



3^a International Conference
on Road Safety

PRIORITIES ACTIONS AGENDA

13|06|2022

Scientific Research in Road Safety Big Data and Artificial Intelligence

George Yannis, Professor



Department of Transportation Planning and Engineering,
National Technical University of Athens

Presentation Outline

1. Road Safety Science
2. Key Road Safety Big Data Challenges
3. Key Road Safety Artificial Intelligence Challenges
4. AI + Big Data = Road Safety
5. Pending barriers
6. Conclusions



Why Science in Road Safety?

- Road Safety is a typical field with high risk of **important investments not bringing results**.
- Absence of **monitoring** and accountability limits seriously road safety performance.
- Decision making in road safety management is highly dependent on appropriate and **quality data analysis**.
- Very often we look where the data are and **not where the problems and solutions are**.



The Need for Scientific Evidence in Decision Making

➤ The **policy making cycle**

- Vision and strategy;
- Problem identification;
- Target Setting and priority setting;
- Development of measures;
- Establishing and implementing the programme;
- Monitoring of and evaluation of outcomes.

The use of road safety data and scientific analysis is involved in each stage

➤ **Necessity** to:

- Consolidate and organize existing data
- Make data and information available
- Provide an holistic approach (analyses, methodologies, benchmarking tools)
- Support road safety decision making at all stages



Road Safety Data to Support Evidence-based Policies

- Fatalities and their evolution
- Exposure
- Safety Performance Indicators
- Causation (in-depth accident investigations)
- Health indicators
- Economic indicators
- Driver behavior, attitudes etc.
- Road safety rules and regulations
- Road safety measures assessment

Do we have the data we need?

Do we need the data we have?

