Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 3 (2015) 2534 – 2542

Procedia MANUFACTURING

6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015

An international crowdsourcing study into people's statements on fully automated driving

Pavlo Bazilinskyy*, Miltos Kyriakidis, Joost de Winter

Department BioMechanical Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

Abstract

Fully automated driving can potentially provide enormous benefits to society. However, it has been unclear whether people will appreciate such far-reaching technology. This study investigated anonymous textual comments regarding fully automated driving, based on data extracted from three online surveys with 8,862 respondents from 112 countries. Initial filtering of comments with fewer than 15 characters resulted in 1,952 comments. The sample consisted primarily of males (74%) and had a mean age of 32.6 years. Next, we launched a crowdsourcing job and asked 69 workers to assign each of the 1,952 comments to at least one of 12 predefined categories, which included positive and negative attitude to automated driving, enjoyment in manual driving, concerns about trust, reliability of software, and readiness of road infrastructure. 46% of the comments were classified into the category 'no meaningful information about automated driving', leaving 792 comments for further analysis. 39% of the comments were classified as 'positive attitude towards automated driving' and 23% were classified as 'negative attitude towards automated driving'. In conclusion, the public opinion appears to be split, with a substantial number of respondents being positive and a significant number of respondents being negative towards fully automated driving.

Peer-review under responsibility of AHFE Conference

Keywords: Survey; Questionnaire; Crowdsourcing; Driverless car; Automated driving

* Corresponding author. Tel.: +31-15-278-7891.
E-mail address: P.Bazilinskyy@tudelft.nl

Peer-review under responsibility of AHFE Conference
doi:10.1016/j.promfg.2015.07.540

1. Introduction

It is generally believed that fully automated driving (FAD), or 'level 5 automation' according to the SAE levels of driving automation [1], will be a common mode of transportation in the (far) future. Automated driving could have large positive influences on society in terms of safety and efficiency of road transport.

Automated driving is currently a much discussed topic in academic institutions [2–8], governmental bodies [9,10], and industries [11–16]. Recently, automated driving has also become a topic of great interest to the public [7,8,17,18]. For example, one blog article on the lane changing capabilities of the Tesla S was read 27,842 times and received a relatively large number of 96 comments (as recorded on 12 December 2014) [19]. A particular comment on this blog illustrates that people have legitimate questions regarding the robustness of automated driving technology in demanding environmental conditions: "*Does an inch-thick crust of mud and salt screw up the sensors' ability to accurately measure the environment around the car?*"

Although the topic of automated driving is widely discussed in public fora, little scientific knowledge is available regarding the international perspective on the foreseen radical change in society and the level of acceptance of this technology. The present study aimed to investigate the public opinion on FAD.

1.1. Collected comments on automated driving

During 2014, in our research group, three surveys were launched on the CrowdFlower online platform (www.crowdflower.com) to poll the public opinion on fully automated driving. In this paper, we analyse the textual comments obtained from these three surveys

Survey 1 (S1) "Research study about driving behavior" [20] is an innovative study that explored the use of the crowdsourcing service CrowdFlower in academia. The 15-item survey focused on respondents' knowledge of automated driving systems and cross-national differences in traffic violations as measured with the Manchester Driver Behaviour Questionnaire (DBQ). In total 1,862 responses were obtained within 20 hours at a cost of $247. The 16th question in the survey "*Any comments?*" invited the respondents to give any comments related to the survey itself and to the topic of the questionnaire – automated driving.

A larger survey (Survey 2; S2) "Opinion on automated driving systems" [21] investigated user acceptance, worries, and willingness to buy partially, highly, and fully automated vehicles. In total 5,000 responses from 109 countries (40 countries with at least 25 respondents) were collected. This study further investigated cross-national differences and assessed correlations with personal variables, such as age, gender, and personality traits as measured with a short version of the Big Five Inventory. The 63rd question "*Please provide any additional comments you may have about the survey*" asked respondents to provide any comments, including their thoughts on automated driving.

