7.2. Artificial Intelligence and Mobility: What Is at Stake for Safety?

Road transport is responsible for the majority of transport fatalities, with 1.35 million fatalities worldwide each year. On a global level, almost 40% of road fatalities occur in urban areas, while vulnerable road users account for 70% of road deaths in urban areas. Indicatively, during 2019, about 22,800 road traffic fatalities were recorded in the 27 EU Member States.[34] Despite significant improvements in road safety, the process of minimising crashes and their respective causal factors has markedly slowed during the last decade, with only a 20% reduction in crash fatalities.[35]

In recent years, the shift from traditional reactive road safety approaches towards the Safe System approach has been pursued. The Safe System approach accepts that all humans inevitably make mistakes. When mistakes do happen, all transport system elements must contribute to the avoidance of fatalities and, if possible, injuries. Innovative data-driven solutions can contribute to a holistic, proactive approach to addressing urban road safety problems, and represent a core principle of the Safe System Approach. A famous manifestation of this approach is Vision Zero, originating from Sweden.[36]

It is therefore clear that transport and road safety researchers, industrial practitioners, authorities and all stakeholders must make concerted efforts to further reduce crash numbers and mitigate crash consequences, with the utmost priority of negating losses of life and limb. It is not only important but imperative to exploit the new capabilities offered by artificial intelligence (AI). The rise and wide market penetration of smartphones, sensors and connected objects (whether mobile

[34] World Health Organization (WHO), Global status report on road safety 2018.

[35] European Transport Safety Council, *Ranking EU Progress on Road Safety*, 12th Road Safety Performance Index Report, 2018.

[36] R. Johansson, "Vision Zero–Implementing a policy for traffic safety", *Safety science*, vol. 47, no. 6, 2009, pp. 826-31.

or infrastructure) has increased the availability of analytical and broad-scope transport-related big data, which can now be effectively interpreted thanks to rapid progress in computational power, data science and computer science developments in the forms of advanced artificial intelligence tools.

The sections that follow outline specific advancements and challenges regarding the implementation of AI and big data to increase the safety levels of mobility and transport activities.

Big Data Developments Relevant To Road Safety

The rapidly increasingly connections that characterise the new transport landscapes have yielded a wealth of big data. The multitude of data sources include the following categories (this is a non-exhaustive list):

- Mobile phone data, including sensor-based data (e.g. Google Maps, Here, Waze)
- Cellular Network Data (e.g. mobile phone operators, etc.)
- Vehicular On-Board Diagnostics data (e.g. OEM industry)
- Camera data, including on-vehicle (internal, dashcam and peripheral) and on the road (cameras of cities, network operators, police)
- Data from car sharing services (e.g. Uber, Lyft, BlaBlaCar)
- Data from bike sharing services (e.g. 8D Technologies, Mobike)
- Social Media data (e.g. Facebook, Twitter)
- Telematics companies (e.g. Oseven, ZenDrive, Octo)
- Private agency sensor data (e.g. INRIX, Waycare)
- Travel Card data (e.g. Oyster card, Opal card)
- Public authority sensor or traffic measurement data (e.g. Ministries, Public Transport Authorities, Cities, Regions)
- Weather data (e.g. OpenWeatherMap, AccuWeather, etc.)

- Census data (e.g. Eurostat, National Statistics)
- Digital map data (e.g. OpenStreetMap, Google Maps, etc.)
- Shared mobility data (e.g. GPS, routing, etc.)
- Research-oriented data (e.g. floating car/instrumented vehicles)

This wealth of data sources provides high granularity for analysis, which in turn allows more precise training, predictions and similar calculations of road safety models, or more targeted and specialised analyses. Indicatively, it is now easier for road safety analysts to perform differentiations by road user category, achieve higher spatial and temporal resolution in the data and focus on niche analyses (e.g. interactions with vulnerable road users, particulars of professional drivers, freight vehicles etc.).

Other new developments in computer science, telematics and telecommunications, combined with the spread of connectivity, are also aiding road safety data collection. Most immediately, the rollout of 5G/6G technologies is facilitating data transmission and manipulation, while the Internet of Things (IoT) is progressively bringing new opportunities and possibilities for data acquisition (cross-device connectivity). Furthermore, on-board diagnostic (OBD) systems have become considerably more affordable in recent years. The widespread use of smartphones and social media allow for more users in an increasingly covered percentage of the road network area. In recent years, drones and satellites have complemented the available range of data, thus providing solutions by capturing interactions that were previously harder to observe.

Social media data can be invaluable for pattern analysis in road safety, and can be an excellent source of first-detection and first-response for crashes. Moreover, a proportion of social media data is publicly available, and thus exploitable for research through text extraction and processing, constituting an immense big data source. Increasingly powerful cloud computing, computer hardware and analysis tools have emerged to facilitate the management and analysis of big data, especially when fused from multiple sources, while technological competition and a wide market enables typically sustainable pricing.