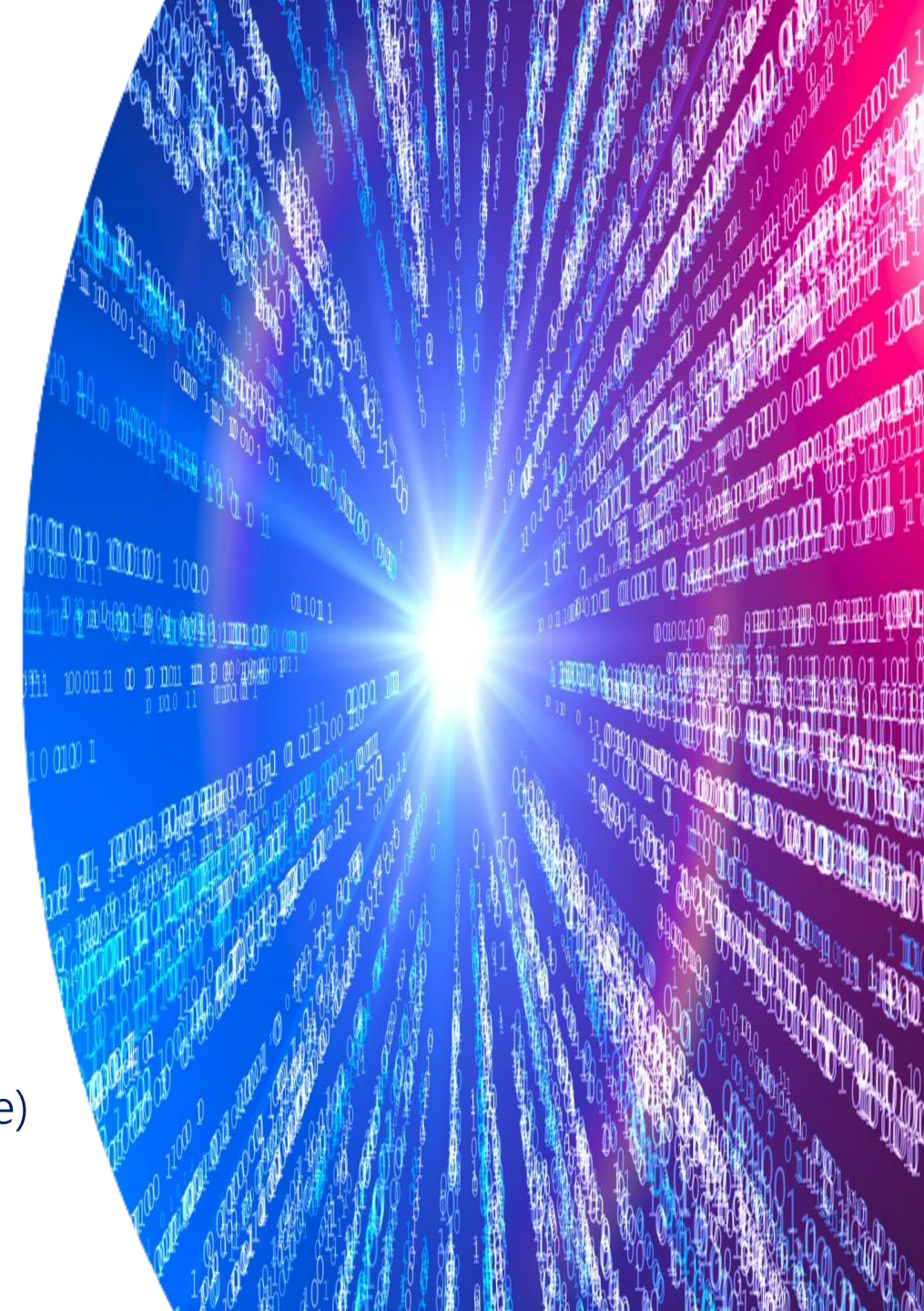


Big Data, Broad Horizons (1/2)

- A wealth of **big data becomes available**
- This enables differentiations per **road user category and focus on niche analyses** (e.g. vulnerable road users, professional drivers, freight vehicles etc.).

A **multitude of data sources:**

- **Mobile Phone data**
 - 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:**
 - On-vehicle (internal, dashcam and peripheral)
 - On the road (cities, 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)



Big Data, Broad Horizons (2/2)

- **Telematics companies** (e.g. OSeven, Floow, 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)



Further New Developments

Several **computer science, telematics and societal advancements** aid road safety data collection:

- Rollout of **5G/6G** technologies facilitates data transmission and manipulation.
- **Internet of Things** (IoT) progressively brings new opportunities and possibilities (cross-device connectivity).
- Affordable **On Board Diagnostic** (OBD) systems.
- Wide **penetration and adoption** of smartphones and social media.
- **Powerful** cloud computing, computer hardware and analysis tools.
- Competition and a wide market offers **sustainable pricing**.



Big Data, Big Issues

- The consequences of using data which are **not always representative** of the whole population (bias towards some user groups) should be assessed and properly corrected.
- It is easy to **wrongly** consider a dataset as unbiased if it covers a specific dimension in detail (e.g. **covering specific road users**) while it can fail in another (e.g. **not covering exposure**).
- **Desired conclusions** should not drive the research approach or outcomes.
- There is a **high risk** for decision makers to be **misled** by the opportunistic analysis of seemingly low-cost data in absence of **qualified data scientists and statisticians**.



How Open are Big Data?

- Fragmentation of data ownership and a **lack of interoperability** between datasets and platforms.
- **Different interests** of the various road safety **stakeholders** in data, creating differing requirements for data access.
- **Data ownership** varies by who generates and collects the data and they may be not willing to share data due to privacy, legal liability, IP, competition, or cost related issues.
- Road safety data are often **ethically or commercially** sensitive.
- The diversity of data sources affecting **data quality**.
- **Variations in hardware and software** used for collecting the data.
- **Lack of expertise** in machine learning, data mining, and data management with a **road safety context**.



