Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK
Département fédéral de l'environnement, des transports, de l'énergie et de la communication DETEC
Dipartimento federale dell'ambiente, dei trasporti, dell'energia e delle comunicazioni DATEC

Bundesamt für Strassen
Office fédéral des routes
Ufficio federale delle Strade

Sensor-based accident research and prevention: Exploring legal and technological opportunities

Sensordatenbasierte Unfallforschung: Rechtliche und technologische Möglichkeiten

Recherche et prévention des accidents basées sur les capteurs : exploration des opportunités légales et technologiques

BSS Economic Consulting AG
Niclas Meyer

HDC
Sylvain Métille
Marie-Laure Percassi

Zürcher Hochschule für Angewandte Wissenschaften (ZHAW)
Markus Hackenfort
Désirée Hagmann

DTC Dynamic Test Center AG
Heinz Reber

École Polytechnique Fédérale de Lausanne (EPFL)
Swiss Data Science Center
Eric Bouillet
Sabrina Ossey
Oksana Riba Grognuz

Forschungsprojekt MFZ_20_07A_02 auf Antrag der Arbeitsgruppe Mensch und Fahrzeug MFZ

Oktober 2024 **1775**

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK
Département fédéral de l'environnement, des transports, de l'énergie et de la communication DETEC
Dipartimento federale dell'ambiente, dei trasporti, dell'energia e delle comunicazioni DATEC

Bundesamt für Strassen
Office fédéral des routes
Ufficio federale delle Strade

Sensor-based accident research and prevention: Exploring legal and technological opportunities

Sensordatenbasierte Unfallforschung: Rechtliche und technologische Möglichkeiten

Recherche et prévention des accidents basées sur les capteurs : exploration des opportunités légales et technologiques

BSS Economic Consulting AG
Niclas Meyer

HDC
Sylvain Métille
Marie-Laure Percassi

Zürcher Hochschule für Angewandte Wissenschaften (ZHAW)
Markus Hackenfort
Désirée Hagmann

DTC Dynamic Test Center AG
Heinz Reber

École Polytechnique Fédérale de Lausanne (EPFL)
Swiss Data Science Center
Eric Bouillet
Sabrina Ossey
Oksana Riba Grognuz

Forschungsprojekt MFZ_20_07A_02 auf Antrag der Arbeitsgruppe Mensch und Fahrzeug MFZ

Oktober 2024

Impressum

Forschungsstelle und Projektteam

Projektleitung
Niclas Meyer

Mitglieder
Eric Bouillet
Markus Hackenfort
Désirée Hackmann
Sabrina Ossey
Marie-Laure Percassi
Heinz Reber
Oksana Riba Grognuz
Sylvain Métille

Begleitkommission

Präsidentin
Bettina Zahnd

Mitglieder
Jörg Arnold
Fabiano Assi
Markus Deublein
Michelle Gruner
Christian Kamenik
Thomas Probst
Mathieu Sinn

Antragsteller

Arbeitsgruppe Mensch und Fahrzeug (MFZ)

Bezugsquelle

Das Dokument kann kostenlos von http://www.mobilityplatform.ch heruntergeladen werden.

Contents

Summary

Background

Modern vehicles are rich sources of data, equipped with numerous sensors monitoring the operation of the vehicle (e.g., engine performance, fuel efficiency, and steering) as well as the surrounding environment (e.g., weather conditions, road features) and road users (e.g., vehicle location, speed). This goldmine of information holds immense promise to revolutionize accident research and prevention, unlocking possibilities for safer roads.

With the introduction of automated driving even more sensor data will be generated in the future. First, original equipment manufacturers (OEMs) need sensor data to train their algorithms for automated driving. Secondly, once the algorithms are developed, vehicles require sensor data to operate autonomously. Therefore, we expect that, over the coming years, more sensors will be installed generating evermore data.

In principle, access to sensor data could open many new avenues for road safety research and prevention. In practice, however, accident researchers and prevention specialists often do not have access to the relevant sensor data. Despite the clear public interest in improving research and prevention efforts, they often struggle to obtain the crucial sensor data needed, like detailed vehicle performance information or anonymized location data. This creates a tension between the public good of safer roads and the limited access currently granted.

Balancing the need to protect investments and promote innovation is complex. On the one hand, the public interest lies in fostering knowledge sharing and technological advancements, potentially leading to safer vehicles and infrastructure. On the other hand, private companies have legitimate interests in protecting their trade secrets and intellectual property, which are often enshrined in various legal norms. Finding the right equilibrium between these competing interests is crucial for sustainable progress.

Much sensor data, including information about vehicle operation and user location, is classified as personal data [1, 2], raising concerns about individual privacy. It's essential to carefully weigh these concerns against the broader goal of preventing accidents. Fortunately, technologies like differential privacy and federated learning offer promising solutions that allow data analysis while safeguarding individual privacy.

Harnessing the potential of sensor data – from vehicle performance to driver behavior – offers a crucial step towards safer roads. Yet, navigating complex commercial, public, and privacy interests poses significant challenges.

Objectives

This project investigates legal and technological solutions to unlock the data's potential, focusing on four key use cases identified through extensive stakeholder engagement.

Methodological approach

The project brings together expertise from multiple scientific disciplines: data science, law, economics, politics and governance, human factors psychology, and engineering. This makes it a truly interdisciplinary project.

Moreover, the project followed a transdisciplinarity approach: It addresses concrete problems and is designed to develop relevant solutions for beneficiaries, which, in the case of this project, means road safety researchers and prevention specialists in academia, government, and private sector. Transdisciplinarity is achieved by a) directly involving the beneficiaries in the project and b) engaging with the beneficiaries from the beginning and involving them in the project design as opposed to confronting them with the results at the end.

Beneficiaries and stakeholders were involved in two ways:

1. **As team members**: Dynamic Test Center (DTC) of Berner Fachhochschule does accident reconstruction and works with sensor data in practice. The team from Zürcher Hochschule für Angewandte Wissenschaften (ZHAW) represents the research community that would like to use sensor data.
2. **As workshop participants**: More than 30 practitioners and beneficiaries were involved in a series of 8 workshops. In the workshops, together with the practitioners and beneficiaries use cases were identified, which then defined the focus of the project.

Each use case addresses a different problem that may be addressed with the help of sensor data. As background for the use case analysis, the project comprises an extensive stock-taking of the sensor data that is available today and might potentially become available in the future. Furthermore, the project provides a discussion of the relevant legal and data science foundations that were then applied to the use cases.

Legal foundations

The project provides an overview of the relevant legal foundations in the context of sensor data. It describes the applicable legal framework, both in Switzerland and in the European Union. In Switzerland, the Federal Act on Data Protection (FADP) is particularly important for the present topic. In Europe, it is interesting to observe that several regulations were adopted about the collection and processing of personal data by connected cars.

Moreover, a detailed discussion is provided on three relevant legal topics: access to data, data protection, and criminal procedure. The project explains which data-sharing models could be used to obtain useful data collected by cars. It provides guidance about the fundamental principles of data protection law. Finally, the project relates to two particular issues that arise in the context of criminal proceedings: the need for a search warrant to use car-collected data and the possibility of using unlawfully obtained evidence (for example data collected by a car in violation of the FADP).

Data science foundations

The project outlines a structured framework for leveraging vehicles' sensor data in accident research and prevention, prioritizing privacy throughout the process. It examines the potential privacy risks associated with this data and proposes a privacy-enhancing data access and processing solution based on the Five Safes framework. By focusing on built-in data-centric measures and tailored privacy techniques, the project offers strategies to mitigate risks effectively. Importantly, it emphasizes that the level of protection should be adapted to the specific needs of different research goals, ensuring responsible data handling.

Moreover, the analysis shows that, in principle, technologies are available (e.g. Federated Learning, Homomorphic Encryption and Secure Multi-Party Computation) to share sensor data with the research and prevention community while meeting data protection requirements and accommodating the commercial interests of the companies and people involved.

Governance Architecture

To fully unlock the potential of sensor data, a comprehensive governance architecture is proposed, aiming to unite a diverse array of stakeholders. This collaborative approach is deemed crucial and necessitates the establishment of a secure and reliable data space. This dedicated environment is envisioned to seamlessly facilitate data access and analysis, fostering transparency and cooperation among all involved parties.

Use case analysis

The legal and data science foundations are applied to four use cases:

Use case 1 - Expanding accident statistics: Can we improve existing accident statistics by leveraging sensor data? This use case explores ways to expand the data pool and gain deeper insights into accidents, ultimately leading to better prevention strategies.

Use case 2 - Exposure: Who drives when and where? This use case delves into sensor data to gain insights into driving behavior, potentially helping identify high-risk areas and tailor safety interventions.

Use case 3 - Hazard warnings and real-time prevention: Imagine receiving real-time alerts about potential hazards on the road. This use case explores how sensor data could be used to predict and prevent accidents in real time.

Use case 4 - Accident reconstruction and determining criminal culpability: Unraveling the truth behind accidents can be complex. This use case examines how sensor data could be used to reconstruct accidents more accurately and determine culpability more effectively.

Discussion and conclusions

In principle, access to sensor data could open many new avenues for road safety research and prevention. However, there are challenges too. To date, not enough threshold values exist to interpret sensor data. There are quality and standardization issues to be expected, too.

Data access, however, appears to be the biggest challenge to date. Our analysis shows that, in principle, the technologies are available (e.g. Federated Learning, Homomorphic Encryption and Secure Multi-Party Computation) to share sensor data with the research and prevention community while accommodating the commercial interests, liability concerns and privacy risk that often stand in the way of sharing sensor data with the research and prevention community. However, the application of these technologies can be costly, and it often requires multiple actors to collaborate. For that purpose, governance architectures are required. Governance architectures define who gets data access under what circumstance, and how costs and benefits are distributed.

Some OEMs and other service providers offer sensor data or analyses based on sensor data for sale. Over the duration of the project, more services have become available. At the same time, state interventions, too, have played an important role in the use cases that we studied. On multiple occasions the EU has adopted legislation the requires firms to share sensor data available where this can be used to improve road safety.

As the accessibility of sensor-data is concerned, FEDRO, the Swiss Federation more generally, but also Cantonal or city governments have several options. First, they can wait for the private sector to possibly make sensor data available to them. Secondly, through legislation they can adopt laws that oblige firms to share sensor data. Thirdly, the state can through financial subsidies or through coordination activities facilitate the emergence of governance architectures that enable the sharing of sensor data.

However, Swiss policymakers should consider the international context as the Federal Council has recognized.