Finally, Survey 3 (S3) [22] examined user acceptance of auditory interfaces in modern cars and their willingness to be exposed to auditory feedback in highly and fully automated driving. This survey obtained 2,000 responses from 96 countries. The 31st question "*Please provide any suggestions which could help engineers to build safe and enjoyable automated cars*" was targeted specifically at receiving feedback on automated driving.

Comments received in these three surveys formed a large amount of text. Text is often analyzed manually by the researchers themselves, a process that can be time consuming and prone to investigator bias. Text mining is a more efficient approach for analyzing large quantities of lexical structures [23–25]. Statistical text mining techniques allow researchers to tag and annotate texts, establish distributions of word frequencies, and extract underlying patterns. There are numerous examples of efficient and fast analyses of text with such approaches. For example, Twitter messages that have no more than 140 characters were processed to receive real-time information on distracted driving messages [26]. Text mining is commonly employed in the field of biomedical research [27,28] and is a widely used technique for analyzing web content [29,30]. Analysis of text with text mining techniques is often faster than manual analyses. However, it requires a deep understanding of the underlying tools. A novel approach of crowdsourcing the task of text analysis was employed in this study: we delegated the classification of comments to dozens of workers from all over the world.

1.2. Using CrowdFlower for classifying comments

A preliminary inspection of the comments led us to conclude that comments consisting of less than 15 characters contained no meaningful information. These comments (N = 5,884, 75%), including empty comments, were therefore removed. Accordingly, a total of 1,952 comments were left for further analysis. The sample consisted primarily of males (1,429 males, 513 females, 10 gender unknown) and had a mean age of 32.6 years (SD = 11.4, N = 1914 with available age data).

The 1,952 comments represented text of considerable size: 175,378 characters (M = 90, SD =101, N = 1,952), or, assuming that the average length of a word in the English language is 5.1 characters [31], about 34,388 words. We reasoned that a manual analysis of such a large number of comments would not be reasonable, since it would take a significant amount of person-hours of work. The option of outsourcing the task among colleagues was discarded, as we suspected it could have led to biased results. Thus, crowdsourcing the job appeared to be a good solution.

The comments were categorized by means of a survey project launched on CrowdFlower. We outlined 12 categories for the classification of the comments. These 12 categories were created through a manual analysis of a random selection of 200 comments. Specifically, the categories were defined based on the frequencies of comments that could be assigned to particular categories. The categories encompassed positive and negative opinions towards automated driving, concerns about the different aspects of automated driving, and the public's vision of automated driving. Table 1 shows the established categories and provides short descriptions.

Table 1. Twelve categories used for classifying the comments.

ID	Code name	Category	Description
1	NEGATV	Negative attitude towards automated driving	Statements that express general negativity towards automated driving.
2	MANUAL	Preference to manual driving (i.e. ability to choose manual driving)	Statements saying that manual driving would be preferred over automated driving. By manual driving we mean a present day situation where cars are controlled by humans.
3	SEMAUT	Preference to semi-automated driving (i.e. ability to choose manual driving)	Statements saying that manual driving would be preferred over automated driving because a driver wants to be in control of his own vehicle.
4	ENJOYM	Enjoyment in manual driving	Statements saying that manual driving would be preferred over automated driving because of the "joy of driving".
5	CCOSTS	Concerns about costs	Statements that express concerns about the cost of automated driving.
6	CTRUST	Concerns about trust	Statements that refer to lack of trust for a vehicle that can drive on its own.
7	CSOFTW	Concerns about security of software (i.e. threats from hackers)	Statements that express concerns about misuse of the software of automated cars (such as threats from computer hackers).
8	CINFST	Concerns about readiness of infrastructure (i.e. unprepared roads)	Statements that express concerns that current modern roads are not prepared to support automated driving.
9	POSITV	Positive attitude towards automated driving	Statements that express general positivity towards automated driving.
10	VISION	Vision of a highly-automated vehicle	Statements regarding the vision of highly-automated driving.
11	NOMEAN	No meaningful information about automated driving	Statements that carry no meaningful information about automated driving.
12	OTHER	Other	All other statements.