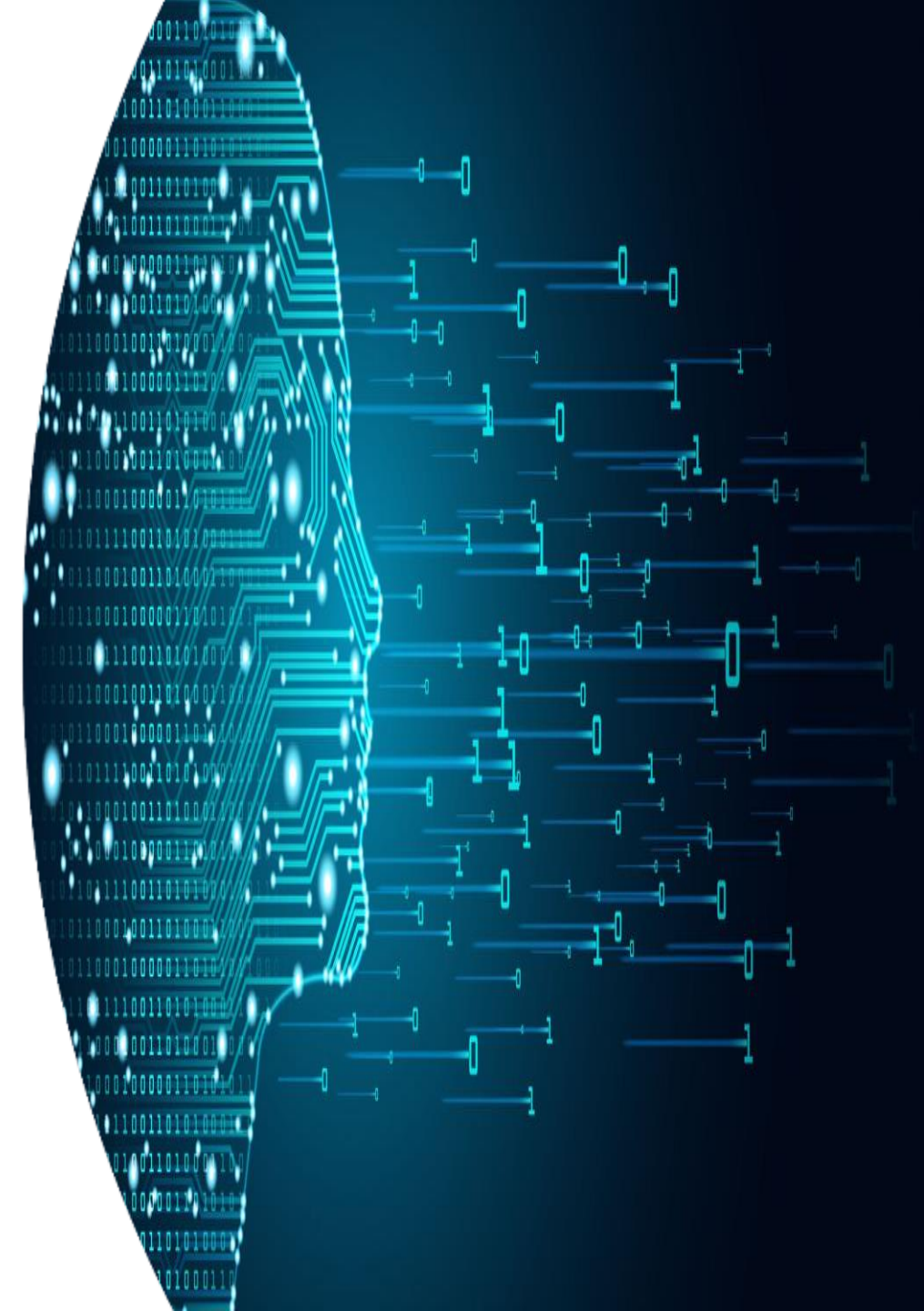
Crash Data Collection

- Automated data collection is possible through:
 - Instrumented/floating vehicles
 - Smartphone sensors (harsh braking, harsh accelerations, harsh cornering, driving distraction via cellphone use, speeding, poor road surfaces).
- 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 (star-rating for roads).
- Drones and satellites complement the range of data, capturing solutions with increased market penetration.
- Active safety system activation can also be considered among surrogate safety metrics, for systems such as:
 - Anti-lock Braking System [ABS]
 - Electronic Stability Control/Program [ESC/ESP]
 - Autonomous Emergency braking [AEB]



AI Advances in Crash Risk Estimation

- An array of new AI methods and machine/deep learning or similar algorithmic models 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 are transformed into **multi-dimension static or dynamic maps** of road risk prediction and **road & driver star rating**.
- Crash datasets are imbalanced, **rare event cases** which find new approaches and venues of analysis through AI methods.
- Infrastructure assessment frameworks start embracing AI methodologies (e.g. the i-RAP **transition** to Ai-RAP).
- A large number of model configurations show **very promising performance**, albeit on specific datasets; **transferability** capabilities are yet uncertain.



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 used in all stages of road safety data collection, transmission, storage, harmonization, analysis and interpretation from **telematics**.
- **Personalized feedback** can be obtained almost instantaneously.
- Algorithm-based route analysis and **personalized hotspot** detection features are actively being examined.
- **During-trip** and **post-trip interventions** are enabled, best administered with gamification and reward systems.