Recommendations

1. Near-accident data: We recommend that FEDRO, Cantonal and city governments start to explore the use of near-accident data to identify risks in the road network. Several firms already provide near-accident data for sale. Given the strong public interests at stake we recommend that they explore legal ways requiring OEMs to share near-accident data free of charge. We also recommend that quality checks and validation tests are done to ensure the quality and comparability of the data. When accident data is used, a privacy risk assessment based on the Five Safes framework should be applied. We recommend using the checklist introduced in Chapter 6.5.
2. We recommend that FEDRO explores ways to integrate EDR data in accident statistics. On its own, EDR data can be difficult to interpret. Therefore, we recommend that FEDRO develops ways to validate and triangulate the EDR data with other sources. We assume that the data can be sufficiently anonymized so that no data protection concerns arise.
3. In the future, when policymakers develop new regulations that concern sensor data, we recommend that they include in the law research and prevention as one of the purposes for which the data can be used. The example of eCall data shows that if this is not defined as the explicit purpose data protection rules may prohibit the research and prevention community from using the concerned data.
4. We recommend that FEDRO monitors and promotes the development of governance architectures to allow for the sharing of more sensor data with accident researchers and prevention specialists in academia, government, and the private sector. We conclude that technically it is possible to set up systems that allow for the sharing of sensor data while both respecting data privacy requirements and accommodating the commercial interests of OEMs.
5. Exposure data: Sensor data has great potential in the context of (risk) exposition. Sensor data, particularly mobility data, due to its current lack of representativeness may not be able to replace the micro census for mobility for now. And it may not yet replace data from traffic counters. However, mobility data may be used where census data or data from traffic counters is not available, for example, regarding specific neighborhoods or sections of road. We recommend FEDRO to support research that explores new ways of generating exposition data from mobility data.
6. Hazard warnings: We recommend that the Swiss Federation joins and supports the Data for Road Safety project. We also recommend that the project is used as a blueprint for other applications that require the collaboration of diverse actors, including OEMs and transport authorities.
7. EDR: Because OEMs are often based in foreign countries, we recommend the introduction of a legal obligation (for example in traffic regulations) requiring importers of vehicles to provide access to EDR data to prosecutors in a readable format.
8. EDR: We consider it to be appropriate that a search warrant is needed to analyze EDR data and do not recommend a modification of the CrimPC. CrimPC shall remain technologically neutral and a specific regime for EDR data is not necessary.

Zusammenfassung

Hintergrund

Moderne Fahrzeuge sind mit zahlreichen Sensoren ausgestattet, die sowohl den Betrieb des Fahrzeugs (z. B. Motorleistung, Kraftstoffverbrauch und Lenkung) als auch die Umgebung (z. B. Wetterbedingungen, Strassenmerkmale) sowie das Verhalten der Verkehrsteilnehmer aufzeichnen. Sensordaten sind eine Goldgrube an Informationen und bergen das Versprechen, die Unfallforschung und -prävention zu revolutionieren.

Mit der Einführung des automatisierten Fahrens werden in Zukunft noch mehr Sensordaten anfallen. Erstens benötigen die Fahrzeughersteller (OEMs) Sensordaten, um ihre Algorithmen für das automatisierte Fahren zu trainieren. Zweitens benötigen die Fahrzeuge, sobald die Algorithmen entwickelt sind, Sensordaten, um automatisiert fahren zu können. Daher erwarten wir, dass in den kommenden Jahren immer mehr Sensoren installiert werden, die immer mehr Daten liefern werden.

Im Prinzip dürften Sensordaten viele neue Wege für die Forschung und Prävention im Bereich der Strassenverkehrssicherheit eröffnen. In der Praxis haben Unfallforscher und Präventionsspezialisten jedoch oft keinen Zugang zu den entsprechenden Sensordaten. Trotz des klaren offensichtlichen öffentlichen Interesses, die Forschungs- und Präventionsbemühungen zu unterstützen, haben sie oft Schwierigkeiten, an die Sensordaten zu gelangen, wie z. B. detaillierte Informationen über das Fahrzeugverhalten oder anonymisierte Standortdaten. So entsteht ein Spannungsverhältnis zwischen dem öffentlichen Interesse, die Verkehrssicherheit zu erhöhen, an sichereren Strassen und dem begrenzten Zugang, der derzeit gewährt wird.

Die Abwägung zwischen der Notwendigkeit, Investitionen zu schützen und Innovationen zu fördern, ist komplex. Einerseits liegt das öffentliche Interesse in der Förderung des Wissensaustauschs und des technologischen Fortschritts, was zu sichereren Fahrzeugen und Infrastrukturen führen kann. Andererseits haben private Unternehmen ein legitimes Interesse am Schutz ihrer Geschäftsgeheimnisse und ihres geistigen Eigentums, das oft in verschiedenen Rechtsnormen verankert ist. Das richtige Gleichgewicht zwischen diesen konkurrierenden Interessen zu finden, ist entscheidend für einen nachhaltigen Fortschritt.

Viele Sensordaten, einschliesslich Informationen über den Fahrzeugbetrieb und den Standort des Nutzers, werden als personenbezogene Daten eingestuft [1, 2], was Bedenken hinsichtlich der Privatsphäre des Einzelnen aufwirft. Diese Bedenken müssen sorgfältig gegen das übergeordnete Ziel der Unfallverhütung abzuwägen. Glücklicherweise bieten Technologien wie Differential Privacy und Federated Learning vielversprechende Lösungen, die eine Datenanalyse unter Wahrung der Privatsphäre des Einzelnen ermöglichen.

Die Nutzung des Potenzials von Sensordaten – von der Fahrzeugleistung bis zum Fahrerverhalten – ist ein entscheidender Schritt zu sichereren Strassen. Allerdings ist es eine grosse Herausforderung, die komplexen kommerziellen, öffentlichen und datenschutzrechtlichen Interessen unter einen Hut zu bringen.

Ziele

Dieses Projekt untersucht rechtliche und technologische Lösungen, mit denen das Potenzial der Daten erschlossen werden kann, und konzentriert sich dabei auf vier Anwendungsfälle, die durch eine umfassende Einbeziehung von Interessengruppen ermittelt wurden.

Methodischer Ansatz

Das Projekt bringt Fachwissen aus verschiedenen wissenschaftlichen Disziplinen zusammen: Datenwissenschaft, Recht, Wirtschaft, Politik und Verwaltung, Psychologie und Technik. Dies macht es zu einem wirklich interdisziplinären Projekt.

Darüber hinaus verfolgt das Projekt einen transdisziplinären Ansatz: Es befasst sich mit konkreten Problemen und zielt darauf ab, relevante Lösungen für die Nutzniesser zu entwickeln, d. h. im Falle dieses Projekts für Verkehrssicherheitsforscher und Präventionsspezialisten in Hochschulen, Behörden und im privaten Sektor. Transdisziplinarität wird erreicht, indem a) die Begünstigten als Projektmitglieder direkt am Projekt beteiligt werden und b) externe Begünstigte von Anfang an in die Projektgestaltung einbezogen werden, anstatt sie erst am Ende mit den Ergebnissen zu konfrontieren.

Die Begünstigten und Interessengruppen wurden auf zwei Arten einbezogen:

1. **Als Teammitglieder**: Das Dynamic Test Center (DTC) der Berner Fachhochschule führt Unfallrekonstruktionen durch und arbeitet in der Praxis mit Sensordaten. Das Team der Zürcher Hochschule für Angewandte Wissenschaften (ZHAW) repräsentiert die Forschungsgemeinschaft, die Sensordaten nutzen möchte.
2. **Als Workshop-Teilnehmer**: Mehr als 30 Praktiker und Nutzniesser waren an einer Reihe von 8 Workshops beteiligt. In den Workshops wurden zusammen mit den Anwendern und Nutzniessern Anwendungsfälle identifiziert, die dann den Fokus des Projekts definierten.

Jeder Anwendungsfall behandelt ein anderes Problem, das mit Hilfe von Sensordaten gelöst werden kann. Als Hintergrund für die Analyse der Anwendungsfälle umfasst das Projekt eine umfassende Bestandsaufnahme der Sensordaten, die heute verfügbar sind und in Zukunft verfügbar werden könnten. Darüber hinaus werden die relevanten rechtlichen und datenwissenschaftlichen Grundlagen erörtert, die dann auf die Anwendungsfälle angewendet wurden.

Rechtliche Grundlagen

Das Projekt gibt einen Überblick über die relevanten rechtlichen Grundlagen im Kontext von Sensordaten. Es beschreibt den geltenden Rechtsrahmen sowohl in der Schweiz als auch in der Europäischen Union. In der Schweiz ist insbesondere das Bundesgesetz über den Datenschutz für das vorliegende Thema von Bedeutung. In Europa wurden mehrere Verordnungen über die Erhebung und Verarbeitung personenbezogener Daten durch vernetzte Fahrzeuge erlassen.

Darüber hinaus werden drei relevante rechtliche Themen ausführlich erörtert: Zugang zu Daten, Datenschutz und Strafverfahren. Das Projekt erklärt, welche Modelle der gemeinsamen Datennutzung genutzt werden könnten, um nützliche Daten zu erhalten, die von Autos gesammelt werden. Es bietet eine Anleitung zu den Grundprinzipien des Datenschutzrechts. Schliesslich bezieht sich das Projekt auf zwei besondere Fragen, die sich im Rahmen von Strafverfahren stellen: die Notwendigkeit eines Durchsuchungsbefehls für die Verwendung von im Auto gesammelten Daten und die Möglichkeit der Verwendung von unrechtmässig erlangten Beweismitteln (z. B. Daten, die von einem Auto unter Verletzung des DSG gesammelt wurden).

Datenwissenschaftliche Grundlagen

Das Projekt skizziert einen strukturierten Rahmen für die Nutzung von Sensordaten aus Fahrzeugen in der Unfallforschung und -prävention, wobei der Datenschutz während des gesamten Prozesses Vorrang hat. Es untersucht die potenziellen Risiken für die Privatsphäre, die mit diesen Daten verbunden sind, und schlägt eine Lösung zur Verbesserung des Datenzugriffs und der Datenverarbeitung vor, die auf dem Five Safes Framework basiert. Durch die Konzentration auf integrierte datenzentrierte Massnahmen und massgeschneiderte Datenschutztechniken bietet das Projekt Strategien zur wirksamen Risikominderung. Dabei wird betont, dass das Schutzniveau an die

spezifischen Bedürfnisse der verschiedenen Forschungsziele angepasst werden sollte, um einen verantwortungsvollen Umgang mit den Daten zu gewährleisten.

Darüber hinaus zeigt die Analyse, dass grundsätzlich Technologien zur Verfügung stehen (z. B. Federated Learning, homomorphe Verschlüsselung und Secure Multi-Party Computation), um Sensordaten mit der Forschungs- und Präventionsgemeinschaft zu teilen und dabei die Datenschutzanforderungen zu erfüllen und die kommerziellen Interessen der beteiligten Unternehmen und Personen zu berücksichtigen.

Governance-Architektur

Um das Potenzial von Sensordaten voll auszuschöpfen, wird eine umfassende Governance-Architektur vorgeschlagen, die darauf abzielt, eine Vielzahl von Interessengruppen zusammen zu bringen. Dieser kooperative Ansatz wird als entscheidend angesehen und erfordert die Einrichtung eines sicheren und zuverlässigen Datenraums. Diese spezielle Umgebung soll den Datenzugriff und die Datenanalyse nahtlos erleichtern und die Transparenz und Zusammenarbeit zwischen allen Beteiligten fördern.