In our crowdsourcing project, we only allowed workers from English speaking countries. CrowdFlower provides the option to select up to 15 countries per project. Hence, our workers were from Antigua and Barbuda, Australia, Bahamas, Barbados, Belize, Canada, Dominica, Grenada, Ireland, Jamaica, New Zealand, Saint Lucia, Trinidad and Tobago, United Kingdom, and the United States. To assure sufficient quality of the categorization, the highest (third) level of performance of contributors was selected. That is, only the most highly ranked workers were invited to perform the categorization. A maximum of 200 randomly selected judgments per contributor and IP address were permitted. In total, 69 workers classified the comments. The total amount to be paid for the crowdsourced categorization of 1,952 comments was $120. Each comment was processed by at least five workers, while the workers were not allowed to review the same comment more than once. The workers received the comments in random order, and they were allowed to classify comments in more than one category.

To control the quality of data, we adopted a *threshold* when analyzing the data. The threshold defines the minimum number of workers that assigned a comment to a particular category for the categorization to be accepted as valid. The cases where *threshold* was equal to 1, 2, 3, 4 and 5 judgments were handled.

Furthermore, countrywide differences were analyzed at the national level by comparing the opinion of people on automated driving as a function of their country's income. Information on the Gross Domestic Product (GDP) per capita of countries involved in the surveys was extracted from the records of the World Bank [32]. The values of GPD per capita of The Bahamas, Barbados, Taiwan, and United Arab Emirates were retrieved from the International Monetary Fund [33]. One comment originated from the Palestinian Territories, and no information on the GDP per capita of that country could be found. Hence, that comment was excluded, leaving 1,951 comments for the cross-national analysis.

Finally, the authors selected three comments from each category as representative examples of the opinions of the respondents. Only comments that were written in English language, were clearly stated and easily interpretable, and were at least 50 characters long.

2. Results

In total, 11,760 reviews (or 'judgments' according to the terminology used by CrowdFlower) of comments were received. That is, on average, each comment was reviewed 6.02 times (SD = 1.52, N = 1,952). The responses were gathered between 19 November 2014 17:05 and 20 November 2014 00:48 (CET). The categorization job received an overall satisfaction rating of 4.5 out of 5.0. The respondents ranked the clarity of the instructions as 4.4 / 5.0, fairness of the questions as 4.2 / 5.0, easiness of the survey as 3.9 / 5.0, and the offered payment ($0.75 for categorization of 100 comments) as 4.2 / 5.0.

We first explored the effect of different values of the *threshold* parameter. If *threshold* equaled 1, no judgments were ignored; if *threshold* equaled 2, then 3,438 judgments were ignored; if *threshold* equaled 3, 4,792 judgments were ignored; if *threshold* equaled 4, 5,581 judgments were ignored; and finally, if *threshold* equaled 5, 6,186 judgments were ignored.