AI Advances in Vehicle Technology (1/2)

- Navigation of **complex road environments** becomes more attainable at an increasing rate, high-end RADAR/LIDAR and sensor technologies at the forefront of developments.
- Several traditional problems are eliminated by RADAR/LIDAR (**e.g. reliance on lighting/obstructions**).
- The decision making process is **improved** and **refined through deep learning**.
- Purpose-made **systems** receive purpose-made **tools and algorithms**, such as grocery delivery or fixed-route public transport.
- Most developers design their systems independently and are not reliant on **infrastructure adaptations**.
- **Over-the-air AI upgrades** become a new reality.



AI Advances in Vehicle Technology (2/2)

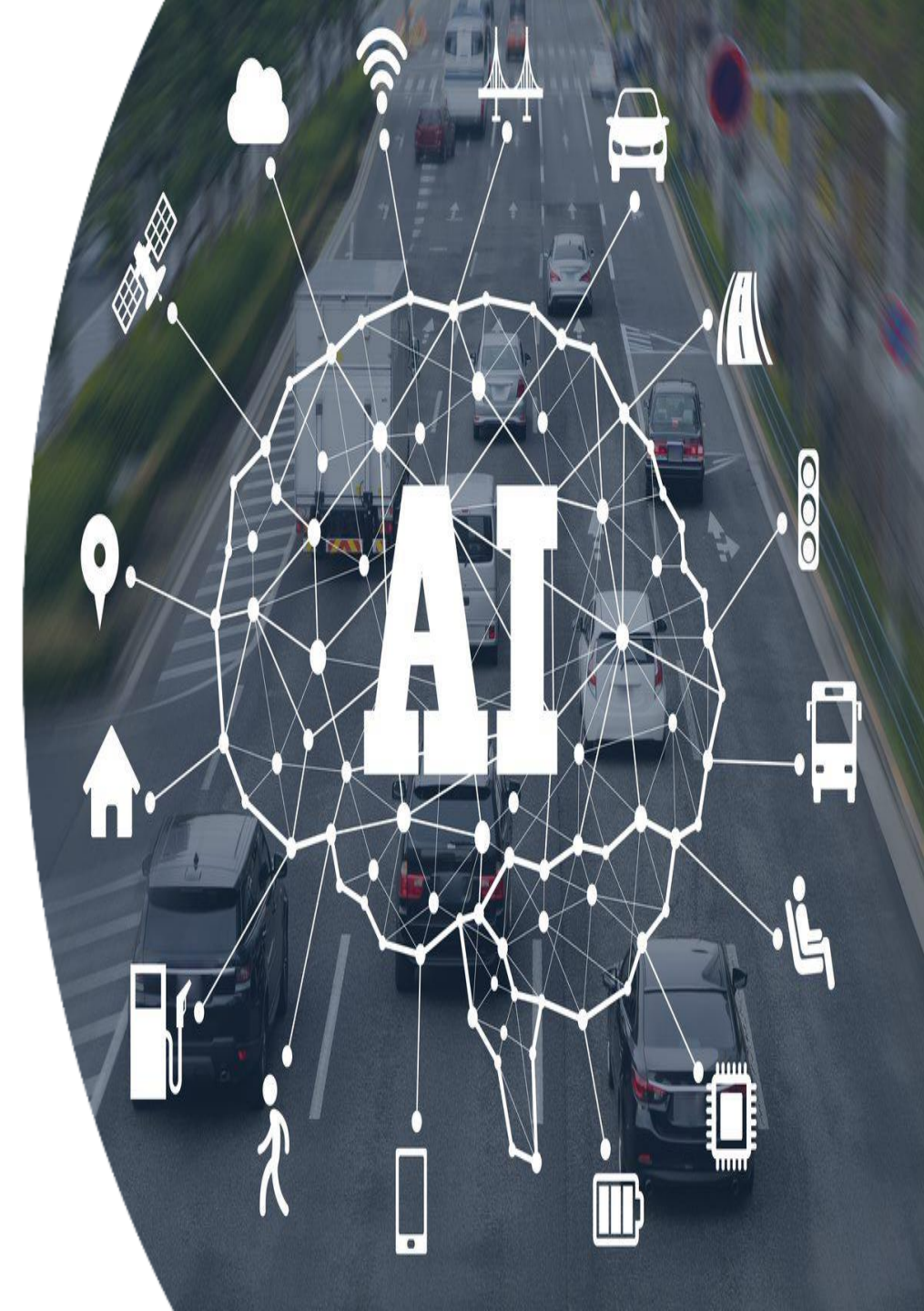
- More **physical test areas** and **virtual testbeds** are provided and examined.
- Software **errors** are gradually contained, **reaction times** are minimized.
- Facial recognition technologies aid **commercial company claims** with insurance carriers (e.g. Nauto).
- **Vehicle cooperation algorithms**: traffic conflict reduction, efficient traffic management.
- **Additional connectivity byproducts**: increased parking availability, increased fuel efficiency, freight vehicle platooning.
- **Flying vehicles (VTOL)** concepts are co-considered.



