea. Just as automation has long been a part of human history, dreams of artificial life suffuse our legends—though as Minsoo Kang rightly notes, these early automata are of diverse types and kinds, some readable as automata only in hindsight due to their similarities with contemporary robots [Kang, 2011, p. 15]. Nonetheless, that we continue to retell these stories should tell us something about the lure of the automated, its power as a "hybrid entity" that can mediate between living and nonliving worlds [Kang, 2011, p. 19]. The myths about Hephaestus and his creations, notably Talos, a golden female automaton, come to us from antiquity [McCorduck, 2004, Ch. 1], but continue to be cited as historical antecedents in literature on autonomous robots and their ethical issues [Lin, 2012, p. 3]. Judah Loew ben Bezalel, a Talmudic scholar, is in legend the creator of the golem, a being animated from clay who functioned as a spy against the Gentiles [McCorduck, 2004, Ch. 1]—the rabbi occupies a special position among the most prominent AI researchers of the 20th century: Marvin Minsky and Joel Moses grew up with a "family tradition that they are descendants of Rabbi Loew," and Joel Moses claims a number of other American scientists (including John von Neumann and Norbert Wiener) also consider themselves among the descendants [McCorduck, 2004, Ch. 1]. This is all to say that ancient myth and legend continue to subtly underpin research in autonomy and artificial intelligence.

As Jessica Riskin chronicles in her studies of eighteenth and nineteenth century automata, clever inventors like Vaucanson, interested in going beyond mere representation, created a variety of impressive simulations of life.[13] But meaningful interaction requires closing the loop between sensing and acting with corrective feedback.[14]

---

[12] They generally pull from major moments that garnered enough popular attention to enter the cultural knowledge-base: ELIZA, Deep Blue, Watson, and Siri.

[13] *Simulation* in its modern sense, meaning "experimental models that can elucidate the natural," rather than its contemporary sense which would have connoted artifice [Riskin, 2003, p. 605–606].

[14] This began in the 17th century in windmills [Hills, 1996], before James Watt's famous steam engine governor (late 18th century). James Clerk Maxwell's 1868 paper on centrifugal governors in steam engines became one of the central papers in early control theory [Mayr, 1971].

Bringing together a number of existing areas including control theory, cybernetics—from the Greek *cybernetes*[15] meaning "steersman" [Wiener, 1956, p. 6]—extended automation's reach to more complex systems: "control and communication in the animal and machine." Cybernetics envisions the world in terms of feedback mechanisms: all systems that pass and respond to messages internally are within its explanatory sphere [Wiener, 1956, p. 10-15].[16] AI is interested in re-creating many of the same self-regulating systems within computers—and claims a similarly broad explanatory role, whether the goal is playing chess, investing in the market, or driving a vehicle.

2.2.1 Logical Beginnings

The Dartmouth Conference in 1956—hosted by John McCarthy, who originated the term "Artificial Intelligence"—assembled many who would continue to be preeminent researchers over the next decades:[17] all were united by "the idea that there was a rigorous and objective way of explaining the human intellect" [McCorduck, 2004, Ch. 5]. General purpose intelligence was the dream motivating "Dartmouth Summer Research Project" work in "neuron" networks, self-improving machines, and computational creativity [McCarthy et al., 1955]. Representative of this first age of AI, Allen Newell and Herb Simon's physical-symbol system hypothesis states that "symbols lie at the root of intelligent action" [Newell and Simon, 1990, p. 109].[18] But a symbolic AI approach rooted in the physical symbol-systems hypothesis did not yield results as quickly as expected.[19] As Pat Winston put it: "Everyone searched for a kind of

---

[15] This is as written in Wiener's contemporary papers. It could more properly be written as *kybernetes*, or κυβερνήτης [Wiener, 1965, p. 11].

[16] Homeostasis, balance, and motion disorders like locomotor ataxia and Parkinsons all are cited by Wiener as within the cybernetic realm.

[17] Marvin Minsky, Nathaniel Rochester, and Claude Shannon also co-hosted, and attendees included Trenchard More, Oliver Selfridge, Ray Solomonoff, Allen Newell, and Herbert Simon.

[18] Such constructs are *symbol systems* in that they contain symbols and processes that act upon symbols. And they are *physical* in that they obey physical laws and are realizable through engineering. They operate via heuristic search, pruning a tree of possibilities in a hopefully "intelligent" way [Newell and Simon, 1990, p. 124].

[19] It did produce expert systems useful in particular domains, but there is good reason to think we do not spend most of our time processing symbols: symbolic logic takes a lot of mental capacity, so we generally use other sorts of shortcut processes to come to decisions. Pollock calls these "quick and inflexible" or "Q&I" models [Pollock, 1989, p. 120] and Dennett refers to them as "habitual methods" or mechanical routines [Dennett, 1990, p. 157]. It may well be that introspection on thought is a

philosopher's stone, a mechanism that when placed in a computer would require only data to become truly intelligent" [Winston, 1987, p. 4]—they did not find one.

The failures of logic-based robotic systems like Shakey—the DARPA-funded, SRI robot named for its tendency to shake when in motion [McCorduck, 2004, Ch. 10]—to achieve intelligent results disillusioned many researchers. While limited processing power made responses and reactions slow, the uncertainty inherent in the real world made them unreliable. Far from the generally accepted story of continued, natural progression (that while AI may not have been possible before, it must certainly be possible now due to continued development and greater computing power), artificial intelligence proceeded in fits and starts. Achieving humanlike intelligence via logic systems was the archetypal dream of early AI, but while logic is part of the puzzle it is not the entirety of it.

2.2.2 Robotics, Embodiment, Probability

Persistence, funding, and new approaches did yield results. DARPA's massive 1983 US Strategic Computing initiative had a 10-year funding plan, including work on "image understanding" and interpretation, an autonomous land vehicle (ARV), a pilot's associate, and computerized battle management software [Roland and Shiman, 2002]. It was canceled without achieving its stated goals, but this was not truly failure: "AI now performs miracles unimagined when SC began," especially in areas of natural language processing, "though it can't do what SC promised," [Roland and Shiman, 2002, p. 328]. Other approaches to building intelligent robots were developed in the 1980s. Rodney Brooks's subsumption architecture [Brooks, 1987, p. 353] attempted to cut out cognition altogether, focusing on sensing and reaction and iteratively building small systems that could cope with uncertainty: insects, for example, lack cognition, but respond more adroitly to the world than Shakey and its contemporaries. Instead of modeling the world, this approach treats the world as "its own best model" and

---

large part of why the logical theory seemed so compelling: researchers like Newell and Simon used "think aloud" experiments to identify problem-solving techniques [McCorduck, 2004, Ch. 10], which seems naturally to suggest a logical mode.

focuses on embodiment and action within the environment [Ekbia, 2008, p. 256].

The layered subsumption approach is modular, building complex behaviors on top of and out of simpler ones, and presents tempting biological similes.[20] But while it could theoretically be used to engineer greatly complex systems (cars), this architecture did not alone manage to allow the creation of such robots. Brooks himself abandoned the project of building intelligence from these humble, subsumptive blocks.[21] He skipped the middle of the evolutionary tree, straight to humanoid forms, because the evolutionary approach was too slow: Brooks reported, it was "starting to look like, if I was really lucky, I might be remembered as the guy who built the best artificial cat," a distinction he apparently did not desire [Brooks, 2002, p. 65]. But to do this, bootstrapped knowledge from other sources was necessary, bringing back some of the world modeling inherent in older, symbolic approaches.

Dealing with uncertainty in a more pragmatic way, statistical AI approaches dominate the modern robotics space, typified by Sebastian Thrun, Wolfram Burgard, and Dieter Fox's book *Probabilistic Robotics*.[22] This research, organized around algorithms based on Bayes' rule, distinguishes itself from prior model-based (logical) or "reactive behavior-based" (subsumption) approaches by dealing gracefully with uncertainties in both the world model and sensor inputs [Thrun et al., 2005, p. 9]. Probabilistic robotics depends on mapping the world and modeling the robot's relation to it.

Localization has been a fundamental issue for robotics: In order to figure out how to act, a robot needs to know where it is; and in general, robotic environments can be expected to change rapidly with motion, and to vary significantly from place-to-place [Fernandez-Madrigal and Claraco, 2013, p. 4]. Motions must continually be reevaluated in new and emerging situations. Key to this puzzle, the robot must be able to determine its own relationship to some target location:

---

[20] One can imagine the human being as a robot with a subsumption architecture, where breathing and heartbeat are lower layers than balance, which is lower than voluntary motion, which is lower than logical thought.

[21] Instead, his later research, for example the Cog robot, shifts focus from "emergence" to "integration," and reversed some of his initial fervor to avoid representation [Ekbia, 2008, p. 258].

[22] The book begins with a nod to the self-driving car project: "Wouldn't it be great if all our cars were able to safely steer themselves, making car accidents a notion of the past?" Utopian vision of technological possibility are not limited to press accounts [Thrun et al., 2005, p. 3].

this "target-robot relation" is an "unavoidable" component of successful navigation [Fernandez-Madrigal and Claraco, 2013, p. 5].[23] While localization is the usage of environmental elements to estimate a robot's position, those elements themselves must be known in order to perform this process. Mapping is the complementary process, the estimation of "*unknown* spatial relations that exist between environment elements" to allow for subsequent navigation [Fernandez-Madrigal and Claraco, 2013, p. 5]. This presents a precedence problem: to build maps with autonomous systems, the systems must know where they are; in order to know where they are, they must have maps to measure against [Fernandez-Madrigal and Claraco, 2013, p. 6].

Though one way to determine robot position in the world is through simultaneous localization and mapping (SLAM)—a serious research area in robotics, as it is ideal for regions that are impractical to map beforehand, such as ones that are always changing—it is easier to localize by comparing measurements to a known map. However, that implies that the map is pre-made, and therefore sets limits on the rate at which the environment can change and still allow the robot to operate. Today, localization and mapping problems alone are both considered "satisfactorily solved in practical situations"—though no single algorithmic approach is ideal for all purposes—given sufficient computational resources and environmental data [Fernandez-Madrigal and Claraco, 2013, p. 5-6]. SLAM techniques, however, are still somewhat less reliable or developed.

2.2.3 Learning and "Knowing"

But despite that probabilistic robotics fundamentally underlies the major automated vehicle innovations reported by the popular press, it is rarely, if ever, mentioned. What receives attention in the self-driving car narrative is a different gloss on statistical systems: today's buzzword AI technique, "deep learning," appears to yield radical new possibilities everywhere it is applied—though it really just means a return of

---

[23] This relation can be expressed in different forms—either quantitative representations such as maps, or logical, prepositional statements—and these representations carry their own techniques of interpretation.

neural networks,[24] which, armed with some improvements in weight generation, more layers of nodes, and more data to train with, have been able to eclipse many of the previous techniques.

Training the neural network involves adjusting the connections' weights so as to be more sensitive in distinguishing different types of inputs. This tuning, traditionally, has been done using supervised methods and backpropagation of errors: the output result is compared to an expected classification result (determined by humans), and the error used to tune the weights more appropriately.[25] Deep learning extends this technique, decreasing the amount of "feature-engineering" (identifying the features the network should use to distinguish inputs) required to train the system. The "ideal" training situation is entirely unsupervised: the network independently "learns" the features of unlabeled and unclassified input, without any human input. This ideal prospers both due to convenience in terms of human effort (less programmer effort required to build labeled sets of training data[26]) and due to the deep-seated ideology, which reappears in the imagined operations of self-driving cars, that machine independence is paramount. Unsupervised learning is often seen as more impressive and valuable research.

Popular claims about the utopic promise of deep learning, to learn about the world "on its own," abound. Google's and Stanford's recent improvements in image recognition, driven by deep learning [Markoff, 2014b], triggered a wave of popular speculation about computer vision meeting or surpassing that of humans. Such advances might seem to translate automatically into the self-driving car space. Though some articles note that current state-of-the-art image recognition is still considerably less capable than people are, many articles still present the uncritical idea that we have solved the

---

[24] These systems, also known as "multilayer perceptrons," still have little or nothing to do with actual neurons: they are brain-like in only the barest toy-model sort of way. Consisting of webs of interconnected nodes, they may have multiple layers (hence "deep" in deep learning) or may be shallow. Each node corresponds to a particular feature or combination of features in the input [Jordan and Bishop, 2004].

[25] These techniques are actually used in the probabilistic robotics field, to train learning algorithms to develop maps out of noisy sensor data [Thrun et al., 2005, p. 284-297].

[26] This labeling is considered "inhumane" work, and contributes to a reluctance to use machine learning in automotive applications, according to Göde Both [Both, 2014b].

image “understanding” problem—or rather, that increases in computing power mean it will necessarily soon be solved. By this narrative, seemingly limitless possibilities are open before us: a new universal problem-solver, like feedback control before it, seems to have broadened the frontiers of the future.

Though computer science and philosophies of AI have been using “intelligence,” “knowledge,” and “understanding,” among other words, to talk about computers since the beginning of the field, these uses should not be taken at face value. “Intelligence” is slippery, and its definition is not constant over time. Weaving was once considered to be a peculiarly human capability, a sign of an advanced, intelligent mind [Riskin, 2003, p. 627], but after Vaucanson’s loom allowed mechanical devices to weave seemingly on their own, this capacity was no longer seen as uniquely human and was no longer a marker of intelligence. The same process occurred with chess in the 1990s: when IBM’s Deep Blue beat Gary Kasparov, chess ceased to be the standard by which intelligence could be judged, precisely because it had been achieved. Real intelligence had to lie elsewhere: for example, in the game Go, mastery of which has continued to elude machines [Riskin, 2003, p. 623].

Nevertheless, machines manage to do things that *seem* intelligent. Our heuristics for understanding the observed behaviors of machines slip slowly over time from self-conscious scare-quoted use into casually accepted statements. While automatic translation may seem “intelligent,” or a system that can define *étoile* as “star” may seem to possess “knowledge,” this intelligence or knowledge is perhaps very different than our own. A deep epistemological question presents itself: how do we know, and how do machines “know”? Many AI systems operate via statistical pattern recognition, so we may ask whether we believe human intelligence is also merely pattern recognition: Does a system that can associate *star* with its definition really know what a star is? Is linguistic association sufficient for knowledge?

H. R. Ekbia and others remind us that we should be skeptical of the applications of cognitive terms to computational processes [Ekbia, 2008]. Ornstein, Smith, and Suchman, in their 1984 article “Strategic Computing,” warn of the difference between domain capabilities and “common sense,” and suggest that “unwarranted optimism” and a particular funding climate (issues also present today) push researchers to mask the shortcomings of AI with "semantic shifts" [Ornstein et al., 1984, p. 14]. We alter the definitions of "knowledge" and "understanding" to fit what our machines can do, and these claims, taken literally, "give rise to unrealistic confidence in the power of the technology" [Ornstein et al., 1984, p. 15]. The two-way process of linguistic and technological change—that intelligence gets applied to describe whatever researchers manage to achieve, while real "intelligence" retreats away from each computational advance—leaves these terms poorly defined. The "understanding" involved in even the most impressive current systems is limited. Image recognition systems, for example, are not able to answer questions about the scenes that have to do with the material properties of the objects, or likely results of various actions [Gomes, 2014c].[27] But as technical stories spiral out of the lab, such subtleties are often lost, and rhetoric or appearance of intelligence outweighs the actual technological substance behind them.

Artificial intelligence has a powerful public narrative of self-evident progress. We would do well to remember that progress in AI has been slow, contingent, and heterogeneous, a product of a wide variety of concepts and techniques. Logical, precise models, reactive-subsumptive control, and probabilistic re-framings of the navigation problem all had their parts to play in this evolution. It is difficult to argue with techniques that "work" and spawn numerous real-world systems, from personal assistants to self-driving test vehicles. However, the limits of each new paradigm are rarely obvious *a priori*, and it remains to be seen how far current algorithms can carry the field, and whether subsequent algorithmic advances will allow progress to continue unabated. Autonomous cars, as a research area of artificial intelligence, find themselves in an environment of surging hopes and interest in the field, driven by new or newly extended techniques. This climate may aid the naturalization of the idea that fully-automated vehicles in the AI/robotic mode are inevitable, or even coming soon. But no one has yet found Winston's philosopher's stone.

---

[27] Another recent slew of articles focuses on the "self-aware" Mario created by researchers at the University of Tübingen. Any pretense to worry about a "self-aware AI . . . with an insatiable desire for material wealth" that knows "how to kill" [Vincent, 2015] is simply journalistic excess, as the researchers themselves well know.

2.3 Aircraft and Autopilots

Autopilots have a strong rhetorical role as models of ground vehicle automation, but autopilot design and use differs significantly from its popular representation. It is not a fully automated system that controls the entire aircraft, but a set of specific tools leveraged by pilots to increase their agency. Autopilots are also complex systems, with many modes that require deep technical knowledge and training to use appropriately [Harris, 2014a]. Stories of accidents due to autopilot mode confusion abound. The official report for the Air France Flight 447 crash cites a failed hand-over of control from the autopilot system to the pilots as the cause: when the autopilot system shut down due to a failure of its airspeed indicators, the pilots were not prepared to so suddenly resume the task of flying a large airliner in poor conditions [BEA, 2012].

But these concerns with cockpit automation are not new, and pilots and aircraft designers have been negotiating human roles in flying machines since the very early days of aviation. The tension between stable and unstable aircraft design—will the aircraft "fly itself" in a given orientation or does it require constant control inputs to maintain course—goes back as far as the Wright Flyer, as do the competing professional identities that accompany them: are you simply a chauffeur, or a "true airman" [Mindell, 2011, p. 21]? The Wright's focus on the operator's skill created the unique profession of the "pilot," and unstable aircraft were the standard for years, until human fatigue over long periods of flight changed the importance of stability to aircraft operation [Mindell, 2011, p. 22-24]. Stability, this subsequent virtue of a well-engineered airframe, had to be again renegotiated in the context of supersonic flight, where stability problems related to supersonic airflow required electronic solutions—pilots were still skeptical of black boxes seemingly out of their control.[28] Today, many airlines have guidelines that require automation systems to be used whenever possible [Flight Deck Automation Working Group, 2013, p. 38], but automation in the cockpit has long been a contested technology.

While the AF 447 example shows the perils of human interaction with automated

---

[28] See J. O. Roberts, "The Case Against Automation in Manned Fighter Aircraft," which argues for emphasis on information display rather automated control [Mindell, 2011, p. 35].

systems, an issue which will affect automated vehicles as well, its significance within aircraft development does not seem to be broadly appreciated outside of that field. It gets taken—including by one of the automated vehicle researchers I interviewed—as meaning that human interaction should *never* be expected, and that until automated vehicles can be fully self-driving in all circumstances, they will not be sufficiently safe. But aircraft companies are learning a different lesson: Boeing and Airbus, the two most prominent manufacturers of commercial airliners, are distancing themselves from complete automation, while increasing the computerization of their cockpits through digital displays and tools to assist pilots in the task of flying [787 Dreamliner..., 2012] [Brown, 2011]. The focus for future development is on adaptive or adaptable automation rather than complete computer control. Both types of systems allow for balancing the operator's load—adaptive systems automatically, adaptable systems by operator request—taking on tasks during high-stress, busy situations, but handing back tasks during periods of limited load in order to keep the operator informed of and engaged in the operation of the system [Liu et al., 2012]. Such systems have their own engineering challenges, but represent a very different answer to the question "what should the role of pilots be?" than does the further complete automation of aircraft operations [Kaber et al., 2001]. Asking people to be mere machine tenders, present only to ensure the continued operation of the machinery, is indeed untenable, as it produces boredom, inattention, and the risk of catastrophic failures like the AF 447 crash. Aircraft experience shows that sustained human engagement with automated systems is possible, but has to be designed into the system from the beginning.

2.4 NASA and High Technology

NASA, as a purveyor of high technology, also gets leveraged in the popular narrative as a source for automated vehicle design inspiration: "And so Google's new vehicle design takes a leaf out of NASA's design book to cope with such eventualities. 'It doesn't have a fallback to human—it has redundant systems,' said Fairfield. 'It has two steering motors, and we have various ways we can bring it to a stop"'

[Simonite, 2014]. This is a reductive view even of NASA's unmanned space systems, let alone manned systems with which the Google vehicles must share the critical characteristic of containing human occupants. Spaceflight, both manned and unmanned, provides myriad reasons to invest in automated systems: time delays prevent or at least greatly inhibit remote control from Earth; conditions where humans are not uniquely equipped to survive suggest the use of mechanical explorers instead; and precise control functions with redundant backup systems allow the use of the unique capacities of computerized systems to monitor constantly and act immediately in the event of an emergency. But the story of automation in spaceflight is not nearly so simple as it appears. Rather than being the ultimate and obvious endpoint of a progression of human engineering in space, the roles and implementations of automation remained fraught and highly contested.

2.4.1 Manned Spaceflight

Manned spaceflight might seem to be a story having little to do with automation, as its public picture has often focused on the skill and bravery of human astronauts. But it did not begin this way: from 1952 to 1954, *Collier's* magazine published a series of articles titled *Man Will Conquer Space Soon* which described Wernher von Braun's vision for manned spaceflight [Collier's Magazine, 1954].[29] Lurking behind this majestic picture of manned spaceflight was a dark realization for prospective astronauts: to von Braun himself, the astronaut would be a mere passenger, ferried into space by automated rockets.[30] Von Braun was not alone in his cautions to test pilots about the limits of their capacities: Richard Horner warned that technology would progress faster than human beings: the "link" that improves the least over time "is the man himself" [Mindell, 2011, p. 19].

Pilots—prospective astronauts—were not about to find themselves cut out of the

---

[29] The illustrations from this magazine are reputed to be some of the "most influential images of the early space age" [Scott and Jurek, 2014, p. 9].

[30] In a speech to the Society of Experimental Test Pilots in August of 1959, a hostile audience for this sort of rhetoric, von Braun emphasized that human control in rocketry is "actually undesirable" because human beings are "outrageously slow and cumbersome" in missile terms [Mindell, 2011, p. 66-67].

control loop. As Mindell chronicles in *Digital Apollo*, their resistance involved professional pride and concern about reliability. Al Blackburn's response to von Braun was particularly empassioned, and recounted his many experiences with "brain-dead autopilots, broken fire control systems, and failed cockpit computers" [Mindell, 2011, p. 68]. Pilots had reason to doubt the capabilities of computerized guidance and control systems. Existing X-15 fly-by-wire control [Jenkins, 2000] was even controversial. Milt Thompson, a test pilot in the X-15 program, said of it: "you would like him [the engineer responsible] to be in the airplane with you to be exposed to any adverse results" [Mindell, 2011, p. 55]. But human resourcefulness was one answer to automated technology's lack of reliability.

Such debates over control did not go away, but lasted through each of the Mercury, Gemini, and Apollo programs. Some of this was political: Apollo was conceived of by Kennedy[31] as a program of diplomacy, which would help to "win the battle that is now going on around the world between freedom and tyranny" by impressing the "minds of men everywhere" [Kennedy, 1961]. Space missions had to be piloted because the Soviets were sure to continue their program. Human roles were important: Soviet spacecraft continued to be more automated than their Western counterparts, and this automation was seen as evidence of a lack of skill [Mindell, 2011, p. 90]. Compared to the highly-automated Mercury capsules,[32] the Gemini craft—intended to demonstrate capabilities necessary for a mission to the moon—engaged human pilots in new kinds of operations. While pilots wanted a role in launch vehicle guidance,[33] no human would fly a rocket off the launch pad. The importance of orbital maneuvers (rendezvous and docking) for the Gemini missions suggested returning the role of piloting to the human, but as spacecraft pilots soon discovered, orbital rendezvous could not be achieved by the traditional manner of flying. "Numbers, equations, and calculations" would be required, and were bootstrapped to the pilot's senses with a

---

[31] Famously "not that interested in space" [Kennedy, 1962].

[32] A change in nomenclature, from space "capsules" to spacecraft, responded to the national importance of human involvement [Kauffman, 1994, p. 85].

[33] Simulator studies were commissioned to identify the capacity of humans to actually guide boosters into orbit. Spun in a centrifuge in Johnsville, PA to simulate the immense acceleration of takeoff, some pilots managed to fly simulated rockets into orbit, serving as evidence for the importance of pilots to the "reliability and flexibility" of launch vehicles [Mindell, 2011, p. 72].

new readout, the IVI or Incremental Velocity Indicator [Mindell, 2011, p. 86-87].[34]

The Apollo program started too soon to learn lessons from Gemini—the designs were largely complete by the time Gemini missions flew—but arguments about human roles also shaped this parallel program. Due to requirements of weight, space, and reliability,[35] the human beings were placed at the nexus of many devices, including a digital computer and manual controls. Supported by ground controllers, astronauts were again counted on for sextant readings, for initiating appropriate program modes and monitoring automation systems [Bennett, 1972, p. 4]. The LM's approach to the surface also involved a complex dance of manual and automated capabilities [Cheatham and Bennett, 1966]. As Allan Klumpp describes his "hybrid" system: "The essence of the approach phase guidance system is that the LM commander can manually steer the LM to the selected landing site, yet the trajectory he flies is produced by an automatic guidance system" [Klumpp, 1968, p. 129–130]. These historical examples are about neither heroic pilots nor automated spacecraft, but a successful combination of the capabilities of both, aided by large-scale networks of people and organizations.

2.4.2 Researchers and "Robot Geologists"

While these manned missions were occurring, scientists pushed for unmanned exploration,[36] which is cheaper by orders of magnitude [Murray and Cox, 1989, p. 66]. But even for unmanned exploration systems in space, autonomy is not the end-all-be-all goal. More important is the link between human operators and scientists on the ground, and the remote science platform responsible for carrying out instructions. While people may think of the Mars rovers *Spirit* and *Opportunity* as autonomous robots remotely carrying out science experiments—Clancey describes them

---

[34] The IVI could be programmed with particular velocity changes in different axes and would show visually when those particular burns had been achieved, so that the human pilot would know when to stop accelerating.

[35] Serious research was being done in the 1950s on simple automated probes to do a fly-by of Mars [Battin, 1989, p. 1]. But self-contained navigation capability was largely removed from Apollo due to memory restrictions [Tindall, 1966].

