Artificial Intelligence in Self-Driving Cars Research and Innovation: A Scientometric and Bibliometric Analysis

Gianluca Biggi [a*] and Jack Stilgoe [b]

[a] Responsible Management Research Center, Department of Economics & Management, University of Pisa, Pisa, Italy.

[b] Department of Science and Technology Studies, University College London, UK

*Corresponding author. E-mail: gianluca.biggi@ec.unipi.it

Artificial Intelligence in Self-Driving Cars Research and Innovation: A Scientometric and Bibliometric Analysis

Abstract

This paper presents a scientometric and bibliometric analysis of research and innovation on self-driving cars. Through an examination of quantitative empirical evidence, we explore the importance of Artificial Intelligence (AI) as machine learning, deep learning and data mining on self-driving car research and development as measured by patents and papers. Alongside the exponential growth in the rate of inventive activities and scholarly efforts, we find evidence for a rapid and meaningful shift in the application of the technologies related to data gathering and processing for the purpose of self-driving cars after 2009. We show that this shift mirrors major changes in the landscape of innovators as well as increasing scholarly attention to the ethical, legal and social aspects of self-driving cars. Research and innovation relating to self-driving seem to be increasingly defined in terms of artificial intelligence, which neglects some aspects of future socio-technical systems that may be required to realise the potential of the technology.

Keywords: self-driving cars; research; innovation; artificial intelligence; bibliometric; scientometric.

JEL Classification: O30 O33

1. Introduction

The possibilities of self-driving cars have attracted substantial investment and attention. In addition to considering questions of technological opportunities and limits, we should also, following the sociology of expectations (Borup et al., 2006), seek to analyse how self-driving futures are being imagined, to improve anticipatory governance (Guston, 2014) of the technology. One way to explore such expectations and the aligning of financial and scientific resources behind them, beyond superficial statements from self-driving vanguards (Hilgartner et al., 2015), is to consider existing scientific publications and patent applications. In mapping these features, we might detect signs, following Dosi (1982) and Kuhn (1962) of an emerging "technoscientific paradigm" that knits together technological and scientific trajectories. This could in turn enable a debate on the social constitution (Szerszynski et al., 2013) of self-driving technologies that allow policymakers to contribute to the shaping of technological means in the service of publicly desirable ends.

According to one calculation[1], "autonomous vehicles" was one of the largest areas of investment in artificial intelligence research in 2020, with $4.4 billion of private investment. AVs, it seems have become a test case for the development of AI. We should therefore pay attention to the role that AI research and development is playing in the development of AVs and ask whether these dynamics are driving innovation in a particular direction, enabling some sociotechnical systems while foreclosing other possibilities.

1.1 The history and politics of self-driving cars

Understanding the possible plural futures of self-driving cars is aided by an understanding of alternative histories that could be told about technology. For many of the most prominent self-driving innovators and early accounts of their prowess, the relevant history centres on robotics and artificial intelligence. The key event for this story is the third DARPA grand challenge competition in 2007. According to one account "the moment… when everything changed" (Burns & Shulgan, 2018) was when a handful of teams of roboticists made machines that were able to complete a drive around an uninhabited town.

[1] Stanford AI Index 2021, https://aiindex.stanford.edu/wp-content/uploads/2021/03/2021-AI-Index-Report_Master.pdf, [Accessed 24 March 2021]

This history is told as part of a dominant "narrative of autonomy" (Tennant & Stilgoe, forthcoming), but it does not represent the full story. Historians such as Wetmore (2020) and Vinsel (2019) describe how the focus on the autonomous *vehicle* is just the most recent phase of a long history of self-driving innovation that, for much of the second half of the twentieth century, concentrated on infrastructures that would enable self-driving. More recently interest in self-driving cars, at least in the US, has downplayed or ignored questions of infrastructure. The assumption is that investment in upgrading infrastructure is unlikely and, if money were available, it would be too slow (Tennant & Stilgoe, forthcoming). The new focus has been on making smart cars rather than smart roads. If this means a prioritisation of software over hardware, and a focus on data for artificial intelligence, this could have profound implications for the political economy of self-driving systems, the structure of markets and the future of mobility. The politics of transport could find itself further entangled in the politics of platform capitalism (Pasquale, 2016). A focus on AI could crowd out alternative models and postpone or externalise consideration of the issues raised by disruptions to current mobility patterns. Our paper offers a first attempt to map the landscape of self-driving R&D to guide anticipatory governance.

Our research follows the work of Gandia et al. (2019) who map relevant scientific publications, and Cho et al. (2021), who map patent activity around self-driving cars by proving the ground for the understanding of antecedents, drivers and consequences of self-driving cars research and innovation. Despite these contributions, there is still a research gap: the absence of scientometric and bibliometric analyses based on the impact of the different areas of AI on self-driving car research and innovation. Our paper follows this methodological approach.

This paper unpacks the potential impact of the different areas of AI on self-driving cars research and innovation. We explore quantitative evidence on the evolution of AI in terms of scientific and inventive outputs as measured by patents and papers from 1994 through 2018. Section 2 presents the research methodology, explaining the necessary steps to perform the scientometric and bibliometric technique. Section 3 presents the exploratory analysis of the patent and article samples. In Section 4, our concluding remarks summarise the main findings, identify limits of the approach and highlight possibilities for future research.

2. Data

This study employs scientometric and bibliometric techniques to identify the main characteristics, evolution and potential future trends of self-driving cars research and innovation with a specific focus on the impact of artificial intelligence (AI). The scientometric technique is a method that refers to knowledge domain visualisation (Pollack & Adler, 2015) which could be applied to detect innovation trajectories using patent data (Narin, 1994; Narin & Hamilton, 1996) while bibliometrics is a quantitative technique applied to published literature (Börner et al., 2003) used to map the structure and evolution of numerous subjects based on a set of scholarly articles. Our analysis draws upon two distinct datasets, one that captures a set of self-driving cars related patents from the Clarivate Analytics Derwent World Patent Index (DWPI) database, and another one that identifies related scientific publications from Clarivate Analytics Web of Science (WoS). Within this section, we provide details in the assembly of these datasets and summary statistics for the variables of our sample. To track the development of various fields and sub-fields of the research and innovation on self-driving cars we began by identifying patents and scientific publications by leveraging self-driving cars related synonyms such as "autonomous vehicle" and "intelligent vehicle" and all possible combinations of those terms. Our search strategy is fully reported in Appendix A.

2.1 Patent Sample and Summary Statistics

Our patent analysis focuses on patent documents through the Clarivate Analytics Derwent World Patent Index (DWPI) database which covers the most relevant issuing patent authorities worldwide from 1994 to 2020. We conducted a keyword search utilizing the keywords described in Appendix A[2]. Through the DWPI we were able to gather detailed information about each patent, including filing, application and grant year, topical information, as well as information on inventors and assignees. This search yields 82,005 patent applications grouped in 35,800 patent families[3] from 1994 to 2018[4]. We decided to focus our level of analysis at the

[2] We tried several variations of the aforementioned keywords and alternative algorithmic approaches, but this did not result in a meaningful difference in the patent set.
[3] A Patent Family is a set of patents filed with different patenting authorities that refer to the same invention. DWPI has a strict family definition in which each member shares exact priorities (except for non-convention equivalents) with each other family member. https://support.clarivate.com/Patents/s/article/Derwent-Innovation-DWPI-and-INPADOC-Family-Criteria?language=en_US.
[4] We decided to analyze patents and related patent families up to 2018 to avoid data truncation problems that may occur in analysing extremely recent data.

patent family level as it is defined as a set of patents taken in various countries to protect a single invention (see Martinez, 2011 for a discussion).