Big Data Challenges for Road Safety Exploitation

Nonetheless, big data can induce big issues for the prospective analyst. To start with, the consequences of using data which is not always representative of the whole population should be assessed and properly corrected. There is undoubtedly bias towards certain user groups as, despite market penetration, younger demographics are more engaged with smartphones and social media interactions. Furthermore, bias can have many dimensions. It is easy to wrongly consider a dataset as unbiased if it covers a specific dimension in detail (e.g. covering different road user categories), while failing in another (e.g. not covering exposure per category). Predominantly, publication biases can also manifest and always need to be considered in research, both in strictly road-safety topics and in the wider economic impact assessment.[37] Even using extensive databases, *a priori* desired conclusions should not drive the research approach or outcomes. Lastly, proper road safety analyses based on big data processing can be costly in terms of data acquisition, overall equipment and human capital. There is a high risk for decision makers to be misled by the opportunistic analysis of seemingly low-cost data in the absence of qualified data scientists and statisticians.

For such applications, the openness of big data is a constant question. A fragmentation of data ownership and a lack of interoperability between datasets and platforms is currently

[37] E.g. R. Elvik, "Effects on road safety of converting intersections to roundabouts: review of evidence from non-US studies", *Transportation Research Record*, vol. 1847, no. 1, 2003, pp. 1-10.; e.g. O. Ashenfelter and M. Greenstone, "Estimating the value of a statistical life: The importance of omitted variables and publication bias", *American Economic Review*, vol. 94, no. 2, 2004, pp. 454-60.

observed, especially in the industry. There are different commercial interests of the various road safety stakeholders in data, creating differing requirements for data access based on acquisition rate, granularity, intended use and so on. An additional layer is introduced by ownership, as several intermediaries have manifested. Data ownership varies depending on which party generates and collects the data. It is possible that they may be not willing to share data due to issues relating to privacy, legal liability, IP, competition, or costs. In other words, road safety data is often ethically or commercially sensitive.

It is important to remember that data is not free, and that all big data-related tasks, from acquisition to processing and provision have several maintenance and physical or digital infrastructure-related costs.[38] The diversity of data sources has undoubtedly been affecting data quality, and that can be discerned in several instances, for instance by examining the frequencies of outliers and/or unreasonable values. This can be quite straightforward to verify, for instance, in cases of traffic volume or weather data. Unavoidably, variations in hardware and software used for collecting data also impacts quality as well, even with well-maintained collection in mind. Last but not least, there is a notable lack of expertise in introducing the road safety context when conducting machine learning, data mining, and data management within the transport domain. A lot of analysts hail strictly from a computer science background and may not necessarily have the essential understanding of proper exposure measurement, road safety analytical design or risk factor and road safety countermeasure causal relationships.

---

[38] International Transport Forum (ITF), *Artificial Intelligence in Proactive Road Infrastructure Safety Management*, ITF Roundtable Reports, no. 187, OECD Publishing, Paris, 2021.

Surrogate Road Safety Measures

When discussing road safety AI applications, consideration must be given to a critical positive trend in road safety analysis in the form of surrogate road safety measures. These measures are alternative measures that can augment or even substitute the rarer (and less appropriately reported) crash and injury data. Examples of surrogate safety measures include traffic conflicts, harsh driving events, spatial/temporal headways, and many others.[39]

A massive advantage of surrogate safety measures is that they can become readily available for proactive analyses before crashes occur or in areas with limited or no crash data availability. In addition, such measures show less under-reporting and can even aid with crash reporting. More research on the validation of surrogate safety metrics is essential, to reveal which metrics not only correlate with reported crashes but also have accurate predictive capabilities. There is also a need to predict the number of fatalities and/or injuries with good utility and to determine how these metrics can integrate crash participant fragility, physical properties and crash type consequences. The adoption of surrogate safety metrics implies that road safety research is now being conducted across several different indicators, instead of just examining crashes and injuries. This new multidimensionality leads to the review of statistical training needs, so that data is not misused/misinterpreted in relation to what exactly constitutes a safety-critical situation.