[36] Jerome Wienser, chair of the Presidential Science Advisory Committee, felt that his duty was to argue against manned missions on scientific grounds [Levine, 1982, Chapter 2].

often showing up anthropomorphized as a "robot geologist" or "explorer" in news coverage [Clancey, 2014, p. 7]—they are in reality telerobotic systems, remotely carrying out human commands. The rovers (MER) are not sent out to wander or find goals by themselves. They drive mostly blind, with only their immediate obstacle detection, following manual waypoints entered by human navigators on Earth: their autonomous path-finding systems are much slower and use more power, and are therefore generally avoided [Clancey, 2014, p. 118]. While the rovers are out of contact for the span of two weeks during solar conjunctions, the instruments lie dormant and the rovers sit still [Clancey, 2014, p. 25]. Though the rovers are technically capable of carrying out pre-planned science sequences during this time, including automatic navigation, this capability is not used: human observation is too important and autonomous operation too risky. Human scientists define sites of interest, locations to take samples, and the paths to reach them most safely and effectively. Scientists contact the rovers at least once per day to relay plans, and again to retrieve results, with more intense schedules during the early parts of the program [Clancey, 2014, p. 58]. Small sets of operations are requested, and the outcomes monitored, with new plans being made by humans to account for new data at each step.

The recent missions step back from the programmatic operation of the earlier *Viking* landers, toward closely coupled human control and supervision. The Mars rovers could certainly have been made more autonomous, but this would have opposed their function. The quality of the mission depends on "aspects of the MER's design that promote the *agency* of the scientists" themselves, rather than automated operation specifically [Clancey, 2014, p. xii]. Scientific work in the field is "opportunistic, serendipitous, and incremental" [Clancey, 2014, p. 32], yielding not so much to *a priori* plans of great detail, but Suchman's "situated action" [Suchman, 1987]. Researchers on the ground actually experience a sense of "telepresence" through these "synergistic" machines, created by virtue of their closely coupled operations and the MER's semi-anthropomorphic bodies [Clancey, 2014, p. 55]. Scientists see as if they are the rover, and have to "retool" their thinking, to become the "mind" of the rover and plan its work in a "symbiotic" way [Clancey, 2014, p. 106, 110, 118]. The system's autonomy does not replace the scientists, but allows them to do more of what they want to do (and are best at), and less of what they don't: some types of automation could reduce the autonomy of the scientists, and their ability to act creatively and spontaneously to capitalize on new findings [Clancey, 2014, p. 118-119]. This is a subtle point about automation design and human agency, one worth considering when evaluating designs of automated cars.

As Clancey suggests, autonomy is not an "inherent property of technology" but "a relation between people, technology, and a task-environment" and should be considered in those terms [Clancey, 2014, p. 119]. We might say, colloquially, that "*Opportunity* encountered" something [Clancey, 2014, p. 8], but this is convenient shorthand for scientists on Earth encountering it through the telerobotic platform. Human knowledge, perception, and common sense are integral to rover missions, and organizations operating expensive technology in a high-risk environment quite reasonably want humans to be responsible for the well-being of the equipment. The same issues play out with robotic vehicles for underwater exploration. The most famous of these, Alvin and Jason, deeply involve human scientists as operators in real time, either directly from within the vehicle or through a tethered link [NOAA, 2015a] [NOAA, 2015b]. Most autonomous underwater vehicles (AUVs) are not truly autonomous. This is largely an issue of risk aversion. While it has often been assumed, even sometimes by system designers, that long duration missions will be performed autonomously, this is not the primary mode in which AUVs are operated: the last thing you want to happen, as a scientist presiding over an expensive piece of equipment, is to lose the vehicle irretrievably or have an instrument fail on the first day and not know about it for weeks. So instead of long, autonomous missions, most operations involve many shorter missions with low-bandwidth acoustic links coupling the device with shipboard researchers.[37]

While self-driving cars would not be much like Mars rovers or underwater robots, telerobotic exploration provides potential lessons to be learned in terms of human and machine roles, and the level of autonomy one wants from a machine in a human-

---

[37] David Mindell, discussion with the author, September 3, 2014.

machine relationship. Automation can be a tool to increase human agency, both for pilots in the air and geologists on Earth. Considering what type of interaction promotes the agency of human owners and occupants is particularly important, especially while recognizing that the answer to what humans want to do is not necessarily "nothing." And while local streets are not as remote and inaccessible as Mars or the bottom of the ocean, risks of loss and failure must still be considered in the human-machine system design. Cars are capital-intensive pieces of equipment, and we generally want to know where they are at all times. Whether vehicles are monitored from within or without, it seems unlikely designers will be able to escape the need to loop the human into the decision-making process.

2.5 From the Driver's Seat

Aerospace experience with automation parallels some developments in the automotive realm. Computerized aids are sites of contradictory feelings and pressures: cruise control, when first introduced commercially as Chrysler's "Auto-Pilot," was described as "faintly ominous" by *Popular Science*, but nevertheless seemed like a "genuine help" for reducing fatigue [Rowsome, 1958]. It is well enshrined within legal principles that drivers using cruise control and even automated driver aids are legally responsible for controlling their vehicles at all times, and yet these same aids may reduce attention and ability to react in an emergency situation—the other side of reducing fatigue. Traction and stability control are now required for all cars in the US market [Villasenor, 2014], while drivers' personal identities still modulate the extent to which such advancements are seen as valuable, or how often the features get turned off in day-to-day use.

Deep controversies exist among automotive enthusiasts about the proper roles of all-wheel-drive, traction control, and automatic transmissions in the driving process. Even ABS, generally accepted today as a positive technology that enhances safety and performance, is in some senses controversial; this skepticism is motivated by the idea that a professional driver utilizing fully manual "threshold braking," developed through deep experience and human skill, can outperform the computerized system.[38] Human and automated capabilities take on a special role in high-performance racing, which, despite involving technology, puts the human being at the center. Modern racing cars depend on much accumulated engineering expertise, but racing is as much about the drivers as the cars, and so inventiveness is often tempered by rules that disallow certain technologies: ABS and traction control are banned from Formula 1 racing, for example. As Carlos Martinez-Vela describes, regarding NASCAR: "as important as engineering science has become ... given that every team runs virtually the same technology, what makes a difference in performance is the 'human element."' The human is the "only data acquisition system" allowed at the track, and the abilities of human beings to sense and describe performance characteristics, and to work together as a team, are paramount in this community [Martinez-Vela, 2007, p. 178]. New consumer performance vehicles—like the highly computerized Nissan GT-R, with a plethora of all-wheel-drive systems—are criticized by some as "too easy to drive quickly" [Nissan GT-R..., 2015]. An overly computerized vehicle is to some soulless, unexciting, too much like a video game: these vehicles may derisively be said to be for the "PlayStation generation," even while others argue for the superior abilities that advanced technologies convey upon humans,[39] turning an average driver into a hero or track-star. These gains and losses are not all easily quantifiable: while accident rates or track times are easy to measure, subjective judgments involving drivers' identities are not, and may be similarly important in adoption and use.

2.6 Conclusion

This historical journey has attempted to shed light on some of the unspoken assumptions behind autonomous vehicles that shape the popular narrative toward particular visions of automation and away from others. From factory automation, through AI

---

[38] There are innumerable forum posts on this topic scattered around automotive forums, and arguments over physical principles and the capabilities of the technology in different road conditions are intertwined with issues of masculinity and implicit driving skill.

[39] This position stumbles upon a sort of hybrid identity: the human gains through interaction with the technology.

and robotics, to applied automation in multiple types of vehicles, we have seen that history does not support an assumed trend toward ever increasing automation and complete replacement of the human being. Instead, automation is partial, contingent, oriented toward particular tasks and the extension of human agency in specific ways. It shifts the role of labor, allowing the concentration of creative control in management and in those who design the automation systems. It emerges from a multiplicity of ideas about designing effective or "intelligent" systems, which provide no simple answer for replacing human abilities. It extends the capacities of pilots, geologists, and astronauts, without replacing the necessity of judgment and decision-making in when and how to apply automated capabilities. And it is already contested within enthusiast communities due to their particular desires and identity politics. These specificities are often absent from our dreams of self-driving vehicles, and these dreams deserve more careful examination for the consequences of the futures they envision.

Chapter 3

Science Fictions, Designed Dreams

Technologies, with rare exceptions, are imagined through design fiction or science fiction before they are made. These future visions serve a number of purposes: they inspire scientists and engineers,[1] they serve as design studies for the possible shapes of technology, and they act as playgrounds to investigate potential cultural impacts. Though created in different contexts, with potentially different levels of scholarly care, these two types of fiction are not neatly delineated, and are unified as sources of insight into what might be possible with technology. "Design fiction is the cousin of science fiction," as Julian Bleecker puts it, and represents a hybrid practice that attempts to negotiate between facts and wild, playful imaginings to bring light to the multiplicity of possible futures [Bleecker, 2009, p. 8].

But it would be wrong to categorize science fiction as categorically less relevant. Science fiction films often involve scientific consultants with real technical knowledge, whose depictions become what David Kirby calls "diegetic prototypes," used to "demonstrate to large public audiences a technology's need, benevolence and viability" [Kirby, 2010, p. 43]. Kirby connects prototypes (primary "driver[s] of technological innovation"—in Suchman's terms "performative artefacts" [Kirby, 2010, p. 45]) to

---

[1] In fact I think it would be exceedingly rare to find an engineer today who was not influenced by science fiction. My interest in AI has been driven by *Star Trek* and *2001: A Space Odyssey*, the latter perhaps more morbidly than the former. A colleague of mine cites Neal Stephenson's *The Diamond Age* as key to shaping his interest in computational linguistics and human-machine interfaces. Wernher von Braun and other architects of the space program were deeply influenced by Jules Verne, and some even wrote science fiction themselves [Scott and Jurek, 2014, p. 2].

these diegetic prototypes, recognizing their similar rhetorical roles and the ability of diegetic prototypes to mobilize funding for real-life prototypes [Kirby, 2010, p. 44-47]. Science fiction, then, has deep and compelling relationships with real engineers and engineering via its depiction of aspirational or cautionary futures.

Science fiction is a perennial source of popular metaphors and ideas about technological change because its images represent compelling and easily-digestible cultural reference points for technological stories in the popular press. From Asimov to *Terminator*, these sorts of pictures are endemic to discussions of artificial intelligence in the press, and are often mobilized specifically to illustrate the possible forms of automated vehicles. Stories and ideas about what autonomous cars will be are influenced by *Minority Report* [Driverless cars..., 2013], *Knight Rider* [Wade, 2014], and *Total Recall* [Pasdirtz, 2015]. Countless articles begin by describing self-driving vehicles as a Hollywood or science-fiction staple [A Window..., 2015].[2] Popularly referenced fictions are not all so contemporary. *The Jetsons*, though the show did not foreground driverless vehicles, also gets recalled as a source of inspiration due to its flying cars, which trigger similar visions of wild futures made possible through technology [Pepitone, 2014]. However, many worlds that look lovely in science fictions or design fictions represent places we may not actually wish to live, and these portrayals can afford to ignore issues of reliability and human interaction that we cannot avoid in the real world.

3.1 The Stakes of Our Stories

Driverless cars are in part works of theater, objects of technological spectacle which are not yet truly real, despite their existence as physical objects which can be photographed and experienced.[3] Such real-world contact is yet limited to a lucky few.

---

[2] *Minority Report* is of particular importance in this context, not only because it is so often referenced but because its depiction of automated vehicles was developed in consultation with Harald Belker, an automotive designer responsible for both numerous Hollywood collaborations as well as real-world products [Melanson, 2012].

[3] And in this way they are similar to concept cars of other types, which are designed to embody company ideals, brand languages, and generally act as advertising as much as design studies for actual future vehicles.

Google hosts special press events to allow select journalists to ride in their automated vehicles. Their capabilities are touted, advertised, but their media picture is still tightly controlled. Not subsumed to the mundane, the quotidian,[4] they still possess that magic that makes science fiction connections obvious and compelling. But this does not mean such connections are unproblematic. By their alignment of these vehicles to science fiction, popular narratives emphasize images of the technology as it has already been envisioned, and obscure the multitude of potential implementations (such as controlled-access personal rapid transit systems, like the French Aramis project so comprehensively covered by Latour in *Aramis, or the Love of Technology* [Latour, 1996]) that can emerge from engineering practice in the moment. Self-driving cars threaten to become natural and obvious, in a particular form, through their associations with existing, known and culturally assimilated media portrayals.

We find ourselves in an era of seemingly continual technological change that is strangely most noticeable in the most mundane parts of our lives—how we shop, how we communicate, how we find partners—almost as if "designed by a bored researcher who kept one thumb permanently on the fast-forward button" [Gibson, 2000, p. 7]. No longer does the future seem to be defined by flying cars and jetpacks. Instead, it is defined by information, and its collection and use. The stakes of autonomous vehicles are thereby deeply intertwined with the stakes of other networked information technologies.[5] To examine the ways in which our conventional design fictions may fail us, I will take as assumed the general form of the autonomous vehicle as present in Google's development work. I seek to take this concept to its logical conclusion and interrogate the social, cultural, and informational norms implicit in it.

---

[4] Some articles from people who have ridden in these vehicles stress the boredom and mundanity that comes with operating this technology [Davies, 2015a]. But these do not represent the preponderance of coverage. It also should be noted that boredom always comes with the potential of its opposite: operating automated systems can involve "hours of boredom punctuated by moments of terror" [Sheridan, 1992, p. 339].

[5] What gives us confidence in these visions? Why don't they seem outlandish, like some of the works of science fiction they are grouped with here? Some of this has to do with their manner of presentation, lent the gravitas of "science" by emanating from scientists rather than from the illusionary powers of Hollywood.

3.1.1 Data Gathering and Monitoring

In May of 2014, the European Union's Court of Justice ruled in the landmark Costeja decision that since Google is processing "personal data," and acting as a "data controller," it may be compelled to remove links to pages containing personal information from its search results [ICO, 2014] [Court of Justice of the European Union, 2014]. There is growing recognition that publicly available data can be highly sensitive and that it may be beneficial to allow individuals certain legal rights to control their own electronic reputations, at least in particular circumstances. But much of our information is even harder to control. Many current data-driven business models[6] are fundamentally united in that increases in functionality are predicated on invasions of or encroachments on what we used to think was private, and represent increasingly invasive data collection and sharing at a massive scale—often this information is used internally to improve services, but it may also be aggregated and sold to third parties, and in either event may be stolen or leaked by disgruntled employees or thieves.

Whether networked with each other or connected only to central servers, automated vehicles will allow large amounts of data to be collected and shared with other entities; but it may be that not all the information that is collected should be shared outside the immediate context of its use. What is sensed, and what can be appropriately transmitted back to servers for processing and storage, is of paramount importance [Nissenbaum, 2010].[7] There may be legitimate uses for certain types of sensitive information, but while providing it to municipal governments specifically to assist in city planning may be legitimate, selling it to advertisers to help them design more effective billboards may not be. Privacy issues involving motor vehicles are likely to become much more complicated as vehicles become able to record more, and potentially "know" more, about their passengers.

With what networks, and for what reasons, will autonomous vehicles be con-

---

[6] See for example Google, which has more accurate or "targeted" advertising as its fundamental revenue stream. Facebook's new ad network competes in a similar space.

[7] Following Nissenbaum, privacy must always be seen in the context of particular users and a particular use [Nissenbaum, 2010, p. 2]. It is not that data about our commuting routes, for example, should never be collected, but that collected data should not be sent uncritically to any third party without our knowledge or consent, a violation of our information norms [Nissenbaum, 2010, p. 3].

nected? Current driverless car concepts depend on networked information for vehicle guidance. Google's vehicles use inertial navigation devices [Knight, 2013] alongside other sources of position data: their current navigation systems depend on global positioning satellites.[8] But autonomous vehicles will be connected to more than just global positioning networks. GM's OnStar service already connects equipped vehicles to central servers for purposes of safety, security, and convenience. The system can automatically alert the authorities in case of an accident or theft, and provides vehicle diagnostics to the owner's tablet or smartphone, along with the ability to configure settings, lock and unlock doors, and remotely operate the lights and horn [OnStar, 2014]. While these vehicles are not (yet) autonomous, OnStar's present capabilities are representative of features that will become more common in highly connected and computerized vehicles, including autonomous ones.

Information about the vehicle and its surroundings, including the locations of cars and pedestrians, precise GPS coordinates of the vehicle itself, and the vehicle's speed and acceleration, not only represent important knowledge for vehicle localization and Intelligent Vehicle-Highway Systems, but new sources of potential revenue for the groups in position to collect them. Uber, which through its GPS-enabled ride-hiring application still collects only a fraction of the data that would be available through a self-driving vehicle, has agreed to share its ride data with municipalities for purposes of city planning [Jardin, 2015]. Though this data is ostensibly being shared for the public good,[9] it also serves private ends: to curry favor with authorities that might otherwise attempt to shut the service down. And it would not be far-fetched, in the current information landscape, to see such information sold to third-party advertisers.

Prevailing data ideology tells us that we can understand ourselves better through

---

[8] Unlike these, guidance systems for particular applications (e.g. intercontinental ballistic missiles) can be inertial only, isolated from the outside world [MacKenzie, 2002].

[9] One potential reason for municipalities to collect such data is dynamic road pricing. This is not fundamentally a bad idea if done in a way that is as non-discriminatory as possible, but putting another corporate entity in between road users and the road is not necessarily good. It strikes me, since I am writing this shortly after the FCC's reclassification of Internet service under common carriage laws, that if we find ourselves in a battle over "road neutrality" in a decade or two, it will have a decided air of irony to it. Pervasive tracking and self-driving allow for all manner of road pricing schemes, some of which will be unduly discriminatory.

collected data, finding meaningful information within it. Out of this ideology come publications like Uber's blog[10] describing customer "insights" gained through their ride data, which are ostensibly interesting to the public. But the uses of data collection—for corporate profit and public benefit—are deeply intertwined, and strongly influenced by the possibility of using statistical analyses to find and exploit patterns. Automated vehicles are part of this general culture, and, without significant effort to resist erosions of privacy and data protection, stand to open our lives to greater scrutiny in terms of the means of and reasons for our personal mobility.

These sorts of security and privacy issues are not unique to autonomous vehicles, nor even to networked vehicles. What Roger Clarke calls "dataveillance" is already possible: electronic tolls allow for a measure of tracking [Nissenbaum, 2010, p. 25]. Networked traffic cameras are already being used to amass large databases of information about (non-networked) cars and their travel patterns by reading passing license plates [Nissenbaum, 2010, p. 26]. Despite their sensitivity, collection and use of this data remains largely unregulated.[11] While contested data collection is clearly already possible, networked vehicles, with their arrays of sensors, provide more avenues of data collection, and therefore stand to increase our present-day problems with mass surveillance and personal privacy.

Setting aside governmental surveillance, networked technologies do not have a great security record, and there is no reason to believe automated cars will be any different.[12] Networked capabilities within vehicles are being used to implement vehicle kill-switches, capable of remotely disabling vehicles on the whim of the controlling organization—for example, if vehicle rental or loan payments are not completed on time—and the existence of these devices presents opportunities for hackers to

---

[10] See blog.uber.com. One particularly noteworthy post, from March of 2012, was about customers' "Rides of Glory"—in other words, one-night-stands. This post was later taken down by the Uber team, but articles about it remain [Harris, 2012]. Another post about the connection between Uber rides and areas of crime—notably prostitution—was also removed [O'Brien, 2014].

[11] Groups such as the EFF and ACLU are attempting to do something about this [Kayyali, 2014]. California police are coming under criticism from civil rights groups about their mass collection of vehicle position information through police license plate cameras [Maass, 2014].

[12] Commercial networked devices routinely lack basic security measures. And for a large number of other systems, security is defeated by being out-of-date or by using default passwords that were never changed. See for example [Zetter, 2014].

cause nuisance or harm [Goodman, 2015]. Existing vulnerabilities in web services and dealership practice allowed Ramos-Lopez to hijack 100 cars from a Texas dealership [Poulsen, 2010]. A high-school student at a security camp in July of 2014 was apparently able to hack into an automobile with $15 worth of components, gaining access to the remote start feature [Bigelow, 2015]. The potential for this kind of access gets more frightening when one considers the effects an integrated, network-connected vehicle that is capable of moving and navigating on its own might have on the reach of a hacker or negligent employee.

Additionally, Google has envisioned vehicles that can determine their number of occupants, and use facial-recognition or other biometric systems to identify them. According to one patent [Zhu et al., 2011], these vehicles could prevent unauthorized persons from putting a child in a car, prevent convicted sex offenders from operating their vehicles within the legally-required distances of schools and playgrounds, or prevent a car's doors from being opened (even from the inside) by a child unless an authorized adult is present. These are only visions,[13] but represent a perspective on safety that posits technological surveillance and enforcement as appropriate measures against potential criminal behavior. Whether or not protecting against these threats is an appropriate use of this information is a matter for societal judgment, but such proposals, if enacted, would require these vehicles to have unprecedented levels of very sensitive knowledge about people and their lives: biometrics, criminal histories, family and trust networks.

3.1.2 Maps and Geography

The information that may be collected and processed by automated vehicles is not only about human inhabitants of the environment: other types of data collection are also implicated in current visions of the driverless car project. In order to drive with us, autonomous systems will have to understand, for at least a practical sense of "understanding," traffic rules and their accompanying signs, signals, lanes, and

---

[13] Patents are notorious for trying to cover as many possible angles of a technology even if they are not intended to be applied in practice.

customs. It is through years of existing as a human being in a particular cultural context that we know to drive on roads but not on sidewalks, and how to tell the difference. But we do not have this luxury in training machines to interpret the world.

Current vehicles build on decades of robotics work in mapping and localization. Though successful road tests have been accomplished without navigational assistance, depending only on visual stimuli (such as the EUREKA PROMETHEUS project in the 1990s) [Ulmer, 1994a], modern systems are tending to use more external stimuli, rather than fewer, in an attempt to increase safety and vehicle capability. Google's vehicles do not operate in a SLAM mode: even as advanced as they are, the vehicles require hyper-detailed 3D maps in order to operate properly on public roadways [Gomes, 2014b]. These maps are generated by human-piloted vehicles outfitted with special sensor arrays, like the LIDAR Google uses for their autonomous vehicles, which drive a route and collect data that can be used to reconstruct the model used for future drives.[14] And just as for map applications on smartphones and computers, those maps would need to be downloaded to the automated car before or during a trip, via a wireless data connection.

Pre-made maps inform the vehicle where stoplights, signs, and curbs are, reducing the computational load on the machine in the crowded visual landscape of driving, and allowing it to focus on elements of the environment that are changing rather than those likely to be static [Gomes, 2014b]. When Google's car was certified for testing in Nevada, Google was allowed to pre-select the route the car would take, so that they could build the comprehensive model the system requires beforehand [Harris, 2014c]. The system would likely not have been capable of passing a test in which the examiner could have added detours on the fly. And though Google claims to have driven more than 700,000 miles with their cars, those are not 700,000 unique miles. A limited, pre-mapped route has been driven many times to achieve those numbers [Gomes, 2014a].

Mapping claims a unique capability to represent the real, objectively and diagram-

---

[14] Nhai Cao (Global Product Line Manager at TomTom), presentation at The Road Ahead Forum on Future Cities 2014, Cambridge, MA, MIT, November 21, 2014.

matically, but also requires that the world remain largely static, at least on the order of how long it takes to update the map. The necessary level of continual mapping is a massive task if the vehicles must be usable everywhere. The United States alone contains almost 8.5 million road miles,[15] and it took years for Google Street View to acquire the level of coverage it currently has. The maps required for driverless car localization are a significantly more difficult project in terms of amount of data, reliability of data, and therefore frequency of updates.[16] The utopian discourse of driverless cars implicitly suggests that such vehicles will be available everywhere, and are the solution to nationwide transit problems. But widespread egalitarian access to maps-based devices depends upon a rapid, widespread mapping initiative. The success of such an initiative is dependent not only on a large amount of human effort and capital investment, but a series of decisions about what regions should take priority.

Most likely, certain areas will be mapped and restricted, or separate rapid transit systems will operate on divided roads that can be carefully monitored.[17] While Mountain View, California may be mapped early, rural West Virginia or Northern Maine may not be mapped as soon or as frequently.[18] Inequalities may be increased if routes frequented by upper-middle class professionals likely to own new autonomous cars are mapped first, while roads around low-income communities are ignored. Such decisions are easy to imagine being justified by market forces, but would cut directly against the utopian narratives of driverless cars.

Furthermore, the seemingly universalizing forces of maps tend to hide issues of geographical and cultural specificity. More than abstract problems of localization and mapping, automated vehicles present the problem of having to exist in an environment that is highly complicated, varied, and cultured. Maps alone are insufficient for anything but the most simplistic view of vehicle operations. Programmed devices must know about speed limits, about traffic lights, about rules of the road that

---

[15] Data as of 2008 [Carney, 2010].

[16] TomTom [Knox, 2015] and Nokia [Lee, 2014] both claim to be attempting such a mapping project.

[17] This point came up in a discussion with a scientist with experience at a UK-based firm working on personal rapid transit, or PRT, systems.

[18] And it is worth noting that 60 percent of road fatalities occur on rural roads, and that the rate of fatalities per mile on rural roads is twice that for urban roads [Broviak, 2007, p. 11].

were never designed for autonomous systems. These devices must respond to human caprices and be adapted to longstanding, ingrained laws and habits. They must include historical knowledge, rooted in the legal and social histories of roadways, which may differ between cities and states, and certainly between countries across the world. Local customs and behaviors vary, and even if maps are available, the same vehicle programming may not work for Los Angeles, Boston, and the rural Midwest, let alone Singapore, Mumbai, and Cairo. The map, for all of its objective standardization, still represents real places subject to cultural histories and vulnerable to socio-economic dynamics. These social and regional issues are often ignored in the driverless vehicle narrative, but nevertheless stand to be critical to the manner in which these technologies could enter everyday life.

