EUROPEAN COMMISSION

RESEARCH EXECUTIVE AGENCY

MOTORIST

7th FRAMEWORK PROGRAMME

Marie Curie Initial Training Network (ITN)

Grant Agreement 608092

MOTOrcycle Rider Integrated Safety

Deliverable no. 2.3

Road User Profiling Using Automated Vehicles: A Discussion

Deliverable no.	2.3
Dissemination level	Public
Work Package	1
Authors	J. C. F. de Winter, D. Dodou, & R. Happee
Status (F: final, D: draft)	D
File Name	Road User Profiling Using Automated Vehicles: A Discussion
Project Start Date and Duration	February 1, 2014 - January 31, 2018
Publication	The title of this deliverable was originally titled "(Rider) profiling and modelling state of the art." While reviewing the relevant literature, we noted that the topic of road user profiling cannot be treated separately from advances in sensor technology and automated driving. Furthermore, the issues that apply to profiling of motorcycle riders also hold for other road users such as cyclists. Accordingly, we re-titled the deliverable as "Road User Profiling Using Automated Vehicles: A Discussion."

Road User Profiling Using Automated Vehicles: A Discussion

J. C. F. de Winter, D. Dodou, R. Happee

Department of BioMechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands

Abstract

Automated vehicles are equipped with increasingly sophisticated sensors that detect and classify other road users. We argue that in the future, the *how*, *where*, and *who* of all errors and violations on the roads, including those of non-automated vehicles, powered two-wheelers, cyclists, and pedestrians, will be known. This information can be exploited to improve automated driving systems, to manage accident-prone sites, and to provide personalised feedback, rewards, and penalties. Although ubiquitous data collection may be necessary for achieving the Vision Zero goal of traffic fatalities and injuries, it also raises questions about ethics and privacy.

1. Introduction

Road traffic accidents are a serious public health concern. The lifetime odds of dying in a motor vehicle crash in the United States has been estimated at 1 in 114, which is high compared to, for example, air transport accidents (1 in 9,821) (National Safety Council, 2017). Young people are overrepresented (World Health Organisation, 2015), which makes the socioeconomic burden expressed in disability-adjusted life years (DALYs) particularly severe.

2. Data Collection On-Board Vehicles

Accidents are caused by both errors and violations. For example, studies have found that inattention (an error) and excessive speed (a violation) contribute to accidents (Treat et al., 1979). Automotive manufacturers devote ample resources to developing advanced driver assistance systems (ADAS), such as adaptive cruise control (ACC), automated emergency braking (AEB), lane keeping assistance (LKA), electronic stability control (ESC), and intelligent speed adaptation (ISA). These systems sense the environment and mitigate the consequences of driver errors or violations. For example, loss-of-control errors are mediated by ESC, whereas speeding and tailgating violations are forestalled by ISA and ACC, respectively.

Sensors onboard automated vehicles can now detect multiple road users in the environment. Ohn-Bar and Trivedi (2016) provided a review of ongoing research activities in three areas: (1) measuring the human in the vehicle (e.g., distracted/attentive, hands on wheel), (2) measuring humans around the vehicle (e.g., cyclists'/pedestrians' intent, trajectory, attention), and (3) measuring humans in surrounding vehicles (e.g., whether the driver in a nearby vehicle is attentive).

We argue that it will become technologically feasible to obtain 'total information' on road users' behaviours. It is important to note that this vision could already be realised if only a portion of road users transmit their data, considering that automated vehicles can detect an entire traffic situation (e.g., NVIDIA, 2017; Waymo, 2017; Fig. 1).