According to the DWPI Class codes which are a classification system developed by Clarivate Analytics to enable effective and precise search in a particular area of technology, out of the 82,005 patent applications included in our sample 67,169 (98.38%) are classified in the Engineering category, 42,755 (62.62%) in Instruments Instrumentation, 34,387 (50.37%) in Computer Science, 32,276 (47.27%) in Transportation 21,492 (31.48%) in Telecommunications, and 12,999 (19.041 %) in Automation Control Systems.[5]

2.2 Article Sample and Summary Statistics

Our analysis of scientific publications uses the Web of Science Social Science Citation Index (WoS SSCI) database which contains the world's leading scientific and technical journals from 1900 to 2020.[6] To compare the sample of scientific publication with our patent sample we run the same keyword search as for the sample of patents within the same period (i.e., 1994-2018). As in the case of the patent sample, we were able to gather topical information about each scientific publication, including title, abstract, keywords, source information as well as the author and institutional affiliations. This search yields 3,326 scientific records.

Of the 3,326 scientific articles in the sample, 1,255 (37.74%) are classified in Engineering Electrical Electronic category according to WoS Categories, 1,173 (36,37%) in Computer Science Artificial Intelligence, 582 (17.52%) in Automation Control Systems, 453 (13.64%) in Transportation Science and Technology and 403 (12,12%) in Robotics.

We identify the most productive academic institutions. 654 (1.17%) of scientific articles in our samples are published by the Massachusetts Institute of Technology (MIT), 554 (0.99%) by the Chinese Academy of Science, 551 (0.99%) by the University of California Berkeley, 468 (0.84%) by the University of Tsinghua and 444 (0.79%) by Carnegie Mellon University.

We also differentiate between authors affiliation country 14,999 (27%) are affiliated in an USA based institution, 9,486 (9.48%) in China, 3,948 (7,17%) in Germany, 2,717 (4.90%) in Japan and 2,454 (4.04%) in UK.

[5] DWPI class codes overlap.
[6] https://clarivate.com/webofsciencegroup/solutions/webofscience-ssci/

3. Exploratory empirical analysis

3.1 Patent analysis

3.1.1 Inventive activity

Figure 1 shows the growth in the number of self-driving cars related to patent families. We can see from 1994 to 2012 a steady growth followed by the exponential growth of patenting activity from 2013 onwards. Compared to the average growth rate for the period 1994-2012, which is around 8.8% per year, the average growth rate for the period 2013-2018 is 56.5%.

FIGURE 1 NEAR HERE

To capture the overall rate of inventive activity among firms, we compute an average score accounting for the total number of patent families per year over the total number of unique firms per year. Compared to gross yearly patent family counts that do not account for firm-level heterogeneity, this measure provides a more fine-grained idea of the overall industry inventive activity across time. The trend line is shown in Figure 2 where we observe a steep increase over the period of observation, indicating a greater number of firms engaged in self-driving cars patents and overall increase in the inventive activity.

FIGURE 2 NEAR HERE

3.1.2 Keyword analysis

The DWPI database includes details on International Patent Classification (IPC), a hierarchical indexing system developed by the World Intellectual Property Organization (WIPO) for patent analysis. This indexing system provides some granular detail about the application field for which the patent is intended. Also, the IPC classes allow us to begin examining the hypothesis that the influential trajectories within self-driving car innovation have shifted from hardware-related technologies to more software-related technologies due to the uptake of AI. To this end, we identify the 10 main fields and sub-fields in which self-driving related patents are intended. Table 1 compares the distribution of patents across the fields and sub-fields. Main IPC Classes from 1994 to 2018 encompass G05D (Systems for controlling or regulating non-electric variables); G08 (traffic control systems); B06W (Conjoint control of vehicle sub-units of different type or different function) which define technologies related to the subjects of information and communication technology (ICT) such as methods for controlling position or course, controlling traffic systems for vehicles, marking, sensing, and conveying data, as well as facilitating vehicle navigation. The most striking result is that, compared to the first period (1994-2012) which is characterized by low and steady inventive activities around self-driving cars, the second period (2013-2018) shows an upsurge of the G08G (Traffic Control Systems) and G06K (Recognition of data) IPCs which describe the operation of computers, calculation of data, in combination with the IPCs that describe vehicles in general denoting a turning toward systems for processing and data recognition, systems that can recognize certain patterns, and perform certain calculations for the purpose of self-learning in vehicles. These systems are commonly used for traffic control, control of vehicle position, automatic piloting and navigation.

TABLE 1 NEAR HERE

3.1.3 Inventors' patterns

The idea that the increase of the inventive activity related to self-driving cars is characterized by a major pattern of change driven by different areas of AI could be corroborated by also looking at the landscape patent assignees or inventors. Table 2 reports the list of the top 20 inventors in self-driving cars related technologies for the period 1994-2012 and the period 2013-2018. The most patent-active companies in the two periods are different as far as their sector of reference is concerned. While in the period 1994-2012, technological leadership is held by vehicle and hardware companies, the second period (2013-2018) is characterized by the emergence of data-driven and platform companies as Google, IBM, Qualcomm, Uber and Valeo. The influence of data within the landscape of self-driving cars can also be seen in recent arrangements between traditional vehicle manufacturers and software companies, as with General Motors' acquisition of Cruise Automation in 2016 and Ford acquiring SAIPS in 2016 to specialise in self-driving machine learning and computer vision technologies. Table 2 provides information on the geographical localization of the leading patent assignees, showing an increasing concentration of activities in the United States and the emergence of South Korea as a technological leader, with Japan losing its position.

TABLE 2 NEAR HERE

To measure an industry's market concentration, we use a standard measure as the Herfindahl–Hirschman Index (HHI) (Martinelli et al., 2020). The closer a market is to a monopoly, the higher the market's concentration (and the lower its competition). If thousands of small firms were competing, each would have nearly 0% market share, and the HHI would be close to zero, indicating nearly perfect competition. Figure 3 shows an increasing concentration in the number of patents related to self-driving cars since 2015. Those results are in line with the results reported in Table 1 since the shift from more hardware-related technologies to more software-related technologies mirrors the shift from a typically less concentrated industry to a more concentrated industry.

FIGURE 3 NEAR HERE

3.2 Scientific publication analysis

3.2.1 Scientific productivity

Figure 4 shows the growth over time of the 3,326 published scientific articles relevant to self-driving cars. It indicates that the observation period 1994–2018 can be divided into three general phases: a low and steady phase from 1994 to 2006 during which no more than ten papers were published annually, an early emergence period from 2007 to 2012 characterized by ups and downs and a period of ferment, or "take-off phase", which started in 2013. Since that year, scholarly activity dealing with self-driving cars has been growing exponentially, hitting almost 480 published articles in 2018.

FIGURE 4 NEAR HERE

3.2.2 Keywords analysis

Our bibliometric analysis starts with the identification of keywords included in our sample of scientific articles. Bibliographic records, as keywords, are useful elements of scientific work (Small et al., 1985) and, recently, have been acknowledged as reliable methods to study the scope and dynamics of scientific fields in technology policy and innovation studies (e.g., Bai et al., 2020; van der Have & Rubalcaba 2016; Biggi & Giuliani, 2021). The use of keywords facilitates more objective and reliable analyses and provides a "big picture" of extant research by revealing the structural and dynamic aspects of scientific research (Noyons et al., 1999; Börner et al., 2003). To perform a fine-grained analysis on the importance of AI in self-driving research, we separately analysed each of the three phases identified in Table 3. Table 3 shows the 20 most used keywords for research on self-driving cars for each of the identified phases (i.e., 1994-2006, 2007-2012, 2013-2018) and their relative importance against the total number of keywords used in each phase.