3.1.3 Seeing and Sensing

For all this, however, vehicle localization and sensing depend upon visual interpretation of the vehicle's surroundings. Interpretation of the world around us is a task that seems particularly easy for human beings, but particularly difficult for machines. The invention of the photocell, early a tool for workplace monitoring and surveillance, provided a simple channel through which electrical systems could respond to the amount of light reaching them [Coopersmith, 2015, p. 44] [Nye, 1990, p. 361], but perceiving detail and depth, identifying shapes, and interpreting expression and motion are all capabilities of human vision that require more sophisticated technologies to reproduce. And machine vision problems, including object recognition and scene interpretation, continue to be difficult, even with increased processing power and new algorithms.

As an engineering discipline, computer vision takes a decidedly practical and reductionist view of what it means to see. The goal is generally to identify particular objects in a scene, to recognize faces, or to differentiate free space from things a robot should not run into. But this so-called objective focus still encodes certain subjective judgments about objects (including people), behaviors, and intent. While it may be an objective question whether or not a particular object is physically present and visible in a scene, it is not necessarily so self-evident which objects are noteworthy or important to detect—these choices depend on applications, and the judgments of designers as to what is worth measuring. And while computer vision is having success with object detection, a wide variety of human knowledge about objects and scenes is missing in current computer models, including propositional understandings ("what would happen if...?"), projections about occlusions ("what is behind that?"), and connections to other sensory modes ("what does that object feel like?") [Gomes, 2014c].

Vision is a particularly attractive sense to use in automated cars as it is integral to how humans drive, and should ideally not require infrastructural changes. The apparent primacy of computer vision, however, only holds for fair weather. Rain, sleet, and snow interfere with vision-guided systems, and currently prevent them from operating safely, as numerous more skeptical news articles are happy to note [Knight, 2013] [Gomes, 2014b]. Precipitation reduces visibility (which decreases the distances at which objects can be detected, and therefore the time the system has to react), decreases image contrast, and presents interference from drops or streaks on the glass [Cord and Gimonet, 2014].[19] "Developing algorithms that work perfectly under all weather conditions appears to be unrealistic," so a modular and cautious approach may be necessary in order to build robust systems [Cord and Gimonet, 2014, p. 50]. Though commercial use of vision-guided self-driving vehicles on sunny days may be possible in the near future, use in less-than-ideal conditions is likely to be limited for some time, with human oversight compensating for algorithmic deficiencies.

Early approaches pioneered by Dickmanns and others were designed primarily for highway operation under constant human supervision, which made their rudimentary strategy of searching for lane markings and vehicles acceptable, since pedestrian interactions were likely to be rare.[20] Some contemporary commercial systems, now using digital video cameras and off-the-shelf consumer hardware, are primarily guided based on such visual sensors [Dickmann et al., 2014]. Roof-mounted LIDAR arrays can supplement these systems, scanning the environment with rotating laser beams

---

[19] These issues also affect people (though we are good at filtering out raindrops on the windshield rather than interpreting them as obstacles), and raindrop detection to help advanced driver-assistance systems compensate for reduced visibility is a current research area.

[20] VITA-II papers mention pedestrian interaction management as future work necessary for changing the operation domain from the highway to urban and country environments [Ulmer, 1994b].

to create a detailed 360-degree representation of objects and their distances. This technology solves some of the difficulties of image interpretation by default, as it can provide highly-sensitive information about free space and obstacles. Shape-detection algorithms can then be used (in addition to vision-based data) to classify obstacles as different types of objects: pedestrians, bicyclists, cars, and trucks.[21]

While much computer vision research uses machine-learning algorithms to detect objects,[22] engineers in automotive applications are justifiably reluctant to rely on machine-learning: as Göde Both has noted in his ethnographic research on developers of driverless cars in Europe [Both, 2014a], machine-learning techniques are brittle and unpredictable [Both, 2014b]: neither characteristic makes them suited for software that must be highly reliable and on which people's lives literally depend.[23] Even discounting these concerns, shape-detection is not a complete solution, as the knowledge it provides about objects is only skin-deep: the sensors cannot differentiate a rock from a crumpled newspaper, and Google's car will swerve to avoid both [Gomes, 2014a]. Further distinguishing objects requires much additional information, including fine-grained information about the object's surface appearance, and interpretation of physical properties from observable behavior (e.g. bouncing or rolling). None of this is fundamentally impossible, but it presents necessary areas for research.

Detected categories allow the system to make statistical predictions about likely types of behavior based on assumptions about those categories [Zhu et al., 2011]. These sorts of predictions are something that human drivers do constantly, and are therefore likely to be important for autonomous vehicles.[24] Categories alone, even if achieved, are also insufficient if detected objects are treated solely as obstacles to be

---

[21] This also works for other kinds of physical objects, like signs [Fukuda et al., 2014].

[22] Pedestrian detection algorithms search for personlike shapes, where "personlike" is determined by, for example, training a classifier using thousands of images previously labeled (by people) as being images of humans.

[23] Though machine learning can be highly effective, it is generally difficult or impossible to know what the system has actually "learned" and therefore how it will react in new and unknown situations [Both, 2014b]; so mixed techniques are used.

[24] The DARPA Urban Challenge crash, the first crash between two autonomous cars, provides an important lesson on the vagaries of object detection and prediction: the classification threshold between moving and stationary, set too high, allowed one vehicle to interpret the other as stationary, leaving no room for unexpected behavior [Fletcher et al., 2009].

avoided. John Leonard shows a collection of photographs and videos of situational edge-cases—such as police directing traffic, especially when combined with sun glare or occlusions of the sight lines—that would make behavior prediction and response difficult.[25] His take-away is that humans will be required to account for these sorts of difficult perceptual situations.

However, because autonomous cars see—putatively "as we see"—their sight can be leveraged as visual evidence of their operation. Computer vision systems that identify pedestrians can be shown to do so via detection boxes that act as diagnostic tools for researchers and direct representations of internal system information.[26] The new technologies of vehicle automation thereby produce through their operation new forms of evidence, which can be presented through electronic media. We can point to these images and identify that the vehicle is operating as it should. Three-dimensional shapes, standing in stark relief against the background in LIDAR scans, bear witness to the sensory operation of the vehicle, and present "transparent" visual proof. These shapes too are demarcated by boxes, which represent their computational transformation from information into an object or artifact of interest. Because the visual detection technologies used by autonomous vehicles are compatible with the visual technologies of media representation, new types of seeing are opened to us.

A coincidence of sensing and representation in the evocative and powerfully persuasive medium of visual representation stands to shift the way we perceive driverless systems in their operation, providing us a new manner of insight and introspection, but also a new level of obfuscation. When considering how these vehicles are presented to us—and the fervor with which some researchers demand that we accept them[27]—it is important to remember the many black boxes of information-processing behind the naturalized image of the sensor readout. As neuroscientists have noted, images, the inheritors of Enlightenment notions of visual evidence *par excellence*, have a way

---

[25] Discussion with the author, December 3, 2014.

[26] See for example Volvo's advertisement for the S60's pedestrian detection capabilities [S60 Advertisement, 2015].

[27] For example, technical speakers at the The Road Ahead conference at MIT, hosted by the SENSEable Cities lab (November 21, 2014) suggested concerned users "get out of the way" of coming fully autonomous technology.

Figure 3-1: A simulated pedestrian detection image, showing the detected figures outlines by yellow boxes. Photo by Marjan Lazarevski [Lazarevski, 2014].

of convincing the viewer of the validity of researchers' claims about physical brain locations and their effects [Lehrer, 2011], especially since those pictures are presented as the direct, transparent products of sensing techniques [Joyce, 2008, p. 76]. The conflation of human and machine vision which makes these images so confusing is as misleading in vehicle navigation as it is in neuroscience.[28] Significant statistical processing is necessary to make sense of data in either realm, processing that is generally beyond the public's gaze, but the end result leaves the viewer with a false sense that something real has been detected, revealed, and affirmed by the image.

3.1.4 Functionalism, Utilitarianism, Ethics

We should likewise not be deceived by an apparent parallelism of human and machine knowledge, or an elision of the difference between human and machine "cognition." Self-driving cars will not be humanlike in understanding, even while they can detect and identify pedestrians as objects of interest within a particular epistemological

---

[28] In her book *Magnetic Appeal*, Joyce argues that "seeing does not equal truth or unmediated access to the human body," but that practices equating these are so common that images are often used to stand for truth despite doctors' awareness of how social practices shape this evidence [Joyce, 2008, p. 76]. Popular narratives are particularly prone to fall prey to the "myth of photographic truth" [Joyce, 2008, p. 75]. These tendencies are of great relevance when considering other complex, technological projects dependent on imaging and which use images rhetorically, to stand for the "truth" of their ability to perform a task—such as detect a pedestrian in a crosswalk.

frame. The robots envisioned by current AI ventures bear little resemblance to those of Asimov or the dreams that grew from the Dartmouth Conference.[29] Those working in the field are well aware of the difficulties of general-purpose AI [Sofge, 2015]; therefore, much current work is fundamentally utilitarian, building systems with clear goals, metrics for success, and market segments.

However, the utilitarian model of AI makes good sense for a number of reasons. First, much can be achieved with current technologies: it is possible to have working prototypes on the road, generating interest and publicity, even in relatively controlled conditions. Current prototypes are highly capable of attracting media interest, despite that describing them as "understanding" would be to use the word in a way that it could no longer credibly mean what it generally does.

Second, not all humanlike characteristics may be helpful for building specific applications. One would likely not want one's self-driving car to be preoccupied or emotional. Much of the discussion around why autonomous vehicles are necessary centers on just such qualities: distractability, sleepiness, lapses in concentration.[30] We would not wish to emulate such characteristics in robotic systems. Though these are human capabilities, however, when presented in this way they exist largely as caricatures of the human. People possess a variety of other capabilities which might be helpful to many AI applications. As AI researcher Doug Lenat wrote in 1997:

> Before we let robotic chauffeurs drive around our streets, I'd want the automated driver to have a general common sense about the value of a cat versus a child versus a car bumper, about children chasing balls into the streets, about young dogs being more likely to dart in front of cars than old dogs (which, in turn, are more likely to bolt than elm trees are), about death being a very undesirable thing [Lenat, 1997, p. 199].

This is a difficult knowledge and perception problem. But even more, it is an issue

---

[29] Chauffeur and Deep Mind, the Google divisions responsible for cars and general-purpose AI, are geographically and organizationally separated. A Google employee I spoke to at MIT said that Deep Mind had only very limited contact with the Chauffeur team, as of late 2014.

[30] A large proportion of articles make a claim like this. For example: "They don't get sleepy or distracted, they don't have blind spots, and there is nothing on their 'minds' except getting safely from point A to point B" [Merrill, 2014].

of selfhood, embodiment, even sentience. While cats, children, and bumpers can be identified as objects, and children chasing balls into the street can be identified as a pattern, a computer programmed to respond to these stimuli may respond correctly without "knowing" anything. While a machine can be programmed to avoid running into people, can it have any understanding of death? Can it be programmed to "feel guilt"? Does it need?

Current approaches, however, assume this kind of deep understanding is unnecessary, both for the technical creation of such vehicles as well as their public acceptance. If the machine behaves in an appropriate way, like an ideal human driver, it will be a "satisfactory social prosthesis" in Collins's terms, and "will not appear alien" even if it cannot truly "understand" [Collins, 1990, p. 31].[31] And yet the field of artificial intelligence, as discussed in chapter 2, is bound up in the idea of "intelligent" machines that can be said to "know." It is an interesting read on the changing times to notice that Lenat's statement seems to suggest humanoid robotic drivers operating regular cars. As well as moving toward functionalist systems, the industry has moved toward embedded systems within devices, systems that make no pretenses to be humanoid, but instead revel in appliance-hood. While a dominant dream might once have been to build a world full of humanoid robots, the conceit of modern consumer AI—IBM's Watson cloud API, "internet of things" approaches—is that we can make everything smart.[32] Nonetheless, despite lacking the facilities for actual ethics, these devices possess an implicit ethics: their behaviors will instantiate the moral and ethical judgments of their human creators, based on human-authored heuristics and statistical predictions. They will not *act* ethically, but must *behave* according to their programmatic ethics. This will be fundamental to how we understand our relationships to such complicated technological systems.

Robot ethics is an issue of growing importance to society at large, given the rapidly-expanding uses of robotics for labor, military, research, entertainment, and

---

[31] Driving, from this perspective, involves only behavior-specific acts, which can be satisfactorily emulated using the behavioral coordinates of action [Collins, 1990, p. 33–37].

[32] Whether this shift in the form of AI systems makes customers more or less nervous about computer-driven vehicles is an empirical question. But it certainly suggests a desire to make the systems more invisible.

healthcare, among other regimes [Lin, 2012, p. 5-6]. It involves a breadth of issues, including safety and errors (how should robots be introduced in order to minimize adverse effects?), law and responsibility (when accidents happen, who is liable?), ethical codes (how, and with what ethical frameworks, can robots be programmed to operate "ethically"?), and social impact (how do we weigh potential for robots to eliminate jobs?) [Lin et al., 2012]. Driverless car ethics issues have spurred a volume of articles, with varying command of the questions involved.[33] Lin has been successful in urging a dialogue within various autonomous vehicle research groups about the ethics of their products, as well as getting media recognition of this push for self-conscious development [Newman, 2014]. However, many articles elide fundamental distinctions in the situations being discussed, distinctions which should be foremost in our minds when we consider the stakes of developing autonomous machines.

When we speak of robotic cars "making decisions" of what to do in a crisis situation, we implicitly accept the idea that such decisions are really being made by programs. Discussing whether or not those behaviors are ethical risks suggesting that our robots have all the capabilities—of sensing, knowing, and processing—necessary to carry out, *in toto*, what we consider to be "ethics." But realistically, what we could create is an *ethical calculus* for autonomous vehicles. Such a calculus would be a quantification of ethics according to some particular formalism, so as to allow a computer program to select a course of action in particular situation. In a very real sense, decisions are not "being made" by amoral vehicles.[34] They are being made by software engineers, self-consciously or not, whether an explicit calculus is used or developers' ethics are only implicitly present as a consequence of coding decisions. Ultimately,

---

[33] One article discusses a notable researcher who has purportedly "stressed the need for driverless cars to be flexible enough in their engineering to be able to make ethical judgments that aren't necessarily written into their programming," a statement whose meaning is difficult to parse [Davies, 2015d]. A program that does things that are not in its programming at all is incoherent—even machine-learning systems are shaped and programmed, though their rules are not explicitly written. This sort of phrasing is unfortunate, as it risks suggesting the robot somehow operates "outside of its program," making ad-hoc ethical decisions based on criteria of its own invention.

[34] Not every device behavior can be predicted, and it would be foolish to place full responsibility on the programmers: there is real autonomy in devices, in that they may do things we do not want. But though all devices have bugs and will be unpredictable in certain circumstances, the first place to look for ethics, for an implicit or explicit ethical calculus, is the human beings that do the design.

if we want to care about how ethically systems operate, we must look at how they are programmed, and what goals that programming is intended to serve. Only with this focus can we agitate against systems that are fully black-boxed—closed-source, protected intellectual property potentially defended by anti-circumvention laws—and that enact an ethics impervious to scrutiny.[35]

3.1.5 Safety and Statistical Risk

The primary force in the current self-driving car push is safety. The total number of vehicle-related deaths is roughly 35,000 deaths per year in the US, but this alone does not tell the story [Table 1103, 2012]: at around 1 death per 67 million miles, humans seem relatively competent in an absolute sense—by comparison, the average American might drive 1 million miles[36] in his or her lifetime.[37] Part of why the autonomous vehicle problem is such a difficult one is that individual human beings can drive a long time without having an accident.[38]

By reducing the impact of human frailties, automation can do better. But especially careful human drivers can also clearly beat the human average. How safe do autonomous vehicles need to be in order to be allowed on our roads? Safer than the average human? Or safer than the very best drivers?[39] Such questions have real impact when it comes to how devices are designed and when they become commercially viable, and are in part ethical questions. The autonomous vehicle enterprise seems to call for using such projected statistics to define policy. But people are also accustomed

---

[35] Digital rights management and anti-circumvention laws are abused routinely to lock out consumers and even muzzle security researchers. For a recent example, see [Higgins, 2013].

[36] This is a conservative estimate, extrapolating from an average yearly mileage of between 13,000 and 15,000 miles [FHWA, 2013].

[37] Looking at non-fatal accidents as well, humans get involved in about one accident per 286,000 vehicle miles, still large on the scale of a driver's overall life experience, which aids our personal exceptionalism.

[38] Yet, around 93 percent of crashes with "critical reasons" involve driver error as one critical reason for the crash, 30 percent involve a roadway factor, and 12 percent involve some vehicle-related reason. There may be many reasons for a crash, and estimating when the human is solely or primarily responsible for a crash is not trivial [NHTSA, 2008].

[39] Chris Gerdes has admitted, as recently as September 2014, that the deep, intuitive experience of race-car drivers to handle emergency situations is something they are still working to match [Carson, 2014].

to the current automobile death rate, and any autonomous vehicle crashes are likely to attract deep scrutiny as to whether a human could have prevented the accident. If, instead, human oversight to otherwise fully automated systems were required, the additional risks of supervision would need to be accounted for—including the potential for risk homeostasis [Wilde, 1998], the tendency to behave less cautiously in situations that appear to be more safe. And any possible framework for quantitatively measuring and regulating improvement (e.g. number of deaths, monetary cost, etc.) hides all manner of assumptions and potentially unforeseen consequences: the best decisions by some metrics will be non-optimal in others.

These are questions of policy, but also questions of human acceptance. Just how safe these vehicles are expected to be has become a point of public contention. Sivak and Schoettle of the University of Michigan Transportation Research Institute suggest that "it is not a foregone conclusion that a self-driving vehicle would ever perform more safely than an experienced, middle-aged driver," due primarily to issues of sensing and predictive knowledge [Sivak and Schoettle, 2015b, p. 7].[40] Most firmly, they attempt to impress that no conceivable implementation of self-driving vehicles will have zero fatalities. One popular response to this type of argument is the following:

> Of course, the researchers are trying to correct what they regard as excessive technological optimism. Still, is it entirely fair of them to compare robocars only to the best drivers? Most accidents are caused by the worst ones, and it is beginning to become clear that those are the people that a robot could outperform with one clanky arm tied behind its back [Ross, 2015].

Others argue that self-driving vehicles should be considered not "as bad as a middling driver," but "as *good* as one," and place great faith in the "pinnacle of human mastery of software" to outperform human drivers [Templeton, 2014]. But these per-

---

[40] This is worth recognizing, even if it is also trivially true that nothing about AI development is *a priori* a foregone conclusion. Even if "computational speed, constant vigilance, and lack of distractibility" are not alone sufficient to beat out all human drivers [Sivak and Schoettle, 2015b, p. 4], I expect AI techniques will likely approach human abilities to use predictive knowledge, given sufficient development time. Still, Sivak and Schoettle provide a valid note of caution here.

spectives are far too simplistic, and miss the point in a significant way. The question of whether an automated vehicle's fatality rate exceeds some people's, matches that of the safest drivers, or bests all human drivers stands to determine which vehicles are legal [Sivak and Schoettle, 2015b, p. 6]. Projecting vehicle risk functions is tricky enough when comparing against a known human standard, and becomes even more difficult when one considers that the "conventional vehicle" risk function is itself going to change with future safety technologies, including ADAS. The fully self-driving vehicle must be compared to a moving target. Meanwhile, academic researchers are often unable to gain access to these systems to test them against exploits [Madrigal, 2014].[41]

In both cases, measurements and predictions of reliability and risk are key to the development of autonomous vehicles. The National Research Council pays significant attention to the fact that "the lack of generally accepted design, implementation, and test practices for adaptive/nondeterministic systems will impede the deployment of some advanced IA [increasingly autonomous] vehicles and systems" in aviation and that "existing V&V [verification and validation] approaches and methods are insufficient for advanced IA systems" for many of the same reasons [National Research Council, 2014, p. 2].[42] As in the automotive space, the core reason for increasing autonomy is to increase reliability, but being assured of this reliability is difficult. Exhaustive testing of every logic path does not scale to more complex systems, and new validation approaches are required to account for the impacts of human operators or supervisors on system behavior [National Research Council, 2014, p. 39–40]. While the NRC panel is concerned about validation of existing aeronautical control systems, automated vehicles may contain 10 times as much code as modern, highly-computerized aircraft (100 million lines versus 5 to 10 million for the F-35 Joint Strike Fighter) [Taylor, 2015]. The NRC panel is not alone in recognizing these issues for automated systems: the Defense Science Board in their task force re-

---

[41] The implication here is quite disturbing: car companies are so far able to set the agenda for their own testing (and Google is apparently lobbying to count simulated virtual tests as road-test miles [Harris, 2014b]), and third-party oversight is not yet possible.

[42] Among the high-priority research projects they identify as most pressing and most difficult, they include both the development of methodologies to "characterize and bound the behavior of adaptive/nondeterministic systems" and the creation of standards for the "verification, validation, and certification of IA systems" [National Research Council, 2014, p. 4].

port *The Role of Autonomy in DoD Systems* attempts to address the importance of the larger environment in which automated systems operate, and within which they can produce "unintended operational consequences" [DSB, 2012, p. 2]. They warn of *brittle* platforms, and emphasize the importance of developing ways to predict and understand the resilience of systems [DSB, 2012, p. 7, 11]. As cars become more automated, these difficult-to-predict risks will gain importance on the road.

However, when it comes to public adoption of automated cars, risk assessment is only part of the story. Does it make us feel better—more comfortable, more likely to get into autonomous taxis and spend our money on autonomous cars—just to know that they are statistically safer than the average driver? Despite the comparative statistical safety of flying,[43] people tend to be more afraid of getting onto an airliner than getting into their cars. While this may have to do with a number of factors, including that aircraft do not remain on the ground during operation, it also represents a situation in which passengers give over their agency to pilots performing a job they do not understand and could not take over in an emergency. Perceived safety—due to accident scale, publicity, trust, or other factors—may be very different than statistical safety.[44] But while research into how to get human beings to trust robotic drivers is being done [Ross, 2014b], the voices pushing for autonomous cars sooner rather than later would suggest that the statistics are all that matters (and indeed, would tend to fabricate those statistics out of mere predictions).

3.1.6 Planning, Policy, Cities

Besides being presented as steps toward road safety, driverless cars also become part of the rhetoric of personal mobility, "public" transportation, and urban planning. Potential or imagined benefits of these vehicles include empowerment for the blind or

---

[43] This information is relatively available and publicly known [Brown, 2010], though that does not necessarily change people's minds about travel.

[44] The CAST working group accounted for these issues in their plan to reduce airline accident rates by 80 percent [Flight Deck Automation Working Group, 2013, p. 28]. If air traffic were to increase with industry predictions but maintain the 1997 rate of 1.5 major crashes per 1 million departures, we could have seen one fatal crash per week by 2005, and one per day by 2025 [Experts Predict..., 1999]. This frequency of accidents was judged to be intolerable to airline customers.

elderly, reduction in traffic congestion, elimination of parking problems, and persistent access to a dense network of hireable point-to-point transport vehicles (essentially, driverless taxis). There is also an environmental case being made for these vehicles, as they will drive predictably and can therefore be tuned to be highly fuel-efficient. The social, environmental, and urban planning implications are at least worth considering in passing before moving on to consider alternate models for how these technologies could be developed.[45]

Developers tend to buy into utopian visions of big cities freed of cars, with no parked cars on the streets and only driverless cabs.[46] This desire to change the way cities operate is admirable: our cities have a difficult hundred-year history of coping with the automobile, and many of their changes have not been for the better when examined holistically. The prevalence of automobiles has literally re-shaped city centers [Burden, 2007, p. 7], widened streets [Burden, 2007, p. 10] [Fernandez, 2007, p. 65], reduced the potential for vibrant civic life [Miara, 2007, p. 60] and made many cities less safe for pedestrians and bicyclists. Many of the streets in the most pedestrian-friendly zones of older cities would no longer be legal to build under current traffic codes.[47] Urban planners have in recent years begun to take the automotive threat seriously, pushing for reduced lane-width requirements and designing urban areas that limit vehicle traffic on purpose [Fernandez, 2007, p. 67]. If self-driving vehicles have a high probability of effecting a positive change on the cityscape, as some proponents claim, they could be an important addition to these developments.

Congestion and other environmental factors, such as air pollution, are other social reasons used to support automated vehicle development. With carefully controlled, automated driving, the fuel consumption of each individual car could be reduced, just

---

[45] Critique of automated vehicle systems on these grounds is not new, and Marcia D. Lowe of the Worldwatch Institute wrote in 1993, in response to the US push for automated highways, that "Even more astonishing [than the level of spending] is the total lack of organized opposition to the idea, despite evidence that smart cars and highways may well exacerbate the very problems they are supposed to solve" [Novak, 2013].

[46] My interviews, and those of a fellow researcher performing ethnographic work, suggest developers see themselves as trying to change society. This self-image is not limited to those working on mobility systems [Naughton, 2015].

[47] In some of the United States' most pedestrian-unfriendly cities, major intersections in the downtown area are now more than nine 12-foot-wide lanes across [Speck, 2014].

as so-called “hypermilers” today use altered—and very conservative—driving techniques in order to increase their fuel efficiency [O’Rourke, 2010]. Vehicle-to-vehicle or vehicle-to-infrastructure communications can share road information between vehicles, reduce gaps between cars, and thereby increase throughput and efficiency of the road system.[48] Singapore is engaging in controlled vehicle tests to work on just such ideas, to improve efficiency with vehicle-to-vehicle communications and predictive traffic patterns.[49] SMART[50] research suggests a significantly smaller number of cars could supply all the mobility needs of Singapore if they were automated and could respond to demand.[51] Paolo Santi of the MIT SENSEable City lab suggests replacing traffic lights with a slot-based system[52] to double the capacity of the roadway.

These proposals involve numerous potential pitfalls. While coordinated fleets of vehicles could increase road throughput by 2 to 20 times by various estimates [Adams and Brewer, 2007, p. 229], the top end of those increases would be necessary to exceed the throughputs of buses or subway trains, which can be 10 times those of cars [Lowe, 2007, p. 222].[53] Capacity estimates that assume 4 people per car would require significant changes in ride-sharing behavior to match in the real world. Much depends on actual usage. Utopian visions of increased capacity forget the complex dynamics of human behavior: city planning has found time and time again

---

[48] Daniela Rus of MIT’s CSAIL speaks of “data-driven mobility,” using the affordances of automated vehicles to improve traffic patterns by collecting transit data “in the cloud” so it can be queried at scale. Speaking at The Road Ahead Conference, hosted by the MIT SENSEable City Lab, Cambridge MA, November 21, 2014.