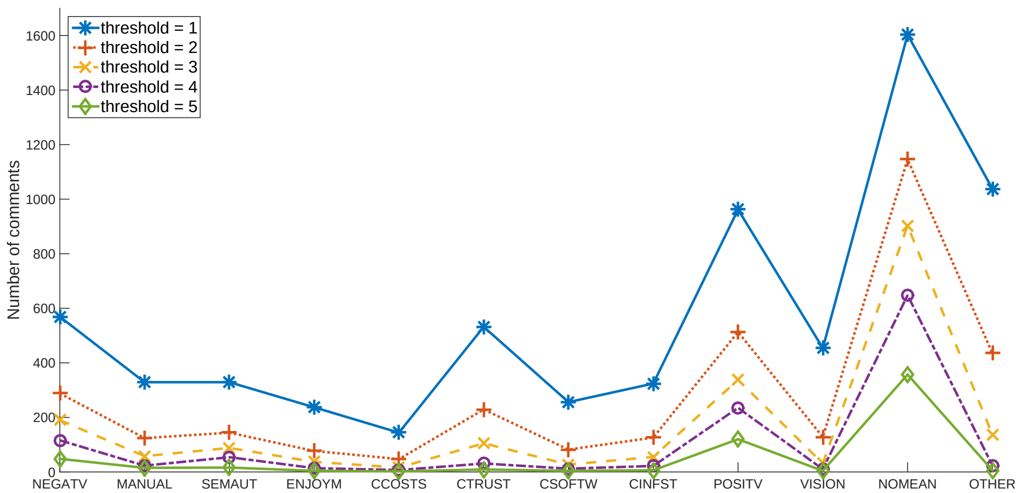

Fig. 1. Classification of comments for the values of *threshold*. N = 1,952.

Figure 1 shows the numbers of accepted comments per category. According to our interpretation, the most valid and robust outcome was achieved when *threshold* equaled 3. When *threshold* equaled 3, 16% (N = 309) of the comments were classified as POSITV, while 9% (N = 185) were classified as NEGATV. In addition, 5% (N = 98) of the comments were marked as CTRUST, whereas 3% (N = 52) expressed CINFST. Furthermore, 4% (N = 83) of the comments expressed a preference for semi-automated driving (i.e., SEMAUT). Finally, 3% (N = 56) of the comments indicated that people would prefer manual driving, that is, they were marked as MANUAL, and 2% (N = 36) of the comments were classified as ENJOYM.

The dominant category in the classification of comments was NOMEAN, with 46% (N = 903) of the comments classified into this category. These comments were seen as statements that carried no meaningful information about FAD and were excluded from the data set. Subsequently, 792 meaningful comments were left for further analysis.

Table 2 shows the numbers of the comments assigned to combinations of categories. Numbers on the diagonal indicate the total numbers of comments that were classified into each category. The off-diagonal values are subsets of the corresponding numbers on the diagonal of the table; they indicate how many comments were tagged into each pair of categories. A total of 185 of 792 comments (23%) were classified as negative towards automated driving (NEGATV), while 309 of 792 (39%) were classified as positive (POSITV).

Table 2. Numbers of comments classified into a category (diagonal) and into two categories (off-diagonal), with *threshold* = 3. (N = 1,952)

NEGATV	185										
MANUAL	32	56									
SEMAUT	6	0	83								
ENJOYM	11	14	7	36							
CCOSTS	0	0	1	0	14						
CTRUST	22	3	7	1	1	98					
CSOFTW	6	0	1	0	0	8	25				
CINFST	7	1	0	1	0	4	1	52			
POSITV	2	2	6	6	2	9	2	7	309		
VISION	0	0	1	0	0	1	0	2	15	28	
OTHER	1	0	0	0	0	1	0	2	1	1	72
	NEGATV	MANUAL	SEMAUT	ENJOYM	CCOSTS	CTRUST	CSOFTW	CINFST	POSITV	VISION	OTHER

A comparatively small number of comments were assigned to more than one category. Specifically, 32 comments were categorized as both MANUAL and NEGATV, which is an expected result, as respondents who have a negative attitude to automated driving also prefer to have manual control of a car. Furthermore, 22 respondents expressed a low level of trust towards automated driving and indicated a negative attitude towards automated driving (CTRUST & NEGATV). Seven people indicated that they would rather use a semi-automated vehicle while also expressing a low level of trust in automated driving (CTRUST & SEMAUT). However, 9 respondents had a positive attitude to automated driving while they also mentioned a low level of trust in automated driving (CTRUST & POSITV). Furthermore, 11 comments were categorized as ENJOYM and NEGATV, 14 comments as ENJOYM and MANUAL, and 7 comments as ENJOYM and SEMAUT.