The first phase (1994-2006) is characterized by a diversity of keywords. For the period 2007-2012, we observe an upsurge of terms related to processes and systems for the gathering, elaboration and transmission of data for AI and automation in vehicles, such as "data mining", "machine learning", "deep learning", "intelligent speed adaptation", and "intelligent transportation system".[7] This pattern of research suggests that the three much-publicised DARPA Grand Challenges in 2004, 2005 and 2007 certainly marked and probably contributed to a renewed interest in self-driving car AI and robotics research and development (R&D) (Lima et al., 2015). The third period (2013-2018), which we label as the 'take off phase' of self-driving car research, shows two main patterns. First, we observed a steep increase in the use of terms such as "data mining", "machine learning", "artificial neural networks" and "deep learning", suggesting consolidation of interest in self-driving as an application, and a prominent test case, for artificial intelligence. It is important to note that, especially in recent years, the use of such keywords has attracted considerable research attention due to the extensive hype

[7] In the process of keyword identification, we decided to cluster the terms "data mining", "machine learning", "artificial neural networks" and "deep learning" as they refer to the processes of data gathering, processing and transmission for the purposes of self-driving cars.

in the media (Naudé, 2021) allowing many companies to use AI terms to repackage their activities (Brennen, 2018)). Second, we observe the increasing frequency of keywords as "safety", "traffic safety", "injury", "pedestrian", and "driver distraction", which reflects both the justification and problematization of self-driving car technologies. A typical argument for the justification for deploying self-driving cars is that, by being removing human fallibility, they will be much safer in terms of preventing crashes. Concurrently, self-driving cars have been problematized as robots in public spaces, raising complex questions of safety assurance, human-machine interaction, and legal accountability. There has been a particular interest in the ethical dilemmas that self-driving cars might introduce (Lin, 2016). Most publications on the ethics of AVs have taken a very narrow view of ethics, treating self-driving cars as a case study for the application of practical ethics thought experiments, rather than offering analysis that would be of use to policymakers or engineers (JafariNaimi, 2018).

TABLE 3 NEAR HERE

To dynamically identify emergent research-front concepts, Figures 5 show the keywords' bursts, according to Kleinberg's burst detection algorithm (Kleinberg, 2003). The use of the burst-detection algorithm is particularly suited for detecting sharp increases of interest within a topic (Chen & Huang, 2012). Figure 5 shows the top 15 keywords with the strongest citation bursts over the period 1994 – 2018. The first column displays the keywords, the second and the third column display the initial and final burst period, and the third column shows the burst strength found by Kleinberg's (2002) algorithm. The strongest keyword bursts are associated with "neural network", "decision support system" and "heuristics". Neural networks have become a particularly important technology for recent self-driving R&D. From 2008 onwards, benefitting from computer power offered by new Graphics Processing Units, neural networks have shown dramatic increases in the performance of a type of AI called "deep learning" (LeCun et al., 2015) that is used for image and voice recognition in technologies that have quickly become ubiquitous. For self-driving cars, many companies rely on neural networks for the computer vision that processes signals from sensors as well as predicting the movements of other road users and planning a safe route. Neural networks have also been problematised for their opacity (it is often hard to know and to trace why a neural network produces a particular output for a given set of inputs), which presents issues for self-driving in the event of crash investigations and safety improvement (Stilgoe, 2018).

FIGURE 5 NEAR HERE

3.2.3 Scientific and public terminology of self-driving cars

Finally, we analyse the terminology used in research on self-driving cars (Figures 6 and 7). For the article sample, we see the increasing use of "autonomous vehicles" as the descriptor for the technology. One might expect this term, rooted in a precise engineering usage of "autonomy", to be a term of art, but we see from Google Trends that it has recently become the dominant term in wider use too. The Google Trends comparison of many commonly used terms suggests an earlier dominance of "driverless cars" in common parlance, a term which is seldom used in the scientific literature. The convergence between scientific and public terminology is worthy of note. Perhaps the narrative of autonomy is persuasive for the public as well as for researchers and innovators.

FIGURE 6 NEAR HERE

FIGURE 7 NEAR HERE

4. Concluding remarks

Our results show rapid growth in patenting and research activity relevant to self-driving cars. We see shifts away from dominant automobile innovators. We also see the clear and growing importance of AI, data and software in self-driving car R&D. Scientific publications are increasingly likely to define themselves according to self-driving cars. In terms of the sociology of expectations, we may be witnessing some self-fulfilling prophecies (Borup et al., 2013). However, we should recognise that much knowledge and inventive activity that will be relevant for self-driving innovation may not be represented in these data. Research may be kept private and intellectual property may be protected through trade secrets rather than patents or open-sourced (Searle, 2020; Calvin & Leung 2020).

Our findings resonate with analyses of AI papers that suggest a narrowing of research (Klinger et al 2020). Much AI research seems to fit a narrative of a "race", with a focus on large datasets and computationally intensive processes (Klinger et al., 2020). If self-driving R&D follows the pattern of "racing" to develop "the" technology, questions of what forms the technology might take and what purposes it might serve may not get asked until commitments and policy positions are set. In this paper, we have attempted to ask what our data mean for the project of reconnecting technological means with ends.

Making sense of these findings as indicators or early warnings of the social constitution of self-driving cars is not straightforward. We conclude that the data reveal a sense not just of the quantity of innovation, but of its quality and direction. We see a concentration of research and innovation on AI and data, which are likely to be key components of systems. But eventual self-driving car systems are likely to also involve infrastructure changes and be constructions as much of policy as of technology. This will demand a breadth of expertise, including local expertise and tacit knowledge. The intellectual property on future locally embedded systems will likely be hard to capture and hard to measure.

For technology developers, the focus on AI is understandable, as this represents a part of the system over which they have direct control, and which might enable competitive advantage. Researchers, however, even those disconnected from corporate interests, seem to be following this narrative of autonomy (Tennant & Stilgoe, forthcoming). The narrative of autonomy may invite hype, and be a poor guide for policy, but it may nevertheless contribute to a de facto governance of the technology and forms of sociotechnical. lock-in (Arthur, 1989). As technology developers realize the limits of AI in mimicking and exceeding the capabilities of human drivers, they could look to draw connections with other parts of the sociotechnical system. Laws, infrastructures etc may come under pressure to accommodate new technologies to which policymakers are already committed. The issues that receive policy attention may be those around which there is already a body of research. We have seen that questions of safety and ethics have received more attention than those of equity, access and future mobility systems (this finding is similar to that of Kassens-Noor and colleagues (2020) in their qualitative review of AV social research). Questions of the politics of technology are, as is often the case, postponed until the technologies are visible (see Guston, 2014 on anticipatory governance).

The data hint at a mode of technological change that configures self-driving as a software problem more than a hardware one. If policymakers want to avoid repeating some of the dynamics of digital platforms, in which developers can scale first and ask questions later, they should understand that current patterns of intellectual property and research may not reflect an eventual sociotechnical reality and should look to diversify research and innovation. They should pay attention to possible concentrations of intellectual property that might lead to oligopolistic market structures.