Analyse von Anwendungsfällen

Die rechtlichen und datenwissenschaftlichen Grundlagen werden auf vier Anwendungsfälle angewandt:

1. **Anwendungsfall 1** – Ergänzung der Unfallstatistik: Können wir die bestehenden Unfallstatistiken durch die Nutzung von Sensordaten verbessern? In diesem Anwendungsfall wird untersucht, wie der Datenpool erweitert und tiefere Einblicke in Unfälle gewonnen werden können, was letztlich zu besseren Präventionsstrategien führt.
2. **Anwendungsfall 2** – Gefährdung: Wer fährt wann und wo? In diesem Anwendungsfall werden Sensordaten ausgewertet, um Einblicke in das Fahrverhalten zu gewinnen, die bei der Identifizierung von Risikobereichen und der Anpassung von Sicherheitsmassnahmen helfen können.
3. **Anwendungsfall 3** – Gefahrenwarnungen und Echtzeit-Prävention: Stellen Sie sich vor, Sie würden in Echtzeit vor potenziellen Gefahren im Strassenverkehr gewarnt. In diesem Anwendungsfall wird untersucht, wie Sensordaten genutzt werden könnten, um Unfälle in Echtzeit vorherzusagen und zu verhindern.
4. **Anwendungsfall 4** – Unfallrekonstruktion und Ermittlung der strafrechtlichen Verantwortlichkeit: Die Aufklärung von Unfällen kann sehr komplex sein. In diesem Anwendungsfall wird untersucht, wie Sensordaten genutzt werden können, um Unfälle genauer zu rekonstruieren und die Schuldfrage effektiver zu klären.

Diskussion und Schlussfolgerungen

Grundsätzlich könnte der Zugang zu Sensordaten viele neue Wege für die Forschung und Prävention im Bereich der Strassenverkehrssicherheit eröffnen. Allerdings gibt es auch Herausforderungen. Bislang gibt es nicht genügend Schwellenwerte für die Interpretation von Sensordaten. Auch sind Qualitäts- und Standardisierungsprobleme zu erwarten.

Die grösste Herausforderung scheint jedoch im Zugang zu den Daten zu liegen. Kommerziellen Interessen, Haftungsbedenken und Datenschutzrisiken stehen einer gemeinsamen Nutzung von Sensordaten mit der Forschungs- und Präventionsgemeinschaft häufig Wege stehen. Unsere Analyse zeigt aber, dass grundsätzlich Technologien zur Verfügung stehen (z. B. Federated Learning, homomorphe Verschlüsselung und Secure Multi-Party Computation), um Sensordaten mit der Forschungs- und Präventionsgemeinschaft zu teilen und die genannten Bedenken und Interessen zu berücksichtigen. Die Anwendung dieser Technologien kann jedoch kostspielig sein und erfordert häufig die Zusammenarbeit mehrerer Akteure. Zu diesem Zweck sind Governance-Architekturen erforderlich. Governance-Architekturen legen fest,

wer unter welchen Umständen Zugang zu den Daten erhält und wie Kosten und Nutzen verteilt werden.

Einige OEMs und andere Dienstanbieter bieten Sensordaten oder auf Sensordaten basierende Analysen zum Verkauf an. Während der Projektlaufzeit sind weitere Dienste verfügbar geworden. Gleichzeitig haben auch staatliche Eingriffe eine wichtige Rolle in den von uns untersuchten Anwendungsfällen gespielt. Die EU hat mehrfach Rechtsvorschriften erlassen, die Unternehmen dazu verpflichten, verfügbare Sensordaten weiterzugeben, wenn diese zur Verbesserung der Verkehrssicherheit genutzt werden können.

Was die Zugänglichkeit von Sensordaten betrifft, so haben das ASTRA, der Bund im Allgemeinen, aber auch Kantons- und Stadtregierungen mehrere Möglichkeiten. Erstens können sie darauf warten, dass der Privatsektor ihnen Sensordaten zum Kauf anbietet. Zweitens können sie Gesetze erlassen, die die Unternehmen verpflichten, Sensordaten zu teilen. Drittens kann der Staat durch finanzielle Subventionen oder durch Koordinierungsmassnahmen das Entstehen von Governance-Architekturen fördern, die die gemeinsame Nutzung von Sensordaten ermöglichen. Die Schweizer Politik sollte dabei jedoch den internationalen Kontext berücksichtigen, wie der Bundesrat erkannt hat.

Empfehlungen

1. Daten zu Beinahe-Unfällen: Wir empfehlen dem ASTRA, den Kantonen und den Städten, die Nutzung von Daten zu Beinahunfällen prüfen, um Risiken im Strassennetz zu identifizieren. Mehrere Firmen bieten bereits unfallnahe Daten zum Verkauf an. In Anbetracht der starken öffentlichen Interessen, die auf dem Spiel stehen, empfehlen wir, dass sie rechtliche Möglichkeiten ausloten, um die OEMs zu verpflichten, diese Daten kostenlos zur Verfügung zu stellen. Wir empfehlen ausserdem, dass Qualitätskontrollen und Validierungstests durchgeführt werden, um die Qualität und Vergleichbarkeit der Daten zu gewährleisten. Wenn Daten zu Beinaheunfällen verwendet werden, sollte eine Risikobewertung zum Schutz der Privatsphäre auf der Grundlage des Five-Safes-Rahmens durchgeführt werden. Wir empfehlen die Verwendung der in Kapitel 6.5 vorgestellten Checkliste.
2. Wir empfehlen dem ASTRA, Möglichkeiten zur Integration von EDR-Daten in die Unfallstatistik zu prüfen. EDR-Daten können für sich allein genommen schwierig zu interpretieren sein. Wir empfehlen deshalb, dass das ASTRA Möglichkeiten zur Validierung und Triangulation der EDR-Daten mit anderen Quellen entwickelt. Wir gehen davon aus, dass die Daten ausreichend anonymisiert werden können, so dass keine datenschutzrechtlichen Bedenken entstehen.
3. Wenn die Politik in Zukunft neue Regelungen für Sensordaten erarbeitet, empfehlen wir, die Forschung und Prävention als einen der Zwecke, für die die Daten verwendet werden können, in das Gesetz aufzunehmen. Das Beispiel der eCall-Daten zeigt, dass, wenn dies nicht als expliziter Zweck definiert wird, die Datenschutzvorschriften der Forschungs- und Präventionsgemeinschaft die Verwendung der betreffenden Daten untersagen können.
4. Wir empfehlen dem ASTRA, die Entwicklung von Governance-Architekturen zu überwachen und zu fördern, um die gemeinsame Nutzung von mehr Sensordaten durch Unfallforscher und Präventionsspezialisten in Hochschulen, Behörden und im privaten Sektor zu ermöglichen. Wir kommen zu dem Schluss, dass es technisch möglich ist, Systeme einzurichten, die die gemeinsame Nutzung von Sensordaten ermöglichen und dabei sowohl die Anforderungen des Datenschutzes als auch die kommerziellen Interessen der OEMs berücksichtigen.
5. Expositionsdaten: Sensordaten haben ein grosses Potenzial im Zusammenhang mit der (Risiko-)Exposition. Sensordaten, insbesondere Mobilitätsdaten, können aufgrund ihrer derzeitigen mangelnden Repräsentativität den Mikrozensus für Mobilität vorerst nicht ersetzen. Und sie können auch noch nicht die Daten von Verkehrszählern ersetzen. Mobilitätsdaten können jedoch dort eingesetzt werden, wo Zählungsdaten oder Daten von Verkehrszählern nicht zur Verfügung stehen, z. B. für bestimmte Quartiere oder Strassenabschnitte. Wir empfehlen dem ASTRA, Forschungsarbeiten

zu unterstützen, die neue Wege zur Generierung von Expositionsdaten aus Mobilitätsdaten erforschen.

6. Gefährdungswarnungen: Wir empfehlen, dass der Bund dem Projekt "Data for Road Safety" beitritt und es unterstützt. Wir empfehlen auch, das Projekt als Vorlage für andere Anwendungen zu nutzen, die die Zusammenarbeit verschiedener Akteure, einschliesslich OEMs und Verkehrsbehörden, erfordern.
7. EDR: Da die OEMs oft im Ausland ansässig sind, empfehlen wir die Einführung einer gesetzlichen Verpflichtung (z.B. in der Strassenverkehrsordnung), die die Importeure von Fahrzeugen verpflichtet, den Staatsanwaltschaften Zugang zu EDR-Daten in einem lesbaren Format zu gewähren.
8. EDR: Wir halten es für angemessen, dass für die Analyse von EDR-Daten ein Durchsuchungsbefehl erforderlich ist und empfehlen keine Änderung des Strafrechts (Strafgesetzbuch, Strafprozessordnung. Die Strafprozessordnung soll technologisch neutral bleiben, eine spezielle Regelung für EDR-Daten ist nicht erforderlich.

Résumé

Contexte

Les véhicules contemporains représentent d'abondantes sources de données, équipés de nombreux capteurs qui surveillent tant le fonctionnement du véhicule (par exemple, les performances du moteur, le rendement énergétique et la conduite) que l'environnement environnant (comme les conditions météorologiques et les caractéristiques de la route) ainsi que les usagers de la route (par exemple, la localisation du véhicule et sa vitesse). Cette richesse d'informations promet de transformer la recherche et la prévention des accidents, contribuant à rendre les routes plus sûres.

Avec l'avènement de la conduite automatisée, une quantité encore plus importante de données de capteurs sera générée à l'avenir. Tout d'abord, les fabricants d'équipements d'origine nécessiteront ces données de capteurs pour entraîner leurs algorithmes de conduite automatisée. Ensuite, une fois les algorithmes développés, les véhicules auront besoin de données de capteurs pour fonctionner de manière autonome. Par conséquent, on peut s'attendre à ce qu'un nombre croissant de capteurs soient installés dans les années à venir, générant ainsi toujours plus de données.

En théorie, l'accès aux données des capteurs pourrait ouvrir de nouvelles perspectives pour la recherche et la prévention en matière de sécurité routière. Cependant, dans la réalité, les chercheurs et les spécialistes de la prévention des accidents se heurtent souvent à des obstacles pour obtenir ces données cruciales. Malgré l'intérêt public évident en faveur de l'amélioration des efforts de recherche et de prévention, l'obtention de données de capteurs pertinentes, telles que des informations détaillées sur les performances des véhicules ou des données de localisation anonymes, reste souvent difficile. Cette situation crée une tension entre le besoin public de routes plus sûres et l'accès actuellement limité aux données.

L'équilibre entre la protection des investissements et la promotion de l'innovation est complexe. D'un côté, l'intérêt public réside dans la promotion du partage des connaissances et des avancées technologiques, susceptibles de conduire à des véhicules et à des infrastructures plus sécurisés. D'un autre côté, les entreprises privées ont des intérêts légitimes à protéger leurs secrets commerciaux et leur propriété intellectuelle, souvent régis par diverses normes juridiques. Trouver un juste équilibre entre ces intérêts divergents est essentiel pour assurer un progrès durable.

Une grande partie des données des capteurs, comprenant des informations sur le fonctionnement du véhicule et la localisation de l'utilisateur, est catégorisée comme des données personnelles [1, 2], soulevant ainsi des préoccupations quant à la protection de la vie privée. Il est impératif d'évaluer attentivement ces préoccupations par rapport à l'objectif plus vaste de prévention des accidents. Heureusement, des technologies telles que la confidentialité différentielle et l'apprentissage fédéré offrent des solutions prometteuses permettant d'analyser les données tout en préservant la vie privée des individus.

L'exploitation du potentiel des données des capteurs, allant des performances des véhicules au comportement des conducteurs, représente une étape cruciale vers des routes plus sûres. Cependant, la navigation complexe entre les intérêts commerciaux, publics et privés pose des défis significatifs.

Objectif du projet

Ce projet explore les solutions juridiques et technologiques visant à libérer le potentiel des données, en mettant l'accent sur quatre cas d'utilisation clés identifiés grâce à la participation active des parties prenantes.