Next, the comments were analyzed at the national level. Figure 2 shows the numbers of comments for three groups created based on the GDP per capita of corresponding countries: low-income countries (with GDP per capita between \$694 and \$3,900, N = 265), medium income countries (with GDP per capita between \$4,403 and \$21,035, N = 263), and high income countries (with GDP per capita between \$21,910 and \$111,162, N = 264). The categories were created automatically by sorting the comments and splitting them into three equally sized groups. Fisher's exact test showed that people from high-income countries were more likely to be negative ($p < .001$) and less likely to be positive ($p = .001$) about automated driving than people from low income countries. People from high-income countries were also more concerned about software issues ($p = 0.048$). The other differences between respondents from low versus high income countries were not statistically significant ($p > .05$ for each of the other 8 categories).

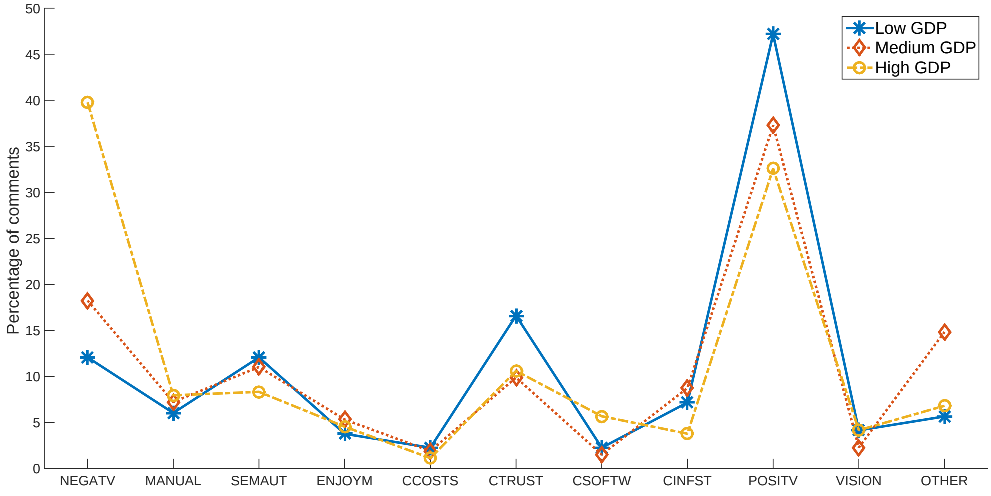

Fig. 2. Percentage of comments assigned to the categories based on the GDP per capita of a country of origin of the comment: low, medium, and high. Groups of GDP per capita were created by sorting the comments based on the GDP per capita and splitting them into three equal groups. N = 792.

Finally, Table 3 introduces examples of comments for all categories. Code names indicate assigned categories.

Table 3. Examples of respondents’ comments per category. The comments are not edited for grammar and spelling.

Code name	Comment
NEGATV1	The idea of fully automated cars scares me even more than other drivers.
NEGATV2	I will never set foot in a fully automated vehicle.
NEGATV3	I think this idea is unsafe and bizarre, actually.
MANUAL1	I will prefer to use manual driving because fully automated will make you lazy mentally.
MANUAL2	I prefer a manually driving 100%.
MANUAL3	I like manual driving.
SEMAUT1	I don’t like the fully automated vehicles because i cant control it, a highly automated vehicle sounds much better.
SEMAUT2	I can't think of any as I don't like the idea of automated cars, as I prefer to have control.
SEMAUT3	Totally automated is giving up total control and some people may not like it.
ENJOYM1	A fully automated car will eliminate driving pleasure. There should be an option for manual driving.
ENJOYM2	I enjoy the manually driving too. Cause I feel I’m the driver :)
ENJOYM3	It is not a wise idea at all. If it really happens than there will be no race driver and no one will enjoy driving. No will ever say "Let's go for a long drive".
CCOSTS1	I hope the automated cars will be sold at a price that is not too expensive.
CCOSTS2	I would buy it for a good price and use it once it is on the market for a while and I'm sure that the system is safe. I would not be a pioneer on that, since safety is evolved.
CCOSTS3	Both price and quality accessible to everyone.
CTRUST1	I think this technology will take a long time to be really reliable and trusted.
CTRUST2	I trust my driving much more than I trust a computer system to do it for me. With the fully automated system, I would not want it because it would not allow for driver control if something major happened that a computer couldn't respond to.
CTRUST3	The cars should never be fully automated, at least the cars should maintain a certain degree of manual system as technology sometimes can fail.
CSOFTW1	I think manual control as a must. If anything malfunctioned or something like virus attack will be really dangerous without