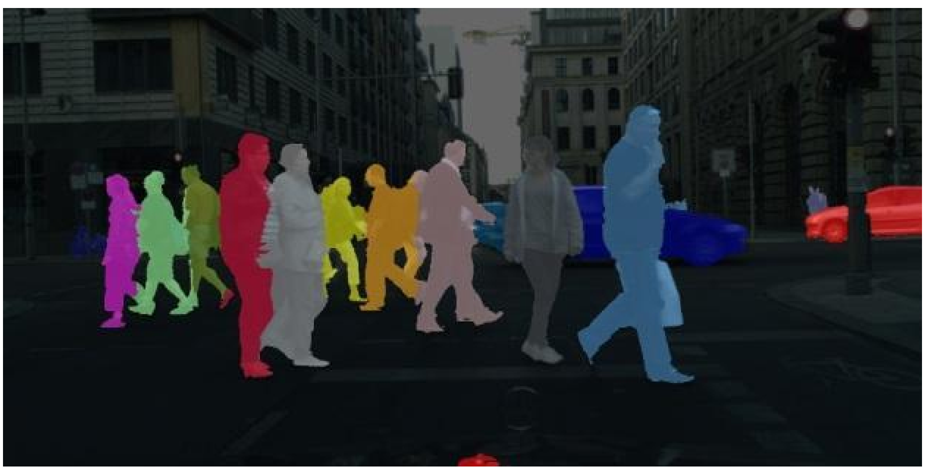

Figure 1. It is now possible to classify multiple road users in the environment (Pixelwise instance segmentation of an image from the Cityscapes dataset; Arnab & Torr, 2017a, 2017b; Cityscapes Dataset, 2017).

3. How Do Errors and Violations Lead to Accidents?

It is our proposition that, in the next half a century, sensors and cameras in vehicles can measure everything on the roads. Consider a situation where a motorcycle encounters an automated car. The rider does not pay attention and crosses the road without having priority. Despite performing an emergency brake, the automated car collides with the rider, resulting in a fatal injury. It is likely that the cameras onboard the car have recorded the sequence of events that led to the accident. Several opportunities exist:

- The car transmits the sensor information to the vehicle manufacturer, which could use this information to improve the safety of its fleet. For example, the automated driving software could be revised with improved motorcycle rider classification, so that riders are detected earlier. Considering the seriousness of the accident, the manufacturer may have the moral obligation to share the information with other manufacturers.

- The sensor data may be useful in court or for accident reconstruction purposes. Dashcams are already used for the same purposes (although the legal status of the use dashcams differs per country; Štitilis & Laurinaitis, 2016).

- The car could automatically produce a horn sound, in an attempt to direct the rider's attention (Fig. 2a). Above, we used a fatal accident as an example. The same principles could be applied to any behaviour, such as errors and violations that result in a near miss or an inefficiency. For example, the automated driving system may signal 'after you' to make a traffic situation more efficient (see also Nissan Motor Corporation, 2015; Fig. 2b).