Approche méthodologique

En termes d'approche méthodologique, le projet tire parti de l'expertise issue de diverses disciplines scientifiques telles que la science des données, le droit, l'économie, la politique et la gouvernance, la psychologie des facteurs humains, et l'ingénierie. Il se positionne ainsi en tant que projet véritablement interdisciplinaire.

De surcroît, le projet adopte une approche transdisciplinaire en abordant des problèmes concrets et en étant spécifiquement conçu pour élaborer des solutions pertinentes pour les bénéficiaires. Dans ce contexte, les bénéficiaires incluent les chercheurs en sécurité routière et les spécialistes de la prévention issus des milieux universitaires, gouvernementaux et privés. La transdisciplinarité est assurée en impliquant directement les bénéficiaires dans le projet et en établissant un engagement dès le début, les intégrant dans la conception du projet plutôt que de les confronter aux résultats à la fin.

L'implication des bénéficiaires et des parties prenantes s'est effectuée de deux manières distinctes :

1. **En tant que membres de l'équipe :** Le Dynamic Test Center (DTC) de la Berner Fachhochschule réalise des reconstructions d'accidents et travaille concrètement avec des données de capteurs. De même, l'équipe de la Zürcher Hochschule für Angewandte Wissenschaften (ZHAW) représente la communauté des chercheurs souhaitant exploiter les données de capteurs.
2. **En tant que participants à l'atelier :** Plus de 30 praticiens et bénéficiaires ont pris part à une série de huit ateliers. Ces sessions ont permis l'identification des cas d'usage en collaboration avec les praticiens et les bénéficiaires, orientant ainsi le projet dans sa trajectoire.

Chaque cas d’usage aborde un problème spécifique résoluble grâce aux données de capteurs. Pour étayer l'analyse de ces cas, le projet comprend un inventaire complet des données de capteurs disponibles actuellement et potentiellement dans le futur. En outre, le projet propose une discussion approfondie sur les fondements juridiques et scientifiques pertinents qui ont ensuite été appliqués aux différents cas d’usage.

Fondements juridiques

Le projet offre un aperçu des fondements juridiques pertinents dans le contexte des données de capteurs, détaillant le cadre juridique applicable tant en Suisse qu'au sein de l'Union européenne. En Suisse, la loi fédérale sur la protection des données (LPD) revêt une importance particulière pour ce sujet. Sur le plan européen, il convient de souligner que plusieurs règlements concernant la collecte et le traitement des données personnelles par les voitures connectées ont été adoptés.

De plus, le projet propose une discussion approfondie sur trois aspects juridiques cruciaux : l'accès aux données, la protection des données et la procédure pénale. Il explore les modèles de partage de données pouvant être utilisés pour accéder aux données collectées par les voitures, offrant des conseils sur les principes fondamentaux de la loi sur la protection des données. Enfin, le projet aborde deux questions spécifiques qui se posent dans le contexte des procédures pénales : la nécessité d'un mandat de perquisition pour l'utilisation des données collectées par les voitures et la possibilité d'utiliser des preuves obtenues illégalement, telles que des données collectées par une voiture en violation de la LPD.

Fondements de la science des données

Ce projet élabore un cadre structuré pour exploiter judicieusement les données des capteurs des véhicules dans la recherche et la prévention des accidents, plaçant la protection de la vie privée au cœur de tout le processus. Une évaluation approfondie des risques potentiels pour la vie privée liés à ces données est effectuée, et une solution d'accès et de traitement des données, renforçant la protection de la vie privée, est proposée en se basant sur le concept des cinq coffres-forts.

En se concentrant sur des mesures intégrées axées sur les données et en utilisant des techniques de protection de la vie privée adaptées, le projet présente des stratégies efficaces pour atténuer ces risques. Il est essentiel de souligner que le niveau de protection doit être ajusté en fonction des besoins spécifiques des divers objectifs de recherche, garantissant ainsi un traitement responsable des données.

Par ailleurs, l'analyse démontre qu'en principe, des technologies telles que l'apprentissage fédéré, le cryptage homomorphe et le calcul multipartite sécurisé peuvent permettre le partage des données des capteurs avec la communauté de la recherche et de la prévention. Ces technologies respectent les exigences de protection des données tout en tenant compte des intérêts commerciaux des entreprises et des personnes concernées.

Architecture de gouvernance

Afin de tirer pleinement parti du potentiel des données des capteurs, une architecture de gouvernance complète est suggérée, visant à rassembler un ensemble varié de parties prenantes. Cette approche collaborative est considérée comme cruciale et requiert l'établissement d'un espace de données sécurisé et fiable. Cet environnement dédié est spécifiquement conçu pour faciliter l'accès aux données et leur analyse, tout en encourageant la transparence et la coopération entre toutes les parties concernées.

Analyse des uses cases

Les principes juridiques et les avancées en science des données sont appliqués à quatre cas d'usage spécifiques :

1. **Cas d'usage 1- Élargissement des statistiques sur les accidents** : Cette première analyse vise à déterminer si les statistiques d'accidents existantes peuvent être améliorées en exploitant les données des capteurs. L'objectif est d'élargir le pool de données pour obtenir des informations plus approfondies sur les accidents, afin de développer des stratégies de prévention plus efficaces.
2. **Cas d'usage 2- Exposition** : En se penchant sur les données des capteurs, ce cas d'utilisation cherche à obtenir des informations détaillées sur le comportement au volant, permettant ainsi d'identifier les zones à risque et d'adapter les interventions en matière de sécurité.
3. **Cas d'usage 3- Alertes en cas de danger et prévention en temps réel** : En imaginant la réception d'alertes en temps réel sur les dangers potentiels de la route, ce scénario explore comment les données des capteurs pourraient être exploitées pour prédire et prévenir les accidents en temps réel.
4. **Cas d'usage 4- Reconstitution des accidents et détermination de la responsabilité pénale** : Face à la complexité à élucider les causes d'un accident, ce cas d'utilisation examine comment les données des capteurs pourraient être utilisées pour reconstituer les accidents avec plus de précision et déterminer la responsabilité de manière plus efficace.

Discussion et conclusions

En théorie, l'accès aux données des capteurs pourrait ouvrir de nouvelles perspectives passionnantes pour la recherche et la prévention en matière de sécurité routière. Cependant, plusieurs défis doivent être relevés. Aujourd'hui, il manque des valeurs seuils suffisantes pour interpréter les données des capteurs, et des problèmes de qualité et de normalisation sont à anticiper.

Le défi le plus significatif semble actuellement résider dans l'accès aux données. Notre analyse indique que, en principe, des technologies telles que l'apprentissage fédéré, le cryptage homomorphe et le calcul multipartite sécurisé sont disponibles pour partager les données des capteurs avec la communauté de la recherche et de la prévention, tout en tenant compte des intérêts commerciaux, des préoccupations en matière de responsabilité et des risques pour la vie privée qui entravent souvent ce partage. Cependant, l'application

de ces technologies peut être coûteuse et nécessite fréquemment la collaboration de multiples acteurs. À cet égard, des architectures de gouvernance s'avèrent indispensables. Elles définissent qui peut accéder aux données et dans quelles circonstances, tout en précisant la répartition des coûts et des avantages.

Certains fabricants et fournisseurs de services proposent la vente de données de capteurs ou d'analyses basées sur ces données. Au cours du projet, de nouveaux services ont également émergé. De manière concomitante, l'intervention de l'État a joué un rôle crucial dans les cas d'utilisation que nous avons examinés. À plusieurs reprises, l'Union européenne a adopté une législation imposant aux entreprises le partage des données de capteurs lorsque celles-ci peuvent contribuer à améliorer la sécurité routière.

En ce qui concerne l'accessibilité des données de capteurs, l'Office fédéral des routes (OFROU), la Confédération suisse, ainsi que les gouvernements cantonaux ou communaux, disposent de plusieurs options. Tout d'abord, ils peuvent attendre que le secteur privé mette éventuellement les données des capteurs à leur disposition. Deuxièmement, ils peuvent adopter des lois contraignantes qui obligent les entreprises à partager les données des capteurs. Enfin, l'État peut faciliter, par le biais de subventions financières ou d'activités de coordination, l'émergence d'architectures de gouvernance permettant le partage des données de capteurs.

Toutefois, les décideurs politiques suisses doivent tenir compte du contexte international, comme l'a reconnu le Conseil fédéral.

Recommandations

1. Données sur les quasi-accidents : Nous préconisons que l'OFROU, ainsi que les gouvernements cantonaux et communaux, commencent à explorer la possibilité d'utiliser les données relatives aux quasi-accidents afin d'identifier les risques sur le réseau routier. Actuellement, plusieurs entreprises proposent déjà à la vente des données concernant les quasi-accidents. Étant donné l'importance des enjeux publics, nous recommandons d'examiner les voies juridiques permettant d'inciter les équipementiers à partager gratuitement les données liées aux quasi-accidents. Il est également suggéré d'instaurer des contrôles de qualité et des tests de validation pour garantir la qualité et la comparabilité des données. Lors de l'utilisation de données d'accidents, il est recommandé d'effectuer une évaluation des risques pour la vie privée en se basant sur le cadre des cinq sécurités. Nous préconisons l'utilisation de la liste de contrôle présentée au chapitre 6.5.
2. Nous recommandons à l'OFROU d'explorer les moyens d'intégrer les données des enregistreurs de données d'événements (EDR) dans les statistiques d'accidents. En soi, les données EDR peuvent s'avérer difficiles à interpréter. Par conséquent, nous conseillons à l'OFROU de développer des méthodes de validation et de triangulation des données EDR avec d'autres sources. Nous partons du principe que les données peuvent être suffisamment anonymisées pour éviter tout problème de protection des données.

3. In the future, when policymakers develop new regulations that concerns sensor data, we recommend that they include in the law research and prevention as one of the purposes for which the data can be used. The example of eCall data shows that if this is not defined as the explicit purpose data protection rules may prohibit the research and prevention community from using the concerned data.

4. Nous recommandons à l'OFROU de surveiller et de promouvoir le développement d'architectures de gouvernance visant à faciliter le partage plus étendu des données de capteurs avec les chercheurs spécialisés dans les accidents et les experts en prévention, tant au sein des universités que dans les secteurs gouvernementaux et privés. Notre conclusion indique qu'il est techniquement possible de mettre en place des systèmes autorisant le partage des données de capteurs tout en respectant les

impératifs de confidentialité et en tenant compte des intérêts commerciaux des fabricants.