[49] Lam Wee Shan, speaking at The Road Ahead Conference, hosted by the MIT SENSEable City Lab, Cambridge MA, November 21, 2014.

[50] Singapore-MIT Alliance for Research and Technology. Singapore seems to have a unique position in shaping this research, due to its existing vehicle restrictions and willingness to leverage itself as a technology test-bed.

[51] Their modeling suggest that 300,000 vehicles could serve all mobility needs during peak times with waits of less than 15 minutes, and even 200,000 vehicles could reduce wait times to about 3 minutes during off-peak hours [Spieser et al., 2014]. This compares to 779,890 passenger vehicles actually owned in Singapore, according to 2011 numbers.

[52] Each vehicle would request a slot to pass through the intersection, and would be instructed to proceed according to the orchestrations of a central control system (a fruition of Vladimir Zworykin’s work at RCA from the 1950s). Speaking at The Road Ahead Conference, hosted by the MIT SENSEable City Lab, Cambridge MA, November 21, 2014.

[53] Adams and Brewer estimate 10 solo drivers per square foot of right-of-way per hour for conventional highways, and between 100 and 200 for modified approaches that use normal subcompact vehicles and special bubble-like micro-cars, respectively. Lowe describes per-hour capacity for a rail line as high as 70,000 people, compared with 30,000 for buses and just 8,000 for private cars.

over the last 50 years that increasing the capacity of roadways does not result in less congestion, only more people driving [Marshall, 2007, p. 219].[54] If a factor of 10 increase in highway throughput were possible, but caused a factor of 10 increase in overall miles driven, congestion would not be ameliorated, and significantly more environmental impact might result. Any large-scale impacts of automated vehicles are dependent on the architecture of the overall transportation system. And therefore it is the transportation system, not automated vehicles alone, that must be remodeled.

Increases in vehicle use are implicit in the visions of those who push for automated cars to serve needs that are presently underserved. Just as increases in the use of cars to commute have decreased the quality of urban spaces, what is to stop cities of automated vehicles from being even more unfriendly to people?[55] These risks are not equally distributed. An intersection that requires an electronic device to request a slot to cross will not gracefully tolerate the poor, or anyone who does not own a smart car or smart phone, whatever his or her age, race, class, or situation. When examined more closely, one researcher described, the visions of driverless car developers often have a male bias: businessmen taking driverless cabs to work, and drunk students using automated vehicles to ferry themselves back to their apartments. These visions may be greatly influenced by the developers themselves—primarily male engineers with masculinist preconceptions—either thoughtlessly or as a marketing strategy. But regardless of the source, they suggest designed solutions may tend to be predisposed to certain types of uses, and less amenable to others that fall too far outside this vision. Alternative uses are still possible, of course—Uber is already being used by parents to ferry their kids to school [Hoder, 2013] [Shapiro, 2013]—but inherent gender bias in design is not a problem that should be ignored when making large-scale changes to infrastructure, as it can deeply affect the ways and frequencies with which people choose to use new technologies.[56] Thinking about the variety in types, uses, and users

---

[54] One researcher I talked to put the problem in terms of Jevon's paradox: at some point, efficiency improvements will saturate, "so the only true mechanism [of reducing congestion], at least the one that's been proven, through research, is pricing, is through taxation."

[55] This question was in fact asked at the MIT SENSEable Cities conference in 2014. It received little in the way of reply from the panel.

[56] Indeed, a recent study by the University of Michigan Transportation Research Institute suggests men may be more amenable to autonomous vehicle technologies, and it is worth considering the role

of road vehicles—from subcompacts to vans to pickup trucks; from solo commuting to carrying kids to hauling construction equipment or moving furniture; from the wealthy to the poor, the urban to the rural—makes clear that cars are multipurpose vehicles, with culturally specific uses. Acknowledging this specificity is vital for developing fair and just technologies, particularly considering the potential for automated taxi fleets to be used to justify decreased investment in—or even, in some simulations, wholesale replacement of [Spieser et al., 2014]—public transportation infrastructure. We must improve, rather than erode, equitable access to transportation.

When the expected user is the commuting, upper-middle-class, working father or the privileged college boy (and the opposers of the technology are branded as "soccer moms," as one informant said), we risk developing vehicles that preferentially serve certain dominant uses and not others. When I asked one researcher whether developers have an idea of what driverless cars would mean for people who have children, he replied: "They don't think about it." Though this blindness to other needs and uses may not be shared by all developers in this space, it is not something we can afford.

3.2 Conclusion

This chapter has investigated six different perspectives on the social, cultural, and informational contexts of the autonomous vehicle: the technologies of vision and of mapping involved in its operation, the data it collects, the statistics with which it is motivated, the ethics and epistemologies with which it is designed, and the civic ends to which it might be mobilized. These contexts show what is implicit in design fictions of the automotive future. We have seen that the automated vehicle so envisioned exists at the nexus of a number of deep, important questions about our relationships to technology: What level of privacy is appropriate? What counts as machine knowledge? How do we design machines that can be genuinely ethical, and do we need to? We have also seen that these vehicles present a number of thorny

---

that gender-biased technological visions may have in that effect [Migliore, 2014].

problems not often recognized: What level of representational work is required to make them function? How do we predict the relative safeties of devices that do not exist? Can we achieve better urban design simply by adding autonomy to vehicles? All of these questions are united by being largely unanswerable—and the problems, unsolvable—by a focus only on the device itself, narrowly defined. Instead, these are system-level problems, questions that must be asked of large networks of human and machine actors with subtle and shifting interrelationships. Prevailing notions about AI do not easily allow space for this perspective, but while issues will persist, alternative paradigms of automation present the possibility to soften some of the more troubling questions of machine capabilities and human-machine relationships.

Chapter 4

Hybrid Controls, Hybrid Possibilities

As we have seen, dominant automated car narratives rest on two primary, interlocking assumptions. First, the nature of the ideal human-machine interaction for vehicle control is assumed to be known. Second, an inevitable progression toward not only greater autonomy but complete autonomy is assumed as the starting point of these arguments.[1] Rarely if ever does the question of how much autonomy or supervision to provide to the automated system enter the discussion as an engineering parameter over which designers have control, and which should be responsive to larger goals. Instead, advanced driver-assistance technologies and fully-self-driving operation with no need for human supervision are recognized as technologically connected, but perceived as fundamentally dichotomous approaches. And through all this, fully automated[2] operation is assumed to be the ultimate end goal: the ideal human-machine interaction is, in a sense, no human-machine interaction. This perspective comes from particular roots in automation history, but as we have also seen is by no means the only way to read this history. We require an alternative paradigm with which to envision automated systems, one that takes joint human-machine operation seriously.

---

[1] Google's claim is that only complete vehicle automation can entirely address road safety issues, but Urmson goes further to say that the incremental approach will never achieve the necessary full automation: "That's like me saying if I work really hard at jumping, one day I will be able to fly" [Fried, 2015]. While layers of assistive technologies cannot be expected to simply become driverless without high-level control and coordination, this perspective wrongly assumes that driverless vehicles, "completely" automated, are the only reasonable end goal.

[2] And as we will see in this chapter, even so-called "full" autonomy would not, in practice, be so simply disconnected from human oversight.

4.1 Human Supervisory Control

No current "self-driving" vehicles designed to operate on public roadways (as opposed to in controlled conditions) operate in a fully autonomous mode. The cars operate—and legally, can only operate—in a supervised mode, wherein a human driver is responsible for overseeing the automated systems. Even if the vehicle is capable of performing a maneuver "on its own," that operation is monitored, at least intermittently, by a person who can theoretically correct errors made by the automation.[3] This point is made not only by numerous articles describing vehicle operations in general terms, but by Google's own job postings searching for "vehicle safety specialists" to join the self-driving car team in Mountain View. The ideal operator will develop "a unique set of operational skills" using the vehicles, and operate "comfortably in a fast-paced environment, sometimes managing up to four communication channels simultaneously via various high- and low-tech mediums" [HireArt, 2015]. This person's primary duties include filing daily reports and monitoring operations of the software with "constant focus."[4] This role is clearly not one for a passive participant, an observer who sits in the seat, lets the software run (and perhaps self-diagnose failures), and only presses an abort button in an emergency. The test driver is instead an active participant in the complex process of vehicle operation. Further details of these drives, however, are difficult to come by. Test drivers are, expectedly, required to "keep all project details confidential" [HireArt, 2015].

However, in May of 2012, Google tested one of their vehicles on public roads in Nevada, in the only government test yet conducted in the United States. The documentation of this test—which occurred with engineers Chris Urmson (the project lead) and Anthony Levandowski in the front seats—exposes fractures in the tradi-

---

[3] Of course there are serious limits to this capability, as HSC research shows.

[4] While this constant focus is likely to be solely an artifact of the car as a device still in testing phases, the development of a "unique set of operational skills" is just the sort of learning one might expect to perform in order to operate a car with a novel human-machine interface. While it may be Google's intention that only test drivers need develop this kind of expertise, requiring it of prospective users is not necessarily a bad idea. While there is a reluctance to require drivers to acquire new skills, people are already licensed before they are allowed to drive, and the terms of that licensing may well need to change with new technologies and new relationships to automotive technology.

tional perspectives on vehicle automation [Harris, 2014c]. Even though the structure of the checklist only breaks down operation into "Autonomous," "Driver Assist," and "Driver Only" modes, it shows that the human driver was required to assist, or to take control, at multiple points during the test drive. The test records a mix of autonomous and driver-assisted operation when the vehicle faced road construction, switching into manual mode and requiring human assistance to continue. These handovers were not limited to construction, however: "Wojcik [the examiner] also recorded that the car needed driver assistance with some turns, although she did not note the circumstances" [Harris, 2014c]. This should not be taken simplistically, as evidence that the system is not sufficiently advanced. Instead, it is evidence that real operations are more nuanced than narratives of them tend to allow for, involving mixes of attention and control that change over time and varying road situations.

This operational mode, then, is more properly a human supervisory, or joint human-machine, mode than an autonomous one. The study of HSC is by no means new, but perhaps because it does not interact with human fears of obsolescence, loss of agency, or "robot apocalypse" the way AI does—it doesn't stand to damage our egos in the same way, in part because it just sounds rather staid and boring—it has not been as commonly recognized or discussed in popular narratives. And yet, supervisory control is implied any time an article mentions a human driver or co-driver monitoring a system, or taking control at a critical moment. These moments are generally implied weaknesses to the device, as the self-driving vehicle narrative is organized around the ultimate goal of fully autonomous robot cars. Supervisory control, however, admits different goals and possibilities.

Thomas Sheridan's *Telerobotics, Automation, and Human Supervisory Control* comes out of previous work at the MIT Man-Machine Systems Laboratory, and is a core text in the HSC field. Despite its age it remains a great introduction[5] and a

---

[5] And in fact, Sheridan is almost prescient in his identification of the vehicle automation technologies (specifically AVCS) that would make it to market: all of the ones he lists we see today, and the last (like automatic lane keeping) are really only just now becoming available [Sheridan, 1992, p. 256]. Many of these technologies were in development at the time, as Sheridan's book came out during the middle of the EUREKA PROMETHEUS project.

much-needed counterpoint to naïve ideas about automation.[6] Many systems involving supervisory control had already entered relatively common use before the field was constituted (including aircraft autopilots, automatic elevators, and even, perhaps arguably, washers and dryers [Sheridan, 1992, p. 8]): supervisory control, according to Sheridan, truly came into its own as part of research on the teleoperation[7] of vehicles under time delay, specifically on the moon. The time delay enforced a fundamental constraint on direct operation, as the results of any action require three seconds to be reported back to Earth, and therefore made apparent the great benefit of having the remotely operated system include its own internal control loop to allow it to perform simple delegated tasks [Sheridan, 1992, p. 9].

Coming out of an idealized picture of supervision and delegation within human management structures, supervisory control, strictly defined, occurs when

> one or more human operators are intermittently programming and continually receiving information from a computer that itself closes an autonomous control loop through artificial effectors and sensors to the controlled process or task environment [Sheridan, 1992, p. 1].

The less-strict definition loosens the requirement that the device close a control loop of its own, simply requiring it to interconnect "through artificial effectors and sensors to the controlled process or task environment": only in the strict case can the computer operate without the human as an autonomous system "for some variables at least some of the time" [Sheridan, 1992, p. 1]. This emphasis on partial, gradiated control[8] is emblematic of the entire HSC project, and represents its fundamental

---

[6] Human supervisory control history shares touch-points with the history of factory automation, coming out of theories of management and human factors engineering. Frederick Winslow Taylor is identified as a key player in this history, less for his "dehumanizing" approach to the worker than as his intent to generate "a new interest in the sensorimotor aspects of human performance"—in other words, the way that human capabilities interact with the tools they use to accomplish the tasks they are set [Sheridan, 1992, p. 7].

[7] *Teleoperation* means the extension of an operator's sensing and control capacity to a remote location, via an artificial assemblage [Sheridan, 1992, p. 4].

[8] The computer system may function primarily on the "efferent or motor side" to actually implement directives from the supervisor, subject to its own sensors [Sheridan, 1992, p. 3]. Or it may act principally "on the display side," processing incoming sensory information into a form digestible for the supervisor, or, as is usual, it may do some of both [Sheridan, 1992, p. 3]. As Sheridan notes,

ideological difference from the AI-focused perspective on automated systems. This is neither a weakness nor an unwillingness to be sufficiently bold, but a well-considered engineering strategy. Sheridan identifies seven motivations to develop supervisory control, of which six[9] are eminently relevant to self-driving cars, and so I will include those here in their entirety:

> (1) to achieve the accuracy and reliability of the machine without sacrificing the cognitive capability and adaptability of the human, (2) to make control faster and unconstrained by the limited pace of the continuous human sensorimotor capability, (3) to make control easier by letting the operator give instructions in terms of objects to be moved and goals to be met, rather than instructions to be used and control signals to be sent, (4) to eliminate the demand for continuous human attention and reduce the operator's workload, (5) to make control possible even where there are time delays in communication between human and teleoperator, (6) to provide a "fail-soft" capability when failure in the operator's direct control would prove catastrophic [Sheridan, 1992, p. 12].

Though supervisory control, as well as human factors engineering by proxy, is interested in mathematically modeling the human operator in her engagement with the control system—itself a fraught project in several ways[10]—the specifics of such modeling are not necessary to understand the concept or its application to "self-driving" vehicles. To provide a specific example of supervisory control, consider the following situation: a highly-automated vehicle is set up to operate in a supervised manner. The vehicle is capable of navigating traffic on its own, but includes a map interface, a digital display, a shifter that includes an autonomous mode "gear" selection, a standard set of wheel and pedals, a turn signal control, and a cruise-control-like control stalk. The user may set a destination on the map interface and engage the automation

---

the human may assume direct control of the entire system or certain variables within the system at various points [Sheridan, 1992, p. 3].

[9] Number 5 is not relevant as stated, but is very relevant in principle to a distracted driver.

[10] People rarely behave as ideal mathematical functions, and respond differently in test conditions and real life. The black-boxing of the operator into a stimulus-response system was an issue on the X-15 [Mindell, 2011, p. 54] and continues to be problematic today.

from a stop, or engage the automation while the vehicle is in motion and optionally select a destination. The user is expected to be available to assist the vehicle with maneuvers, and oversee its behavior: she may take the wheel at any time to direct the vehicle, or use the pedals to force its speed to alter; she may request lane changes using the lane signal stalk; she may use the cruise control stalk to subtly alter the vehicle's speed to suit traffic and conditions. At any time when the automation is engaged, she may alter the destination on the map. When the automation is engaged, the vehicle will warn if it is encountering a situation it cannot handle, and will revert to a minimal-risk condition if the operator does not intervene (e.g. pull to the shoulder and slowly come to a stop). Whether or not the automation is engaged, certain ADAS or AVCS are operating in closed-loop mode, including pedestrian collision detection. And vehicle data including detected objects and planned paths through the environment are always being presented on the digital display to assist the user in evaluating the environment and determining vehicle intent.

This hypothetical vehicle functions in a clearly supervisory mode, since high-level commands can be provided by the human operator, to be carried out by the automation in accordance with its sensors. Information from the environment is processed by the vehicle and returned to the user via the display, providing the user with more cues as to the environment and indications as to the status of the automated system. Such a vehicle bears little resemblance to the self-driving vehicle envisioned by Google, but looks quite like a current vehicle might after a decade of evolutionary development of driver assistance systems. Furthermore, it manages to address all of Sheridan's supervisory control motivations. It combines the constant vigilance of automated sensing with the judgment and perception of the human operator (1 and 2); it allows the operator to provide high level controls in autonomous mode, instructing the vehicle which lane to be in or which turns to make without having to maneuver manually (3); it allows the operator to pay intermittent attention to the vehicle by including fail-safe modes and the ability to handle most driving situations (4, 5); in the case of a time delay in response (e.g. due to drowsiness) or an operator's failure to control, the vehicle will attempt to behave safely (5 and 6). While actually designing and engineering such a system is by no means as simple as sketching it briefly as I have, it should be clear that a vehicle of this description is at least as plausible as a fully autonomous robotic car. It presents unique problems of human interaction and attention which cannot go unremarked upon, but it also presents unique opportunities for blended capabilities that may not only compensate for deficiencies in computer vision, mapping, or automated sensing, but may do much to address human discomfort with automation systems and concern with being outside of the control loop of their automated vehicle.

Issues with attention have always been implicit within the discussion, since inattention is precisely the human quality that makes self-driving vehicles most necessary, at least rhetorically. Experience shows that human operators who become reliant on automation to perform a task are ill-equipped to take that task back in a crisis situation, and task hand-over is likely to be catastrophic [BEA, 2012] [Chow et al., 2014]. An often-used justification for entirely eliminating the human from the vehicle control loop is related to this tendency not to monitor or be cognitively involved with an operating automation system.[11] Therefore, companies like Google seem to dream of minimizing the interface as much as possible, rather than more broadly considering what sort of interface makes sense [Markoff, 2014a]. While the interface/UX-first approach of tech companies entering this space seems as if it would put interaction at the center, their focus on simplicity often presupposes a certain kind of highly automated operation. So though supervisory control is implicit in much popular writing about autonomous systems, it is not generally acknowledged as an important and developed field.[12]

---

[11] Larsson suggests a less perfect automation system might be better by some metrics, as experience with it is empirically related to greater awareness of its weaknesses, and an expectation of having to take control intermittently [Larsson, 2012]. Notably, this paper and others like it are representative of human factors or ergonomics research.

[12] Human supervisory control is not even that well considered in many engineering fields; even today, the human factors engineer is often low in status within the engineering hierarchy, and is too often brought in at the last minute to design an interface for an already specified device rather than actually being involved in device design from the beginning, as effective joint-cognitive systems engineering requires. David Mindell, discussion with the author, September 8, 2014.

4.2 Cognitive Networks, Networked Subjectivities

With our attention on supervisory control, however, new questions and processes come into focus. If the machine is clearly no longer the sole component of analysis, what formerly neglected pieces must be considered? How do we assess system design or performance, or think about the ways these newly expanded devices operate? As Sheridan notes, much study of HSC has gone on elsewhere under different names. Coming instead from cognitive science, Edwin Hutchins's anthropological work on group cognition processes expands from the traditional focus on the individual agent to systems of interacting agents and technologies.[13] Hutchins uses the example of an airline cockpit [Hutchins, 1995] to show how the cognitive properties of a system as a whole may be very different than those of individual human actors within that system. Studying the humans alone, or the automation alone, does not suffice to explain the overall behavior of a joint human-machine cognitive system.

Hutchins is very particular that cockpit speed bugs do not naïvely increase pilot memory.[14] Instead, they are information-processing tools, responsible for changing the form of information, giving it a new representation and altering the interaction with that information [Hutchins, 1995, p. 282]. The whole system can be said to have a "memory" that is distinct from the pilot's memory: the "*cockpit system* remembers speeds" by virtue of the pilots within that system judging a needle on the airspeed indicator against the position of the bug [Hutchins, 1995, p. 283].[15] Current vehicle

---

[13] And, notably, Hutchins responds to the canonical focus of cognitive science on the "mental processes that organize the behavior" of an individual, a position he cites as having its standard statement in Newell and Simon's 1972 *Human Problem Solving* [Hutchins, 1995, p. 265-266]. These two early AI pioneers show up again here, and their individualist focus sheds some light on why the narratives of automated vehicle technology that are inflected by AI are so different than they would be if told instead through the lens of supervisory control.

[14] The speed bugs—which note on the airspeed indicator important speeds at which the flight characteristics of the aircraft need to be altered to maintain safe operations—are rudimentary, but the issues they present do not disappear with increased automation. Hutchins and Klausen co-wrote another article [Hutchins and Klausen, 1996] which focuses on the interactions of the Captain, first officer, second officer, and air-traffic control. This exchange, though it involves control yokes, altitude alerts, and other technological actants, is primarily focused on the joint capabilities of the human crew, and presents fundamentally related cognitive systems issues.

[15] It is worth noting here that these are not the remarks of a maverick. The research involved was supported by a grant from the NASA Ames Research Center, as part of the Aviation Safety/Automation Program.

automation systems are more complex than the cockpit's visual-mechanical aides, but are appropriate subjects of the same kind of analysis.[16] Complex jobs involving spatial processing and reasoning (driving between the lines and avoiding other cars) are likewise transformed through technological interfaces into other tasks. It is precisely these changes in task to which we should attend if we would like to properly understand (and design) the role of the operator within an automated vehicle.

A more nuanced view of the cognitive tasks involved in interacting with automated systems also goes quite far toward clarifying some of the more difficult parts of system automation. Ironically, increasing automation can increase the load on the operator, rather than decreasing it, and lead to increased failure rather than increased reliability [Parasuraman and Wickens, 2008]. A distributed, cognitive attention to automation highlights the problem: an automation system may transform a task from a complex visual task for the operator (e.g. driving) to another complex visual task (e.g. monitoring a control panel with many displays and lights, and switches to press) while the overall processes within, and results of, the system remain fundamentally congruent. Looking at which tasks are allocated where, and what the demand on human perceptual and decision-making systems is at any particular time, improves the analysis. As one should expect, people drive largely based on experience and instinct, [Knapton, 2015] not logical thought; and similarly, system performance should be expected to shift over time as interactions with automation are internalized. As Hutchins and Klausen describe, "It is possible to design computer systems with open interfaces (Hutchins, 1990) that support learning in joint action but this can only be done when the designer goes beyond the conception of the isolated individual user" [Hutchins and Klausen, 1996, p. 13]. Appropriately recognizing interactions is critical to successful design.

Complex dynamics of information transfer are critical for the safe operation of large-scale systems. For automobiles, the propagation of a "plan" of each vehicle's path

---

[16] As an aside, I would argue that if a number of modern information technologies, like smartphones, seem to have done relatively little to alter human beings and their capacities, it is because the individual is still the unit of analysis. When the unit of analysis is instead set at the individual plus his or her immediate technological or biological surroundings, much greater differences in capability become apparent.

on the road (either within the minds of drivers or the code of an algorithmic system) into the state of the vehicle's controls keeps vehicles physically apart, parallel to the safe operation of fleets of aircraft.[17] There is a difference in the professionalization of pilots compared with the "ordinary" status of an everyday driver, and this difference in status and roles changes the amount of training we expect operators to have and the sorts of interactions that may appeal to them, but there is no fundamental reason why supervisory or joint-cognitive systems approaches and experience are not applicable to automated cars, if appropriately moderated for the role and status of the driver. However, the levels of automation formulations generally applied to describe and classify automated vehicles do not attend to this concept in sufficient detail.

4.3 Engineering Standards and Policy Documents

The engineering standards and policy documents set up to guide the industry's development—particularly levels-of-automation formulations, coming from a long history of such taxonomies [Parasuraman et al., 2000] [Proud et al., 2003] [Huang, 2007] [Parasuraman and Wickens, 2008][18]—strongly shape automated vehicle narratives. Developed to guide researchers and public agencies toward an appropriate understanding of automation, these documents are easily taken as evidence for how automated systems must or will develop in practice. The primary formulations of levels of autonomy for self-driving cars have been published by the National Highway Traffic Safety Administration (NHTSA) and SAE International (formerly, Society of Automotive Engineers).[19] These reports exhibit some resonances and contradictions in how these organizations represent autonomous vehicles and their human drivers—and speak to different ideas of who the "driver" will be in automated systems. However, both make assumptions about the way autonomy is or will be implemented that may

---

[17] Hutchins and Klausen state: "if we step back and look at the entire aviation system and ask how it is that aircraft are kept separated from each other, we see that it is through the propagation of representational state of descriptions of flight paths into the state of the aircraft controls themselves" [Hutchins and Klausen, 1996, p. 14].

[18] Perhaps traceable, in a way, to Sheridan's work [DSB, 2012, p. 23].

[19] The German Federal Highway Research Institute, BASt, has also produced a taxonomy that predates the SAE's work, and influenced it in key ways.

not ultimately be valid and certainly foreclose on alternative ways in which systems could be designed.[20] When closely interrogated, they do not sufficiently heed the complexities of joint human-machine systems operation.

4.3.1 NHTSA: Reacting to the Industry Narrative

The NHTSA's levels of automation focus largely on the human driver and human costs, and seem to represent an inability or unwillingness to think beyond human-machine oppositions which parallels the general narrative tendencies we have already examined in automation and AI history. The agency identifies three "distinct but related streams of technological change":

> (1) in vehicle crash avoidance systems that provide warnings and/or limited automated control of safety functions; (2) V2V communications that support various crash avoidance applications; and (3) self-driving vehicles [NHTSA, 2014, p. 3].

Though the agency positions these technologies "as part of a continuum of vehicle control automation" [NHTSA, 2014, p. 3], aligning themselves with Sheridan, this belies their three-part dissection of the technological landscape and their following five-level taxonomy. They correctly recognize that today's "crash avoidance and mitigation technologies" are the "building blocks for what may one day lead to a driverless vehicle," but incorrectly assume that an easy line can be drawn between driven and driverless [NHTSA, 2014, p. 3]. In truth, vehicle automation technologies present a spectrum of possibilities that are not clearly delineated, but stretch from the "fully manual" control of early automobiles to current amounts of automation, to potential future technologies that are remotely monitored. Like the mythical personless factory, the driverless vehicle can only exist by dismissing its new forms of labor.