	manual control. It will be like knight rider huh.
CSOFTW2	Well, my main concern with automated cars is the possibility of someone hacking the car and being able to take over the car. So I would think that security would be extensively tested to prevent such cases.
CSOFTW3	It’s a good survey to take but in my opinion fully automated will be a sure shot risk because computers are also justified as devil at bad hacking times.
CINFST1	The driving conditions in our country (country name) needs a lot of improvement.
CINFST2	In my country, the road infrastructure is very bad, and i think cars will have a tough time becoming automated.
POSITV1	I think the concept of Fully Automated Driving is very fascinating and it could be possible in the near future as technology develops and human beings advance.
POSITV2	Hopefully this wonderful technology will occur in my lifetime.
POSITV3	I hope this becomes available in the very near future.
VISION1	Since I cannot drive manually due to my bad vision, a fully automated car would be great for me, as long as I can see the instructions in the car and program it to get where I want to go, I should be able to get a licence for it.
VISION2	Opportunities to disabled.
VISION3	I am a disabled person and I have really bad eyes and limited field of view. I would be really happy if we had fully automated vehicles here in (name of country) so I wouldn't need second person to drive a car for me. It is sad though because it will take a lot of years for our country to introduce such vehicles on large scale.
NOMEAN1	I have no additional comments.
NOMEAN2	Don't judge me.
NOMEAN3	Comfy chair, spacious.
OTHER1	Correct sensors should be installed.
OTHER2	Don’t you have any other thing to do? Like finding a solution for global warming.
OTHER3	Correct sensors should be installed.

3. Discussion

In this study, free-response comments from three crowdsourced surveys involving 1,952 respondents were categorized by means of another project submitted to CrowdFlower. The decision to involve external people in the categorization was taken after determining that the classification of all comments would be cumbersome and time-consuming to do ourselves, and it could be biased. Our approach to the categorization of a large amount of text proved to be efficient and successful. Moreover, a threshold for accepting categorization of comments was developed, and robust results were obtained when this variable was set to 3.

The main finding was that the public opinion appears to be split, with a significant number of respondents being positive (POSITV) and a significant number of respondents being negative (NEGATV) towards FAD. This result is consistent with a previous survey study [21], which analyzed the public opinion on automated driving using five-point Likert items. A portion of the population does not appear to trust automated vehicles (CTRUST) and prefers to drive manually (MANUAL). A small number of comments were categorized into multiple categories. A dual categorization indicates obvious connections between categories. For example, 32 of 1,952 comments were categorized with both a negative attitude towards automated driving and a preference for manual control of cars in the future.

The comments were also analyzed at the national level, where they were grouped by GDP per capita of the respondents’ country. The results revealed an association between income level and the number of comments per category. People from high-income countries were more likely to express a negative comment and less likely to express a positive comment about automated driving. In one of our previous surveys using five-point Likert items, we found that people from countries with a higher GDP were more concerned about automated vehicles transmitting data than people from low income countries [21].