AI + Big Data = Road Safety

AI facilitates the **proactive management of traffic safety** in various ways:

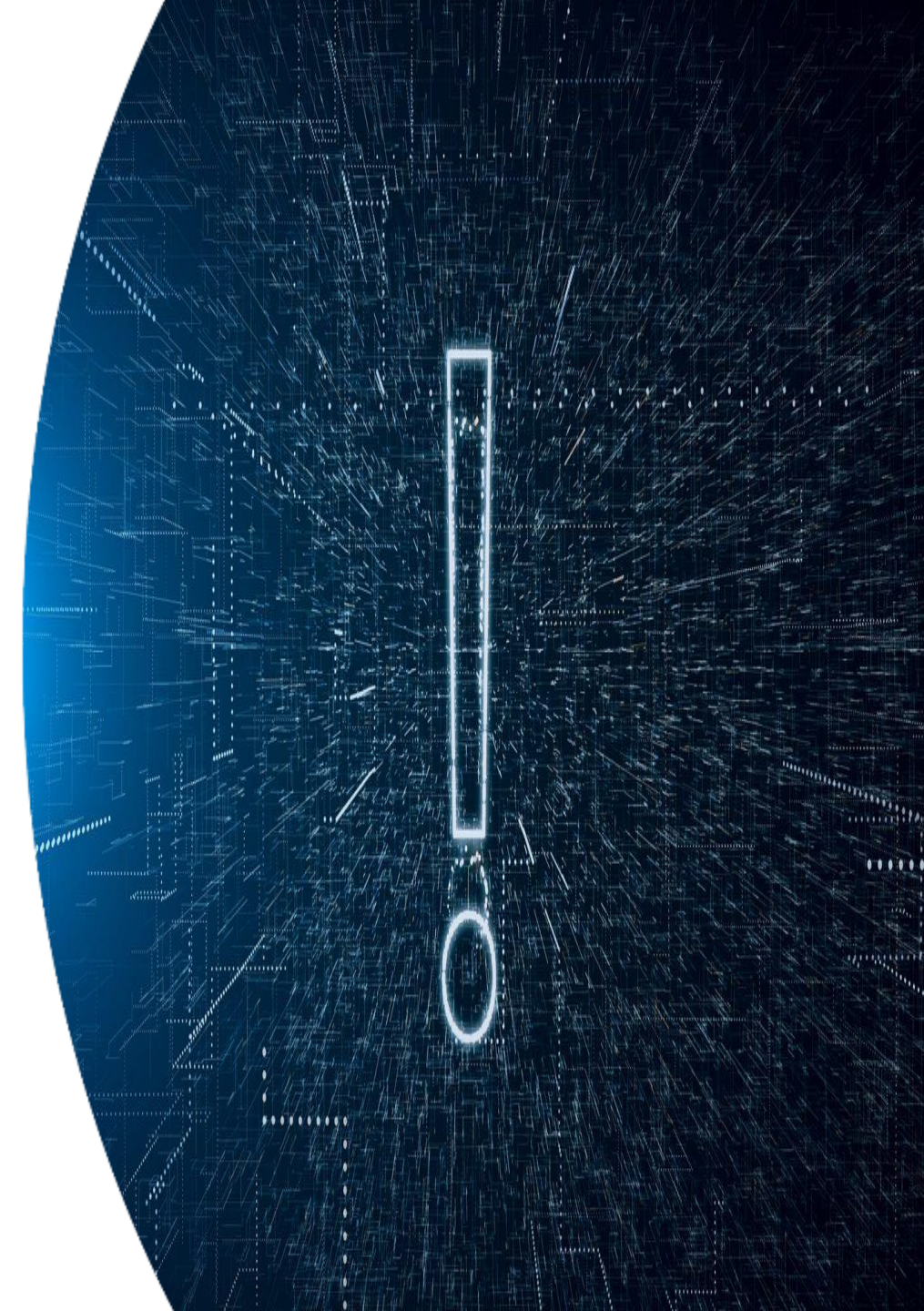
- Collection of data on road infrastructure conditions and traffic events through wide and broad-scale **sensors and systems** such as real-time computer vision.
- **Identification of high risk locations** proactively, through predictive multi-layer models.
- Enabled by multiparametric big data, AI pushes the limits of **pattern recognition** and **reaction times** beyond human capabilities and may thus uncover **new crash-prone** road configurations.
- Recent developments in the field of so-called “**explainable AI (XAI)**” begin to cope with the challenge of the “black box” phenomenon.