According to the NHTSA definition, automated vehicles are those in which "aspects of a safety-critical control function (e.g., steering, throttle, or braking) occur

---

[20] The SAE in particular made an attempt to avoid prescriptive definitions of autonomy in their work. But even a descriptive set of levels contains an explicit hierarchy, which can be read as an implicit narrative of progress or roadmap for development.

without direct driver input" [NHTSA, 2014, p. 3]. This definition *excludes* warning systems that nevertheless still use automation in varying degrees, drawing a suspect line between monitoring and action. Though systems that only provide information and systems that directly affect the mechanical state of the vehicle do differ, making that distinction part of the definition of "automation" perpetuates an inaccurate view that automation is only meaningful or important to consider when it affects a physical mechanism. Information automation (road condition warnings, traffic notifications, etc.) is written out of the NHTSA's definition. But information automation is one of the historically dominant modes of automotive automation, even in the eyes of the NHTSA [Wetmore, 2003, p. 11], and has important safety implications.

The agency's definition also excludes non safety-critical "control functions."[21] This neglects both the vital roles automation plays in tertiary components as well as the delicate dance of human and automated control that happens in current vehicles. Automated wipers and hazard lights may well be "safety critical" in at least some situations. The state of the transmission and drivetrain, controlled via automatic (read: automated) or manual shifting, is fundamentally important to throttle control and the ability to safely control the vehicle. Shifting represents one of several areas in which complex relations of human-requested and automatically provided operating criteria are already visible: both the driver and automation may be concurrently monitoring the engine RPMs, and the driver may request shifts that reinforce or override the preferences of the automation (or which, in the case of a shift that would over-rev the engine for example, are rejected by the automation). A guide for the continued automation of automobiles could (and, I would argue, should) make an honest attempt to reason about such preexisting examples and build from them. But instead, they are ignored, and thus the terms of the agency's definition of "automation" itself are incoherent and self-contradictory. By writing out current systems from the field of "automated vehicles," the NHTSA has hobbled their terminology, and limited its usefulness to make sense of the broad spectrum of automation approaches that

---

[21] The agency is likely attempting to make this distinction so that cars with currently banal automatic technologies like automatic transmissions do not count as "automated vehicles," but this is the wrong footing on which to rest a definition of automation.

are actually being used in existing vehicles.

The NHTSA's levels of automation only compound these issues with simplistic distinctions that do not account for the majority of ways these systems could be engineered. Level 0, or "no-automation," represents precisely the illustrated issues inherent in the NHTSA's definition of *automation*. Level 1 is the only automation level beyond 0 that does not have the ability to apply all types of mechanical input to the system: it is limited to "one or more specific control functions" that operate "independently," and only for certain periods of time [NHTSA, 2014, p. 4]. The necessity of independence between automated "control functions" is especially hard to fathom: in a modern automobile, few if any systems are truly independent, instead coordinated via electronic management systems.[22]

The report is even confused on the number of control functions, though it ultimately means that both steering and brakes/throttle cannot be automated at the same time in a level 1 system:

> there is no combination of vehicle control systems working in unison that enables the driver to be disengaged from physically operating the vehicle by having his or her hands off the steering wheel AND feet off the pedals at the same time [NHTSA, 2014, p. 4].

This continues the agency's curious focus on the physical state of the driver's body, repeated in level 2 automation to make the distinction that now the operator can be "disengaged from physically operating the vehicle by having his or her hands off the steering wheel AND foot off pedal [sic] at the same time" [NHTSA, 2014, p. 5]. This description of bodily state is the most coherent part of the levels definitions, but makes little practical sense for defining automation. One can drive hands-and-feet free in a non-automated vehicle, for short periods of time, in the right conditions. Does the car become a level 2 system during this time?

While this question seems fatuous, it becomes important if the physical state of the driver's body will define system-level automation—and, given the NHTSA's power

---

[22] For example, see systems like Bosch's comprehensive electronic energy management, EEM.

as a regulatory body, come to define legal aspects of automated vehicle policy, such as how vehicles are registered. (It is also worth noting that this clarification is unnecessarily normative, and does not transfer to other types of bodily input like GM's abandoned "Unicontrol" or controls designed for amputees.) Though this stipulation attempts to address what the human driver is actually doing, it is insufficiently granular to account for the various types of mental and physical effort exerted: from using control stalks, to monitoring for vehicles, to consulting a GPS or mentally planning a route, to monitoring and adjusting an automation system via steering-wheel based controls. While Hutchins's cockpit remembers its speeds through a combination of human activity and physical cognitive aids, the NHTSA's "vehicle" may sometimes "not perform a control function" [NHTSA, 2014, p. 3] and nevertheless remain on the road. But this preoccupation with the body results from the NHTSA's charter: the human is a safety liability, something to be protected. The operations of hands and feet represent readily visible and naïve distinctions between human and automated, even if the distinctions do not turn out to be particularly useful when faced with technical scrutiny. For their stated audience, primarily state lawmakers who are not likely to be human factors engineers or well-versed in human-machine interaction, these markers of human action are persuasive but likely to contribute to bad policies.

Further levels extend the functions which are automated. Levels 2-4 have automated systems capable of making all the electromechanical inputs necessary to drive; they differ only in the extent to which human stewardship is necessary [NHTSA, 2014, p. 4–5]. This exposes naïve assumptions about how vehicles will be automated: one whole system at a time—and these assumptions appear implicitly, again and again, in press accounts. How much control is handed back when the human has to take over more vehicle operation tasks is not clear: One primary control function? All primary control functions? The NHTSA is trying to define an overall automation level for a vehicle, but their definition masks system specificities. The report is not clear on the level of automation for a vehicle in which some control systems require continuous monitoring and others can transition to manual control on an appropriate timescale. And the NHTSA's taxonomy rests on the assumption that the work of driving a vehicle will remain basically the same, some tasks simply shuffled to the computer—system control first, monitoring of control second—with no new cognitive loads placed on the human as a result. This too is a naïve position that is not supported by current examples of automation technology, which often result in the generation of new types of human work (monitoring the automation system itself). This labor is implicit in the NHTSA's taxonomy, but not examined in great detail. The NHTSA suggests special training to "authorize the operation of self-driving vehicles," [NHTSA, 2014, p. 11] and positions the operator as a subject of education and regulation while calling upon him or her to be the ultimate decision authority:[23] the automated functions should always defer to driver input to the wheel and pedals (with the sole exception of already-proven technologies such as traction and electronic stability control) [NHTSA, 2014, p. 13].

The NHTSA framework overlooks a large proportion of the actual work of driving, and is therefore a poor model for evaluating and regulating automated vehicle systems. Its bias toward full-automation, structuring the hierarchy of levels around fully self-driving vehicles as the technological peak, slants the development narrative—in popular attention as well as in potential regulation and legislation—toward one particular approach to vehicle automation while ignoring or discounting alternatives. I do not mean to imply that the NHTSA are myopic, but their framework is deeply influenced by their institutional culture as well as the circumstances of its creation. Their preliminary policy document was reactionary, an attempt to regain ground and provide some sort of guidance to an industry already testing on public roads. The automation-first approach, which wants to jump directly to an NHTSA level 4 system, must be tempered for liability reasons since, as the NHTSA is very aware (given their research interest in level 2 and 3 systems), that technology is not yet ready.[24] When trying to make space for technologies that are not yet mature, keep-

---

[23] They mention "Several State automated vehicle laws consider the person who activates the automated vehicle system to be the 'driver' of the vehicle even if that person is not physically present in the vehicle": their only commentary on this point is to say that they know of no current systems (level 4) capable of this operation [NHTSA, 2014, p. 5].

[24] They continually restate the point that only few level 3, and no level 4, systems currently exist, and that neither of these technology levels is yet ready for unfettered operation [NHTSA, 2014, p. 10, 14].

ing a human driver in a position to recover when things go wrong, as the NHTSA does, makes sense as a normative strategy. However, the NHTSA implicitly buys into the prevailing public-facing narrative in industry, that fully-autonomous vehicles are the obvious, natural end product of the evolution of automated systems. Their document, therefore, reinforces that narrative.

4.3.2 SAE: The Human as Engineered Component

Contrasting the NHTSA's model with the SAE's provides a deep look into the priorities of the organizations. The SAE document provides a full taxonomy of levels of automation for "on-road motor vehicles," and the presumed audience consists of engineers not policymakers [SAE, 2014]. It focuses on the three "highest" automation levels ("conditional, high and full automation"), implicitly because these are the newest areas of research and therefore the most important. The report is careful to distance itself from the terms "autonomous" or "self-driving" as used in the media, preferring instead its own carefully defined terminology [SAE, 2014, p. 5-6].

The SAE report does try to square itself with NHTSA recommendations, however approximately, and represents the hidden difference between NHTSA's levels 2 and 3 as the monitoring of the environment by the vehicle (in level 2 the vehicle does not monitor, and in level 3 it does). In contrast, the NHTSA report does not make this plain, and indeed a level 2 system can be said to "relinquish control with no advance warning" which should be impossible without either knowledge of the environment or the occurrence of a bug: the system needs something on which to base its decision to "relinquish control" if such a decision can be said to be made [NHTSA, 2014, p. 5]. Rather than overloading the word "control," the SAE considers the "dynamic driving task" to have a number of components, including the detection and classification of objects and events, the response to such events, planning of maneuvers, steering and turning (including lane-holding and changing), acceleration and deceleration, and "enhancing conspicuity" (referring to lighting, signaling, gesturing, etc.) [SAE, 2014, p. 6]. Similar to the NHTSA's restricted focus, the SAE's taxonomy focuses primarily on four major aspects of this task: steering, acceleration/deceleration, monitoring of the environment, and fallback performance, but is somewhat clearer about how tasks are defined—in particular, it is clearer about the separation of the driving task into longitudinal and lateral components. Also like the NHTSA, the SAE's definitions show all systems above level 1 having the full electromechanical capability to drive the vehicle (to execute both longitudinal and lateral driving tasks) which makes it difficult to account for complex hybrid systems in which some tasks may be highly automated and others not.

The primary distinction for the SAE report is whether the "human driver" or "automated driving system" monitors the driving environment [SAE, 2014, p. 5]. The report deals appropriately with human psychology, stating specifically that higher levels of automation are based on the expectation that the human need not "and therefore will not" continuously monitor the environment [SAE, 2014, p. 9]. But their picture is still binary, and organized around monolithic tasks, rather than considering the moment-to-moment distribution of human cognitive effort. This is implicitly connected to their inclusion of the lower levels of automation only "as points of reference to help bound the full range of vehicle automation" [SAE, 2014, p. 2]: they are not within the document's focus, which favors an implicit—to the NHTSA's explicit[25]—rhetoric of progress toward the higher levels of autonomy. Not only are hybrid task monitoring systems classified as level 2 regardless of their complexity or capabilities, the SAE's taxonomy by definition leaves out whole classes of systems where the default execution of the dynamic driving task is up to the human but monitoring/fallback performance are computerized, or where the dynamic driving task is shared in a complex way, as is monitoring and fallback performance. It considers only systems in which longitudinal and lateral control is handed over to the system earlier in the hierarchy than monitoring or fallback performance.

The SAE is better about identifying new types of human work, such as the monitoring of the automated system, generated by partial automation strategies.[26] They

---

[25] The NHTSA explains this "progressive" focus by deferring to existing laws: "Generally, these laws seem to contemplate vehicle automation at Levels 3 and 4, as discussed above, i.e., some form of self-driving operation. Accordingly, these recommendations are tailored to Levels 3 and 4 automation" [NHTSA, 2014, p. 10].

[26] For instance, the human driver "constantly supervises dynamic driving task executed by driver

are also more conscious of issues of the handover of control, and specifically discuss delayed-release of particular tasks “when immediate human takeover could compromise safety,” slowly transitioning control of a system from fully autonomous to fully manual though never lingering in between [SAE, 2014, p. 4]. It should be noted that this directly contrasts with the NHTSA’s driver-focused approach, as it places fundamental decision authority in the hands of the computer system. When development is not necessarily oriented toward achieving fully autonomous operation quickly—or when a taxonomy is not designed around making space for Google’s self-driving car project and its ideological orientation—and is instead focused on incremental improvements, such electronic overrides seem more acceptable.

The SAE mentions the true complexities of driving at the report’s end, describing how the *dynamic driving task* is distinct from “driving”: “Driving entails a variety of decisions and actions, which may or may not involve the vehicle being in motion or even being in an active lane of traffic” [SAE, 2014, p. 12]. Driving is split into Strategical, Tactical, and Operational components, where strategical includes trip planning and route selection, tactical includes maneuvering, and operational involves “split-second reactions that can be considered pre-cognitive or innate” [Michon, 1985]. The SAE is explicit about their exclusion of Strategical effort, presumably the human’s task, from the definition of the dynamic driving task. However, this admission does expose an inconsistency within the SAE’s taxonomy: level 5 automation (and some types of level 4 automation) definitionally require Strategical effort (route selection is implicit within GPS navigation), which is not separately mentioned as a system capacity in the document.

Despite their differences, these taxonomies are united in lending credence to the teleological narrative of automation: further technological development implies “higher levels” of automation, which imply a decreasing role for human beings. (How different these taxonomies might appear if structured or numbered differently!) The tendency, endemic to these taxonomies, to predicate work on jumping to the flashy upper tiers of automation results from preconceptions about the value of level 4–5 au-

---

assistance system” [SAE, 2014, p. 3].

tonomy and the intermediate stages by which it will be achieved. The SAE's greater facility with the technologies involved seems likely to have contributed to their ability to achieve a somewhat more nuanced view of vehicle automation than the NHTSA. But the technology-first perspective that shapes these sorts of formulations comes from a deep history of engineering practice, and a tendency to treat the user as a black-boxed entity, another signal processor with defined inputs and outputs, subject to mathematical modeling—a set of qualities that is highly useful for engineering work, as it is amenable to descriptive standards and interface specifications.

Both taxonomies ignore the external labor involved in that automation: human action displaced in space and time. Where do remotely monitored systems fit? What about human preparatory work that allows for automated systems to operate? These are questions with serious implications that are not answered by either of these autonomy formulations, but should be addressed by any document that attempts to describe, normalize, or regulate automated vehicle operations.

It should be noted that levels formulations, and their assumptions, are not without existing criticism. The Defense Science Board identifies that these levels formulations are "often incorrectly interpreted as implying that autonomy is simply a delegation of a complete task to a computer, that a vehicle operates at a single level of autonomy and that these levels are discrete and represent scaffolds of increasing difficulty" [DSB, 2012, p. 23-24], all interpretations made by both the NHTSA and SAE taxonomies—ones I have been attempting to critique. Preoccupation with autonomy "deflects focus from the fact that all autonomous systems are joint human-machine cognitive systems" and "reinforces fears about unbounded autonomy" while obscuring the fact that no systems are fully autonomous [DSB, 2012, p. 24]. Humans are always involved somewhere along the line, in the programming, production, and use of automated systems.[27] These levels formulations lose sight of the fact that autonomy and control are already fractured and contingent: what "level" is my vehicle operating in when I drive an automatic transmission in semi-manual mode, or engage cruise con-

---

[27] It makes sense that the military, with significant automation experience, as well as a culture of human command hierarchies, would find it especially important to make this apparent.

trol, or allow a turn signal stalk to automatically return to neutral after a turn of the wheel is detected, or stand on the brake pedal to engage ABS in an attempt to bring the vehicle to a stop as quickly as I have requested. These currently everyday situations already defy levels, and represent overlapping, shifting delegations of authority that may change on a second-by-second basis. I also claim, to add to their critique, that the levels seem to suggest scaffolds of increasing *desirability*, which, as we have seen in historical examples in factories, air, and space, is a matter of perspective.

In a sense, these autonomy formulations pull together three previous strands of this thesis: (1) they are influenced by science and design fictions, and seek to respond to the narratives of technological revolution as transmitted through the popular marketing of what nevertheless remain laboratory objects; (2) they repeat many of the mistakes of the dominant narrative of automation history, and neglect some of the complexities of human-machine labor relationships evidenced by their historical antecedents; and (3) they represent models for the continued development of cutting-edge AI technologies, which perpetuate the vision of an unbroken chain of development toward (chimerical) "full" autonomy.

4.4 Alternative Futures

"Full autonomy" will not be so simply *full*: at the very least, a human-rated[28] system will need to have a stop or abort button. And having such an option implies some human oversight, which means a number of additional questions soon arise: How much monitoring? When does it happen? And how is it regulated or supported by the device? New tools and interfaces may be necessary to facilitate this process, to allow for the human to take on new roles in the human-machine relationship: currently, engineers co-piloting Google's driverless vehicles sit with laptops to monitor the vehicle telemetry [Bilger, 2013]. Data displays like this, or like the pretty LIDAR visualizations that get used to explain and dramatize sensor data, could form the

---

[28] This is a term of art in engineering: a human-rated (or man-rated) system must meet stricter safety standards than one that is unmanned.

core of the instrumentation for the automated vehicle. More discussion should be happening around the dashboard of the future; in the dominant narrative of the driverless vehicle, the dashboard is vestigial, but it could be a vibrant space presenting unparalleled new possibilities for human-machine interaction.

While it may be just as difficult, or even more difficult, to engineer an appropriate joint human-machine system that takes the capabilities and desires of people seriously, rather than trying to minimize human involvement to the greatest extent possible, that approach might result in vehicles that come to be more socially acceptable in certain ways—notably, in terms of autonomy and ethics. These ideas present distinct possible futures worth investigating.

Attention, and how to preserve, regulate, monitor, and transact it, must be at the center of vehicle development.[29] And the parameters of these interactions are in part defined by the end goal. Much popular discussion centers around texting as a key motivator of the need for automated cars. But as David Mindell describes, "wanting to text is different from needing a fully automated vehicle,"[30] and it seems quite possible to create a vehicle that will allow distractions on the order of the time span of writing and sending a text even if it is not feasible, or not culturally acceptable, to build one that is fully automated.[31] Entirely different time-scales of automated operation are involved, and represent different engineering situations.[32] While people do panic in crisis situations [Wise, 2015], it may be more appropriate and more useful to consider systems that scaffold and support human capacity and attention, rather than attempting to eliminate it altogether. In all this, we should avoid the clichéd question "People: sinners or saints?" which seems so popular in

---

[29] I must add here that having to attend to something is not naïvely a good thing. Who gets what kind of agency, and exerts what *kind* of attention (creative or menial), is at the center of the human-machine dynamic in automation technology. Not focusing on the role of human attention risks that these decisions get made by fiat, on the basis of other factors.

[30] David Mindell, discussion with the author, September 8, 2014.

[31] Recent AAA research on teen accidents shows distraction to be a factor in 58% of crashes, and that drivers were not paying attention to the road for an average of 4.1 out of the final 6 seconds before the crash [Green, 2015]. This suggests that a system able to maintain control of the vehicle for times on the order of 10 seconds, and stop it safely if necessary, could do a lot to address this kind of distraction.

[32] Time scale of human involvement is a good way to think about autonomy: greater autonomy means longer time scales between human interventions.

media discussions: it represents a "false position" that does not adequately reflect the complex nature of operating automated systems [Woods and Hollnagel, 2006, p. 1].[33] Success in automation is not a given that the human presence degrades,[34] but something to be achieved through careful engineering of automated components and their human interactions.[35]

It seems highly unlikely that a supervisory control perspective will fix the potential contribution of automated vehicles to the surveillance state: data gathering technologies will still be available, and comprehensive mapping may still be necessary for operations, though human supervision might reduce reliance on remote processing of data in the short term and thereby improve privacy. In any case, the potential still exists to use hardware and software to try to restrain and control people's actions and movements. Supervisory control will also certainly not fix the reliance on projections of risk for decision-making about vehicles: while reducing certain risks, this approach may increase others, and any estimates of such risk will still be estimates only, derived from models that may be incomplete.

However, human supervisory control stands to alter the dynamics of use and therefore change the sort of urban planning necessary to account for these vehicles, as well as the environmental effects of their use. Systems designed with supervision in mind may operate in a different way than those designed primarily with autonomy in mind. Operational differences will change the behaviors favored by these systems: Can they operate on their own and drive to remote parking garages? Or be sent to pick up children from school? How much attention is required to operate the vehicle, and if that attention comes from the human within the vehicle, how will that affect

---

[33] Ultimately, as Woods and Hollnagel point out, all cognitive systems are finite and will therefore also have errors, lapses, and failures [Woods and Hollnagel, 2006, p. 1–2].

[34] The PARC/CAST Flight Deck Automation Working Group found that human pilots are actively involved in mitigating risk from other system components, and states that "an exclusive focus on pilot errors will not take into account the positive actions and decisions that pilots do on a frequent basis" [Flight Deck Automation Working Group, 2013, p. 30].

[35] For example: "fully automatic systems [e.g. airbags, ABS] as well as fully autonomous systems depend on humans to define and limit the scope of their authority and the range of possible actions . . . the human's role becomes more rather than less important when moving toward the autonomous end of the spectrum because it is so important to assure that the systems are properly designed, tested, deployed, and monitored" [National Research Council, 2014, p. 14-15].

how likely people are to drive rather than take other forms of transport that could reduce, rather than increase, road congestion?

At the most radical, we are facing not a system of autonomous robots but fleets of telerobotic vehicles with closed-loop control systems providing limited autonomy, that are supervised by remote operators within the datacenters of multinational IT corporations.[36] But even this, though it presents certain advantages, is by no means our obvious future. Such a networked system worsens issues of privacy and data security. Longer time periods of unsupervised operation, and more limited involvement of the people within the vehicle, require automation technologies with greater reliability and situational capability in order to reduce the risk of catastrophe for operators—whether that catastrophe is measured by personal death or system failure leading to financial ruin. Human beings are successful operators and risk-mitigators, who are capable of performing complex and difficult tasks. And it would be wrong to categorize human involvement as purely a temporary requirement, until computers are sufficiently advanced: it may be that whatever the ultimate capabilities of machines, we may find it socially unacceptable to surrender certain areas of operation completely to computers.

Where supervisory control considerations come to totally change the landscape is in the field of ethics. The ethical operation of vehicles involving computer code will always depend to an extent on that code, and the previous discussion about programmatic ethics still holds. But what changes here is that the vehicle is no longer free to operate based only on its programmed ethics for long periods of time. Instead, the human operator is involved in decision-making processes at a much more granular level. Will some emergency situations likely be handled in a largely automated fashion, just as today's pedestrian detection systems can perform an emergency stop to avoid pedestrians a driver does not see? Almost certainly. But these issues,

---

[36] I have encountered few companies who seem to be honest about this supervisory arrangement, but the founder and CEO of the mini-cab app company Kabbee envisions just such a thing: "There's no reason why a taxi driver couldn't have something like three screens in front of him . . . and sit in a studio directing three or four or five cars through the roads of London where they are driverless cars with a human controlling the GPS" [Dawson, 2015]. Whether the appropriate number is 5 or 5000 cars, this vision seems plausible, and bears similarities to current Predator UAV operations. It also sounds quite a bit like GM's 'Autoline' concept from 1964.

though perhaps ethically tricky, are much closer to situations we are already working through. Ironically, one positive effect of the autonomous vehicle might be to estrange us from vehicle operations, and thereby draw us to consider more seriously the ethics programmed into more ordinary software systems, but the effect in the dominant narrative seems to be drawing a clear distinction between current and future automation systems. This distinction, evident in the NHTSA's taxonomy, is artificial. But consistent human involvement in vehicle operations seems to open new doors for human acceptance. What the NHTSA should do is consider carefully why it feels that ABS and traction control can be allowed to override human inputs, and what separates them from systems that should not be allowed to do so.

One of many contradictions at the heart of automated vehicle research is that these vehicles promise to provide to us a greater measure of creativity[37] in the act of driving, to remove some of the "menial" and routine tasks of manual control in favor of strategic decision-making. But despite a focus on relieving tedium, these systems have not primarily been envisioned as providing people creative control. Instead, in the process of following the dream of fully automated operation—where the human labor has been entirely removed from the shoulders of the person in the vehicle, and displaced to the invisible labors of mapping, programming, and monitoring—engineers are designing systems where the "driver" seems present largely to ensure the operation of the machinery, burdened with new but perhaps even more menial tasks of machine tending.[38] Troublingly, these visions ignore the skilled, rewarding parts of driving, and disregard the shift in the automated vehicle from a family space to an extension of the working environment, which threatens its own sort of potentially uncreative labor—there is, however, an alternative.

Supervisory control presents the possibility for a renegotiation of car culture and its troubled relationship with the automated vehicle. In my interviews and research

---

[37] They claim to do so largely by allowing people to do something other than sit in traffic: to read, to sleep, to eat, to work.

[38] See for example Tom Simonite, "Lazy Humans Shaped Google's New Autonomous Car" [Simonite, 2014], which discusses the human role within Google's test vehicle, and the company's response. This is purely speculation, but due to the way the system operated previously, it is possible at least one Google employee fell asleep at the wheel, which was the catalyst for their concern and change in focus.

work at conferences, I noted especially the contradictions within common characterizations of human drivers and their future. One interviewee basically asserted that humans are bad drivers and that perhaps we should "all have our licenses taken away," but later stated that "car nuts" should welcome self-driving cars because they will free up the roads for enthusiast use. This combination is self-contradictory. The question of whether people will be allowed to drive also looms large in industry conferences, with similarly unhelpful responses.[39] While human driving is not going away in the near future, the idea that autonomous vehicles will free up the roads for people who want to drive manually seems more like a platitude than a thought-through design strategy. These questions of use and licensing are deeply embedded in the idea of progress. "Does the potential for greater safety imply we should be taking human licenses away?" is a question that only makes sense with a particular vision of autonomy, one that does not involve human supervision. Perhaps a better question is what sort of licensing and training is necessary, or should be necessary, to operate new kinds of vehicles, which must be defined by the roles that the human operator must perform. People are already licensed to drive, and there is no reason to say that new types of training may not be necessary or good.[40] With continued human involvement in driving, there is the possibility of a middle-ground in the hybrid narrative: future cars, using increasing amounts of automated safety technologies, may be vulnerable to the same critiques as the "overly" computerized Nissan GT-R. But these vehicles could still provide both an ethical role for the human and a place for enthusiasts who want to be involved in vehicle operations.