One of the categories presented to the workers of CrowdFlower for the categorization was “*preference to manual driving*”. This category was created to indicate that the preference is given to manual driving over automated driving. However, it could also be understood as “I prefer automated driving to manual driving”, which indicates that the preference is given to automated driving. Another issue was that the category “*vision*” could have been understood by some crowdworkers literally as ‘eyesight’ instead of ‘imagination’ or ‘prospect’ (as we intended). The examples in Table 3 (e.g., VISION1 & VISION3) reveal that FAD is preferred among people having bad eyesight or physical disabilities. Nonetheless, the workers indicated that instructions were well defined by ranking their clearness as 4.4 / 5.0.

A limitation of the present results is that CrowdFlower respondents are not representative of the entire population of stakeholders of future FAD vehicles. It is expected that such vehicles will initially be purchased by wealthy people, whereas jobs on CrowdFlower are performed mostly by people with relatively low income [21]. One recommendation is to launch more questionnaires, possibly using the ‘classic’ approach of asking people in person, in an attempt to target more diverse layers in the society. The questions that were asked in the surveys may be re-launched on CrowdFlower a few years in the future. This would create the opportunity to elucidate temporal trends in the public opinion.

Acknowledgements

The authors are involved in the Marie Curie ITN: HFAuto (PITN-GA-2013-605817).

References

[1] SAE, SAE Levels of Driving Automation, (2013). http://www.sae.org/misc/pdfs/automated_driving.pdf (accessed December 19, 2014).
[2] J.C.F. De Winter, R. Happee, M.H. Martens, N.A. Stanton, Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence, Transp. Res. Part F Traffic Psychol. Behav. 27 (2014) 196–217. doi:10.1016/j.trf.2014.06.016.
[3] R. Hoeger, A. Amditis, M. Kunert, A. Hoess, F. Flemisch, H.-P. Krueger, et al., Highly Automated Vehicles for Intelligent Transport: Have-It Approach, in: 15th World Congr. Intell. Transp. Syst. ITS Am. 2008 Annu. Meet., 2008.
[4] S. Kato, S. Tsugawa, K. Tokuda, T. Matsui, H. Fujii, Vehicle control algorithms for cooperative driving with automated vehicles and intervehicle communications, IEEE Trans. Intell. Transp. Syst. 3 (2002) 155–161. doi:10.1109/TITS.2002.802929.
[5] A.H. Jamson, N. Merat, O. Carsten, F.C.H. Lai, Fully-automated driving: The road to future vehicles, in: Proc. Sixth Int. Driv. Symp. Hum. Factors Driv. Assessment, Train. Veh. Des., 2011: pp. 2–9.
[6] D. Begg, A 2050 vision for London: What are the implications of driverless transport, 2014.
[7] S. Casley, A. Jardim, A. Quartulli, A Study of Public Acceptance of Autonomous Cars, Worcester Polytechnic Institute, 2013.
[8] KPMG, Self-Driving Cars: Are We Ready?, 2013.
[9] European Commision, Roadmap to a single European transport area – towards a competitive and resource efficient transport system, 2011.
[10] NHTSA, Preliminary Statement of Policy Concerning Automated Vehicles, 2013.
[11] Volvo, Volvo Car Group’s first self-driving Autopilot cars test on public roads around Gothenburg, (2014). https://www.media.volvocars.com/global/en-gb/media/pressreleases/145619/volvo-car-groups-first-self-driving-autopilot-cars-test-on-public-roads-around-gothenburg (accessed December 20, 2014).
[12] Volvo, Volvo Cars conducts research into driver sensors in order to create cars that get to know their drivers, (2014). https://www.media.volvocars.com/global/en-gb/media/pressreleases/140898/volvo-cars-conducts-research-into-driver-sensors-in-order-to-create-cars-that-get-to-know-their-driv (accessed December 20, 2014).
[13] A. Davies, A System That Any Automaker Can Use to Build Self-Driving Cars, (2014). http://www.wired.com/2014/11/delphi-automated-driving-system (accessed December 20, 2014).
[14] J. Walker, Jaguar Land Rover pledges to produce a self-drive car within ten years, (2014). http://www.birminghampost.co.uk/business/manufacturing/jaguar-land-rover-pledges-produce-8148485 (accessed December 20, 2014).
[15] Daimler, Mercedes-Benz sends autonomous automobiles onto the USA’s most extensive testing ground, (2014). http://www.daimler.com/dccom/0-5-658451-1-1745411-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0.html (accessed December 20, 2014).
[16] K. Sommer, Mobility Study 2013, 2013.
[17] B. Schoettle, M. Sivak, A survey of public opinion about autonomous and self-driving vehicles in the U.S., the U.K., and Australia, Michigan, USA, 2013.
[18] D. Howard, D. Dai, Corridors to the Future: Transportation and Technology - Opportunities and challenges presented by connected–automated vehicles, in: Transp. Res. Board 93rd Annu. Meet., Washington DC, 2014.
[19] D. Lavrinc, Tesla Model S Auto Pilot Does Lane Changes For You, (2014). http://jalopnik.com/tesla-model-s-auto-pilot-does-lane-changes-for-you-1644644652 (accessed December 12, 2014).
[20] J.C.F. De Winter, M. Kyriakidis, D. Dodou, R. Happee, Using CrowdFlower to study the relationship between self-reported violations and traffic accidents, in: AHFE, 2015.
[21] M. Kyriakidis, R. Happee, J.C.F. De Winter, Public opinion on automated driving: Results of an international questionnaire among 5,000 respondents, Transp. Res. Part F Traffic Psychol. Behav. (2014) 127–140. doi:http://dx.doi.org/10.1016/j.trf.2015.04.014.
[22] P. Bazilinskyy, J.C.F. De Winter, Auditory interfaces in automated driving: an international survey, PeerJ Comput. Sci. (2015).