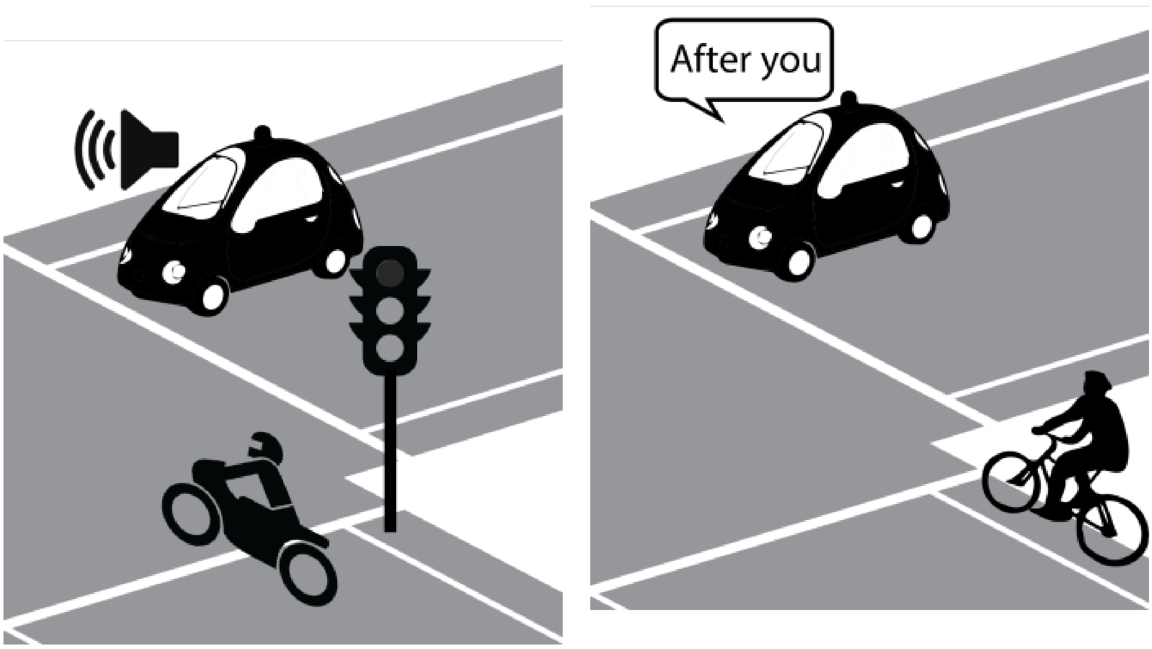

Figure 2. Possibilities for giving real-time feedback if an automated vehicle detects a hazard (left) or an inefficiency (right).

4. *Where* Do Errors and Violations Occur?

If the sensory information is sent via a wireless connection and stored in a database, researchers can use this information to pinpoint *where* errors and violations regularly occur. Hotspots can already be identified from geo-specific accident data. Adding errors, violations, and near accidents would improve the predictive power of such analyses (e.g., Ryder et al., 2016). For example:

- If a particular intersection yields a high amount of errors or violations (e.g., close encounters with vulnerable road users), then the intersection could be redesigned (e.g., traffic lights, roundabout, lane markings), already before an accident has occurred at that site. Similarly, if it is known at which road sections drivers drive close to other vehicles or exceed speed limits, then extra police enforcement could be applied at these sites.

- Adapting the infrastructure in real time is possible as well. For example, traffic lights could be scheduled depending on current traffic flows, an approach that is already commonplace (Göttlich et al., 2017; Steingrover et al. 2005). If road authorities have established which locations are error/violation-prone, it will also be possible to warn drivers in real time that they are entering a traffic situation/location that is statistically hazardous (Ryder et al., 2017). Similar approaches are already being used by route navigation companies to measure where traffic jams occur in real-time (e.g., Google Developers, 2017; TomTom N.V., 2015).

5. Who Makes Errors and Violations?

Accident proneness is a theory stating that certain drivers are overinvolved in accidents because of their clumsiness or personality traits, and which has often been discredited (Burnham, 2009; Elvik, 2011; Haight, 1964). The typical reasoning against accident proneness is as follows: Drivers perform a psychometric test, and their accident records are collected, either retrospectively or prospectively. Usually, it is found that the correlations between test scores and accident are small (e.g., $r < 0.10$), leading to the conclusion that the notion of accident proneness should be abandoned (Ranney, 1994).

Recent research has shown that these low correlations are because accidents are rare and largely random events: some drivers may never be involved in an accident, whereas others may be involved in an accident due to bad luck. However, simulation studies and empirical data suggest that if enough data are collected, accident data are highly reliable, up to $r = 0.8$ (Af Wåhlberg, 2009; De Winter, 2014).

While accidents are rare events, driver behaviour itself appears to be stable. For example, Groeger (2000) found high test-retest correlations of driving speed. Multi-day driving simulator research has shown that it is possible to calculate error-scores and violations-scores per driver (De Winter et al., 2007; Fig. 3), and various on-road studies have also shown that it is possible to create a driver risk profile using sensors in smartphones (e.g., accelerometers, GPS) and vehicle sensors (e.g., Castignani et al., 2015; Johnson & Trivedi, 2011; Júnior et al., 2017).