5. **Données d'exposition** : Les données de capteurs présentent un potentiel considérable dans le contexte de l'exposition au risque. Toutefois, en raison du manque actuel de représentativité, notamment des données de mobilité, elles pourraient ne pas encore être en mesure de remplacer intégralement le micro-recensement de la mobilité. De même, elles ne peuvent pas se substituer aux données fournies par les compteurs de trafic. Néanmoins, les données de mobilité demeurent utiles dans les situations où les données de recensement ou les informations des compteurs de trafic ne sont pas disponibles, notamment pour des quartiers ou des tronçons de route spécifiques. Nous recommandons à l'OFROU de soutenir la recherche qui explore de nouvelles méthodes de génération de données d'exposition à partir des données de mobilité.
6. **Avertissements de danger** : Nous préconisons que la Confédération suisse rejoigne et appuie le projet Data for Road Safety. Nous recommandons également que ce projet serve de modèle pour d'autres applications nécessitant la collaboration de divers intervenants, notamment les fabricants et les autorités de transport.
7. **EDR** : Pour faciliter l'accès aux données des enregistreurs de données d'événements (EDR), nous recommandons l'établissement d'une base juridique obligeant les importateurs de véhicules à rendre accessibles les données EDR au ministère public. Étant donné que les frabricants établis dans d'autres pays ne peuvent être directement soumis à une réglementation suisse, les obligations légales imposées aux importateurs peuvent constituer un levier efficace pour inciter les fabricants à rendre les données EDR accessibles.
8. **EDR :** Nous estimons qu'il n'est pas nécessaire de modifier le droit pénal (Code de procédure pénale - CrimPC) pour faciliter l'accès de la police et du ministère public aux données de conservation des données EDR. Il est essentiel que le CrimPC CPP demeure neutre sur le plan technologique, et l'instauration d'un régime spécifique pour les données EDR n'apparaît pas comme une nécessité.

Pour l'avenir, lors de l'élaboration de nouvelles réglementations sur les données des capteurs, nous recommandons aux décideurs politiques d'inclure la recherche et la prévention comme l'un des objectifs légaux pour lesquels les données peuvent être utilisées. L'exemple des données eCall illustre que si cette finalité n'est pas explicitement définie, les règles de protection des données peuvent empêcher la communauté de la recherche et de la prévention d'utiliser les données concernées.

1 Introduction

1.1 Background

The data recorded by an increasing number of sensors installed in cars has the potential to revolutionize the field of accident research and prevention [1]. The new data collected by vehicle sensors promises, for instance, to:

- Replace costly naturalist driving studies;
- Identify near-accidents by using stochastic models;
- Provide insights into the chain of events leading up to (near) accidents,
- Facilitate the reconstruction of accidents etc.

The opportunities are limitless. But so are the challenges:

1. Lack of access: Accident researchers and prevention specialists often do not have access to the sensor data. For the most part, car and original equipment manufacturers (OEMs) control access to the data by means of encryption technologies (technically) and patent protections (legally). Often, researchers - and even prosecutors - have to rely on the goodwill of companies (OEMs, service providers, etc.) to share sensor data with them.
2. Public and private interests: There is a serious public interest to provide researchers and prevention specialists with access to sensor data. However, private companies, too, have legitimate interests not to share such data. They have invested in technologies and have a right to protect their investments. And if companies were forced to share their data, this could potentially jeopardize their business model and undermine their incentives to invest in the necessary research and development.
3. Protecting privacy: Most sensor data are personal data. Therefore, legal protections apply. [1, 2]. The data may not only provide information on the accident, but also the private life and habits of individuals. Therefore, the privacy rights of the individual need to be balanced carefully against the broader public interest of generating new knowledge about road safety, preventing accidents, reconstructing accidents and supporting accident investigations.
4. Lack of validated threshold values: Even when access is granted there are no validated threshold value in many cases. This makes the data difficult to interpret.

In sum, access to data means negotiating a complex set of commercial, public and private interests (see top part of Fig. 1) and find a balanced approach.

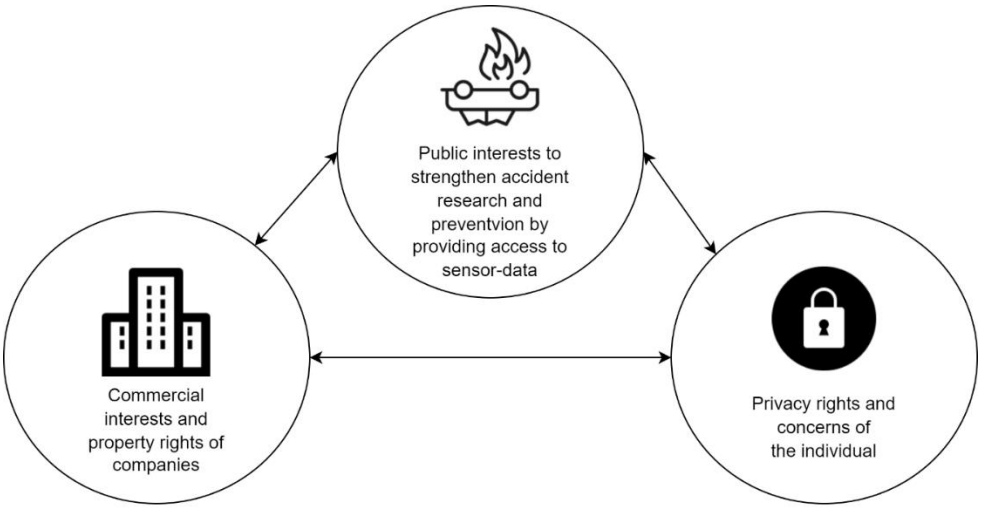

***Fig. 1** Public, private and commercial interests need to be considered*

1.2 Objectives

In the project, we explore legal and technological solutions for balancing:

1. the strong public interests in favor of sharing data to make our roads safer with
2. the legitimate commercial interests of vehicle manufacturers, equipment and service providers, as well as
3. the legitimate privacy rights of individuals.

The objective is to find and to assess solutions for making sensor data available to accident researchers and prevention specialists and respect the above-mentioned interests. These solutions could for example include laws that require the sharing of sensor data when public interests (improving road safety) justify this or data science systems (e.g., secure data spaces, clearing house models) that allow for the secure sharing of sensor data.

Given the transdisciplinary nature of the project, however, we seek to maximize the relevance of the project to the beneficiaries and stakeholders by engaging with them and by applying the findings to concrete use cases.

1.3 Structure of the report

After a brief description of the methodological approach (Chapter 2), we provide an overview of the available sensor data (Chapter 3), the legal foundations (Chapter 5) and the data science foundations (Chapter 6) that need to be considered, when dealing with sensor data. Based on these foundations, we present a governance architecture to be applied for some use cases (Chapter 7). Then, we study the potential of using sensor data in four use cases (Chapters 8-11) and conclude with our recommendations (Chapter 12).

2 Methodological approach

2.1 Project design

The project is interdisciplinary and transdisciplinary. Interdisciplinarity means that multiple scientific disciplines are brought together in the project: data science, law, economics of innovation, politics and governance, human factors, psychology, engineering.

Transdisciplinarity means that the project seeks to address concrete problems and to develop relevant solutions to beneficiaries, which, in the case of this project, means road safety researchers and prevention specialists in academia, government and private sector. Transdisciplinarity is achieved by a) directly involving the beneficiaries in the project and b) by engaging with the beneficiaries from the beginning and involving them in the project design as opposed to confronting them with the results at the end. Thereby, the project design assures that the project's focus is of actual relevance to the beneficiaries and stakeholders.

In this project, beneficiaries and stakeholders were involved in two ways.

1. They were involved as team members. DTC do accident reconstruction and work with sensor data in practice. The team from ZHAW represents the research community that would like to use sensor data.
2. Additional practitioners and beneficiaries were contacted and involved in the second work package (see below). We began with a systematic mapping of the landscape of stakeholders and beneficiaries. These stakeholders were then contacted and involved in a series of workshops, in which they were asked about their experiences using sensor data and the opportunities and challenges they see with regard to the use of sensor data. Based on the results of the workshops, we selected specific use cases to be analyzed in our project.

Thereby, the project design follows the principles and recommendations defined by the Network for Transdisciplinary Research (td-net), which is an initiative of the Swiss Academies of Arts and Sciences that was launched in 2000. [3]

2.2 Work packages

Given the transdisciplinary nature of the project, the work packages (WP) were designed to allow for the integration of different disciplinary perspectives (WP3 und WP4) and the engagement with beneficiaries and stakeholders (WP2). To ensure relevance to the beneficiaries and stakeholders, the project centers on an analysis of four use cases, in which we apply the legal foundations and the data-science foundations established in WP3 and WP4. The following figure summarizes the sequence and interactions of the work packages.

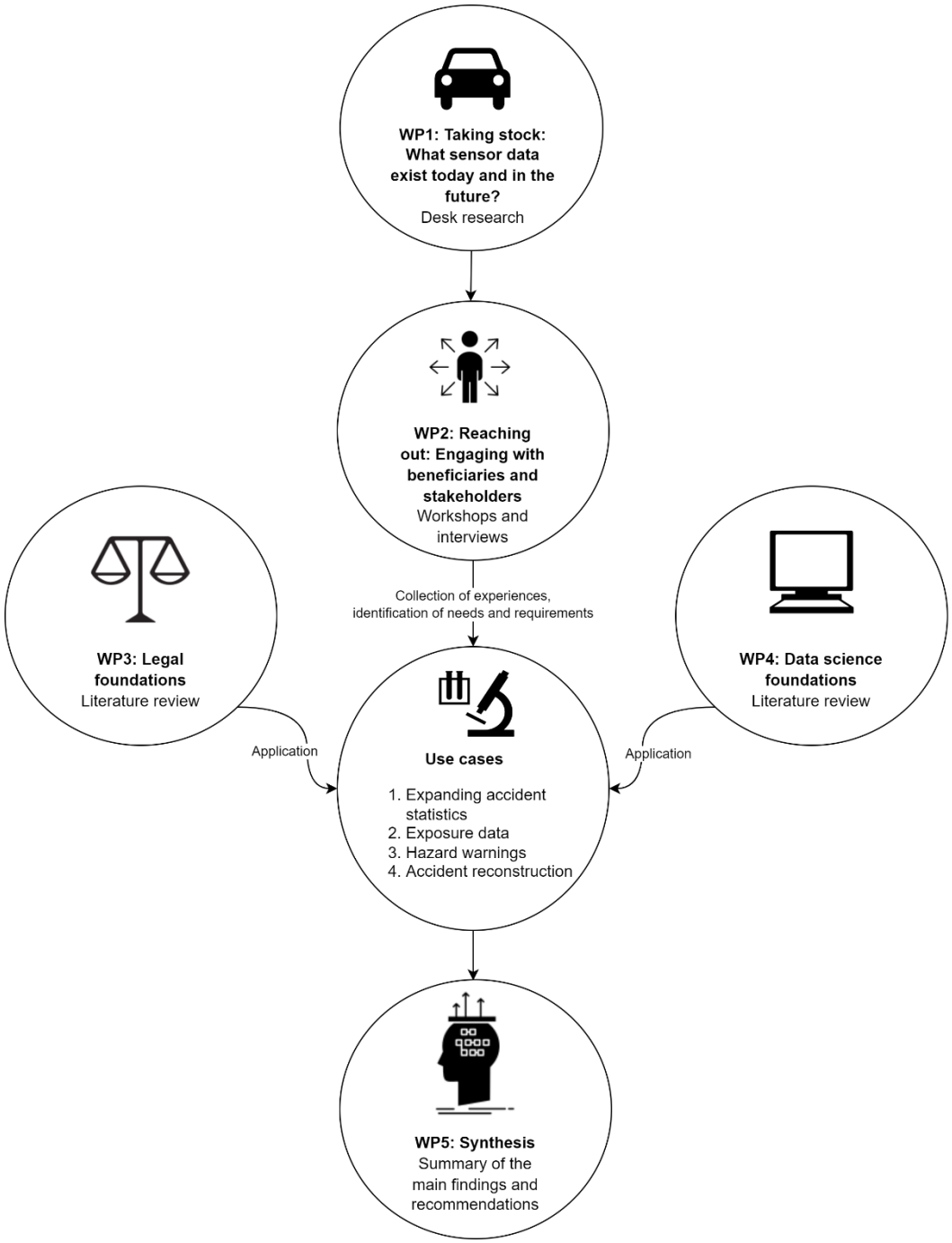

***Fig. 2** Work packages*

2.2.1 Taking stock

As a foundation for the project, we began by taking stock of what vehicle sensor data is available today and what could potentially become available in the future. This work package was based on:

1. Desk research and review of the scientific and policy literature;
2. Practical experience of Dynamic Test Center (DTC);
3. Interviews with other experts

The results of the stock-taking are presented in chapter 3.