4.5 Conclusion

In the end, whether an operator is in a vehicle trying to monitor that it is on the right path and not about to hit an undetected obstacle; or is waiting for a car to arrive,

---

[39] Computer scientist Daniela Rus predicted, at an MIT conference, that people will still be able to drive normally, sharing the roads with automated vehicles in the near term. This is correct, but does little to allay fears for the far future.

[40] Some continuing education could even be a positive influence on current traffic safety.

wanting to ensure that it is en route and not in an accident; or is working as a system-technician at a multinational corporation administering a fleet of automated vehicles, attempting to ensure they are not being hijacked or stolen, the operator functions as a supervisory controller, enmeshed in a complex network of human and machine systems involving, at in its broadest view, designers, producers, owners, lawmakers, and multitudes of others. In a more restricted sense, however, the operator is directly involved in operating a joint human-machine system. Not recognizing this situation as such does not make it not so: it only means we are likely to neglect the most important parts of the system design—in terms of its adoption and long-term use—by blindly following a narrative based on a warped view of history. And we are likely to ignore new skills and competencies that may be important for operation (or even vital to operation), or which may be valuable in other ways, for human reasons. Cars—and transportation systems as a whole—are designed to serve human needs, and must ultimately answer to society's ideas about those needs, even those that might seem to some to be mere ethical squeamishness.

Chapter 5

Conclusion: Driving the Future

Our journey through ideologies of automation and driverless vehicle development has taken us from an explication of the dominant narrative of the “self-driving” vehicle and its histories, through an examination of what it might mean for our society, to an investigation of an alternative paradigm for thinking about automation. What should be clear from this thesis, but which must be made clearer in the public discourse about increasingly automated technologies, is that a movement from human systems, through hybrid systems, to fully autonomous systems is not inevitable. It is not a requirement of technological progress, but one narrative among many which depends on technological, and often infrastructural, progress as well as a mischaracterization of what “full” autonomy is. However, this narrative presents a convenient cover story for the economic objectives of organizations involved in automating the car.

5.1 Reshaping the Road: A Last Lesson in History

We often take for granted that cars have separate spaces on the roadway and interact with pedestrians in a controlled fashion.[1] However, there was a time when city streets were not primarily the venue of the automobile, but were mixed-use areas

---

[1] This is not, however, to minimize the human interaction component between drivers and pedestrians at crosswalks and lights. Such interactions are a critical part of navigating city streets, and some interesting research work has been done building machines that take them seriously [Pennycooke, 2012].

where children played, vendors sold goods, and adults walked, talked, biked, and gathered socially. This caused problems for the drivers and manufacturers of fast moving, dangerous vehicles, which all too frequently inflicted bodily harm on those people with whom they shared the environment. The regimentation of public space into crosswalks, where pedestrians are legally protected, and other parts of the roadway from which pedestrians are supposed to be excluded, is a direct outcome of an early 20th century campaign to reduce pedestrian deaths. But accidents nevertheless occurred, and the victims were primarily children and youths [Norton, 2008, p. 11]; the deaths of children came to have a new social meaning that made them particularly abhorrent. After Mary Miner's death in 1903, the streetcar driver "had a narrow escape from violence at the hands of a mob estimated by the police . . . to have been 3,000 strong" [Zelizer, 1994, p. 22]. Accidental deaths of children were an alarming problem, with a significant public response: mobs attacked the killers, acts of public mourning memorialized the lost, and a national safety campaign began to attempt to reduce these deaths [Zelizer, 1994, p. 23]. Public outrage cast automobiles as "frivolous playthings" or "pleasure cars" [Norton, 2008, p. 12], magnified by a transformation in the sentimental worth of children.

In an attempt to change this, street games were turned into criminal offenses by around 1914, but fatalities kept increasing nonetheless [Zelizer, 1994, p. 38]. The deaths of children at the hands of automobiles were not solely placed on the shoulders of drivers. As the death rate became a national crisis, the press "pinned most of the blame on parents": modern life, it was said, "cannot be retarded to enable heedless children to get out of the way" [Zelizer, 1994, p. 37]. This was not by accident.[2] The homogenization of the road for "transit" [Jain, 2004, p. 73] involved not only pressures on families to keep their children out of the street, but a concerted legal and public relations campaign for automobile-friendly traffic laws organized by automakers concerned about automotive speed governors [Stromberg, 2015]. To attempt to

---

[2] Articles ghostwritten by the National Automobile Chamber of Commerce shifted the blame for traffic accidents to pedestrians; the AAA sponsored safety campaigns in schools; and police and citizens were called upon to shame transgressors in order to set new public standards—even the name for the infraction, "jay-walking," was intended to create public opprobrium for the supposed "hicks" who did not know how to behave in cities, and to shift blame to them [Stromberg, 2015].

compensate for laws that were rarely followed and enforced, auto-friendly groups worked to change the public dialogue. Automobility was an enabling force, which provided auto companies and their supporters the impetus to shift public standards in a particular direction, and to shape the street to their own advantage.

New technologies become entwined with new social standards and legal principles, and technological narratives and epistemologies are deeply involved in this process: What is a road for? What is a vehicle's proper role? We know that a road is intended for driving precisely because we have been taught according to a social code that was designed to foster automobility—greatly responsible for the inhumanity of the modern city that urban designers are interested in reversing. How should we regard automated vehicles in our environment? Automobiles were long subject to disputes over their nature: were they fundamentally safe vehicles misused by people, or fundamentally dangerous technologies that required careful licensing and use to make safe?[3] Machine agency, even under supervision, presents the possibility that our laws and intuitions must change. And what new re-shapings of public space will happen as part of the popularization of these technologies? Will these vehicles increase the segregation of the road space, and require new lanes that are even more insulated from pedestrians? Or will faster reaction times and always-attentive automated safety features foster tighter, mixed-use environments? Which types and classes of users will the new environments and standards benefit? Will environments improve for those with access to technology, while being degraded for those without?

New epistemologies may be necessary at multiple levels. At the level of the individual, what is the role or status of the human and machine? How do we classify and make sense of users and devices, and their relationships of supervision and cooperation? At the level of the system, what new understandings of the role and purpose of the road itself are needed if we want to make use of automated technologies to make cities more humane? At the legal level, what is the status of a vehicle acting autonomously on a certain time scale? And how can responsibility be apportioned

---

[3] In one such battle over cars as "dangerous instrumentalities," a court held that "Until human agency intervenes, they are usually harmless" [Jain, 2004, p. 70]. This raises questions about vehicles that can be said to "operate themselves" for some periods of time.

between supervisors and the supervised?[4] The question of what we can actually ask the human to do, and hold them to task for, is critical both for automation system design and the legal handling of cases involving such automation. And it requires a broader recognition of our networked, cyborg nature.

5.2 Cyborgs in Traffic

Without a recognition of our technological hybridity, we are caught in a false choice: either humans should drive (blind ludditism), or machines should drive (equally blind technophilia). The idea that humans drive through machines, or machines drive via humans, or that human and machine drive in combination, are practically incomprehensible. But once we recognize that we *already* drive through machines, and that they may already drive through us,[5] multiple, multifaceted futures of automated vehicle development open to us. We are not building robotic chauffeurs, but rather designing ourselves, in some fashion, into more hybrid, more cyborg, entities. Fostering an appreciation of this is paramount, one of the great challenges for the public understanding of technology in the near future.[6]

Philosopher Andy Clark—drawing from a deep intellectual history that began in 1960 [Clynes and Kline, 1960]—contends that humans are "*natural-born* cyborgs" [Clark, 2004, p. 6]. Cyborgs are a powerful cultural force,[7] but rarely taken seriously

---

[4] Legal scholars note that torts may already be distributed across human and nonhuman actors (operator, owner, seller, manufacturer, distributors) in legal rulings [Smith, 2013]. The distribution of torts to different actors in the system may vary from situation to situation, given the capacities of the vehicle and the driver [Gurney, 2013, p. 267]. But the current liability framework has problems, including the potential to scapegoat the human being, as well as the expense and difficulty of pursuing litigation.

[5] Consider blind-spot warning lights, which are an automated system designed to affect the human and produce a response, but have no independent capacity to drive the vehicle.

[6] To return to science fiction, Phillip K. Dick said in a 1972 speech: "Someday a human being, named perhaps Fred White, may shoot a robot named Pete Something-or-Other, which has come out of a General Electrics factory, and to his surprise see it weep and bleed. And the dying robot may shoot back and, to its surprise, see a wisp of gray smoke arise from the electric pump that it supposed was Mr. White's beating heart. It would be rather a great moment of truth for both of them" [Dick, 1972]. While "Pete Something-or-Other" does not yet exist, "Fred White" is already here. Fully comprehending this is a critical task for the maintenance of cogent, productive public conversations—and policymaking—about automated vehicles and automated technologies in general.

[7] As Donna Haraway, one of the most recognized philosophers of cyborg culture, describes: "Contemporary science fiction is full of cyborgs—creatures simultaneously animal and machine, who

as a state of existence despite the proliferation of cyborg technologies [Ekbia, 2008, p. 65]. Clark's idea is particularly compelling because it need make no distinction between the analog and the digital, the material and the virtual. A cyborg nature is not an artifact of the present moment, but a fundamental component of human experience. To Clark, our "ability to enter into deep and complex relationships with nonbiological constructs, props, and aids" is what best explains our distinctive intelligence [Clark, 2004, p. 5].[8] This therefore holds not just for cars, or information and computer technologies, but for all kinds of technology. Human technological augmentation goes back into prehistory: our engagement with technological tools and artifacts—whether the tool is a smartphone, a loom, a wheel, or fire—is what makes us, fundamentally, cyborg beings.

This concept is particularly powerful and valuable because it challenges hundreds of years of assumptions about human beings and human achievement—even some that are deeply ingrained in supervisory control. We are possessed by the idea that technologies change, but that the human remains fundamentally the same. However, this assumption rests on a long held dichotomy of human and machine, and the idea that the object of interest is the purely biological human [Ekbia, 2008, p. 327, 331]. Cognitive anthropology starts to break down these divides, with the inspection of a human-machine cognitive system as a whole, but there is more to be done. Ekbia asks, what happens if we abandon the outmoded question of whether machines can reach our level of intelligence or capability, and instead ask how "humans and machines [are] mutually constituted through discursive practices" [Ekbia, 2008, p. 328]?

The problem facing AI, according to Lucy Suchman, "is less that we attribute agency to computational artifacts than that our language for talking about agency, whether for persons or artifacts, presupposes a field of discrete, self-standing entities" [Suchman, 2006, p. 263]. Instead, we should be asking how intelligent behavior

---

populate worlds ambiguously natural and crafted" [Haraway, 2006, p. 117].

[8] While Ed Hutchins looks at these tools and stresses the importance of analyzing at a system level, Clark suggests these tools "are best conceived as proper parts of the computational apparatus that constitutes our minds" [Clark, 2004, p. 6]. These conflicting views are united in their careful consideration of the properties of nonhuman artifacts, and both have something to add to nuanced examination of automation systems.

emerges in the interactions of humans and machines—involving networked, rather than individual, subjectivities—refusing to fix *a priori* the category of the human.

Questions of the appropriate role of the human being involve not only the supervisor, within the vehicle or in a remote data center, but other users in the environment. If streets must be remade in order to make certain technological configurations viable, is that remaking something the public is willing to accept? And who will benefit from it? As Dieter Zetsche, chairman of Daimler AG describes, "anyone who focuses solely on the technology has not yet grasped how autonomous driving will change our society" [Davies, 2015b]. Changes that favor the users or manufacturers of automated vehicles may not favor other users of city spaces, who may find their freedoms foreclosed upon. The history of the remaking of the city and street is part of why I hold that vehicle automation is not an independent factor to be maximized, but a variable that is firmly intertwined with the design of the whole transport system. As I have described, these vehicles sit in a much broader network of social relationships: the city and the street have changed before, and will change again, to accommodate new technologies, assuming sufficient social and economic pressure to catalyze that change.[9] The automated vehicle represents another possible nexus for change, but how the city will be reshaped is an open question.

Until vehicles construct themselves, monitor themselves, pay for themselves, and are intended only to drive themselves around—a strange world indeed—human oversight is inescapable, and complete autonomy impossible. While certain types of labor may be eliminated, and certain laborers marginalized, a system-level view of automation continues to uncover vast amounts of human labor—perhaps at the periphery of the car as an individual object, but unavoidably implicated in the day-to-day reality of vehicle operations. The method and timescale of automated operation may greatly impact public acceptance, to the point that certain types of human involvement may be required to produce marketable technologies.[10] Though this may be distasteful to

---

[9] Technologies do not alone bring about those changes, but their presence, availability, and market viability provide incentives for groups to encourage broad social and infrastructural changes.

[10] Travel sickness is another potential barrier to public acceptance. A recent University of Michigan study suggests car sickness could be a problem for automated vehicle systems, significantly reducing their convenience factor [Sivak and Schoettle, 2015a]. The authors suggest fixes, including maximiz-

some technologists invested in furthering technological solutions to social problems, it is not fundamentally a failure. It is a societal choice to which engineering solutions should be pliable.

Many examples of hybrid systems exist, and joint-cognitive systems design presents an alternative way of looking at problems, and valuing human and machine contributions to their solution, that opens the way for other futures. As we have seen, these approaches and perspectives are not those of "mere" luddites or humanists on the margins of engineering practice, but of an important segment of engineers themselves who are vital to the design and construction of successful real-world systems. The cyborg human-machine system is a very difficult engineering problem, but even the Department of Defense recognizes the value of serious attention to hybrid control for systems that are tasked with keeping us safe, whether in space, in the air, or on the ground. In applying this approach, however, it is important to keep in mind the extent to which we are re-making ourselves into managers or into machine tenders: Are we maintaining a sense of creative agency? Or becoming mere tools of larger organizations? This is especially important as we will use our vehicles every day, and cannot afford to alienate ourselves from them. A future in which our transportation system is indifferent to us is not a productive future to be lauded, but a destructive one to be avoided.

When we use factory automation, artificial intelligence, "NASA engineering," or autopilot to frame ideas about self-driving vehicles, we carry forward certain ideas about these technologies that narrow and constrain our vision, irrespective of actual historical developments. We risk thinking about autopilots as a small step away from being full vehicle operators, rather than as tools used by vehicle operators to support their needs. We risk thinking about factory automation as a teleological process toward the elimination of labor, rather than a deeply contingent process that changes the forms of certain kinds of labor into other forms, or shifts that labor in

---

ing passengers' view with large windows and forcing them to face forward—but a PRT researcher I spoke to described that his group had found minimizing or eliminating forward view made riders in automated PRT systems most comfortable. Certain levels of human involvement could be an alternative way to ameliorate travel sickness issues.

time and space. We risk thinking about NASA as an organization responsible for engineering highly automated, redundant systems, rather than an organization full of people whose jobs include constant monitoring, supervision, and interactions with the same automated systems. We may consider new automation through the lens of the artificial, of building systems to replace people, or through a hybrid lens, building systems to support humans in their tasks. Both ways of considering automated systems can often be applied to the same technological artifact, but result in very different ways of thinking.[11] Throughout, I have emphasized the contingency of current automated vehicle plans on specific ideas about the appropriate and necessary role of technology, and stressed the presence of other ways to regard vehicle automation technology that could bring about different futures. In contrast to the simplistic depictions which form the visible picture of this technology today for those not involved in its engineering—of pro-and-con, for-and-against, but all focused around a particular assumed object—I have described a range of alternative, cyborg, narratives that complicate our ideas of what automated vehicles can and will be.

5.3 Automation and Social Goals

Ultimately, from a design and policy perspective, my point is that we cannot achieve positive social effects by naïvely adding autonomy to existing vehicles. Autonomy and automation are not natural goods to be fostered wherever they can be, but technological tools and strategies for achieving particular goals. The manner and extent to which we build our vehicles to be autonomous stands to produce very different social impacts, to the point where going from the idea of the "self-driving vehicle" to social and cultural impacts is fundamentally backward, since the vehicle itself is not fixed and small details may matter a great deal in real-world use. Instead, social

---

[11] A speech-to-text system can be seen as either a technology to transcribe the human voice into text, instead of using a person, or a tool to assist a human with transcription. But the choice of perspective will likely do much to change how such a technology is designed. I would hypothesize that the focus on what computerized personal assistants can do, for example, as opposed to the process of human and machine jointly carrying out a task and accomplishing something together, does much to explain the repeated failure of personal assistant technology to gain users.

changes have to be at the center of design and implementation. What are the goals a vehicle is designed to achieve, and which goals are more important than others? Increasing statistical safety by a factor of 100? Providing mobility for the elderly or disabled? Reducing the environmental impact of vehicles and their emissions? Increasing throughput of the roadway? Reducing traffic delays for commuters? Equalizing the disparities in mobility between the wealthy and the homeless? It is fantastical thinking to believe that the just-add-autonomy approach will automatically achieve all these goals. Each presents a system-wide design problem, involving cars, people, and infrastructures, that puts different demands on possible solutions. Some goals may be directly conflicting, such as increasing mobility while reducing environmental impacts. Others contradict the motives of organizations responsible for providing the solutions: while automobile companies can promulgate ideas about how automated vehicles can provide greater mobility, help the environment, and reduce traffic, they are unlikely to economically back concepts that are against their interests—anything that reduces the number of vehicles on the road,[12] or reduces the overall cost of those vehicles, without providing a new revenue stream to compensate, represents a deeply suspect investment.

Problematically, the automated car—like the "smart city" which has already been critiqued in this way [Greenfield, 2014]—presents the opportunity for successful technology companies to wrest greater control over everyday life, to worm their way more deeply into our existence and thereby make themselves indispensable. Automation can be a tool to enable one group of people—technologists running multinational organizations—at the expense of the rest of society. But it also does not have to be employed toward those ideological ends; given sufficient willpower, it can be made to serve others. Changes to the city should not be arbitrary, at the whims of a certain set of producers and organizations, but part of a large-scale system design strategy to address the needs of all transportation users. Complex changes to complex systems require significant modeling work, experiments, or even trial and error to get right.

---

[12] Thanks to Joe Dumit for eloquently making this observation during my recent visit to UC Davis.

Old models and habits, like the rules of the AASHTO[13] "Green Book" [Swope, 2007, p. 181, 183], need to be broken, and new ones built to replace them.

Narratives that claim undue caution in the rollout of autonomous vehicles will directly "cost" lives—since people are killed by human drivers every day [Hars, 2015] [Bailey, 2014] [Howard, 2013]—may be used as a lever to argue for the moral imperative of speedy legalization and acceptance, pitted against risk-averse regulators who face public censure if but one child is killed by a "robot car." But the statistics about driverless cars and risk, though partly a function of mathematical modeling, are also, fundamentally, matters of faith. There is no precise answer, as humans and machines both err alone or in combination, and the safety of increasingly autonomous vehicles is a moving target against which more radical approaches must be compared. Results of our changes to complex, high-stakes systems may be unpredictable. So how we implement automated vehicle technology is about narratives, the stories we believe and tell about technology, as much as it is about mathematics and modeling the world. When those stories come to define major policy documents, as they have defined the NHTSA's preliminary levels of automation, we are in a dangerous position moving forward into battles over regulations and infrastructural investments that will come to define the next era of the roadway.

Heterogeneous approaches are valuable, since there is no one-size-fits-all automation solution, no magic technological fix to social and transit problems. But despite a hesitance to prescribe particular technologies, current taxonomies and policy documents normalize certain ideas about automation, while ignoring others. I am not saying that the science-fiction dreams of driverless vehicles are possible or impossible; moreover, I believe in the potential of computer technologies to achieve groundbreaking results. But our choices should be between options, between visions—how we integrate humans and computers to make our driving safer, our cities more livable, our lives better. The Google model which has captured so much press attention is one among innumerable imagined or not-yet-imagined possibilities. Automation is not a take-it-or-leave-it phenomenon, and the choice to accept one particular model

---

[13] American Association of State Highway and Transportation Officials

of vehicle automation or accept tens-of-thousands of prospective road fatalities is a false choice. There is more on heaven and earth than is dreamt of in this philosophy; and it is time we ask of autonomy not the naïve questions of whether it will keep us safe and improve our cities, but the deep, engaged, and difficult questions of how it can be leveraged to achieve such aims. The real question we need to ask is what form technologies of vehicle automation should take, and what benefits and acceptable sacrifices adhere to that choice. Answering this question is a truly multidisciplinary venture, involving engineering, law, policy, design, and ethics. When we automate, do we choose to do so with the philosophy of "do no harm" or "save more lives"?[14] There may be hidden consequences to these philosophies—Who is saved? How? Who bears the cost, or is counted out in order to save the greater number?—that come to deeply affect how systems engineered from them alter the world.

5.4 Lingering Questions

The purpose of this thesis has been to sweep away some lazy thinking about automated systems, and to try to reorient a more nuanced dialog around different sets of issues. As we consider the future of vehicle automation—whether as designers, lawmakers, interested technologists, or concerned citizens—what are the questions we should be asking? What should we grapple with instead of accepting common framings of the issues involved? To conclude, let me summarize the main open questions around which a more productive dialogue about automation could be based.

First, what are the social goals around which automated technology should be designed and implemented? The relative importance of safety, convenience, mobil-

---

[14] These are deep and subtle questions, but proclaiming them meaningless from an engineering perspective is to miss the importance of ideas to guide technological development, and the social importance of developed technologies. I am grateful to Evan Donahue in particular for a discussion of how the implications of the Hippocratic oath differ from those of a philosophy specifically oriented around saving the greatest number. These two basic premises, extrapolated out to their logical conclusions, may diverge rather than converge, in ways that remind us of the dystopic potential of quantification and optimization to lead to solutions that are fundamentally inhumane. It would be foolish to assume that medical systems, themselves a collection of technologies and people, are not affected by these kinds of distinctions. As these logics are applied elsewhere, their uncertain outcomes will remain important.

ity, social equality, and other factors must be matters of public debate, as they are absolutely central to the project of automation.[15]

Second, for any particular design, what are the appropriate roles of human and machine in the human-machine system that automated vehicles will necessarily be? These roles will not be a pure function of technological capabilities, but will involve numerous other factors, and will alter the skills that human operators will be expected to acquire, and the types of licensing and control that are therefore necessary.

Third, as part of this question, how do humans respond to AI systems that cannot simply be said to "think," "know," or "understand" in the same ways that we do (at least not yet) but nevertheless hold our lives in their hands? And how can human decision-making in the design of such systems be reasonably and successfully audited? These concerns may be ameliorated by human involvement and supervision in operations, which means we must consider what types and time-scales of supervision are socially or legally valuable, and therefore should be built into that technology.

Fourth, how should we value different automation approaches, or estimate their risks, when the technologies involved require complex changes to complex systems? This involves balancing the statistical benefits of different automation approaches, and coming to conclusions about how to reasonably estimate risks, test and certify systems, and set engineering priorities based on this information.

Fifth, relatedly, what is the appropriate role of statistics and risk projections in governing policy on automated technologies? How safe is safe enough, for whom, and how is this regulated? Assuming we can come to reasonable predictions about costs and benefits, we must close the loop between this information and our broader social goals. We need to decide how to weigh uncertainties in our evaluation of which goals matter and which are achievable in the near term or long term.

The "autonomous" vehicle is a similar sort of myth to that of the "personless" factory or the self-teaching computer program: always a product of people, responsible

---

[15] To the list of major questions facing us as a result of this technology, given the interest in full-system design, we might wish to add: how do we make transportation technology and urban planning sensitive to social justice and responsive to a broad base of citizens, and inoculate it against takeover by corporate entities that have only their own best interests in mind? But this question is larger than the autonomous vehicle arena.

to people, and involving people, however far in the margins of our vision. Writing this off as anomaly or failure, and focusing on "autonomy" or "self-driving," obscures real complexities of operation that encounters with the world will inevitably involve. The important considerations of when and how people are involved should be addressed carefully and empirically, not by knee-jerk techno-utopianism and blind ideology. There is hope in this altered narrative to ameliorate some of the issues facing automated vehicles today. Rather than "to automate or not to automate," we should ask "how and why do we automate?" Even in the most automated technologies, their autonomy is largely an artifact of the lens through which we engage with them—these systems involve delicate dances of human and machine components, dispersed through time and space, and only pushed to the margins of a point of view that takes the technological object itself, not its sociopolitical and cognitive contexts, as the object of study. Recognizing that continued human involvement will occur as a matter of principle is the most important first step we can take toward rebalancing the narrative of automation toward something more productive, which takes seriously the technological hybridity of our cyborg past, present, and future. In addition, this recognition paves the way for a political critique of automated vehicles that asks, knowing that humans remain involved, who they are, where they are, and how the system is being designed to serve their needs or impede their agency.

Bibliography

[787 Dreamliner..., 2012] 787 Dreamliner... (March 14, 2012). 787 Dreamilner vs. A380: cockpit comparisons. *Aerospace Technology.* http://www.aerospace-technology.com/features/featuredreamliner-vs-a380-cockpit-comparisons.

[A Window..., 2015] A Window... (January 5, 2015). A Window Into a Driverless Future: From Science Fiction to Reality. *Seeking Alpha.* http://seekingalpha.com/article/2798325-a-window-into-a-driverless-future-from-science-fiction-to-reality.

[Adams and Brewer, 2007] Adams, R. and Brewer, T. (2007). Roadways of the Future. In [Kemp, 2007], pages 227–232.

[Aitken, 1960] Aitken, H. G. J. (1960). *Taylorism at Watertown Arsenal: Scientific Management in Action.* Harvard University Press, Cambridge, MA.

[Arnheim, 2006] Arnheim, R. (2006). *Film as Art.* University of California Press, Berkeley, CA.

[Badger, 2015] Badger, E. (March 18, 2015). Elon Musk says we'll outlaw human drivers in a world of driverless cars. Really? *The Washington Post.* http://www.washingtonpost.com/blogs/wonkblog/wp/2015/03/18/should-we-outlaw-human-drivers-in-a-world-of-driverless-cars/.