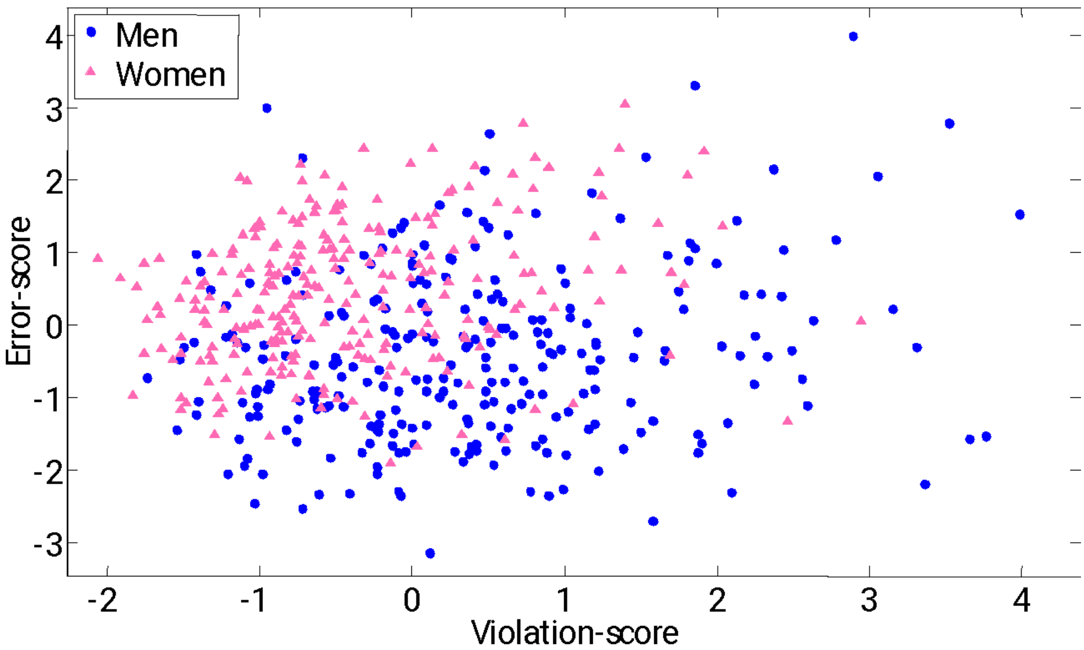

Figure 3. Objective error score and violation score, as extracted from of 520 participants in a simulator-based driver training (data from De Winter et al., 2007).

If automated vehicles transmit their data, several opportunities arise:

- It becomes possible to keep track of *who* makes errors and violations. This information collection would allow for automatic traffic fines, an approach which is more thorough than current local and incidental speed enforcement measures. If automated vehicles are equipped with identification features (e.g., license plate recognition, facial recognition), it will also be possible to record violations of other road users, such as pedestrians, cyclists, and motor riders. It has already been reported that the Shanghai police uses facial recognition to fine cyclists who cycle on the wrong bake path (That's Shanghai, 2017), whereas O'Malley (2014) reviewed the use of facial recognition and other 'telemetric policing' techniques.

- If error and violation data are stored per road user, it becomes possible to calculate a person-specific 'violations score' and 'errors score'. For example, it should be possible to define a cyclists' violations score as a composite of how often he or she runs a red light and ends up in dangerous encounters with other road users. This approach has already been applied on data obtained from driving simulators (De Winter & Kuipers, 2016).

- Besides using the data for enforcement purposes, the information could also be communicated to road users directly (e.g., in driver training). For example, road users may receive feedback via the Internet which they can use to improve their behaviours (e.g., Dijksterhuis et al., 2015).

- The recorded errors and violations could be communicated to insurance companies. For example, car drivers and motorcycle riders may receive a reduction in their insurance premium if they have a low errors and violations score. In 2016, ten insurance companies in the Netherlands offered such policies using a dongle plugged into the ODB port and transmitting velocity, accelerations, and GPS position (Consumentenbond, 2016).