2.2.2 Reaching out: Identifying and defining use cases

Accident research and prevention are a large field. It involves diverse disciplines (e.g., biomechanics, forensics, psychology, accident reconstruction, engineering, etc.).[1] Therefore, the project cannot possibly focus on everything that the use of sensor data might mean for research and prevention. To lend more focus to the analysis and to make sure that the analysis is relevant to the beneficiaries and stakeholders, we have decided to focus on specific use cases. Each use case addresses a different problem that may be addressed with the help of sensor data. In the use cases analysis, based on the legal and data science foundations, we evaluate the possibilities for accessing and using sensor data.

To identify relevant use cases, we held eight workshops with policymakers, representatives of associations and interest groups, legal experts, industry representatives, accident reconstruction specialists and researchers (a list of workshop participants is provided in the appendix).

Each workshop followed the following structure:

1. Introduction of the participants, tour de table (10 minutes)
2. Presentation of the project and the role of the workshop in the project (5 minutes)
3. Discussion: (60 minutes)
    1. Prior experience of using sensor data
    2. Collection of potential use cases
    3. Inputs for the development of legal and technological solutions
4. Wrap-up (5 minutes)

2.2.3 Legal and data science foundations

As background for the use case analysis, we have summarized the legal foundations and the data science foundations. Our summaries are presented in Chapters 5 and 6.

2.2.4 Use case analysis

Finally, we apply the legal and data-science foundations to four use cases. For the use cases, additional interviews were conducted, as a document and literature review.

[1] And the potential use of sensor data becoming is likely to open up many new avenues of research.

3 Stock-taking

3.1 Overview of the (potentially) available data sources

In this chapter, we take stock of what sensor data already exists today and could potentially be used for the purpose of research and prevention. We also estimate what data could become available in the future.

An inventory of currently available and future sensor data sources is provided by a C-ITS study (Cooperative Intelligent Transport Systems in the EU) [86], which is listed in Annex 11 ("Use cases - data to be harmonized - status Nov. 2015"). The list contains around 160 parameters, categorized into the following 21 groups:

1. Vehicle sensor data,
2. Vehicle-identifying data,
3. Vehicle status,
4. Vehicle speed / acceleration,
5. Vehicle position / direction,
6. The fuel,
7. Electric vehicles only,
8. Battery status,
9. Brake,
10. Tires,
11. Oil,
12. Cooling water,
13. Air conditioning,
14. Fault indicator,
15. Electronic control unit,
16. ABS,
17. Security status (anti-theft system),
18. Mileage/total odometer reading,
19. PTI-relevant data,
20. Wiper blades,
21. Other identified data requirements,

The parameters listed in Annex 11 of the C-ITS-study do not only relate to accident research and prevention. The parameters also relate to sensor data that is relevant for repair and maintenance, usage-based insurance, theft notification and recovery, traffic management, vehicle information für fleet vehicle operations, car rental, route planning, city toll and so on. This list is not exhaustive. In our report, however, we concentrate on data that could be useful for accident research and prevention.

Many of the parameters mentioned in the list above are managed, stored and analyzed in various systems in the vehicle, with the Event Data Recorder (EDR) being the most relevant to this project. Others are eCall (emergency Call) and DSSAD (Data Storage System of Automated Driving). In the following subsections we discuss the parameters recorded by vehicle sensors in the context of these systems, starting with the EDR, eCall and DSSAD. We then discuss the potential of floating car data (data collected from vehicles in motion, such as speed, location, and direction, used primarily for traffic monitoring and management, as well as for enhancing navigation systems.) and floating personal data (data about an individual that is collected and disseminated through various sources and devices such as cell phones, often without the individual's explicit consent or knowledge) as additional categories. The table below provides an overview. The subsequent sections provide more detailed elaborations.

The most important abbreviations are explained below in the text and in the glossary in the appendix.

***Tab. 1** Overview of data sources*

Category	Brief description, explanation (not exhaustive)	Data quality characteristics
EDR	**E**vent **D**ata **R**ecorder: The EDR stores the following data (list not exhaustive): - Longitudinal and lateral Delta-V, - Vehicle speed, - Steering, angle, speed, steering, - Angle, yaw rate, brake or accelerator operation, - Brake pressure, ABS activity, - Engine rpm, - Airbag and pretensioner deployment, seat track position etc.	Requirements for recording time, sample rate, minimum range, accuracy and resolution, which differ depending on the directive and parameters, EDR is a circular buffer that records data -5 s to 0 s before the crash (at 2 Hz) and 0 s to 300 ms during the crash (at 100 Hz). In the EU five to six storage areas for deployment and non-deployment events.
eCall	**E**mergency Call: In case of a corresponding accident event, the eCall system sends the following data [81] (list not exhaustive), but based on the DIN EN 17522:2021 standard): - Time Stamp - Geoposition (last and the two previous), - Vehicle type (category) - Driving direction - Vehicle Propulsion Storage Type - Number of belted occupants, rollover	Mandatory in the EU since 2014, eCall also works in Switzerland, previously via the emergency call systems of the vehicle manufacturers, now directly via the nearest police operations center.
DSSAD	The Data Storage System for Automated Driving needs to record the following parameters: - Activation of the automation system - Deactivation of the automation system and Reason for deactivation - Request of the automation system to the driver to take over the vehicle and the reason for the takeover request - Suppression or mitigation of driver intervention by the automation system - Initiation by the automation system of a maneuver to minimize risk - Initiation of an emergency maneuver by the automation system - Occurrence of safety-relevant technical faults.	The legal framework is currently being developed, therefore no details about the data quality are assured.[2] [82]
Floating car data	Camera, lidar, radar	
Floating personal data	Mobile Phone, Wearables GPS Tracker, motions sensors	
Infrastructure data	Traffic cameras, counters	

In the following chapters, we provide a brief overview of the categories presented in the table above.

3.2 EDR

EDR Event Data Recorder: This is a data storage system which is integrated in the airbag control module (ACM). The EDR measures continuously, but only stores data when there is a deployment event (triggering of non-reversible systems, such as airbag and seatbelt pretensioner triggering, etc.) or a non-deployment event (triggering criterion Delta-V>8 kph within 150 ms). Deployment events can overwrite non-deployment events. A deployment event cannot be overwritten. The trigger thresholds for EDR triggering are currently not suitable for detecting collisions between passenger cars and vulnerable road users (VRU) such as cyclists and pedestrians, unless a pedestrian protection system is activated Collisions with pedestrians typically do not lead to a collision-related speed change of more than 8 kph. Thus, no EDR data is stored after such collisions with pedestrians or cyclists.

[2] However, product safety rules already exist and apply to automated driving. See Bundesgesetz über die Produktesicherheit (PrSG, SR 930.11).

For vehicles homologated in the USA and Canada, the EDR has been legally mandatory since 2012. Title 49 of the United States Code of federal regulations (Part 563 – Event data recorders) specifies which data set must be recorded and in which accuracy, resolution and duration it must be stored. The regulation also stipulates that vehicle manufacturers must make tools and/or methods for data retrieval commercially available so that accident investigators and researchers can retrieve the digital traces from the EDR. Most vehicles can be read out with the Bosch CDR system. Individual brands require their own readout system. These include Kia, Hyundai, Tesla etc.

In Europe, EDR has only becoming mandatory in 2022. Regulation (EU) 2019/2144 states that in the EU passenger cars (M1) and light commercial vehicles (N1) with new type approval will have to comply with the EDR requirements from 7 July 2022 (new models) and for all new passenger car registrations from 7 July 2024 (Tab. 2). The EU regulation was drafted based on the UN Regulation No 160 [91] – Uniform provisions concerning the approval of motor vehicles with regard to the Event Data Recorder [2021/1215], which is abbreviated as UN ECE R 160. Supplementary information on Regulation (EU) 2019/2144 can be found in the Delegated Regulation Vehicle safety - technical requirements & test procedures for EU type-approval of event data recorders (EDRs), which has been available since January 2022. This, in turn, refers to UN Regulation No. 160, in which the characteristics of the EDR were harmonized.

The EU regulations go further than American law (CFR 49) as more parameters need to be recorded. In a first step, 41 parameters need to be recorded (Tab. 2).

***Tab. 2** EDR - Data elements and format*
date of entry into force 30 September 2021

Data elements	**Characteristics or Phase**
Delta-V, longitudinal	Postcrash, 0 to 250 ms
Maximum delta-V, longitudinal	Postcrash, 0 to 300 ms
Time, maximum delta-V, longitudinal	Postcrash, 0 to 300 ms
Speed, vehicle indicated	Precrash, -5 to 0 sec
Engine throttle, % full	Precrash, -5 to 0 sec
Service brake, on/off	Precrash, -5 to 0 sec
Ignition cycle, crash	Precrash, -1.0 sec
Ignition cycle, download	At time of download
Safety Belt status, driver	Precrash, -1.0 sec
Airbag warning lamp	Precrash, -1.0 sec
Frontal airbag deployment, time to deploy, driver	Event
Frontal airbag deployment, time to deploy, front passenger	Event
Multi-event crash, number of events	Event
Time from events 1 to 2	As needed
Complete file recorded	Following other data
Lateral acceleration	Postcrash, 0 to 250 ms
Longitudinal acceleration	Postcrash, 0 to 250 ms
Normal acceleration	Postcrash, -1 to 5 sec
Delta-v, lateral	Postcrash, 0 to 250 ms
Maximum delta-V, lateral	Postcrash, 0 to 300 ms
Time maximum delta-V	Postcrash, 0 to 300 ms
Time maximum delta-V, resultant	Postcrash, 0 to 300 ms
Engine rpm	Precrash, -5 to 0 sec
Vehicle roll angle	Pre-/Postcrash -1 up to 5 sec
ABS activity	Precrash, -5 to 0 sec
Stability Control	Precrash, -5 to 0 sec
Steering Input	Precrash, -5 to 0 sec
Safety belt status front passenger	Precrash, -1.0 sec
Passenger Airbag suppression status	Precrash, -1.0 sec
Front Airbag deployment time to nth stage, driver	Event
Front Airbag deployment time to nth stage, front passenger	Event
	Event

***Tab. 2** EDR - Data elements and format*
date of entry into force 30 September 2021

Side Airbag deployment, time to deploy, driver	Event
Side Airbag deployment, time to deploy, front passenger	Event
Side curtain/tube Airbag deployment, time to deploy, driver	Event
Side curtain/tube Airbag, time to deploy, front passenger	Event
Pretensioner deployment, time to fire, driver	Event
Pretensioner deployment, time to fire, front passenger	Precrash, -1.0 sec
Seat track position switch, foremost status driver	Precrash, -1.0 sec
Seat track position switch, foremost status front passenger	Precrash, -1.0 sec
Occupant size classification, driver	Precrash, -1.0 sec
Occupant size classification, front passenger	

In a second step, another 23 parameters will need to be recorded (Tab*. 3*).