A comprehensive analysis of innovation requires also the exploration of the other side of the coin in innovation studies (Coad et al., 2021; Coad, Biggi & Giuliani, 2021; Biggi & Giuliani, 2021), where innovation is often conceptualized as an engine of economic growth and prosperity (Mulgan, 2016). Research aimed at leveraged the use of bibliometric or scientometric techniques to investigate antecedents, dynamics, and consequences of a scientific and or technical field in terms of scientific and technical outputs (e.g. in terms of patent count, articles count, etc.) often overlooking wider social consequences and governance aspects. In this article, besides providing a landscape of the research and innovation around self-driving cars in relation to the impact of the different areas of AI and we shed light on important questions about where science and technology are directed and who is benefitting, with a view to improving the (anticipatory) governance of emerging technology. The attention to the limits and potential downsides of science and technology is timely based on the current unique historical moment characterized by the increased public scrutiny and criticism attracted by emerging technologies.

Finally, we would highlight some limitations of this study. Although the use of scientometrics and bibliometric techniques should reduce subjective bias (Kovacs et al., 2015), our work should be interpreted with some caution. First, we acknowledge that we have opted for a top-down selection of keywords, which may have led to the exclusion of some technical and scientific work that may come to be important for some types of self-driving mobility. We hope that future research will be able to identify new terminologies to continue mapping scientific and technical contributions and their evolution over time. Second, there are considerable drawbacks in using patent data to measure technical efforts (see e.g., Debackere, 2000). With respect to this study, one important limitation is the relates to the underestimation of the magnitude of the research and innovation in the software which cannot be entirely captured through patent data (Hullmann & Meyer, 2003). Third, the use of databases, such as Derwent World Patent Index and Web of Science, might have introduced a bias towards technical and scientific outputs by being more restrictive than other databases (see, e.g., Google Patents or ArXiv.org), may have underestimated the extent of the research and innovation in each subcommunity or did not appreciate the emergence of new topics which are yet be included in the Derwent/Web of Science indexed patents and articles.

5. References

Arthur, W. B. (1989). Competing Technologies, Increasing Returns, and Lock-In by Historical Events. *Economic Journal, 99*(394), 116-131. https://EconPapers.repec.org/RePEc:ecj:econjl:v:99:y:1989:i:394:p:116-31

Bai, X., Wu, J., Liu, Y., & Xu, Y. (2020, 2020/08/01). Research on the impact of global innovation network on 3D printing industry performance. *Scientometrics, 124*(2), 1015-1051. https://doi.org/10.1007/s11192-020-03534-1

Biggi, G., & Giuliani, E. (2021). The noxious consequences of innovation: what do we know? Industry and Innovation, 28(1), 19-41. https://doi.org/10.1080/13662716.2020.1726729

Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology, 37*(1), 179-255. https://doi.org/https://doi.org/10.1002/aris.1440370106

Borup, M., Brown, N., Konrad, K., & Van Lente, H. (2006, 2006/07/01). The sociology of expectations in science and technology. *Technology Analysis & Strategic Management*, 18(3-4), 285-298. https://doi.org/10.1080/09537320600777002

Brennen, J. (2018). An industry-led debate: How UK media cover artificial intelligence.

Burns, L. D., & Shulgan, C. (2018). *Autonomy: The Quest to Build the Driverless Car—And How It Will Reshape Our World*. HarperCollins.

Chen, C.-Y., & Huang, P.-H. (2012). RETRACTED: Review of an autonomous humanoid robot and its mechanical control. *Journal of Vibration and Control, 18*(7), 973-982. https://doi.org/10.1177/1077546310395974

Cho, R. L.-T., Liu, J. S., & Ho, M. H.-C. (2021). The development of autonomous driving technology: perspectives from patent citation analysis. *Transport Reviews*, 1-27. https://doi.org/10.1080/01441647.2021.1879310

Coad, A., Biggi, G., & Giuliani, E. (2021). Asbestos, leaded petrol, and other aberrations: comparing countries' regulatory responses to disapproved products and technologies. *Industry and Innovation*, 28(2), 201-233. https://doi.org/10.1080/13662716.2020.1830041

Coad, A., Nightingale, P., Stilgoe, J., & Vezzani, A. (2021). Editorial: the dark side of innovation. *Industry and Innovation*, 28(1), 102-112. https://doi.org/10.1080/13662716.2020.1818555

Debackere, K. (2000). Managing academic R&D as a business at K.U. Leuven: context, structure and process. *R&d Management, 30*(4), 323-328. https://doi.org/https://doi.org/10.1111/1467-9310.00186

Dosi, G. (1982). Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change. *Research Policy, 11*(3), 147-162.

Gandia, R. M., Antonialli, F., Cavazza, B. H., Neto, A. M., Lima, D. A. d., Sugano, J. Y., Nicolai, I., & Zambalde, A. L. (2019, 2019/01/02). Autonomous vehicles: scientometric and bibliometric review. *Transport Reviews, 39*(1), 9-28. https://doi.org/10.1080/01441647.2018.1518937

Guston, D. H. (2014). Understanding 'anticipatory governance'. *Social studies of science, 44*(2), 218-242.

Hilgartner, S., Miller, C., & Hagendijk, R. (2015). *Science and Democracy; Making Knowledge and Making Power in the Biosciences and Beyond*.

Hullmann, A., & Meyer, M. (2003, 01 Nov. 2003). Publications and patents in nanotechnology. *Scientometrics, 58*(3), 507. https://doi.org/10.1023/b:scie.0000006877.45467.a7

JafariNaimi, N. (2018). Our Bodies in the Trolley's Path, or Why Self-driving Cars Must *Not* Be Programmed to Kill. *Science, Technology, & Human Values, 43*(2), 302-323. https://doi.org/10.1177/0162243917718942

Kassens-Noor, E., Wilson, M., Cai, M., Durst, N., & Decaminada, T. (2020). Autonomous vs. Self-Driving Vehicles: The Power of Language to Shape Public Perceptions. *Journal of Urban Technology*, 1-20. https://doi.org/10.1080/10630732.2020.1847983

Kleinberg, J. (2003, 2003/10/01). Bursty and Hierarchical Structure in Streams. *Data Mining and Knowledge Discovery, 7*(4), 373-397. https://doi.org/10.1023/A:1024940629314

Klinger, J., Mateos-Garcia, J., & Stathoulopoulos, K. (2020). A narrowing of AI research? *arXiv preprint arXiv:2009.10385*.

Kovács, A., Looy, B. V., & Cassiman, B. (2015, /). Exploring the scope of open innovation: a bibliometric review of a decade of research. *Scientometrics, 104*(3), 951-983. https://doi.org/10.1007/s11192-015-1628-0

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.

LeCun, Y., Bengio, Y., & Hinton, G. (2015, 2015/05/01). Deep learning. *Nature, 521*(7553), 436-444. https://doi.org/10.1038/nature14539

Lima, P. F., Trincavelli, M., Mårtensson, J., & Wahlberg, B. (2015, 15-17 July 2015). Clothoid-based model predictive control for autonomous driving. 2015 European Control Conference (ECC),

Lin, P. (2016). Why Ethics Matters for Autonomous Cars. In M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Eds.), *Autonomous Driving: Technical, Legal and Social Aspects* (pp. 69-85). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-48847-8_4

Martinelli, A., Mina, A., & Moggi, M. (2020). The enabling technologies of industry 4.0: Examining the seeds of the fourth industrial revolution. *Industrial and Corporate Change, Forthcoming.*

Martínez, C. (2011). Patent families: When do different definitions really matter? [Article]. *Scientometrics, 86*(1), 39-63. https://doi.org/10.1007/s11192-010-0251-3

Mulgan, G. (2016). Good and bad innovation: what kind of theory and practice do we need to distinguish them. *London: Nesta*.