6. Ethical and Privacy Considerations

Severe issues exist on the ethical side, especially regarding the driver profiling:

- How to ensure that all road users are treated fairly? This fairness criterion may require standardisation and certification of measurement devices. The computed measures also need to be standardised for vehicles of different dynamics (e.g., cars, motorcycles, trucks).

- If our proposed technique suggests that it is statistically likely that a person will be involved in a future accident, should a licensing authority prevent that person from driving, and should a transportation company find a replacement job for this person?

- Should all violations and other illegal acts be prevented? For example, motorised vehicles may be designed to function only within a given envelope of speed and accelerations so that some types of violations become impossible. This question is akin to the question whether intelligent speed adaptation (ISA) should be enforced or not.

- Is it legally possible and ethically acceptable to revoke driver licenses and issue fines to drivers based on their accident proneness (i.e., a statistical index of risk), or will it be preferred to rely on observed behaviours only (i.e., speed or headway), as it is currently done?

- How should thresholds be set? An overly tolerant criterion will be detrimental to road safety, whereas setting a too strict criterion may harm the quality of life of those prevented from driving. Figure 4 illustrates the results of a simulation where the correlation between a violation score and an accident involvement score is 0.6. The figure shows that if screening out the 10% poorest drivers, then 39% of crashes can be prevented. Naturally, there are also misses (people with a low violation score who are still involved in a crash) and false positives (people with a high violation score who are not involved in a crash). Although the threshold level is an ethically challenging concern, the same can be said about thresholds in contemporary driving testing and enforcement (e.g., tests of visual acuity, speed limits, demerit point system in driver licensing).

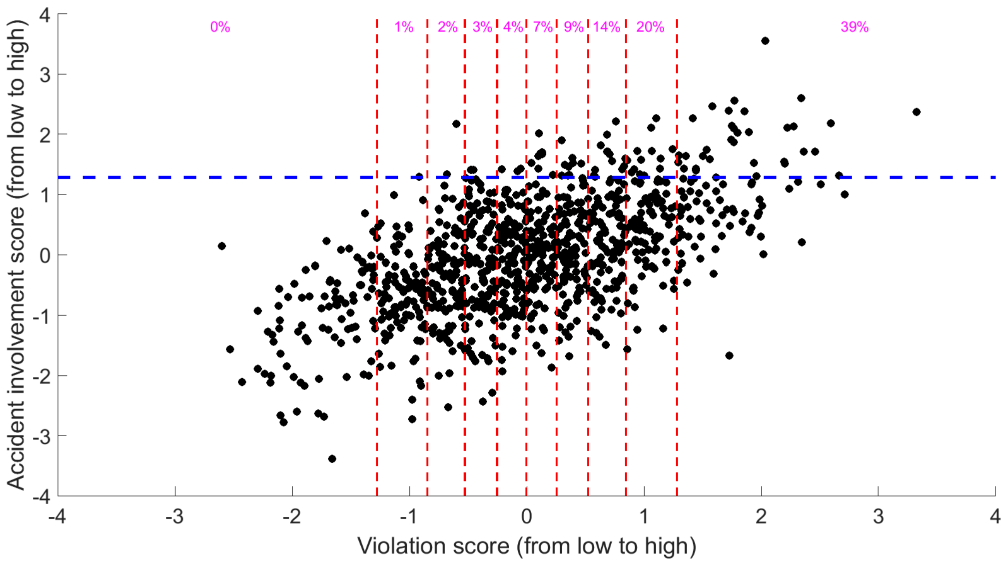

Figure 4. Scatter plot of 1,000 drivers with a correlation $r = 0.6$ between drivers' accident involvement scores and violation scores. The horizontal blue line is drawn at the 90th percentile of the accident involvement score; it is assumed that drivers with an accident involvement score higher than the 90th percentile will be involved in a crash (i.e., the overall accident rate is 10%). The vertical red lines are drawn at each 10th percentile of the driving style score, and the percentage at the top shows the expected accident involvement rate of the drivers within that percentile range of the violation score.