***Tab. 3** EDR – revision of the data elements and format*
01 series of amendments - date of entry into force: 8 October 2022 – in addition to Tab. 2

Data elements	**Characteristics or Phase**
Safety Belt Status, rear passengers	Precrash, -1.0 sec
Tire Pressure Monitoring System TPMS Warning Lamp Status	
Longitudinal acceleration	Precrash, -1.0 sec
Lateral acceleration	Precrash, -5 to 0 sec
Yaw Rate	Precrash, -5 to 0 sec
Traction Control Status	Precrash, -5 to 0 sec
AEBS Status	Precrash, -5 to 0 sec
Cruise Control System	Precrash, -5 to 0 sec
Adaptive Cruise Control Status	Precrash, -5 to 0 sec
VRU secondary safety system deployment, time to deploy	Precrash, -5 to 0 sec
VRU secondary safety system warning indicator status	Event
Safety belt status midposition front	Precrash, -1.1 sec
Far side impact center airbag	Precrash, -1.0 sec
Lane departure warning system status	Event
Corrective Steering function status	Precrash, -5 to 0 sec
Emergency steering function status	Precrash, -5 to 0 sec
Automatically commanded steering function category A status	Precrash, -5 to 0 sec
Automatically commanded steering function category B1 status	Precrash, -5 to 0 sec
Automatically commanded steering function category B2 status	
Automatically commanded steering function category C status	Precrash, -5 to 0 sec
Automatically commanded steering function category D status	
Automatically commanded steering function category E status	Precrash, -5 to 0 sec Precrash,
Accident emergency call system status	-5 to 0 sec
	Precrash, -5 to 0 sec
	Precrash, -5 to 0 sec
	Event

Tab. 3 already contains data elements of the DSSAD (Automatically commanded steering function category status). These must be available for newly homologated models from 7 July 2024 and for new registrations from 7 July 2026.

The data stored in the EDR does not allow the identification of the vehicle occupants [92]. At present, the EDR does not store audio and video files or the driver’s habits and behavior

of the driver. According to EU Regulation 2019/2144, the last 4 digits of the VIN need to be omitted when reading out the EDR.

How can the EDR data be accessed? The EDR data remains in the vehicle and cannot be deleted. Reading out the EDR data requires access to the vehicle via the standardised onboard interface or directly via the airbag control unit. Based on DTC's professional experience of dealing with EDRs for more than ten years, there are four options for accessing data:

1. Direct readout using Bosch CDR via the standardized OBD 2 interface (On Board Diagnostics 2). The EDR function (Event Data Reader) is integrated in the Airbag Control Module and the power supply to the control unit is intact.
2. Remove the control unit and read it out in the laboratory (also using the Bosch CDR tool). The EDR function (Event Data Reader) is integrated in the Airbag Control Module. If the manufacturer has not provided access via Bosch CDR or other devices in the Airbag Control Module, the data can only be read out by the manufacturer of the control unit. That means that the EDR unit has to be physically removed and send to the manufacturer.
3. In older vehicles, the Airbag control Module often does not contain any usable data relevant to the accident.

Most manufacturers of airbag control units now require orders from the investigating authorities as well as a declaration of consent from the vehicle owner and, in some cases, the vehicle manufacturer. While some manufacturers provide the EDR readings free of charge, others charge high fees or refuse to pass on the data.

For illustration purposes we present two anonymized copies of CDR reports below. Both CDR are still based on American legislation (CFR 49) The first report (xxx_RAV4) is from a vehicle whose driver, while turning left into a side street, overlooked an oncoming cyclist who had the right of way. As a result, the cyclist collided with the right side of the vehicle. The Toyota driver stated that she stopped before turning left. The collision speed of 22 km/h at the time of impact could also have been reached if the driver had stopped before turning. However, the EDR data log shows that the driver only slowed down before turning but did not stop. The vehicle speeds in the second row are highlighted by the red box. Had the driver stopped, the cyclist might have had enough time to react. Since the driver did not stop, the cyclist did not have enough time to react. Therefore, the EDR data suggests that the responsibility for the accident rests with the driver of the Toyota, not the cyclist.

Pre-Crash Data, -5 to 0 seconds (Most Recent Event, TRG 1)

Time (sec)	-4.75	-4.25	-3.75	-3.25	-2.75	-2.25	-1.75	-1.25	-0.75	-0.25	0 (TRG)
Vehicle Speed (MPH [km/h])	30.4 [49]	29.2 [47]	28 [45]	26.1 [42]	23 [37]	19.9 [32]	17.4 [28]	15.5 [25]	14.3 [23]	13.7 [22]	13.7 [22]
Accelerator Pedal, % Full (%)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	27.0	28.5	28.5
Percentage of Engine Throttle (%)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.5	7.5	7.5
Engine RPM (RPM)	1,800	1,800	1,800	1,800	1,700	1,800	1,800	1,800	1,800	1,900	2,000
Motor RPM (RPM)	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid
Service Brake, ON/OFF	OFF	OFF	ON	ON	ON	ON	ON	OFF	OFF	OFF	OFF
Brake Oil Pressure (Mpa)	0.00	0.00	0.05	0.86	1.49	0.77	0.24	0.00	0.00	0.00	0.00
Longitudinal Acceleration , VSC Sensor (m/sec^2)	0.000	-0.502	-0.646	-1.723	-2.584	-1.938	-1.364	-1.005	-0.718	0.144	-0.646
Yaw Rate (deg/sec)	-2.93	-3.42	-3.42	-3.42	-2.93	0.98	11.71	24.89	30.74	31.23	30.26
Steering Input (degrees)	-9.0	-13.5	-13.5	-13.5	-10.5	15.0	85.5	168.0	222.0	226.5	225.0
Shift Position	D	D	D	D	D	D	D	D	D	D	D
Sequential Shift Range	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
Cruise Control Status	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF
Drive Mode, PWR	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF
Drive Mode, ECO	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF
Drive Mode, Sport	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF
Drive Mode, Snow	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF	OFF
Drive Mode, EV	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid	Invalid

***Fig. 3** EDR readout. Example 1*

Source: EDR readout obtained by DTCs.

The following diagram provides a visual representation of the accident.

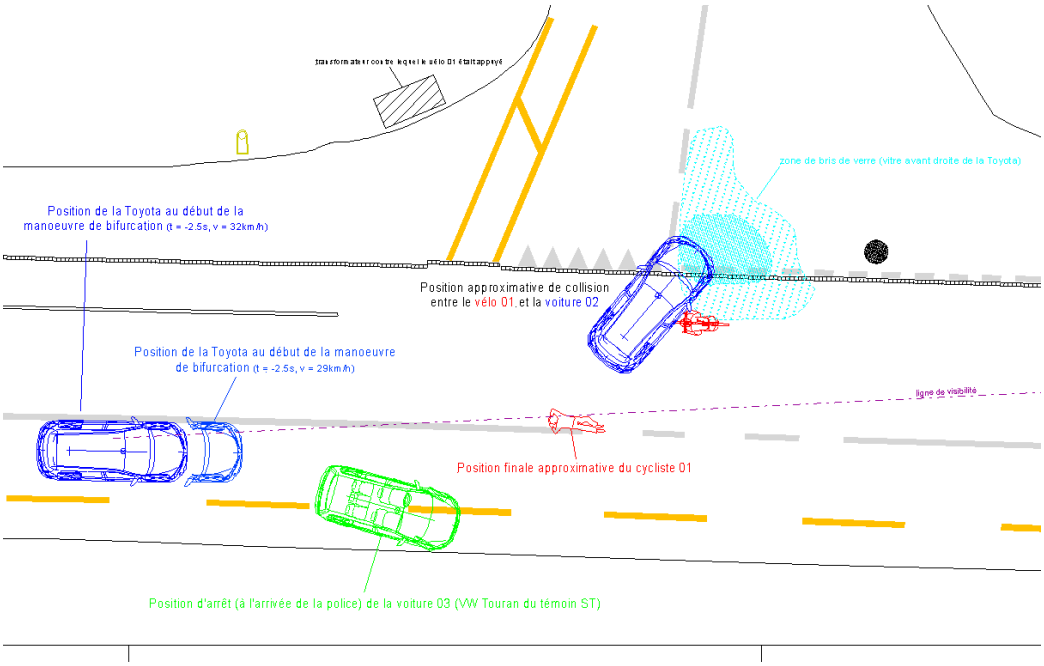

***Fig. 4** Pre-crash data and accident plan of a collision between a left-turning vehicle and an oncoming cyclist*

Source: EDR readout and accident plan obtained by DTCs.

The second report (XXX_T-Cross) is from a rear-end collision during a crash test at the DTC Dynamic Test Center AG, in which a VW T-Cross collided with the rear of a delivery van (Peugeot Expert). This example is meant to show that the EDR data can be imprecise. According to EDR data, the collision speed was 34 kph. However, external measurements

showed that the real speed was actually at 27.5 kph. We will return to such imprecisions later on.

Pre-Crash Data -5 to 0 sec (Record 1, Most Recent)

Time (sec)	Engine RPM (Combustion Engine) (RPM)	ABS Activity	Stability Control	Steering Input (deg)	Speed, Vehicle Indicated (MPH [km/h])	Accelerator Pedal (%)	Service Brake Activation
-5.0	960	Non-Engaged	On	-2	0 [0]	0	Off
-4.5	960	Non-Engaged	On	-2	0 [0]	0	Off
-4.0	960	Non-Engaged	On	-2	0 [0]	0	Off
-3.5	960	Non-Engaged	On	-2	2 [4]	0	Off
-3.0	960	Non-Engaged	On	-2	6 [9]	0	Off
-2.5	960	Non-Engaged	On	-2	9 [15]	0	Off
-2.0	960	Non-Engaged	On	0	12 [20]	0	Off
-1.5	960	Non-Engaged	On	0	16 [25]	0	Off
-1.0	960	Non-Engaged	On	0	19 [30]	0	Off
-0.5	960	Non-Engaged	On	0	21 [33]	0	Off
0.0	960	Non-Engaged	On	0	21 [34]	0	Off

Fig. 5 *EDR readout Example 2*

As the speed in the airbag control unit comes from the averaged value of the four ABS sensors, the accuracy of the values stored in the vehicle depends on the wheel slip (braking) and the tire diameter (influenced by air pressure and tire dimension).

The two examples demonstrate that the EDR readouts always need to be analyzed carefully because they do not necessarily provide a correct representation of the accident. The data needs to be triangulated with other sources.

More broadly, the case of the EDR is also interesting because from the data stored in the EDR it is possible to infer what sensors exist in modern vehicles.

3.3 eCall

EU regulation 2015/785 requires all new cars to be equipped with eCall-technology from April 2018.[3] In the event of a serious accident detected by the vehicle's sensors or processors, eCall automatically dials 112 - Europe's single emergency number. The emergency center operator communicates directly with the driver of the vehicle and tries to find out what happened and what kind of help is needed. Although there is no permanent transmission of position data, eCall communicates in case of an accident the vehicle's exact location to emergency services. Furthermore, the time of the accident, the direction of travel (previous three positions), the type of vehicle, the propulsion type, and the vehicle identification number are transmitted. The minimum set of data described above is defined in the EN 15722:2011 standard. eCall can also be triggered manually by pushing a button in the car, for example, by a witness of a serious accident. Data is transmitted via mobile networks.