Narin, F. (1994, 1994/05/01). Patent bibliometrics. *Scientometrics, 30*(1), 147-155. https://doi.org/10.1007/BF02017219

Narin, F., & Hamilton, K. S. (1996, 1996/07/01). Bibliometric performance measures. *Scientometrics, 36*(3), 293-310. https://doi.org/10.1007/BF02129596

Naudé, W. (2021, 2021/01/02). Artificial intelligence: neither Utopian nor apocalyptic impacts soon. *Economics of Innovation and new technology, 30*(1), 1-23. https://doi.org/10.1080/10438599.2020.1839173

Noyons, E. C. M., Moed, H. F., & Luwel, M. (1999). Combining mapping and citation analysis for evaluative bibliometric purposes: A bibliometric study. *Journal of the American Society for Information Science, 50*(2), 115-131. https://doi.org/https://doi.org/10.1002/(SICI)1097-4571(1999)50:2<115::AID-ASI3>3.0.CO;2-J

Pasquale, F. (2016). Two narratives of platform capitalism. *Yale L. & Pol'y Rev., 35*, 309.

Pollack, J., & Adler, D. (2015, 2015/01/01/). Emergent trends and passing fads in project management research: A scientometric analysis of changes in the field. *International Journal of Project Management, 33*(1), 236-248. https://doi.org/https://doi.org/10.1016/j.ijproman.2014.04.011

Small, H., Sweeney, E., & Greenlee, E. (1985, 1985/11/01). Clustering the science citation index using co-citations. II. Mapping science. *Scientometrics, 8*(5), 321-340. https://doi.org/10.1007/BF02018057

Stilgoe, J. (2018). Machine learning, social learning and the governance of self-driving cars. *Social studies of science, 48*(1), 25-56. https://doi.org/10.1177/0306312717741687

Szerszynski, B., Kearnes, M., Macnaghten, P., Owen, R., & Stilgoe, J. (2013). Why Solar Radiation Management Geoengineering and Democracy Won't Mix. *Environment and Planning A: Economy and Space*, 45(12), 2809-2816. https://doi.org/10.1068/a45649

Tennant, G., & Stilgoe, J. Forthcoming. The Attachments of 'Autonomous' Vehicles

van der Have, R. P., & Rubalcaba, L. (2016). Social innovation research: An emerging area of innovation studies? *Research Policy, 45*(9), 1923-1935. https://EconPapers.repec.org/RePEc:eee:respol:v:45:y:2016:i:9:p:1923-1935

Vinsel, L. (2019). American adventure and autonomous cars. *Science, 365*(6449), 130-131. https://doi.org/10.1126/science.aax8083

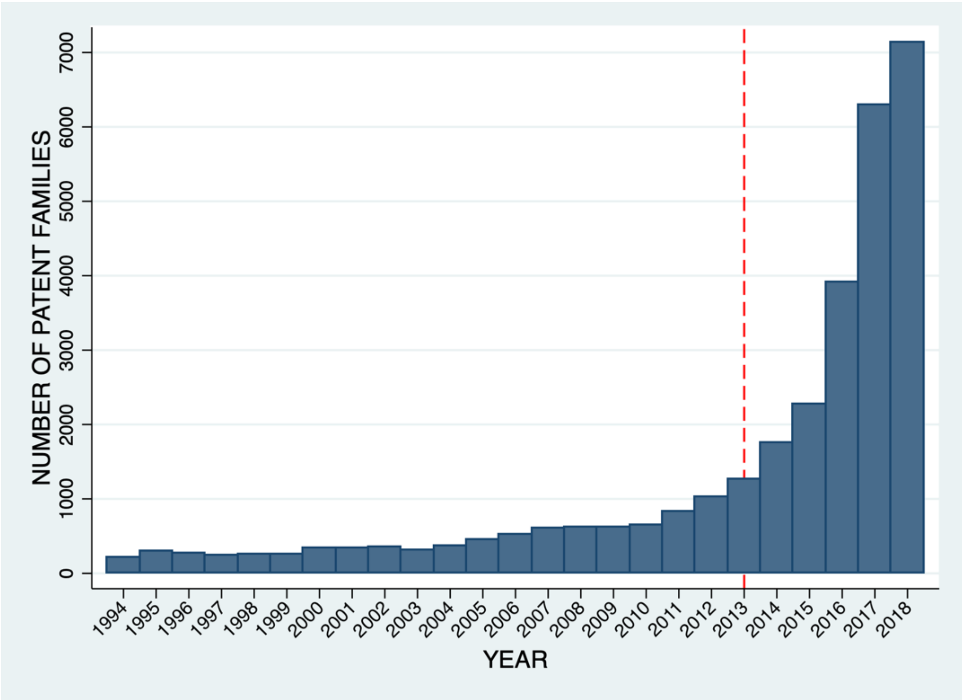
Wetmore, J. M. (2020). Reflecting on the Dream of Automated Vehicles. Visions of Hands Free Driving over the past 80 years / Überlegungen zum Traum vom automatisierten Fahrzeug. Visionen selbsttätigen Fahrens aus den letzten 80 Jahren. *TG Technikgeschichte, 87*(1), 69-94. https://doi.org/10.5771/0040-117X-2020-1-69

Appendix A – Search Strategy

TS=("data mining" AND van) OR TS=("data mining" AND truck) OR TS=("data mining" AND vehicle) OR TS=("data mining" AND automobile) OR TS=("data mining" AND car) OR TS=("support vector machine" AND van) OR TS=("support vector machine" AND truck) OR TS=("support vector machine" AND vehicle) OR TS=("support vector machine" AND automobile) OR TS=("support vector machine" AND car) OR TS=("SMV" AND van) OR TS=("SMV" AND truck) OR TS=("SMV" AND vehicle) OR TS=("SMV" AND automobile) OR TS=("SMV" AND automobiles) OR TS=("SMV" AND car) OR TS=("supervised learning" AND van) OR TS=("supervised learning" AND truck) OR TS=("supervised learning" AND vehicle) OR TS=("supervised learning" AND automobile) OR TS=("supervised learning" AND car) OR TS=("neural network" AND van) OR TS=("neural network" AND truck) OR TS=("neural network" AND automobile) OR TS=("neural network" AND "car") OR TS=("deep learning" AND van*) OR TS=("deep learning" AND truck*) OR TS=("deep learning" AND vehicle*) OR TS=("deep learning" AND automobile*) OR TS=("deep learning" AND car) OR TS=("machine learning" AND van*) OR TS=("machine learning" AND truck*) OR TS=("machine learning" AND vehicle*) OR TS=("machine learning" AND automobile*) OR TS=("artificial intelligence" AND vehicle*) OR TS=("artificial intelligence" AND car) OR TS=("machine learning" AND car) OR TS=(artificial intelligence AND van*) OR TS=(artificial intelligence AND truck*) OR TS=(artificial intelligence AND automobile*) OR TS=("robo* van*") OR TS=("robo* truck*") OR TS=("robo* vehicle*") OR TS=("robo* automobile*") OR TS=("robo* car") OR TS=("connected van*") OR TS=("connected truck*") OR TS=("connected vehicle*") OR TS=("connected automobile*") OR TS=("connected car") OR TS=(self-driving AND van*) OR TS=(self-driving AND truck*) OR TS=(self-driving AND vehicle*) OR TS=(self-driving AND automobile*) OR TS=(self-driving AND motorcar*) OR TS=(self-driving AND car) OR TS=(autonomous AND van*) OR TS=(autonomous AND truck*) OR TS=(autonomous AND vehicle*) OR TS=(autonomous AND automobile*) OR TS=(autonomous AND car) OR TS=(driverless AND van*) OR TS=(driverless AND truck*) OR TS=(driverless AND vehicle*) OR TS=(driverless AND automobile*) OR TS=(driverless AND motorcar*) OR TS=(driverless AND car) OR TS=(automated AND van*) OR TS=(automated AND truck*) OR TS=(automated AND vehicle*) OR TS=(automated AND automobile*) OR TS=(automated AND motorcar) OR TS=(intelligent AND van*) OR TS=(intelligent AND truck*) OR TS=(intelligent AND vehicle*) OR TS=(intelligent AND automobile*) OR TS=(automated AND motorcar)