- Suppose that data from automated driving systems start being used by companies or governments, then who should be in control of this process, who should be the owner of the data, and who should have access to these data? (see also Fagnant & Kockelman, 2015; Kitchin, 2016). Privacy laws are stringent, but some voices argue that the 'end of personal privacy' is near in our digitised society (Madan et al., 2009). Privacy issues become especially severe when driving skill and driving style indicators are combined with other databases, such as databases of medical information and polygenic scores (Krapohl et al., in press).

7. Discussion

We argued that, in the future, all behaviours on the roads will be recordable via onboard cameras and sensors. This data collection, in turn, can contribute substantially to reducing the number of road traffic accidents. Once it is known how errors and violations contribute to accidents, where those behaviours occur, and who is involved, it will be possible to prevent accidents from happening.

The ideas outlined in this document may sound farfetched. However, it is worth noting that early versions of the required technology are already available. For example, road-based surveillance cameras and automated license plate recognition are already commonplace, vehicles are increasingly equipped with dashcams, and automated vehicles are being equipped with semantic scene interpretation. Similarly, the idea of use-based insurance (pay-how-you-drive) is already applied by insurance companies worldwide (Husnjak et al., 2015), whereas live traffic information is available from various telematics applications (e.g., TomTom N.V., 2015). Considering that Facebook's Deepface has been used to identify people from pictures with high accuracy (Bohannon, 2015), it should also be possible to classify road users using facial recognition, provided that camera resolution is sufficient.

The present discussion differs from similar debates regarding the privacy of Internet use (e.g., Facebook, Google), electronic patient files, and genetic screening, because the former involves *public* roads. Although it is possible to keep one's medical records and Internet search history private, it is unlikely that a road user can keep his or her errors and violations shielded from sensors and cameras. Indeed, it is implausible that in twenty years, vehicles are still private units that drive around without being tracked or connected in some way (and see Gerla et al., 2014). In the future, a significant portion of motorised vehicles will either be (highly) automated or manually driven but 'connected,' and capture information on non-motorized vehicles in the vicinity, such as cyclists and pedestrians. This increased connectivity creates ample opportunities for making use of these data. At the same time, these opportunities involve privacy and ethical concerns, and may, therefore, face some backlash from the public. We hope that our paper contributes to critical thinking on this topic in the domain of transportation research.

8. Acknowledgments

This paper is previously published as deliverable of the Motorist project (Marie Curie ITN, Grant Agreement 608092). Joost de Winter is supported by the research programme "How should automated vehicles communicate with other road users?" (project number 016.Vidi.178.04), which is financed by the Netherlands Organisation for Scientific Research (NWO). The authors declare no conflict of interest.

9. References

Af Wåhlberg, A. E. (2009). *Driver behaviour and accident research methodology: Unresolved problems.* Surrey: Ashgate.

Arnab, A., & Torr, P. H. S. (2017a). Pixelwise instance segmentation with a dynamically instantiated network. Retrieved from http://www.robots.ox.ac.uk/~aarnab/projects/cvpr_2017/Instances_CVPR_slides.pptx

Arnab, A., & Torr, P. H. S. (2017b). Pixelwise instance segmentation with a dynamically instantiated network. Retrieved from https://arxiv.org/abs/1704.02386

Bohannon, J. (2015). Unmasked. *Science, 327*, 492–494. https://doi.org/10.1126/science.347.6221.492

Burnham, J. C. (2009). *Accident proneness. A history of technology, psychology, and misfits of the machine age.* Chicago, IL: University of Chicago Press.