-

Pending Barriers for AI (2/2)

- Research and innovation efforts on the use of AI in **computer vision** and **risk prediction** needs **more support**.
- **Thorough 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 may cause several implications (vehicle manufacturers, software engineers, vehicle owners, automated fleet operators).
- **AI dynamic feedback loops on road safety interventions and countermeasures:**
A completely unexplored field!



Pending Barriers for Big Data (1/2)

- Road safety practitioners can rapidly gain by **copying best practices** for data sharing and privacy protection from other fields.
- **More secure alternatives** to data exchange, such as the exchange of queries and responses can be explored, instead of raw information.
- **Multiple-criteria based exploration** and **decision analysis** to determine the most efficient Key Performance Indicators that can be mined or obtained from the available Big Data.
- Establishment of **data harmonization and fusion protocols** – Investigation of the best approach to **reconcile different data scales** (e.g. country, city, city block, road segment, road user).



Pending Barriers for Big Data (2/2)

On a high-level, **Governments** and **Road Safety Authorities** can:

- Mandate the sharing of **aggregate vehicle data**.
- Define a **minimum dataset** for all vehicle manufacturers to report in an anonymous standard aggregate format.
- Collect data on **traffic volume, speed distribution, and locations** where vehicles' **active safety** systems (ABS/ESP/AEB) are engaged.
- Clarify **regulatory frameworks** for data protection.
- Governments should also examine how **Freedom of Information** laws articulate with data protection laws.



Pending Barriers for Road Safety (1/2)

- Road safety cannot be improved based on accurate forecasting alone – **causal factors must be determined**.
- Therefore, priorities should include the development of “**explainable AI (XAI)**” – “white box” techniques.
- Emphasis should be placed on **collaborations across countries** for the integration of all road realities and road safety cultures.
- Funding must also be available to **road safety multi-disciplinary professionals** to conduct post-intervention assessments and validate or re-calibrate the risk prediction tools.



Pending Barriers for Road Safety (2/2)

- New tools need to be aligned with **precise policy** objectives.
- Stakeholders should **commission research** to assess the capability of proxy data and risk mapping tools.

Researchers and practitioners have to:

- **Develop new skills** and a digital infrastructure.
- **Promote a multi-disciplinary approach** to road safety that combines expertise from the fields of data science, technology and safety.
- **Design user-friendly** risk mapping tools to be adopted by road users.
- Update estimates of the benefit/cost ratio of interventions **in a dynamic manner**, along with accessible **user-friendly interfaces**.



Conclusions

- Great potential for **seamless big data driven procedures** from safety problem identification to selection and implementation of optimal solutions.
- Newfound **net present value in road safety data**, available for (real-time) early problem detection and prompt and customized decision support on every level.
- **Considerable ground remains to be covered** for existing road safety AI applications (vehicle, telematics crash analysis).
- **Completely unexplored directions remain** in several road safety aspects (crowdsourcing options, measure effectiveness, data harmonization).
- Big Data and Artificial Intelligence can become **efficient catalysts** for achieving Vision Zero road fatalities by 2050.





3^a International Conference
on Road Safety

PRIORITIES ACTIONS AGENDA

13|06|2022

Scientific Research in Road Safety Big Data and Artificial Intelligence

George Yannis, Professor



Department of Transportation Planning and Engineering,
National Technical University of Athens