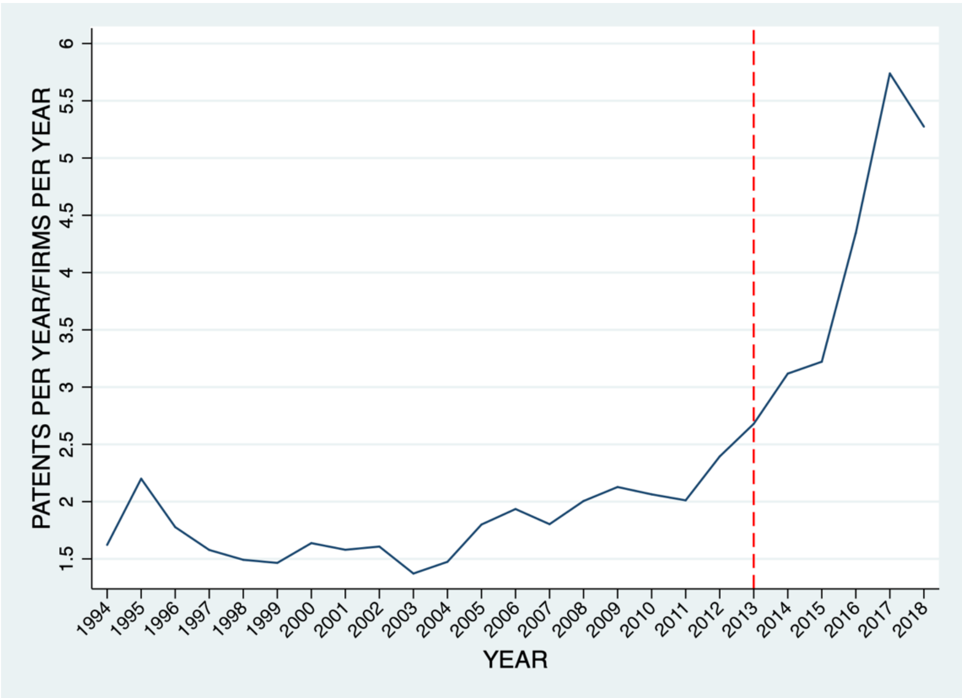
LIST OF FIGURES

Figure 1: Growth of self-driving cars patent families



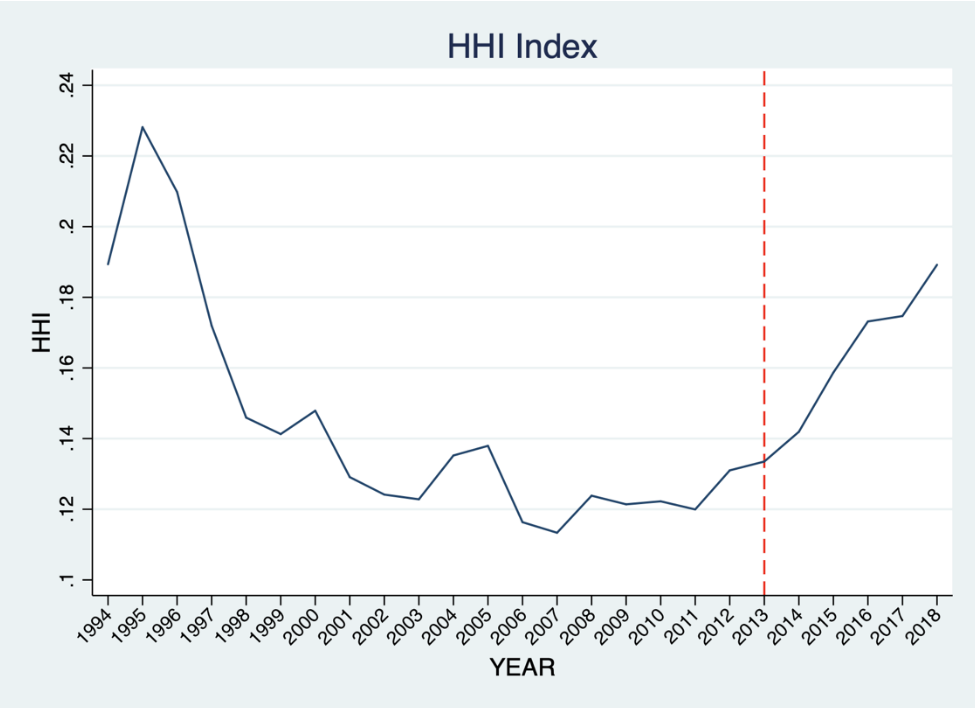
Source: Authors' calculations

Figure 2: patents per year/firms per year



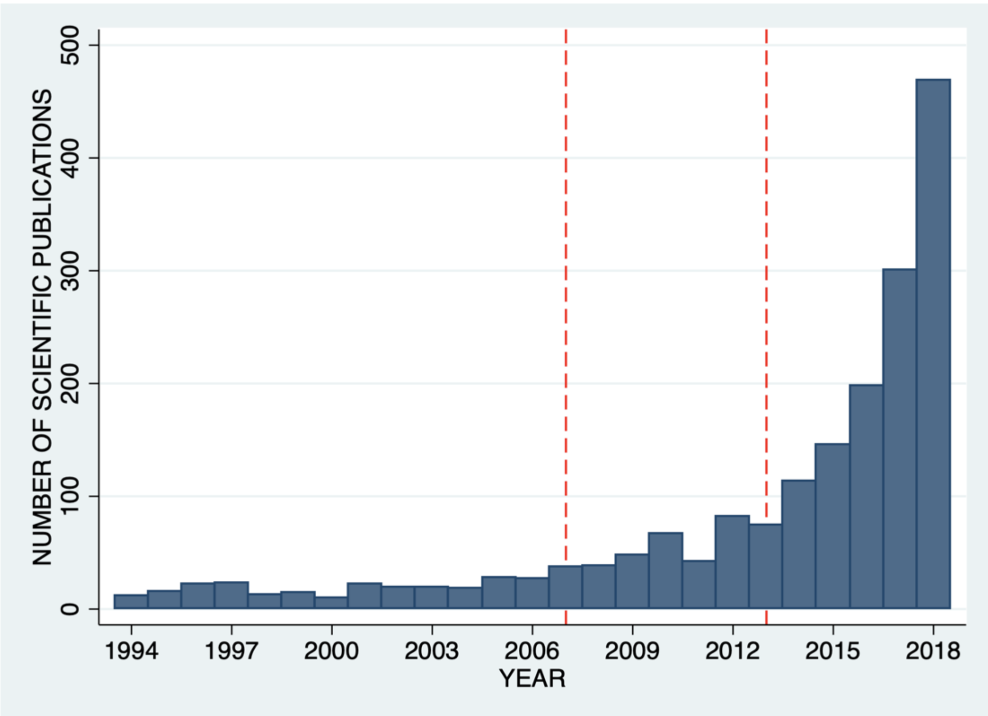
Source: Authors' calculations

Figure 3: Evolution of the concentration of innovative activities across patent assignees