Castignani, G., Derrmann, T., Frank, R., & Engel, T. (2015). Driver behavior profiling using smartphones: A low-cost platform for driver monitoring. *IEEE Intelligent Transportation Systems Magazine, 7*, 91–102. https://doi.org/10.1109/MITS.2014.2328673

Cityscapes Dataset (2017). Dataset overview. Retrieved from https://www.cityscapes-dataset.com/dataset-overview/

Consumentenbond (2016). Rijstijlpolis. Goedkoop voor keurige rijders [Driving style policy. Cheap for neat drivers]. Retrieved from https://www.consumentenbond.nl/binaries/content/assets/cbhippowebsite/gidsen/geldgids/2016/nummer-7---november/gg201611p20-rijstijlverzekeringen.pdf

De Winter, J. C. F. (2014). Why person models are important for human factors science. *Theoretical Issues in Ergonomics Science, 15*, 595–614. https://doi.org/10.1080/1463922X.2013.856494

De Winter, J. C. F., & Kuipers, J. (2016). Relationships between years of licensure and driving style measured with a short simulator-based test (*N* = 650). *Proceedings of the 7th International Conference on Applied Human Factors and Ergonomics*. Orlando, FL. https://doi.org/10.1007/978-3-319-41682-3_54

De Winter, J. C. F., Wieringa, P. A., Kuipers, J., Mulder, J. A., & Mulder, M. (2007). Violations and errors during simulation-based driver training. *Ergonomics, 50*, 138–158. https://doi.org/10.1080/00140130601032721

Dijksterhuis, C., Lewis-Evans, B., Jelijs, B., De Waard, D., Brookhuis, K., & Tucha, O. (2015). The impact of immediate or delayed feedback on driving behaviour in a simulated Pay-As-You-Drive system. *Accident Analysis & Prevention, 75*, 93–104. https://doi.org/10.1016/j.aap.2014.11.017

Elvik, R. (2011). Book Review: Anders af Wåhlberg: Driver behaviour and accident research methodology. Unresolved problems. Ashgate Publishing (2009). *Safety Science, 49*, 751–752. https://doi.org/10.1016/j.ssci.2011.01.011

Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations, *Transportation Research Part A, 77*, 167–181. https://doi.org/10.1016/j.tra.2015.04.003

Gerla, M., Lee, E. K., Pau, G., & Lee, U. (2014). Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds. *Proceedings of the 2014 IEEE World Forum on Internet of Things (WF-IoT)*, Seoul, 241–246. https://doi.org/10.1109/WF-IoT.2014.6803166

Google Developers (August 19, 2017). Web > Maps JavaScript API > Samples > Traffic layer. Retrieved from https://developers.google.com/maps/documentation/javascript/examples/layer-traffic

Göttlich, S., Potschka, A., & Ziegler, U. (2017). Partial outer convexification for traffic light optimization in road networks. *SIAM Journal on Scientific Computing, 39*, B53–B75. https://doi.org/10.1137/15M1048197

Groeger, J. A. (2000). Fast learners: once a speeder, always a speeder? *Proceedings of the 10th Seminar on Behavioural Research in Road Safety*, Esher, 144–151.

Haight, F. A. (1964). Accident proneness, the history of an idea. *Automobilismo e Automobolismo Industriale, 12*, 534–546.

Husnjak, S., Peraković, D., Forenbacher, I., & Mumdziev, M. (2015). Telematics system in usage based motor insurance. *Procedia Engineering*, *100*, 816–825. https://doi.org/10.1016/j.proeng.2015.01.436

Johnson, D. A., & Trivedi, M. M. (2011). Driving style recognition using a smartphone as a sensor platform. *Proceedings of the 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Washington DC, 1609–1615. https://doi.org/10.1109/ITSC.2011.6083078

Júnior, J. F., Carvalho, E., Ferreira, B. V., De Souza, C., Suhara, Y., Pentland, A., & Pessin, G. (2017). Driver behavior profiling: An investigation with different smartphone sensors and machine learning. *PLOS ONE, 12*, : e0174959. https://doi.org/10.1371/journal.pone.0174959

Kitchin, R. (2016). The ethics of smart cities and urban science. *Philosophical Transactions of the Royal Society A, 374*, 20160115. https://doi.org/10.1098/rsta.2016.0115