[Bailey, 2014] Bailey, R. (June 10, 2014). Robot Cars Are a Moral Imperative. *Reason.com Hit & Run Blog.* http://reason.com/blog/2014/06/10/robot-cars-are-a-moral-imperative.

[Battin, 1989] Battin, R. H. (January 9-12, 1989). Some Funny Things Happened On the Way to the Moon. Theodore von Karman Lecture AIAA-89-0861, American Institute of Aeronautics and Astronautics.

[BEA, 2012] BEA (2012). Final report on the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight AF-447 Rio de Janiero – Paris. Crash Report, Bureau d'Enquetes et d'Analyses. English Edition.

[Bennett, 1972] Bennett, F. V. (1972). Apollo Experience Report - Mission Planning for Lunar Module Descent and Ascent. NASA Technical Note NASA TN D-6846. http://www.hq.nasa.gov/alsj/nasa-tnd-6846pt.1.pdf.

[Bigelow, 2015] Bigelow, P. (February 18, 2015). A 14-year-old hacker caught the auto industry by surprise. *AutoBlog*. http://www.autoblog.com/2015/02/18/14-year-old-hacker-caught-industry-by-surprise-featured/.

[Bilger, 2013] Bilger, B. (November 25, 2013). Auto Correct: Has the self-driving car at last arrived? *The New Yorker*. http://www.newyorker.com/magazine/2013/11/25/auto-correct.

[Bleecker, 2009] Bleecker, J. (March 17, 2009). Design Fiction: A Short Essay on Design, Science, Fact and Fiction. *Near Future Laboratory*. http://blog.nearfuturelaboratory.com/2009/03/17/design-fiction-a-short-essay-on-design-science-fact-and-fiction/.

[Boden, 1990] Boden, M. A., editor (1990). *The Philosophy of Artificial Intelligence*. Oxford University Press, Oxford, UK.

[Bogost, 2014] Bogost, I. (October 14, 2014). The Secret History of the Robot Car. *The Atlantic*. http://www.theatlantic.com/magazine/archive/2014/11/the-secret-history-of-the-robot-car/380791/.

[Both, 2014a] Both, G. (April 1, 2014a). What Drives Research in Self-Driving Cars? (Part 1: Two Major Events). *CASTAC Blog*. http://blog.castac.org/2014/04/what-drives-research-in-self-driving-cars-part-1-two-major-events/.

[Both, 2014b] Both, G. (April 3, 2014b). What Drives Research in Self-Driving Cars? (Part 2: Surprisingly, not Machine Learning). *CASTAC Blog*. http://blog.castac.org/2014/04/what-drives-research-in-self-driving-cars-part-2-surprisingly-not-machine-learning/.

[Brooks, 1987] Brooks, R. A. (1987). Autonomous Mobile Robots. In [Grimson and Patil, 1987], pages 343–362.

[Brooks, 2002] Brooks, R. A. (2002). *Flesh and Machines: How Robots Will Change Us*. Pantheon Books, New York.

[Broviak, 2007] Broviak, P. (2007). Cars, People, and Safety. In [Kemp, 2007], pages 11–14.

[Brown, 2011] Brown, D. P. (March 18, 2011). Will the Cockpit of the Future Have Less Pilot Interaction? *Aol Travel*. http://news.travel.aol.com/2011/03/18/will-the-cockpit-of-the-future-have-less-pilot-interaction/.

[Brown, 2010] Brown, D. P. (September 7, 2010). By the Numbers, Flying is a lot Safer than Driving or Taking the Train. *Airline Reporter*. http://www.airlinereporter.com/2010/09/flying-is-safe-and-i-am-going-to-prove-it/.

[Burden, 2007] Burden, D. (2007). Cities, Streets, and Cars. In [Kemp, 2007], pages 7–10.

[Carney, 2010] Carney, K. (February 9, 2010). Road Miles by State. *Cubit's Blog.* http://blog.cubitplanning.com/2010/02/road-miles-by-state/.

[Carson, 2014] Carson, E. (September 15, 2014). 8 truths and myths of driverless cars. *Tech Republic.* http://www.techrepublic.com/article/8-truths-and-myths-of-driverless-cars/.

[Cheatham and Bennett, 1966] Cheatham, D. C. and Bennett, F. V. (1966). Apollo Lunar Module Landing Strategy. Manned Spacecraft Center Report, NASA. http://ocw.mit.edu/courses/science-technology-and-society/sts-471j-engineering-apollo-the-moon-project-as-a-complex-system-spring-2007/readings/4_2_lunr_landing.pdf.

[Chow et al., 2014] Chow, S., Yortsos, S., and Meshkati, N. (2014). Asiana Airlines Flight 214: Investigating Cockpit Automation and Culture Issues in Aviation Safety. *Aviation Psychology and Applied Human Factors*, 4(2):113–121.

[Clancey, 2014] Clancey, W. J. (2014). *Working on Mars: Voyages of Scientific Discovery with the Mars Exploration Rovers.* MIT Press, Cambridge, MA.

[Clark, 2004] Clark, A. (2004). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence.* Oxford University Press, New York, NY.

[Clynes and Kline, 1960] Clynes, M. E. and Kline, N. S. (September, 1960). Cyborgs and space. *Astronautics*, pages 26–27, 74–76.

[Collier's Magazine, 1954] Collier's Magazine (1952–1954). Man Will Conquer Space Soon (series). *Collier's Magazine.* Some scans available at http://www.scribd.com/doc/118710867/Collier-s-Magazine-Man-Will-Conquer-Space.

[Collins, 1990] Collins, H. M. (1990). *Artificial Experts.* MIT Press, Cambridge, MA.

[Coopersmith, 2015] Coopersmith, J. (2015). *Faxed: The Rise and Fall of the Fax Machine.* Johns Hopkins University Press, Baltimore, MD.

[Cord and Gimonet, 2014] Cord, A. and Gimonet, N. (March, 2014). Detecting Unfocused Raindrops: In-Vehicle Multipurpose Cameras. *IEEE Robotics and Automation Magazine*, pages 49–56.

[Costello, 2007] Costello, P. (2007). San Francisco Region, Other Areas, Use Communications System to Manage Traffic. In [Kemp, 2007], pages 168–170.

[Court of Justice of the European Union, 2014] Court of Justice of the European Union (May 13, 2014). Press Release No. 70/14.

[Coxe, 1814] Coxe, T. (1814). *Statement of the Arts and Manufactures of the United States of America for the year 1810.* A. Cornman, Philadelphia, PA.

[DARPA, 2007] DARPA (November 3, 2007). Photo of the MIT Urban Challenge Vehicle. From the DARPA image gallery: archive.darpa.mil/grandchallenge/gallery.html, http://archive.darpa.mil/grandchallenge/images/photos/11_3_07/D2X_0946.jpg.

[Davies, 2015a] Davies, A. (January 7, 2015a). I Rode 500 Miles in a Self-Driving Car and Saw the Future. It's Delightfully Dull. *Wired.* http://www.wired.com/2015/01/rode-500-miles-self-driving-car-saw-future-boring/.

[Davies, 2015c] Davies, C. (March 19, 2015c). Tesla Model S Autopilot update due in 3-4 months. *SlashGear.* http://www.slashgear.com/tesla-model-s-autopilot-update-due-in-3-4-months-19374458/.

[Davies, 2015b] Davies, C. (March 20, 2015b). How Mercedes' F 015 self-driving car is shaping smart cities. *SlashGear.* http://www.slashgear.com/how-mercedes-f-015-self-driving-car-is-shaping-smart-cities-20374602/.

[Davies, 2015d] Davies, J. (February 21, 2015d). Philosopher tackles issue of whether driverless cars make ethical judgments. *The Drum.* http://m.thedrum.com/news/2015/02/21/philosopher-tackles-issue-whether-driverless-cars-make-ethical-judgements-0.

[Dawson, 2015] Dawson, F. (March 31, 2015). How Far Away Is Driverless Technology Really? *Forbes.* http://www.forbes.com/sites/freddiedawson/2015/03/31/how-far-away-is-driverless-technology-really/.

[Dennett, 1990] Dennett, D. C. (1990). Cognitive Wheels: The Frame Problem of AI. In [Boden, 1990], pages 147–170.

[Dick, 1972] Dick, P. K. (1972). The Android and the Human. Speech given at the Vancouver Science Fiction Convention, University of British Colombia. Originally published in SF Commentary #31. Full text at http://genius.com/Philip-k-dick-the-android-and-the-human-annotated#note-4386783.

[Dickmann et al., 2014] Dickmann, J., Appenrodt, N., and Brenk, C. (August, 2014). Making Bertha: Radar is the key to Mercedes-Benz's robotic car. *IEEE Spectrum*, pages 44–49.

[Dickmanns, 2007] Dickmanns, E. D. (2007). *Dynamic Vision for Perception and Control of Motion.* Springer-Verlag, London.

[Dickmanns et al., 1994] Dickmanns, E. D., Behringer, R., Dickmanns, D., Hildebrandt, T., Maurer, M., Thomanek, F., and Schiehlen, J. (1994). The Seeing

Passenger Car 'VaMoRs-P'. *Proceedings of the Intelligent Vehicles '94 Symposium*, pages 68–73.

[Diebold, 1953] Diebold, J. (1953). Automation–The New Technology. *Harvard Business Review*, 31(6):63–71.

[Diebold, 1959] Diebold, J. (1959). *Automation: Its Impact on Business and Labor*. National Planning Association, Washington, D.C.

[Driverless cars..., 2013] Driverless cars... (November 22, 2013). Driverless cars: From Hollywood to real world. *CNN*. http://edition.cnn.com/2013/11/21/business/renault-nissan-ceo-wants-driverless/index.html.

[DSB, 2012] DSB (2012). The Role of Autonomy in DoD Systems. Technical report, Department of Defense, Defense Science Board.

[Ekbia, 2008] Ekbia, H. R. (2008). *Artificial Dreams: The Quest for Non-Biological Intelligence*. Cambridge University Press, Cambridge, UK.

[EUREKA, 2014] EUREKA (Accessed December 15, 2014). Eureka project > 45 prometheus. *EUREKA Website*. http://www.eurekanetwork.org/project/-/id/45.

[Evans, 1848] Evans, O. (1848). *The Young Mill-Wright and Miller's Guide*. Lea & Blanchard, Philadelphia, PA, 12th edition.

[Experts Predict..., 1999] Experts Predict... (November 7, 1999). Experts Predict More Plane Crashes. *Orlando Sentinel*. http://articles.orlandosentinel.com/1999-11-07/news/9911070029_1_plane-aviation-safety-type-of-accident.

[Fernandez, 2007] Fernandez, J. (2007). Boulder Redesigns Residential Streets to Focus on People Rather than Cars. In [Kemp, 2007], pages 65–70.

[Fernandez-Madrigal and Claraco, 2013] Fernandez-Madrigal, J.-A. and Claraco, J. L. B. (2013). *Simultaneous Localization and Mapping for Mobile Robots: Introduction and Methods*. IGI Global, Hershey, Pennsylvania. eBook Collection (EBSCOhost), EBSCOhost (accessed February 26, 2015).

[FHWA, 2013] FHWA (2013). Table 1103. Motor Vehicle Accidents–Number and Deaths: 1990 to 2009. 2013 Status of the Nation's Highways, Bridges, and Transit: Conditions & Performance, Federal Highway Administration. Exhibit 1–5, http://www.fhwa.dot.gov/policy/2013cpr/chap1.htm#5.

[Fisher, 2013] Fisher, A. (September 18, 2013). Inside Google's Quest to Popularize Self-Driving Cars. *Popular Science*. http://www.popsci.com/cars/article/2013-09/google-self-driving-car.

[Fletcher et al., 2009] Fletcher, L., Teller, S., Olson, E., Moore, D., Kuwata, Y., How, J., Leonard, J., Miller, I., Campbell, M., Huttenlocher, D., Nathan, A., and Kline, F.-R. (2009). The MIT–Cornell Collision and Why it Happened. In *The DARPA Urban Challenge*, pages 509–548. Springer Berlin / Heidelberg.

[Flight Deck Automation Working Group, 2013] Flight Deck Automation Working Group (2013). Operational Use of Flight Path Management Systems. Technical report, Performance-Based Aviation Rulemaking Committee.

[Fried, 2015] Fried, I. (March 17, 2015). Google Self-Driving Car Chief Wants Tech on the Market Within Five Years. *re/code*. http://recode.net/2015/03/17/google-self-driving-car-chief-wants-tech-on-the-market-within-five-years/.

[Fukuda et al., 2014] Fukuda, Y., Kobayashi, K., Watanabe, K., and Kinoshita, T. (2014). Target object classification based on a fusion of LIDAR range and intensity data. *2014 Proceedings of the SICE Annual Conference*, pages 1312–1317.

[Gibson, 2000] Gibson, W. (2000). *Neuromancer*. Penguin, New York.

[Gomes, 2014b] Gomes, L. (August 28, 2014b). Hidden Obstacles for Google's Self-Driving Cars. *MIT Technology Review*. http://www.technologyreview.com/news/530276/hidden-obstacles-for-googles-self-driving-cars/.

[Gomes, 2014c] Gomes, L. (October 20, 2014c). Machine-Learning Maestro Michael Jordan on the Delusions of Big Data and Other Huge Engineering Efforts. *IEEE Spectrum*. http://spectrum.ieee.org/robotics/artificial-intelligence/machinelearning-maestro-michael-jordan-on-the-delusions-of-big-data-and-other-huge-engineering-efforts.

[Gomes, 2014a] Gomes, L. (October 27, 2014a). Driving in Circles: The autonomous Google car may never actually happen. *Slate*. http://www.slate.com/articles/technology/technology/2014/10/google_self _driving_car_it_may_never_actually_happen.single.html.

[Goodman, 2015] Goodman, M. (February 24, 2015). The time a hacker remotely bricked cars in Texas. *BoingBoing*. http://boingboing.net/2015/02/24/the-time-a-hacker-remotely-bri.html.

[Green, 1925] Green, H. (November, 1925). Radio-Controlled Automobile. *Radio News*, pages 592, 656. http://www.americanradiohistory.com/Archive-Radio-News/20s/Radio-News-1925-11-R.pdf.

[Green, 2015] Green, M. (March 25, 2015). Distraction and Teen Crashes: Even Worse than We Thought. *AAA NewsRoom*. http://newsroom.aaa.com/2015/03/distraction-teen-crashes-even-worse-thought/.

[Greenfield, 2014] Greenfield, A. (December 22, 2014). The smartest cities rely on citizen cunning and unglamorous technology. *The Guardian.* http://www.theguardian.com/cities/2014/dec/22/the-smartest-cities-rely-on-citizen-cunning-and-unglamorous-technology.

[Grimson and Patil, 1987] Grimson, W. E. L. and Patil, R. S., editors (1987). *AI in the 1980s and Beyond.* MIT Press, Cambridge, MA.

[Guizzo, 2011] Guizzo, E. (October 18, 2011). How Google's Self-Driving Car Works. *IEEE Spectrum.* http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/how-google-self-driving-car-works.

[Gurney, 2013] Gurney, J. K. (2013). Sue my car not me: Products liability and accidents involving autonomous vehicles. *U. Ill. J. L. Tech & Pol'y*, (247). Available at SSRN: http://ssrn.com/abstract=2352108.

[Haraway, 2006] Haraway, D. (2006). A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late 20th Century. In Weiss, J., Nolan, J., Hunsinger, J., and Trifonas, P., editors, *The International Handbook of Virtual Learning Environments*, pages 117–158. Springer.

[Harris, 2014a] Harris, D. (2014a). Improving aircraft safety. *Psychologist*, 27(2):90–94.

[Harris, 2012] Harris, D. (March 26, 2012). The one-night stand, quantified and visualized by Uber. *Gigaom.* https://gigaom.com/2012/03/26/uber-one-night-stands/.

[Harris, 2014b] Harris, M. (August 21, 2014b). Google lobbies to test self-driving cars in Matrix-style virtual world. *The Guardian.* http://www.theguardian.com/technology/2014/aug/21/google-test-self-driving-cars-virtual-world-matrix.

[Harris, 2014c] Harris, M. (September 10, 2014c). How Google's Autonomous Car Passed the First U.S. State Self-Driving Test. *IEEE Spectrum.* http://spectrum.ieee.org/transportation/advanced-cars/how-googles-autonomous-car-passed-the-first-us-state-selfdriving-test/.

[Hars, 2015] Hars, A. (2015). Accident rates of self-driving cars: A critique of the Sivak/Schoettle study. *Driverless Car Market Watch.* http://www.driverless-future.com/?p=720.

[Higgins, 2013] Higgins, P. (November 13, 2013). DRM in Cars Will Drive Consumers Crazy. *Electronic Frontier Foundation.* https://www.eff.org/deeplinks/2013/11/drm-cars-will-drive-consumers-crazy.

[Hills, 1996] Hills, R. L. (1996). *Power From the Wind: A History of Windmill Technology.* Cambridge University Press, Cambridge, UK.

[HireArt, 2015] HireArt (Accessed May 3, 2015). Self-Driving Vehicle Operator / Operations Associate (job posting). https://www.hireart.com/jobs/6ba4ebe8/apply.

[Hoder, 2013] Hoder, R. (November 14, 2013). For More Parents, It's Uber to the Rescue. *Time*. http://healthland.time.com/2013/11/14/for-more-parents-its-uber-to-the-rescue/.

[Hounshell, 1984] Hounshell, D. A. (1984). *From the American System to Mass Production*. Johns Hopkins University Press, Baltimore, MD.

[Howard, 2013] Howard, D. (November 7, 2013). Robots on the Road: The Moral Imperative of the Driverless Car. *University of Notre Dame Blog Network: Science Matters*. http://donhoward-blog.nd.edu/2013/11/07/robots-on-the-road-the-moral-imperative-of-the-driverless-car/.

[Huang, 2007] Huang, H.-M. (2007). Autonomy Levels for Unmanned Systems (ALFUS) Framework Volume II: Framework Models Version 1.0. Special Publication, NIST. http://www.nist.gov/customcf/get_pdf.cfm?pub_id=823618.

[Hutchins, 1995] Hutchins, E. (1995). How a Cockpit Remembers Its Speeds. *Cognitive Science*, 19:265–288.

[Hutchins and Klausen, 1996] Hutchins, E. and Klausen, T. (1996). Distributed Cognition in an Airline Cockpit. In Engstrom, Y. and Middleton, D., editors, *Cognition and Communication at Work*, pages 15–34. Cambridge Universty Press, Cambridge, UK. Text available at http://hci.ucsd.edu/media/uploads/hci_papers/EH1996-1.pdf.

[ICO, 2014] ICO (2014). Summary of Costeja, Information Commissioner's Office. https://iconewsblog.files.wordpress.com/2014/05/key-points-of-cjeu-case.pdf.

[In 50 Years ..., 1959] In 50 Years ... (April 26, 1959). In 50 Years: Cars Flying Like Missiles! *Chicago Daily Tribune*. Via Paleofuture: http://paleofuture.com/blog/2007/4/28/in-50-years-cars-flying-like-missiles-chicago-daily-tribune.html.

[Jain, 2004] Jain, S. S. L. (2004). 'Dangerous Instrumentality': The Bystander as Subject in Automobility. *Cultural Anthropology*, 19(1):61–94.

[Jardin, 2015] Jardin, X. (January 15, 2015). In an effort to suck up to local governments, Uber plans to share your ride data. *Boing Boing*. http://boingboing.net/2015/01/15/in-an-effort-to-suck-up-to-loc.html.

[Jenkins, 2000] Jenkins, D. R. (2000). Hypersonics Before the Shuttle: A Concise History of the X-15 Research Airplane. Monographs in Aerospace History SP-2000-4518, NASA.

[Jordan and Bishop, 2004] Jordan, M. I. and Bishop, C. M. (2004). Neural Networks. In Tucker, A. B., editor, *Computer Science Handbook*, section 7. Chapman & Hall/CRC Press LLC, Boca Raton, FL, 2nd edition.

[Joyce, 2008] Joyce, K. A. (2008). *Magnetic Appeal: MRI and the Myth of Transparency.* Cornell University Press, Ithaca, NY.

[Kaber et al., 2001] Kaber, D. B., Riley, J. M., Tan, K.-W., and Endsley, M. R. (2001). On the Design of Adaptive Automation for Complex Systems. *International Journal of Cognitive Ergonomics*, 5(1):37–57.

[Kang, 2011] Kang, M. (2011). *Sublime Dreams of Living Machines: The Automaton In the European Imagination.* Harvard University Press, Cambridge, MA.

[Kauffman, 1994] Kauffman, J. L. (1994). *Selling Outer Space: Kennedy, the Media, and Funding for Project Apollo, 1961-1963.* University of Alabama Press, Tuscaloosa, AL.

[Kayyali, 2014] Kayyali, N. (February 18, 2014). EFF Submits Letter Opposing Oakland's Domain Awareness Center. *Electronic Frontier Foundation.* https://www.eff.org/deeplinks/2014/02/eff-submits-letter-opposing-oaklands-domain-awareness-center.

[Kemp, 2007] Kemp, R. L., editor (2007). *Cities and Cars: A Handbook of Best Practices.* McFarland & Company, Jefferson, NC.

[Kennedy, 1961] Kennedy, J. F. (May 25, 1961). Special Message to the Congress on Urgent National Needs. Via The American Presidency Project, http://www.presidency.ucsb.edu/ws/?pid=8151.

[Kennedy, 1962] Kennedy, J. F. (November 21, 1962). Transcript of Presidential Meeting in the Cabinet Room of the White House: Supplemental appropriations for the National Aeronautics and Space Administration. http://history.nasa.gov/JFK-Webbconv/pages/transcript.pdf.

[Kessler, 2015] Kessler, A. M. (March 19, 2015). Elon Musk Says Self-Driving Tesla Cars Will Be in the U.S. by Summer. *The New York Times.* http://mobile.nytimes.com/2015/03/20/business/elon-musk-says-self-driving-tesla-cars-will-be-in-the-us-by-summer.html.

[King, 2015] King, L. (March 19, 2015). Google's Gargantuan Push For Cars With No Steering Wheel By 2020. *Forbes.* http://www.forbes.com/sites/leoking/2015/03/19/googles-gargantuan-push-for-cars-with-no-steering-wheel-by-2020/.

[Kirby, 2010] Kirby, D. (2010). The Future is Now: Diegetic Prototypes and the Role of Popular Films in Generating Real-world Technological Development. *Social Studies of Science*, 40(1):41–70.

[Klumpp, 1968] Klumpp, A. R. (1968). A Manually Retargeted Automatic Landing System for the Lunar Module (LM). *Journal of Spacecraft and Rockets*, 5(2).

[Knapton, 2015] Knapton, S. (February 13, 2015). Driverless car beats racing driver for first time. *Telegraph*. http://www.telegraph.co.uk/news/science/science-news/11410261/Driverless-car-beats-racing-driver-for-first-time.html.

[Knight, 2013] Knight, W. (October 22, 2013). Driverless Cars Are Further Away Than You Think. *MIT Technology Review*. http://www.technologyreview.com/featuredstory/520431/driverless-cars-are-further-away-than-you-think/.

[Knox, 2015] Knox, J. (Accessed April 25, 2015). Real-time maps are game changers. *TOMTOM - Connected Car Report*. http://www.connectedcar.org.uk/tomtom/4584538397.

[Larsson, 2012] Larsson, A. F. L. (2012). Driver usage and understanding of adaptive cruise control. *Applied Ergonomics*, 43(3):501–506.

[Latour, 1988] Latour, B. (1988). *Science in Action: How to Follow Scientists and Engineers Through Society*. Harvard University Press, Cambridge, MA.

[Latour, 1996] Latour, B. (1996). *Aramis, or the Love of Technology*. Harvard University Press, Cambridge, MA.

[Lazarevski, 2014] Lazarevski, M. (June 14, 2014). 'Suicidal' pedestrians (photo). *flickr*. https://www.flickr.com/photos/mlazarevski/14239886699/.

[Lee, 2014] Lee, T. (May 16, 2014). Nokia's HERE Maps Powers Self-Driving Mercedes S-Class Vehicle. *Ubergizmo*. http://www.ubergizmo.com/2014/05/nokias-here-maps-powers-self-driving-mercedes-s-class-vehicle/.

[Lehrer, 2011] Lehrer, J. (August 11, 2011). The Rhetoric of Neuroscience. *Wired*. http://www.wired.com/2011/08/the-rhetoric-of-neuroscience/.

[Lenat, 1997] Lenat, D. (1997). From *2001* to 2001: Common Sense and the Mind of HAL. In Stork, D. G., editor, *HAL's Legacy: 2001's Computer As Dream and Reality*, pages 193–210. MIT Press, Cambridge, MA.

[Levine, 1982] Levine, A. S. (1982). Managing NASA in the Apollo Era. The NASA History Series SP-4012, NASA. http://history.nasa.gov/SP-4102/sp4102.htm.

[Light Beams ..., 1936] Light Beams ... (April, 1936). Light Beams Steer Super Racing Cars. *Modern Mechanix*. http://blog.modernmechanix.com/light-beams-steer-super-racing-cars/.

[Lin, 2012] Lin, P. (2012). Introduction. In [Lin et al., 2012], pages 3–15.

[Lin et al., 2012] Lin, P., Abney, K., and Bekey, G. A., editors (2012). *Robot Ethics: The Ethical and Social Implications of Robotics*. MIT Press, Cambridge, MA.

[Liu et al., 2012] Liu, D., Guarino, S. L., Roth, E., Harper, K., and Vincenzi, D. (2012). Effect of novel adaptive displays on pilot performance and workload. *International Journal of Aviation Psychology*, 22(3):242–265.

[Look, no hands..., 2013] Look, no hands... (April 20, 2013). Look, no hands: One day every car may come with an invisible chauffeur. *The Economist*. http://www.economist.com/news/special-report/21576224-one-day-every-car-may-come-invisible-chauffeur-look-no-hands/.

[Lowe, 2007] Lowe, M. D. (2007). Cars, Their Problems, and the Future. In [Kemp, 2007], pages 221–226.

[Lytton, 2014] Lytton, C. (November 27, 2014). Do driverless cars need to make their passengers feel like drivers? *CNN*. http://www.cnn.com/2014/11/27/tech/innovation/gigaom-storify-future-of-transport/index.html.