Source: Authors' calculations

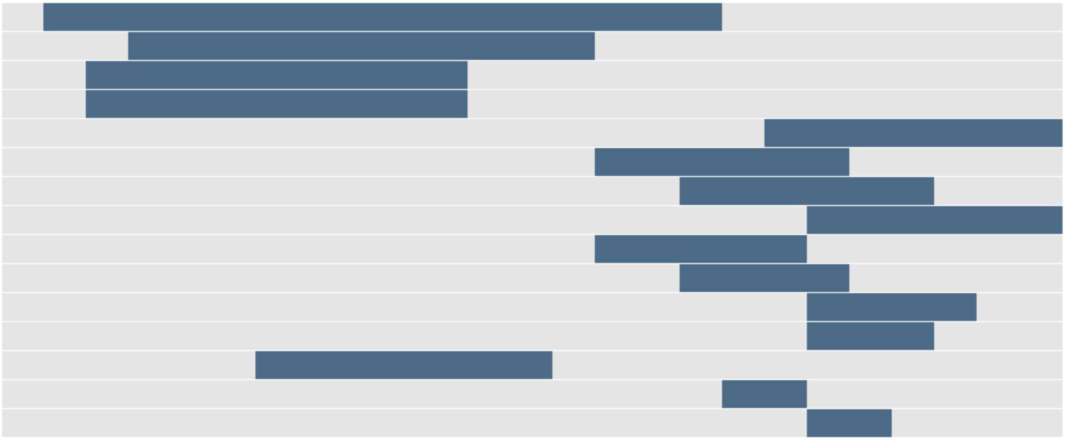
Figure 4: Growth of self-driving car scientific publications



Source: Authors' calculations

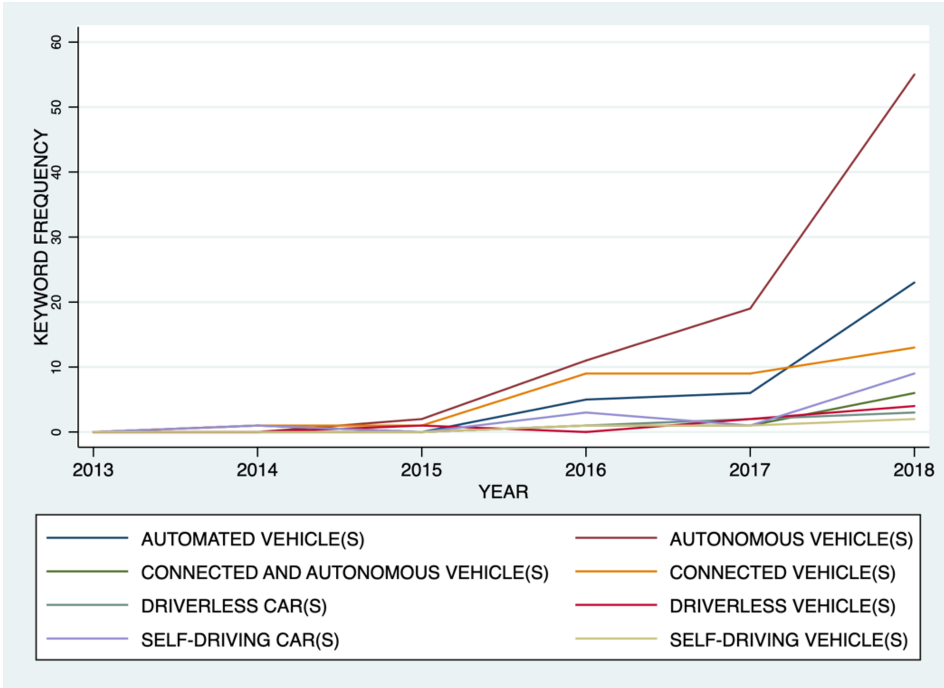
Figure 5: Top 15 keywords with the strongest citation bursts over time (1994 – 2018)

KEYWORD	BEGIN	END	STRENGHT
NEURAL NETWORK	1995	2010	3,3652
DECISION SUPPORT SYSTEM	1997	2007	1,7664
AUTOMATED GUIDED VEHICLE	1996	2004	2,8356
HEURISTICS	1996	2004	2,8356
DRIVER BEHAVIOR	2012	2018	3,1756
CRASH	2008	2013	1,6593
ACTIVE SAFETY	2010	2015	4,1100
INJURY	2013	2018	3,0405
CAR OWNERSHIP	2008	2012	2,3321
AUTONOMOUS BRAKING	2010	2013	1,7636
COLLISION MITIGATION	2013	2016	1,8116
INTELLIGENT SPEED ADAPTATION	2012	2014	2,9121
ACCIDENT	2005	2006	2,5249
DATA MINING	2011	2012	3,3778
ADVANCED DRIVER ASSISTANCE SYSTEM	2013	2014	2,4628



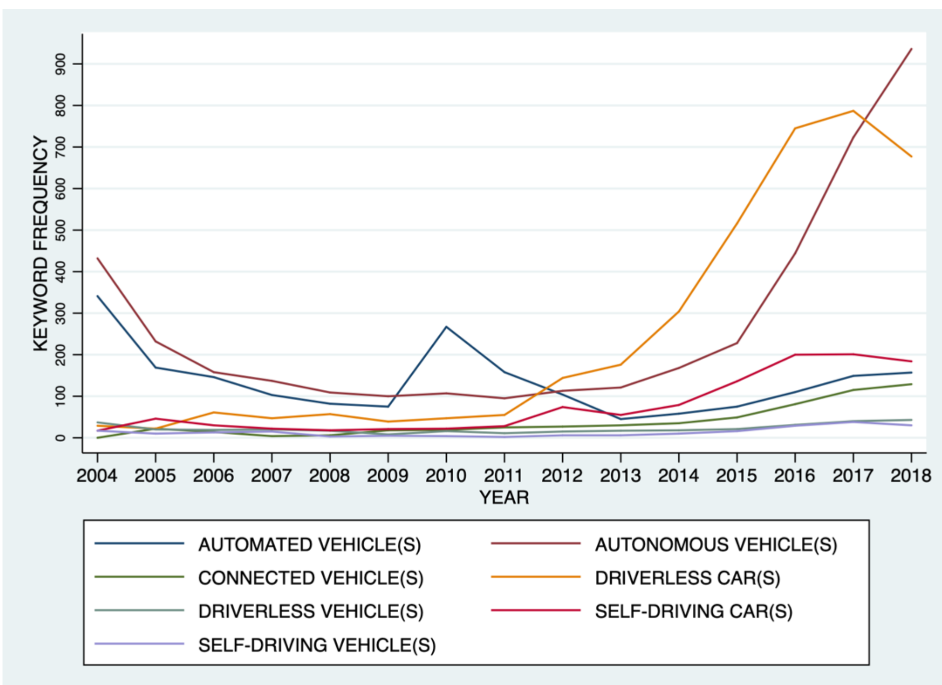
Source: Authors' calculations

Figure 6: Evolution of Self-Driving car related synonyms over time (Article sample)



Source: Authors' calculations

Figure 7: Evolution of Self-Driving car related synonyms over time (Google Trends)



Source: Authors' calculations

Table 1: Distribution of patents across main IPC Classes (1994 – 2012 and 2013 – 2018)