Krapohl, E., Patel, H., Newhouse, S., Curtis, C. J., Von Stumm, S., Dale, P. S., et al. (in press). Multi-polygenic score approach to trait prediction. *Molecular Psychiatry*. https://doi.org/10.1038/mp.2017.163

Madan, A., Waber, B., Ding, M., Kominers, P., & Pentland, A. (2009). Reality mining: The end of personal privacy [presentation]. Retrieved from http://senseable.mit.edu/engagingdata/presentations/ED_SIII_Madan_Waber_et_al.pdf

National Safety Council (2017). Odds of dying. Retrieved from http://www.nsc.org/learn/safety-knowledge/Pages/injury-facts-odds-of-dying.aspx

Nissan Motor Corporation (October 27, 2015). Nissan IDS Concept: Nissan's vision for the future of EVs and autonomous driving. Retrieved from http://nissannews.com/en-US/nissan/usa/releases/nissan-ids-concept-nissan-s-vision-for-the-future-of-evs-and-autonomous-driving

NVIDIA (March 16, 2017). NVIDIA and PACCAR developing self-driving trucks. Retrieved from https://www.youtube.com/watch?time_continue=30&v=Wweoh7WJNUw

O'Malley, P. (2014). Telemetric policing. In G. Bruinsma & D. Weisburd (Eds.), *Encyclopedia of Criminology and Criminal Justice* (pp. 5135–5145). Springer New York. https://doi.org/10.1007/978-1-4614-5690-2_262

Ohn-Bar, E., & Trivedi, M. M. (2016). Looking at humans in the age of self-driving and highly automated vehicles. *IEEE Transactions on Intelligent Vehicles, 1*, 90–104. https://doi.org/10.1109/TIV.2016.2571067

Ranney, T. A. (1994). Models of driving behavior: a review of their evolution. *Accident Analysis and Prevention, 26*, 733–750. https://doi.org/10.1016/0001-4575(94)90051-5

Ryder, B., Gahr, B., & Dahlinger, A. (2016). "An in-vehicle information system providing accident hotspot warnings". *Prototypes, 3*. http://aisel.aisnet.org/ecis2016_prototypes/3

Ryder, B., Gahr, B., Egolf, P., Dahlinger, A., & Wortmann, F. (2017). Preventing traffic accidents with in-vehicle decision support systems-The impact of accident hotspot warnings on driver behaviour. *Decision Support Systems, 99*, 64–74. https://doi.org/10.1016/j.dss.2017.05.004

Steingrover, M., Schouten, R., Peelen, S., Nijhuis, E., & Bakker, B. (2005). Reinforcement learning of traffic light controllers adapting to traffic congestion. *Proceedings of the Belgium-Netherlands Artificial Intelligence Conference*, Brussels, 216–223.

Štitilis, D., & Laurinaitis, M. (2016). Legal regulation of the use of dashboard cameras: Aspects of privacy protection. *Computer Law & Security Review, 32*, 316–326. https://doi.org/10.1016/j.clsr.2016.01.012

That's Shanghai (September 21, 2017). Police now using facial recognition to bust cyclists in Shanghai. Retrieved from http://www.thatsmags.com/shanghai/post/20684/police-now-using-facial-recognition-to-bust-cyclists-in-shanghai

TomTom N. V. (2015). TomTom real time traffic information [white paper]. Retrieved from https://www.tomtom.com/lib/img/REAL_TIME_TRAFFIC_WHITEPAPER.pdf

Treat, J. R., Tumbas, N. S., McDonald, S. T., Shinar, D., Hume, R. D., Mayer, R. E., et al. (1979). *Tri-level study of the causes of traffic accidents: executive summary* (Report No. DOTHS034353579TAC(5)). Washington, DC: National Highway Traffic Safety Administration.

Waymo (2017). Waymo Lidar. Retrieved from https://x.company/media/projects/waymo/lidar.jpg

World Health Organisation (2015). World Health Organisation (2015). Global status report on road safety 2015. Retrieved from http://www.who.int/violence_injury_prevention/road_saf ety_status/2015/en/