Naturally, the collection of surrogate crash measures is becoming increasingly automated and can augment more general-purpose big data. This automated connection is made possible by smartphone sensors (which can be used to obtain data on harsh braking, harsh acceleration, harsh cornering, driving distraction due to cellphone use, speeding, poor road surfaces)

---

[39] A.P. Tarko, "Surrogate measures of safety", in D. Lord and S. Washington (Eds.), *Safe mobility: challenges, methodology and solutions*, Emerald Publishing Limited, 2018.

or instrumented/floating vehicles. Technologies like automatic crash notification (eCall) and event data recorders enable data-driven responses to post-crash problems. Street imagery, also collected by floating vehicles, supports the assessment of road safety performance, such as the star-rating for roads.[40] With ever smarter vehicles, active safety system activation can also constitute a surrogate safety metric. By monitoring the activation of systems such as Anti-lock Braking System (ABS), Electronic Stability Control/Program (ESC/ESP) and Autonomous Emergency Braking (AEB), reliable information about safety-critical events will flow from increasingly connected vehicles, regardless of their level of automation.

Key Road Safety AI Aspects

With these new options and respective challenges unfolding, AI enters the field to open new advances in all aspects of mobility. It is very difficult to predict all AI uses, or even categorise them using well-defined and distinct labels, but it is reasonable to outline advances in three major areas: **(i)** vehicle technology, **(ii)** driver monitoring and **(iii)** crash risk estimation.

Regarding AI Advances in Vehicle Technology, several new systems have been rolling out and continuously improving. The navigation of complex, non-straightforward road environments becomes more attainable at an increasing rate, as high-end RADAR/LIDAR and sensor technologies stand at the forefront of developments. Through the development of connected and automated vehicles, several traditional road safety risk factors and similar problems are eliminated by RADAR/LIDAR, such as exclusive reliance on lighting and limitations caused by obstructions. On the algorithmic front, the decision-making process is improved and refined by deep learning and complex

---

[40] Ai-RAP and Automated Coding for iRAP in SuM4All (Sustainable Mobility for All), *Sustainable Mobility: Policy Making for Data Sharing*, Washington DC, License, Creative Commons Attribution CC BY 3.0, GRA in action series, 2021.

algorithms such as advanced convolutional neural networks for perception, localisation, prediction and decision-making. As is typical with high degrees of development, high degrees of specialisation follow, as purpose-made systems are starting to receive purpose-made tools and algorithms, such as grocery delivery or (initially) fixed-route public transport. It is worth noting that most developers design their systems independently and are not reliant on infrastructure adaptations, while over-the-air AI upgrades become a new reality.

Meanwhile, more physical test areas and virtual testbeds are provided and examined and software errors are gradually contained and reaction times are minimised overall. Facial recognition technologies aid commercial company claims with insurance carriers (e.g. Nauto). Vehicle cooperation algorithms are introduced to fleets, aimed at traffic conflict reduction and efficient traffic management. Furthermore, increased connectivity means additional connectivity byproducts: increased parking availability and freight vehicle platooning can mean reduced road safety exposure indicators, as well as increased fuel efficiency. Ambitious flying vehicle (VTOL) concepts are co-considered.

Regarding AI advances in crash risk estimation, an array of new AI methods and machine/deep learning or similar algorithmic models are available to road safety researchers, stakeholders and authorities for real-time crash risk estimates. Big data on crash occurrence and road and traffic characteristics from infrastructure sensors is transformed into multi-dimensional static or dynamic maps of road risk prediction and road & driver star ratings. Crash datasets are imbalanced, rare event cases which give an incentive for the creation of new approaches and venues of analysis through AI methods. Infrastructure assessment frameworks are starting to embrace AI methodologies, such as the i-RAP transition to Ai-RAP.[41] A large number of model configurations show very promising

---

[41] Ibid.

performance, albeit on specific datasets. Much more research is required on the transferability of AI capabilities to new study areas. The successful performance of a model in one suburb case-study does not guarantee that it will work in another suburb or at city-wide level, so performance is still uncertain.

Regarding AI Advances in telematics & driver monitoring, the insurance industry is heavily investing in telematics, offering reduced premiums for safer driving. AI and data fusion technologies are actively being used in all stages of road safety data collection, transmission, storage, harmonisation, analysis and interpretation from telematics. Personalised feedback can be created and obtained almost instantaneously for participant drivers. Algorithmic route analysis and personalised hotspot detection features based on surrogate safety measures are actively being examined. As far as driver behaviour is concerned, during-trip and post-trip interventions are now possible, and are best administered with gamification and reward systems.

With so many promising new developments, AI and big data applications can be expected to unlock critical road safety advancements, such as attaining Vision Zero safety levels. The most notable venue is that AI facilitates truly proactive management of traffic safety in various ways, such as the following:

- The collection of data on road infrastructure conditions and traffic events through widespread, real-time and broad-scale sensors and systems such as connected vehicle operations and computer vision.
- The identification of high-risk locations through predictive multilayer models before crashes occur.
- Enabled by multiparametric big data, AI pushes the limits of pattern recognition and reaction times beyond human capabilities and may thus uncover new crash-prone road configurations, risky driving behaviour or critical interactions. Essentially multiple in-depth examinations can be conducted per second of analysis.