1994 - 2012				2013 - 2018			
IPC Class	**IPC Sub-class**	**Description**	**Number of patent families**	**IPC Class**	**IPC Sub-class**	**Description**	**Number of patent families**
G05 - CONTROLLING; REGULATING	G05D	SYSTEMS FOR CONTROLLING OR REGULATING NON-ELECTRIC VARIABLES	1131	G05 - CONTROLLING; REGULATING	G05D	SYSTEMS FOR CONTROLLING OR REGULATING NON-ELECTRIC VARIABLES	6145
G06 - COMPUTING; CALCULATING OR COUNTING	G06F	ELECTRIC DIGITAL DATA PROCESSING	1087	B60 - VEHICLES IN GENERAL	B60W	CONJOINT CONTROL OF VEHICLE SUB-UNITS OF DIFFERENT TYPE OR DIFFERENT FUNCTION	4147
G01 - MEASURING; TESTING	G01C	PHOTOGRAMMETRY OR VIDEOGRAMMETRY	728	G08 - SIGNALLING	G08G	TRAFFIC CONTROL SYSTEMS	3336
G08 - SIGNALLING	G08G	TRAFFIC CONTROL SYSTEMS	591	G06 - COMPUTING; CALCULATING OR COUNTING	G06K	RECOGNITION OF DATA	2812
B62 - LAND VEHICLES FOR TRAVELLING OTHERWISE THAN ON RAILS	B62D	MOTOR VEHICLES; TRAILERS	465	G01 - MEASURING; TESTING	G01S	DETERMINING DISTANCE OR VELOCITY BY USE OF RADIO WAVES	2019
B25 - MANIPULATORS	B25J	MANIPULATORS: CHAMBERS PROVIDED WITH MANIPULATION DEVICES	389	H04 - ELECTRIC COMMUNICATION TECHNIQUE	H04W	WIRELESS COMMUNICATION NETWORKS	1420
B06 - VEHICLES IN GENERAL	B60R	VEHICLES, VEHICLE FITTINGS, OR VEHICLE PARTS, NOT OTHERWISE PROVIDED FOR	342	B62 - LAND VEHICLES FOR TRAVELLING OTHERWISE THAN ON RAILS	B62D	MOTOR VEHICLES; TRAILERS	1323

Table 1 (Continued): Distribution of patents across main IPC Classes (1994 – 2012 and 2013 – 2018)

1994 - 2012				2013 - 2018			
IPC Class	**IPC Sub-class**	**Description**	**Number of patent families**	**IPC Class**	**IPC Sub-class**	**Description**	**Number of patent families**
H04 - ELECTRIC COMMUNICATION TECHNIQUE	H04L	TRANSMISSION OF DIGITAL INFORMATION	275	H04 - ELECTRIC COMMUNICATION TECHNIQUE	H04N	PICTORIAL COMMUNICATION	844
B65 - CONVEYING; PACKING; STORING; HANDLING THIN OR FILAMENTARY MATERIAL	B65G	TRANSPORT OR STORAGE DEVICES	247	B64 - AIRCRAFT; AVIATION; COSMONAUTICS	B64C	AEROPLANES; HELICOPTERS	783
H02 - GENERATION, CONVERSION, OR DISTRIBUTION OF ELECTRIC POWER	H02J	SYSTEMS FOR STORING ELECTRIC ENERGY	235	B25 - MANIPULATORS	B25J	CHAMBERS PROVIDED WITH MANIPULATION DEVICES	734

Source: Authors' calculation

Table 2: Top inventors over time (1994 – 2012 and 2013 – 2018)

1994 - 2012				2013 - 2018			
Assignee name	**Country**	**Number of patent families**	**Share over total patent families**	**Assignee name**	**Country**	**Number of patent families**	**Share over total patent families**
TOYOTA	JP	127	1,48%	FORD	US	621	2,74%
DAIMLER	DE	104	1,21%	TOYOTA	JP	512	2,26%
BOSCH	DE	93	1,08%	GENERAL MOTORS	US	427	1,89%
GOOGLE	US	85	0,99%	UBER	US	369	1,63%
SIEMENS	DE	82	0,95%	BOSCH	DE	356	1,57%
PANASONIC	JP	78	0,91%	DAIMLER	DE	342	1,51%
HONDA MOTOR	JP	73	0,85%	SAMSUNG	KR	328	1,45%
IROBOT	US	72	0,84%	IBM	US	322	1,42%
IHI	JP	70	0,81%	GOOGLE	US	291	1,29%
BOEING	US	69	0,80%	BAIDU	CN	282	1,25%
GENERAL MOTORS	US	68	0,79%	INTEL	US	222	0,98%
MITSUBISHI	JP	60	0,70%	HYUNDAI	KR	198	0,87%
HONEYWELL	US	58	0,67%	QUALCOMM	US	172	0,76%
US SEC OF NAVY	US	56	0,65%	BMW	DE	160	0,71%
HITACHI	JP	55	0,64%	AUDI	DE	153	0,68%
TOSHIBA	JP	54	0,63%	VALEO	FR	150	0,66%
CATERPILLAR	US	53	0,62%	HONDA	JP	134	0,59%
JERVIS WEBB COMPANY	US	52	0,60%	LG ELECTRONICS	KR	132	0,58%
KOMATSU	JP	50	0,58%	YANMAR	JP	111	0,49%
SONY	JP	49	0,57%	HERE GLOBAL	NL	110	0,49%

Source: Authors' calculation

Table 3: Keyword dynamics by periods (1994-2006, 2007-2012, 2013 - 2019)

1994 - 2006			2007 - 2012			2013 - 2018		
Keyword	**Frequency**	**Share over the total**	**Keyword**	**Frequency**	**Share over the total**	**Keyword**	**Frequency**	**Share over the total**
Data mining; AI; Artificial neural networks; Machine learning; Deep learning	22	3,54%	Data mining; AI; Artificial neural networks; Machine learning; Deep learning	33	1,56%	Data mining; AI; Artificial neural networks; Machine learning; Deep learning	315	7,65%
Intelligent transportation systems	10	1,61%	Intelligent transportation systems	14	0,66%	Autonomous vehicle; Connected vehicle; Self driving car	90	2,19%
Simulation	6	0,96%	Automation	8	0,38%	Intelligent transportation systems	26	0,63%
Automated guided vehicles; Intelligent vehicles	6	0,96%	Intelligent speed adaptation	8	0,38%	Driving simulator	24	0,58%
Heuristics	5	0,80%	Simulation	7	0,33%	Safety	23	0,56%
Classification and regression trees cart	4	0,64%	Support vector machine (SVM)	6	0,28%	Road safety; Traffic Safety	18	0,44%
Accidents	3	0,48%	Driver behaviour	5	0,24%	Adaptive cruise control	15	0,36%
Decision support system	3	0,48%	Traffic safety; Road safety	5	0,24%	Human automation interaction	15	0,36%
Driver behaviour	3	0,48%	Car ownership	4	0,19%	Autonomous emergency braking	11	0,27%
Human factors	3	0,48%	Optimization	4	0,19%	Injury	11	0,27%
Intelligent speed adaptation	3	0,48%	Decision support systems	3	0,14%	Pedestrian	11	0,27%
Risk assessment	3	0,48%	Distraction	3	0,14%	Public transport	11	0,27%
Traffic safety	3	0,48%	Driving performance	3	0,14%	Sustainability	11	0,27%
Case based reasoning	2	0,32%	Electric vehicles	3	0,14%	Driver distraction	9	0,22%
Geographical information systems (GIS)	2	0,32%	Human factors	3	0,14%	Driving performance	9	0,22%
Injury severity	2	0,32%	Mathematical programming	3	0,14%	Electric vehicles	9	0,22%
Integer programming	2	0,32%	Motor vehicle crashes	3	0,14%	Ethics	9	0,22%
Inventory control	2	0,32%	Pedestrian	3	0,14%	Situation awareness	9	0,22%
Optimisation	2	0,32%	Photo radar	3	0,14%	Smart cities	9	0,22%
Simulated annealing	2	0,32%	Speed	3	0,14%	Eye tracking	7	0,17%

Source: Authors' calculation