[23] A.-H. Tan, Text mining: The state of the art and the challenges, in: Proc. PAKDD 1999 Work. Knowl. Disocovery from Adv. Databases, 1999: pp. 65–70.
[24] A. Hotho, A. Nürnberger, G. Paaß, A Brief Survey of Text Mining., in: Ldv Forum, 2005: pp. 19–62.
[25] M.W. Berry, Survey of Text Mining : Clustering, Classification, and Retrieval, 2004. doi:10.1007/978-1-84800-046-9.
[26] S.C. Roberts, J.D. Lee, Deciphering 140 Characters: Text Mining Tweets On #DriverDistraction, Proc. Hum. Factors Ergon. Soc. Annu. Meet. 58 (2014) 2195–2199. doi:10.1177/1541931214581461.
[27] L. Tanabe, U. Scherf, L.H. Smith, J.K. Lee, L. Hunter, J.N. Weinstein, MedMiner: an Internet text-mining tool for biomedical information, with application to gene expression profiling, Biotechniques. 27 (1999) 1210–1214.
[28] A.M. Cohen, W.R. Hersh, A survey of current work in biomedical text mining, Brief. Bioinform. 6 (2005) 57–71.
[29] R. Cooley, B. Mobasher, J. Srivastava, Web mining: Information and pattern discovery on the world wide web, in: Tools with Artif. Intell. 1997. Proceedings., Ninth IEEE Int. Conf., 1997: pp. 558–567.
[30] S.K. Pal, V. Talwar, P. Mitra, Web mining in soft computing framework: relevance, state of the art and future directions, Neural Networks, IEEE Trans. 13 (2002) 1163–1177.
[31] WolframAlpha, Average word length in English, (2014). http://www.wolframalpha.com/input/?i=average word length in English (accessed December 17, 2014).
[32] The World Bank, GDP per capita, (2014). http://data.worldbank.org/indicator/NY.GDP.PCAP.CD (accessed December 16, 2014).
[33] International Monetary Fund, World Economic Outlook Database, (2013). http://www.imf.org/external/pubs/ft/weo/2013/01/weodata/weoselgr.aspx (accessed December 16, 2014).