- Recent developments in the field of explainable AI (XAI) begin to cope with the challenge of the "black box" phenomenon, shedding light on the causal relationships of risk factor, road safety countermeasure and crash causation.

Current Barriers

Needless to say, there are several pending barriers to AI-related developments. Safe, road-worthy AI systems face significant challenges that are only hesitantly tackled. There needs to be a concentration of effort to achieve AI systems with high interface ability with each other, high interoperability across different road networks, timelessness and resilience to ensure a smooth transition of operations and also considerable scalability for reproduction and functionality across areas.

Currently, the absence of monitoring and accountability seriously limits road safety performance. To counter this, public acceptance and trust must be meticulously built and increased by monitoring and reporting AI progress, and conveying the message that AI in road safety is not only for-profit, but also for-society. The neutral ground must be established by the operation of independent tools, such as the AI Incident Database.[42] In addition, legal and operational frameworks are considerably lagging compared to technical developments; self-updating mechanisms are urgently required for them.

While research and innovation efforts on the use of AI in computer vision and risk prediction are very much in the spotlight at present, they require more peripheral support. Thorough arbitration, as well as evaluation and assessment criteria, must be established across platforms, in research and industry, to deliver robust AI vehicular systems that will actively contribute to fatality reductions. Cybersecurity/malicious hacking concerns

---

[42] S. McGregor, *Preventing repeated real world AI failures by cataloging incidents: The AI incident database*, arXiv preprint arXiv:2011.08512, 2020.

may cause several implications (vehicle manufacturers, software engineers, vehicle owners, automated fleet operators). All these different AI aspects will lead to improvements of road safety interventions and countermeasures. Measuring them via dynamic feedback loops through crash records and surrogate safety measures remains a completely unexplored field.

Naturally, big data applications face their own challenges. At present, large margins remain for road safety practitioners to rapidly gain in terms of data flow by copying best practices for data sharing and privacy protection from other, more free-sharing fields. More secure alternatives to data exchange, such as the exchange of queries and responses can be explored, instead of raw information. On a similar note, multiple-criteria based exploration and decision analysis are needed to determine the most efficient Key Performance Indicators that can be mined or created from the available big data.

In a manner parallel with AI cross-platform operation, the establishment of commonly accepted data harmonisation and fusion protocols would be very beneficial. More investigation is needed on the best approaches to reconciling different data scopes and scales (e.g. country, city, city block, road segment, road user). On a high level, Governments and Road Safety Authorities can mandate the sharing of aggregate vehicle data, or provide financial or similar incentives to industrial partners. Indicatively, a minimum-required dataset can be defined for all vehicle manufacturers to report in an anonymous standard aggregate format. More attention needs to be given on the collection of data on traffic volume, speed distribution, and locations where vehicles' active safety systems (ABS/ESP/AEB) are engaged. Of course, regulatory frameworks for data protection need to be clarified to encompass all aspects of operations in a non-prohibitive manner, while governments should also examine how Freedom of Information laws articulate with data protection laws. Throughout the process, it is essential to observe ethical data sharing.[43]

[43] E.g. Accenture, *Building data and AI ethics committees*, Northeastern University

Road safety remains a complex multifaceted science with its own particularities. It is a given that road safety levels cannot be improved on the basis of on accurate forecasting alone, as causal factors must be determined. Therefore, priorities should include the development of explainable AI (XAI) – "white box" techniques that are more transparent.[44] Emphasis should be placed on collaborations across countries for the integration of all road realities and road safety cultures, which are often overlooked by straightforward analysis. Funding must also be made available to road safety multi-disciplinary professionals to conduct post-intervention assessments and validate or re-calibrate the risk prediction tools.

The creation of new road safety tools must not be a self-serving purpose, but rather a precisely coordinated process, and these new tools need to be aligned with precise policy objectives. Stakeholders should aim to exploit the new technological landscape by commissioning research to assess the availability of surrogate measure data and the generation of risk mapping and road safety assessment tools.

For their part, researchers and practitioners have to develop new skills and a digital infrastructure mentality, and promote a multi-disciplinary approach to road safety that combines expertise from the fields of data science, technology and safety. Estimates of the benefit/cost ratios of interventions can be set to update in a dynamic manner, along with accessible user-friendly interfaces, so that they are readily usable when decisions must be made. Ultimately, all research outputs must be usable, however advanced: for instance, risk mapping tools need to be user-friendly if they are to be adopted by road users.

– Ethics Institute, 2019.

[44] E.g. W. Samek, G. Montavon, A. Vedaldi, L.K. Hansen, and K. Müller (Eds.), *Explainable AI: Interpreting, explaining and visualizing deep learning*, vol. 11700, Springer Nature, 2019.