[Maass, 2014] Maass, D. (October 15, 2014). Civil Liberties Groups Appeal Ruling Over Automatic License Plate Reader Data. *Electronic Frontier Foundation*. https://www.eff.org/deeplinks/2014/10/civil-liberties-groups-appeal-ruling-over-automatic-license-plate-reader-data.

[MacKenzie, 2002] MacKenzie, D. (2002). *Inventing Accuracy: A Historical Sociology of Nuclear Missile Guidance*. Pantheon Books, New York.

[Madrigal, 2014] Madrigal, A. (September 8, 2014). When Cars Are as Hackable as Cell Phones. *The Atlantic*. http://www.theatlantic.com/technology/archive/2014/09/when-cars-are-as-hackable-as-cell-phones/379734/.

[Markoff, 2014a] Markoff, J. (May 27, 2014a). Google's Next Phase in Driverless Cars: No Steering Wheel or Brake Pedals. *The New York Times*. http://www.nytimes.com/2014/05/28/technology/googles-next-phase-in-driverless-cars-no-brakes-or-steering-wheel.html.

[Markoff, 2014b] Markoff, J. (November 17, 2014b). Researchers Announce Advance in Image-Recognition Software. *New York Times*. http://www.nytimes.com/2014/11/18/science/researchers-announce-breakthrough-in-content-recognition-software.html?_r=0.

[Marshall, 2007] Marshall, A. (2007). The Future of Transportation. In [Kemp, 2007], pages 215–219.

[Martello, 2010] Martello, R. (2010). *Midnight Ride, Industrial Dawn: Paul Revere and the Growth of American Enterprise*. Johns Hopkins University Press, Baltimore, MD.

[Martinez-Vela, 2007] Martinez-Vela, C. (2007). The Duality of Innovation: Implications for the Role of the University in Economic Development. Technical report, Doctoral Thesis MIT Technology, Management and Policy.

[Mayr, 1971] Mayr, O. (1971). Maxwell and the Origins of Cybernetics. *Isis*, 4(62):424–444.

[McCarthy et al., 1955] McCarthy, J., Minksy, M. L., Rochester, N., and Shannon, C. E. (August 31, 1955). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. From the Norbert Wiener Papers, MC 22. Box 31: Folder 798, Institute Archives and Special Collections, MIT Libraries, Cambridge, MA.

[McCorduck, 2004] McCorduck, P. (2004). *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. A K Peters via Books24x7.

[McFarland, 2015] McFarland, M. (March 16, 2015). Driverless cars: A tremendous innovation with a glaring Achilles' heel. *The Washington Post*. http://www.washingtonpost.com/blogs/innovations/wp/2015/03/16/driverless-cars-a-tremendous-innovation-with-a-glaring-achilles-heel/.

[Melanson, 2012] Melanson, D. (June 21, 2012). Minority Report at 10: A look at technology from today to 2054. *Engadget*. http://www.engadget.com/2012/06/21/minority-report-at-10-a-look-at-technology-from-today-to-2054/.

[Merrill, 2014] Merrill, B. (December 10, 2014). How Self-Driving Cars Will Change Transportation Forever. *MakeUseOf*. http://www.makeuseof.com/tag/self-driving-cars-will-change-transportation-forever/.

[Miara, 2007] Miara, J. (2007). Boston Uses Public Transit Lines to Stimulate Inner-City Development. In [Kemp, 2007], pages 59–64.

[Michon, 1985] Michon, J. A. (1985). A critical view of driver behavior models: what do we know, what should we do? In Evans, L. and Schwing, R. C., editors, *Human Behavior and Traffic Safety*, pages 485–502. Plenum Press, New York.

[Migliore, 2014] Migliore, G. (August 1, 2014). Women more skeptical about autonomous cars than men, research says. *AutoBlog*. http://www.autoblog.com/2014/08/01/women-skeptical-self-driving-cars-more-than-men/.

[Mindell, 2011] Mindell, D. (2011). *Digital Apollo: Human and Machine in Spaceflight*. MIT Press, Cambridge, MA.

[Modern Mechanix, 1936] Modern Mechanix (April, 1936). Front cover. Accessed December 15, 2014: http://www.coverbrowser.com/covers/modern-mechanix/2.

[Murray and Cox, 1989] Murray, C. and Cox, C. B. (1989). *Apollo: The Race to the Moon*. Simon and Schuster, New York, NY.

[Nader, 1965] Nader, R. (1965). *Unsafe at Any Speed.* Grossman, New York, NY.

[National Research Council, 2014] National Research Council (2014). *Autonomy Research for Civil Aviation: Toward a New Era of Flight.* The National Academies Press, Washington, DC.

[Naughton, 2015] Naughton, J. (February 21, 2015). Google and tech's elite are living in a parallel universe. *The Guardian.* http://www.theguardian.com/commentisfree/2015/feb/22/google-tech-elite-living-in-a-parallel-universe-john-naughton.

[Newell and Simon, 1990] Newell, A. and Simon, H. A. (1990). Computer Science as Empirical Enquiry: Symbols and Search. In [Boden, 1990], pages 105–132.

[Newman, 2014] Newman, J. (June 6, 2014). How to Make Driverless Cars Behave. *Time.* http://time.com/2837472/driverless-cars-ethics-morality/.

[NHTSA, 2008] NHTSA (2008). National Motor Vehicle Crash Causation Survey: Report to Congress. DOT HS 811 059, National Highway Traffic Safety Administration.

[NHTSA, 2014] NHTSA (2014). Preliminary Statement of Policy Concerning Automated Vehicles. Policy statement, National Highway Traffic Safety Administration. http://www.nhtsa.gov/staticfiles/rulemaking/pdf/Automated_Vehicles_Policy.pdf.

[Nissan GT-R..., 2015] Nissan GT-R... (Accessed April 16, 2015). Nissan GT-R Review. *Edmunds.* http://www.edmunds.com/nissan/gt-r/.

[Nissenbaum, 2010] Nissenbaum, H. (2010). *Privacy in Context: Technology, Privacy, and the Integrity of Social Life.* Stanford Law Books, Stanford, CA.

[NOAA, 2015a] NOAA (Accessed April 28, 2015a). Alvin. *NOAA Website.* http://oceanexplorer.noaa.gov/technology/subs/alvin/alvin.html.

[NOAA, 2015b] NOAA (Accessed April 28, 2015b). Jason. *NOAA Website.* http://oceanexplorer.noaa.gov/technology/subs/jason/jason.html.

[Noble, 1984] Noble, D. F. (1984). *Forces of Production: A Social History of Industrial Automation.* Knopf, New York, NY.

[Norman, 1993] Norman, D. (1993). *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine.* Addison-Wesley Pub. Co., Reading, MA.

[Norton, 2008] Norton, P. D. (2008). *Fighting Traffic: The Dawn of the Motor Age in the American City.* MIT Press, Cambridge, MA.

[Novak, 2010] Novak, M. (December 9, 2010). Driverless Car of the Future (1957). *Paleofuture.* http://paleofuture.com/blog/2010/12/9/driverless-car-of-the-future-1957.html.

[Novak, 2011] Novak, M. (February 22, 2011). Flying Carpet Car (1958). *Paleofuture.* http://paleofuture.com/blog/2011/2/22/flying-carpet-car-1958-1.html.

[Novak, 2013] Novak, M. (May 16, 2013). The National Automated Highway System That Almost Was. *Smithsonian.* http://www.smithsonianmag.com/history/the-national-automated-highway-system-that-almost-was-63027245/.

[Novak, 2012] Novak, M. (October 9, 2012). Recapping 'The Jetsons': Episode 03 – The Space Car. *Smithsonian.* http://www.smithsonianmag.com/history/recapping-the-jetsons-episode-03-the-space-car-67174086/.

[Nye, 1990] Nye, D. E. (1990). *Electrifying America: Social Meanings of a New Technology, 1880-1940.* MIT Press, Cambridge, MA.

[Nye, 2013] Nye, D. E. (2013). *America's Assembly Line.* MIT Press, Cambridge, MA.

[O'Brien, 2014] O'Brien, C. (November 24, 2014). Uber removed blog post from data science team that examined link between prostitution and rides. *VentureBeat.* http://venturebeat.com/2014/11/24/uber-removed-blog-post-from-data-science-team-that-examined-link-between-prostitution-and-rides/.

[O'Connell and Myers, 1966] O'Connell, J. and Myers, A. (1966). *Safety Last: An Indictment of the Auto Industry.* Random House, New York, NY.

[OnStar, 2014] OnStar (Accessed October 19, 2014). Home. *OnStar Website.* https://www.onstar.com/us/en/home.html.

[Ornstein et al., 1984] Ornstein, S. M., Smith, B. C., and Suchman, L. A. (1984). Strategic computing. *Bulletin of Atomic Scientists*, pages 11–15.

[O'Rourke, 2010] O'Rourke, M. (2010). Every Gallon Counts: Save Gas With Hypermiling. *Mother Earth News*, (240):16–17.

[Parasuraman et al., 2000] Parasuraman, R., Sheridan, T. B., and Wickens, C. D. (2000). A Model for Types and Levels of Human Interaction with Automation. *IEEE Transactions on Systems, Man and Cybernetics*, A 30(3):286–297.

[Parasuraman and Wickens, 2008] Parasuraman, R. and Wickens, C. D. (2008). Humans: Still Vital After All These Years of Automation. *Human Factors*, 50(3):511–520.

[Pasdirtz, 2015] Pasdirtz, G. (January 31, 2015). A Solution to Driverless Car Problems. *George Pasdirtz Random Noise Blog.* http://georgepasdirtz.blogspot.com/2015/01/a-solution-to-driverless-car-problems.html.

[Pennycooke, 2012] Pennycooke, N. (2012). AEVITA: Designing Biomimetic Vehicle-to-Pedestrian Communication Protocols for Autonomously Operating & Parking On-Road Electric Vehicles. Master's thesis, MIT Media Arts and Sciences.

[Pepitone, 2014] Pepitone, J. (February 11, 2014). Jetsons Age Beckons, But Opinion Split on Driverless Car, Drones. *NBC News.* http://www.nbcnews.com/tech/innovation/jetsons-age-beckons-opinion-split-driverless-car-drones-n27221.

[Peseri, 2013] Peseri, A. (August 25, 2013). The Google Driverless Car: A Cool Thing that Matters. *Humanity Centered Robotics Initiative.* http://hcri.brown.edu/2013/08/25/the-google-driverless-car-a-cool-thing-that-matters/.

[Petersen, 2014] Petersen, B. (January 19, 2014). Just around the bend: A brave new driverless world. *CBS News.* http://www.cbsnews.com/news/just-around-the-bend-a-brave-new-driverless-world/.

['Phantom Auto' to Be ..., 1932] 'Phantom Auto' to Be ... (June 17, 1932). 'Phantom Auto' to Be Operated Here. *The Free-Lance Star.* Via the Google News Archive.

['Phantom Auto' will ..., 1926] 'Phantom Auto' will ... (December 8, 1926). 'Phantom Auto' will tour city. *The Milwaukee Sentinel.* Via the Google News Archive.

[Pollock, 1989] Pollock, J. L. (1989). *How to Build a Person: A Prolegomenon.* MIT Press, Cambridge, MA.

[Poulsen, 2010] Poulsen, K. (March 17, 2010). Hacker Disables More Than 100 Cars Remotely. *Wired.* http://www.wired.com/2010/03/hacker-bricks-cars/.

[Proud et al., 2003] Proud, R. W., Hart, J. J., and Mrozinski, R. B. (2003). Methods for Determining the Level of Autonomy to Design into a Human Spaceflight Vehicle: A Function Specific Approach. Technical report, NASA. http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20100017272.pdf.

[Ramsey, 2015] Ramsey, M. (April 10, 2015). When Red Means Go, What's An Autonomous Car to Do? *The Wall Street Journal.* http://blogs.wsj.com/digits/2015/04/10/when-red-means-go-whats-an-autonomous-car-to-do/.

[Reed, 2014] Reed, B. (March 14, 2014). Bill Gates: Yes, robots really are about to take your jobs. *BGR.* http://bgr.com/2014/03/14/bill-gates-interview-robots/.

[Riofrio, 2015] Riofrio, M. (January 6, 2015). Mercedes-Benz's F015 concept is a self-driving, hydrogen-powered living room. *PCWorld.* http://www.pcworld.com/article/2865478/mercedes-benzs-f015-concept-is-a-self-driving-hydrogen-powered-living-room.html.

[Riper, 2013] Riper, A. B. V. (2013). Imagining Tomorrow: Why the Technological Future We Imagine Is Seldom the One We Get. *Synesis: A Journal of Science, Technology, Ethics and Policy*, 4:T:21–T:28.

[Riskin, 2003] Riskin, J. (2003). The Defecating Duck, or, the Ambiguous Origins of Artificial Life. *Critical Inquiry*, 29(4):599–633.

[Robot Vehicles, 2014] Robot Vehicles (Accessed December 11, 2014). Robot Vehicles. *RobotWorx.* http://www.robots.com/articles/viewing/robot-vehicles.

[Roland and Shiman, 2002] Roland, A. and Shiman, P. (2002). *Strategic Computing: DARPA and the Quest for Machine Intelligence, 1983–1993.* MIT Press, Cambridge, MA.

[Rosenblueth et al., 1943] Rosenblueth, A., Wiener, N., and Bigelow, J. (1943). Behavior, Purpose and Teleology. *Philosophy of Science*, 10:18–24.

[Ross, 2014b] Ross, P. E. (April 25, 2014b). Self-driving Cars You Can Trust. *IEEE Spectrum.* http://spectrum.ieee.org/cars-that-think/transportation/human-factors/selfdriving-cars-need-to-inspire-trust.

[Ross, 2015] Ross, P. E. (January 20, 2015). Has Robocar Safety Been Hyped? *IEEE Spectrum.* http://spectrum.ieee.org/cars-that-think/transportation/safety/has-robocar-safety-been-hyped.

[Ross, 2014a] Ross, P. E. (May 29, 2014a). Driverless Cars: Optional by 2024, Mandatory by 2044. *IEEE Spectrum.* http://spectrum.ieee.org/transportation/advanced-cars/driverless-cars-optional-by-2024-mandatory-by-2044.

[Rowsome, 1958] Rowsome, Jr., F. (April, 1958). 1958 Imperial: What It's Like To Drive An Auto-Pilot Car. *Popular Science.* Available at http://www.imperialclub.com/ imperialclub/Articles/58AutoPilot/index.htm.

[S60 Advertisement, 2015] S60 Advertisement (Accessed January 28, 2015). Volvo Cars. https://www.youtube.com/watch?v=UdTQfegCxF8.

[SAE, 2014] SAE (2014). Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. Surface Vehicle Information Report J3016, SAE International.

[Schiehlen and Dickmanns, 1994] Schiehlen, J. and Dickmanns, E. D. (1994). A Camera Platform for Intelligent Vehicles. *Proceedings of the Intelligent Vehicles '94 Symposium*, pages 393–398.

[Science..., 1925] Science... (August 10, 1925). Science: Radio Auto. *Time Magazine.* http://content.time.com/time/magazine/article/0,9171,720720,00.html.

[Scott and Jurek, 2014] Scott, D. M. and Jurek, R. (2014). *Marketing the Moon: The Selling of the Apollo Lunar Program.* MIT Press, Cambridge, MA.

[Self-Driving Cars Moving..., 2014] Self-Driving Cars Moving... (January, 2014). Self-Driving Cars Moving into the Industry's Driver's Seat - IHS Automotive Study. *Automotive Industries*, pages 507–508.

[Shapiro, 2013] Shapiro, B. (September 25, 2013). Mom's Van Is Called Uber. *The New York Times*. http://www.nytimes.com/2013/09/26/fashion/moms-van-is-called-uber.html?_r=1.

[Sheridan, 1992] Sheridan, T. (1992). *Telerobotics, Automation, and Human Supervisory Control*. MIT Press, Cambridge, MA, 2nd edition.

[Shontell, 2015] Shontell, A. (March 15, 2015). Top Silicon Valley investor explains why he is skeptical about driverless cars. *Business Insider*. http://www.businessinsider.com/bill-gurley-is-skeptical-of-driverless-cars-2015-3.

[Simonite, 2014] Simonite, T. (May 30, 2014). Lazy Humans Shaped Google's New Autonomous Car. *MIT Technology Review*. http://www.technologyreview.com/news/527756/lazy-humans-shaped-googles-new-autonomous-car/.

[Simpson, 2015] Simpson, J. (January 12, 2015). Driver-less car design: Are we sleep-walking into the future? *Pocket-Lint*. http://www.pocket-lint.com/news/132343-driver-less-car-design-are-we-sleep-walking-into-the-future.

[Sivak and Schoettle, 2015a] Sivak, M. and Schoettle, B. (April, 2015a). Motion Sickness in Self-Driving Vehicles. Technical Report UMTRI-2015-12, The University of Michigan Transporatation Research Institute.

[Sivak and Schoettle, 2015b] Sivak, M. and Schoettle, B. (January, 2015b). Road Safety with Self-Driving Vehicles: General Limitations and Road Sharing with Conventional Vehicles. Technical Report UMTRI-2015-2, The University of Michigan Transporatation Research Institute.

[Smith and Anderson, 2014] Smith, A. and Anderson, J. (August 6, 2014). Views from Those Who Expect AI and Robotics to Have a Positive or Neutral Impact on Jobs by 2025. *Pew Research Center*. http://www.pewinternet.org/2014/08/06/views-from-those-who-expect-ai-and-robotics-to-have-a-positive-or-neutral-impact-on-jobs-by-2025/.

[Smith, 2013] Smith, B. W. (2013). Proximity driven liability. *Georgetown Law Journal*, (1777). Available at SSRN: http://ssrn.com/abstract=2336234.

[Smith, 1983] Smith, M. L. (1983). Selling the Moon: The U.S. Manned Space Program and the Triumph of Commodity Scientism. In Fox, R. W. and Lears, T. J. J., editors, *The Culture of Consumption: Critical Essays in American History, 1880-1980*, chapter 6, pages 175–236. Pantheon, New York.

[Smith, 1977] Smith, M. R. (1977). *Harper's Ferry Armory and the New Technology*. Cornell University Press, Ithaca, NY.

[Smith, 2014] Smith, M. R. (2014). Yankee Armorers and the Union War Machine. (Author's draft, July 1, 2014.) To be published in Hacker, B., editor (2016). *Astride Two Ages: Technology and the Civil War*. Smithsonian Institution Press.

[Sofge, 2015] Sofge, E. (January 30, 2015). Bill Gates Fears A.I., but A.I. Researchers Know Better. *Popular Science*. http://www.popsci.com/bill-gates-fears-ai-ai-researchers-know-better.

[Speck, 2014] Speck, J. (October 6, 2014). Why 12-Foot Traffic Lanes Are Disastrous for Safety and Must Be Replaced Now. *The Atlantic, CityLab*. http://www.citylab.com/design/2014/10/why-12-foot-traffic-lanes-are-disastrous-for-safety-and-must-be-replaced-now/381117/.

[Spieser et al., 2014] Spieser, K., Treleaven, K., Zhang, R., Frazzoli, E., Morton, D., and Pavone, M. (2014). Toward a Systematic Approach to the Design and Evaluation of Automated Mobility-on-Demand Systems. In *Road Vehicle Automation*, Lecture Notes in Mobility. Springer International Publishing. Via MIT dSpace.

[Stromberg, 2015] Stromberg, J. (January 15, 2015). The forgotten history of how automakers invented the crime of 'jaywalking'. *Vox*. http://www.vox.com/2015/1/15/7551873/jaywalking-history.

[Suchman, 1987] Suchman, L. (1987). *Plans and Situated Actions: The Problem of Human Machine Communication*. Cambridge University Press, Cambridge, UK.

[Suchman, 2006] Suchman, L. (2006). *Human Machine Reconfigurations: Plans and Situated Actions*. Cambridge University Press, Cambridge, UK.

[Swope, 2007] Swope, C. (2007). Trenton Redesigns Its Freeways to Slow Down Traffic. In [Kemp, 2007], pages 181–185.

[Table 1103, 2012] Table 1103 (2012). Table 1103. Motor Vehicle Accidents–Number and Deaths: 1990 to 2009. Statistical Abstract of the United States 2012, U.S. Census Bureau. Transportation, p. 693. https://www.census.gov/compendia/statab/2012/tables/12s1103.pdf.

[Taylor, 2015] Taylor, E. (March 27, 2015). Testing of software adds to urgency in race for driverless cars. *Reuters*. http://mobile.reuters.com/article/idUSKBN0MN1E820150327?irpc=932.

[Templeton, 2014] Templeton, G. (September 15, 2014). Google's autonomous car gets a 'B' in driving test: Not great, but better than most of us. *ExtremeTech*. http://www.extremetech.com/extreme/187438-googles-autonomous-car-gets-a-b-in-driving-test-not-great-but-better-than-most-of-us.

[Thrun et al., 2005] Thrun, S., Burgard, W., and Fox, D. (2005). *Probabilistic Robotics*. MIT Press, Cambridge, MA.

[Tindall, 1966] Tindall, B. (1966). Spacecraft computer program requirements for AS-207/208. Memo from FM1/Assistant Chief, Mission Planning and Analysis Division 66-FM1-59. http://tindallgrams.net/66-FM1-59.

[Tovey, 2015] Tovey, A. (March 26, 2015). Driverless cars to create 320,000 UK jobs and save 2,500 lives. *The Telegraph.* http://www.telegraph.co.uk/finance/newsbysector/industry/engineering/11495706/Driverless-cars-to-create-320000-UK-jobs-and-save-2500-lives.htm.

[Ulmer, 1994a] Ulmer, B. (1994a). VITA - An Autonomous Road Vehicle (ARV) for Collision Avoidance in Traffic. *Proceedings of the Intelligent Vehicles '94 Symposium*, pages 36–41.

[Ulmer, 1994b] Ulmer, B. (1994b). VITA II - Active Collision Avoidance in Traffic. *Proceedings of the Intelligent Vehicles '94 Symposium*, pages 1–6.

[Villasenor, 2014] Villasenor, J. (April 24, 2014). Products Liability and Driverless Cars: Issues and Guiding Principles for Legislation. *Brookings Institute.* http://www.brookings.edu/research/papers/2014/04/products-liability-driverless-cars-villasenor.

[Vincent, 2015] Vincent, J. (January 19, 2015). Scientists are teaching Super Mario to think and feel. *The Verge.* http://www.theverge.com/2015/1/19/7852975/artificially-intelligent-mario-learns-feel.

[Volvo Car..., 2014] Volvo Car... (April 29, 2014). Volvo Car Group's first self-driving Autopilot cars test on public roads around Gothenburg. *Volvo Cars of North America Newsroom.* https://www.media.volvocars.com/us/en-us/media/pressreleases/145619/volvo-car-groups-first-self-driving-autopilot-cars-test-on-public-roads-around-gothenburg.

[Wade, 2014] Wade, P. (October 22, 2014). Is That You, KITT? Audi Claims to Have Set Drivelress Car Speed Record. *FastCompany.* http://www.fastcompany.com/3037455/is-that-you-kitt-audi-claims-to-have-set-driverless-car-speed-record.

[Walsh, 2013] Walsh, T. (December 22, 2013). Google vs. Tesla vs. Ford: Who Has the Best Self-Driving Car to Get You Home Safely. *The Motley Fool.* http://www.fool.com/investing/general/2013/12/22/google-vs-tesla-vs-ford-who-has-the-best-self-driv.aspx.

[Weber, 2014] Weber, M. (May, 2014). Where To? A History of Autonomous Vehicles. *Computer History Museum.* www.computerhistory.org/atchm/where-to-a-history-of-autonomous-vehicles/.

[Wetmore, 2003] Wetmore, J. (2003). Driving the Dream: The History and Motivations Behind 60 Years of Automated Highway Systems. *Automotive History Review.* Via PDF from Consortium for Science, Policy, & Outcomes at ASU.

[Wiener, 1960] Wiener, N. (1960). On The Technical Development of Automatization and Some of Its Moral Consequences. From the Norbert Wiener Papers, MC 22. Box 33A: Folder 865, Institute Archives and Special Collections, MIT Libraries, Cambridge, MA.

[Wiener, 1965] Wiener, N. (1965). *Cybernetics Or Control and Communication in the Animal and the Machine.* MIT Press, Cambridge, MA.

[Wiener, 1956] Wiener, N. (ca. 1956). The Main Ideas of Cybernetics. From the Norbert Wiener Papers, MC 22. Box 31: Folder 798, Institute Archives and Special Collections, MIT Libraries, Cambridge, MA.

[Wiener, 1953] Wiener, N. (December 13, 1953). The Machine as Threat and Promise. *St. Louis Post-Dispatch.* From Norbert Wiener Papers. MC 22. Box 30C, Folder 732. Institute Archives and Special Collections, MIT Libraries, Cambridge, MA.

[Wilde, 1998] Wilde, G. (1998). Risk homeostasis theory: an overview. *Injury Prevention*, (4):89–91. http://injuryprevention.bmj.com/content/4/2/89.full.

[Winston, 1987] Winston, P. H. (1987). Artificial Intelligence: A Perspective. In [Grimson and Patil, 1987], pages 1–10.

[Wise, 2015] Wise, J. (April 9, 2015). Fear in the cockpit: The tragic plane crash in taipei was the result of mechanical and psychological failures. *Nautilus*, (23). http://m.nautil.us/issue/23/dominoes/fear-in-the-cockpit.

[Woods and Hollnagel, 2006] Woods, D. D. and Hollnagel, E. (2006). *Joint Cognitive Systems: Patters in Cognitive Systems Engineering.* CRC Press, Taylor & Francis, Boca Raton, FL.

[Zelizer, 1994] Zelizer, V. A. (1994). *Pricing the Priceless Child: The Changing Social Value of Children.* Princeton University Press, Princeton, NJ.

[Zetter, 2014] Zetter, K. (April 30, 2014). Hackers Can Mess With Traffic Lights to Jam Roads and Reroute Cars. *Wired.* http://www.wired.com/2014/04/traffic-lights-hacking.

[Zhu et al., 2011] Zhu, J., Ferguson, D. I., and Dolgov, D. A. (2011). System and method for predicting behaviors of detected objects. US Patent Application Publication 2012/0083960 A1, US Patent Office.