Since 2023, the system also works in Switzerland in such a way that the emergency call is forwarded to the nearest police operations center.

eCall is an interesting case because it shows what is technically possible. If the eCall system is able to communicate information about the time and place of an accident, it should in principle be able to communicate other information obtained from the vehicle sensors, too. It should be possible that information like speed and delta-v (see section on EDR above) is also communicated. Potentially, the transmission could also include, whether there has been an impact with vulnerable road users (pedestrians and cyclists). Some vehicles are already able to detect this. When detecting an impact with vulnerable

[3] Regulation (EU) 2015/758 concerning type-approval requirements for the deployment of the eCall in-vehicle system based on the 112 service https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015R0758

road users, some vehicles lift the motor hood to dampen the impact. This could be used as an indicator for the impact with vulnerable road users and communicated with the emergency services.

The DTC Dynamic Test Center AG carried out a crash test (VW T-Cross at a speed of 27.5 km/h against a stationary Peugeot Expert) on 17.08.2023, 15.10 hrs (see EDR protocol above). The example was already mentioned above. In the test both front airbags on the VW T-Cross were deployed. This triggered an automatic eCall. Shortly after the crash (approx. one minute), a response was received via the eCall system. A police officer tried to contact the "person" in the vehicle (the VW). After our feedback that we had carried out a crash attempt, the person replied that he had now seen that the vehicle was in Vauffelin, the site of DTC.

We later had the opportunity to talk to a police officer who gave the following answers.

1. In the past, the eCalls went via the European headquarters of the respective vehicle manufacturer and from there they were the respective countries.
2. Since the summer of 2023 (date unknown) the emergency call is made via the international emergency number 112 directly to the closest cantonal police control centre - in this case to Biel.
3. The protocol of the emergency call must not be released. The respondent could only tell me that only a GPS position was sent from the VW T-Cross and where the vehicle was located according to geodata.

The distance between the location of the vehicle after the crash test and the location of the vehicle was 90 metres.

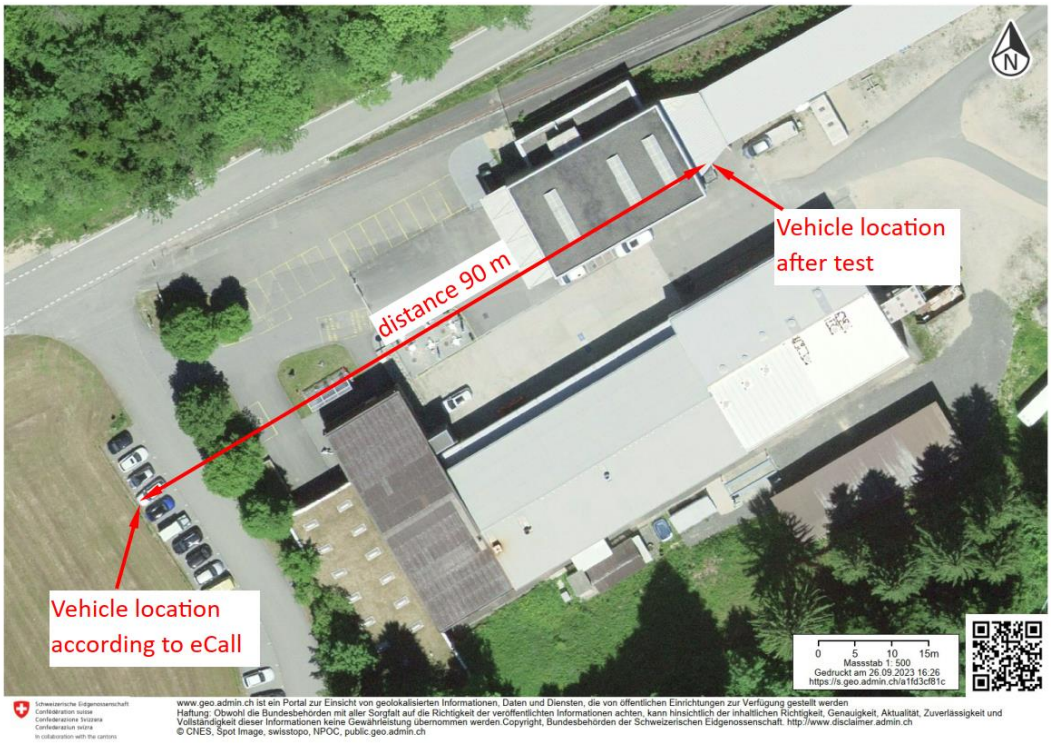

Fig. 6 *Aerial image showing the distance between the real vehicle position and the vehicle position reported via eCall*

The difference can be explained by the fact that the vehicle was in a tunnel with aluminium planking for several hours before and during the crash and the GPS signal may have been disturbed and only one geo position was sent.

3.4 DSSAD

Data Storage System for Automated Driving is a device that records and stores a set of data (“timestamped flags”) during the automated driving sequences of any vehicle equipped with Level 3 / Level 4 / Level 5 Automated Driving Systems (ADS), in order that whenever a significant safety-related event occurs, it can provide a clear picture of the interactions between the driver and the system, before and after (whenever possible) the event [87]. The legal framework is being developed by the Working Party on Automated/Autonomous and Connected Vehicles (GRVA). The latest draft on the properties of the DSSAD dates from 13 January 2024 and is entitled "DSSAD Performance Elements GUIDANCE DOCUMENT" [88]. These topics are currently being developed by the DSSAD subgroup.

There is currently no adopted legal directive on the nature and implementation of the DSSAD [82, 93]. Such a guideline is in preparation and must be introduced as soon as vehicles with assistance systems are approved, according to SAE Level 3. The currently valid version 01.01.2024 of the Road Traffic Act (loi fédérale sur la circulation routière, LCR) does not yet contain any provisions on DSSAD (so-called driving mode memory). However, an amendment of the LCR regarding automated driving was adopted on 17 March 2023[4]. In addition, a project of Ordinance on automated driving was presented in October 2023 [94]. The revised legislation is expected to come into force in the first half of 2025 [95].

The data recorded by the driving mode memory must not be modifiable. When the memory capacity is reached, the oldest data is overwritten. Within the scope of the inspection, the data of the driving mode memory can be read out by the registration authorities.

The vehicle owner must be able to access the data of the driving mode memory via a standard interface. This data must be available to him in an easily readable form. For the clarification of accidents or the assessment of offenses against road traffic regulations, the data in the DSSAD can be read and processed by the competent authorities, such as the police or accident analysts, provided they have the necessary expertise and infrastructure. The precondition for this is a disclosure order from the public prosecutor and the fact that the driver of the vehicle does not exercise their right to privacy. As soon as the data read out is no longer required for any criminal or administrative proceedings, the authority must delete it, but no later than six months after the legally binding conclusion of the proceedings.

3.5 Floating car data

Connected vehicles record a broad range of data. First, there is geolocation data. Cars record their movements and so do satellite navigation systems and cell phones in the cars. Some OEMs, mobile telecommunications providers (Swisscom Ltd) and third-party data providers (Wejo Ltd, Here Technologies) already collect and sell such mobility data [96, 97]. Typically, this involves location, speed and direction of travel for cars – and sometimes of bicycles, too. Depending on the data provider, traffic participants can be geolocalized. Some data providers, for instance, assign a unique identifier to each traffic participant. To address data protection concerns, however, the identifier is changed every 20 minutes [98, 99].

Secondly, there are emergency and safety enhancing systems. Modern assistance systems such as distance cruise control, lane departure warning and emergency brake assist work with the help of cameras, radar and lidar systems. Other cameras are used as parking aids and to monitor the vehicle’s surroundings. As a rule, no image data is stored in the vehicle, although there are exceptions. Automatic transmission of camera data is used by at least one major vehicle manufacturer. Lidar and radar data are uploaded at best in the context of field tests.

[4] See FF 2023 p. 791.

Increasingly, the cameras of the parked vehicle can be accessed via a smartphone (e.g. Tesla Sentry Mode, BMW Remote 3D View or Mercedes Geographical vehicle monitoring) [100, 101].

3.6 Floating personal data

The project focuses sensor-data from vehicles. However, there are also floating personal data and infrastructure data that may be of relevance to accident researchers and prevention specialists.

Personal data is mainly collected via smartphones carried in the vehicle and digital watches. Movement analysis via smartphones carried in the vehicle is of increasing importance for the information enhancement of dynamic navigation systems in the vehicles. Movement profiles created by the smartphones might be evaluated and transmitted back to the vehicles online as congestion information. If a minimum number of smartphones in vehicles move slowly forward, for example, this is interpreted as a traffic jam and displayed accordingly (GPS tracker) [102].

3.7 Infrastructure data

Infrastructure data is collected by traffic cameras (tunnel, intersection and red-light monitoring) and automatic traffic counters. In the future, cameras in the intersection areas could be used to record so-called near-accidents. Currently, video data from stationary and mobile traffic cameras are mainly used to investigate traffic accidents and red-light offenses.

4 Reaching out: Results from the workshops

As described in Section 2.2.2 above, we conducted a series of eight workshops with policymakers, representatives of associations and interest groups, legal experts, industry representatives, accident reconstruction specialists and researchers (a list of workshop participants is provided in the Appendix of this study). In total, more than 30 stakeholders participated in our workshops.

In the workshops, together with the participants, we identified and 11 use cases:

1. Exposure data: Who drives when and where? To determine risk, it is necessary to be able to compare the number of cases in which an accident occurs to the number of cases in which accidents do not occur. For this, exposure data is necessary. Mobility data may potentially be used to generate exposure data.
2. Expanding and improving accident statistics: The workshops participants pointed to several challenges with the existing accident statistics. For instance, they suggested that there is misreporting and underreporting. The workshop participants suggested that the accident statistics could be expanded with vehicle sensor data.
3. Near-accidents: These are defined as situations in which there is an imminent crash danger that can be eliminated only through a successful avoidance maneuver. The workshop participants pointed out that in aviation near-accidents are routinely analyzed. In the road safety domain such analyses are not systematically conducted. Industry representatives pointed out that with sensor data, however, it becomes possible to detect and analyze near-accidents.
4. Real time data for on-the-spot prevention: The workshop participants suggested that vehicles could record and communicate data that could be used by road authorities for prevention measures.
5. Testing the proper functioning of sensors: Several workshop participants pointed to the challenge that to date there are no systems in place to ensure the proper functioning of sensors over the lifecycle of a vehicle. They suggested that new ways need to be developed to test the functioning of sensors and to recalibrate these if necessary.
6. Accident reconstruction: In the case of an accident, sensor data can be used to reconstruct what happened.
7. Determining culpability and liability: Sensor data may also help to determine criminal culpability and potential civil liability.
8. Improving and evaluating prevention measures: Several workshop participants also suggested that sensor data should increasingly be used to evaluate prevention measures. They pointed out that prevention measures are often difficult to evaluate.
9. Who is speeding when and where? Some workshop participants also suggested that sensor data could also be used to monitor if and where vehicles commonly exceed the speed limit. Once speeding hotspots are identified, preventative measures could be taken.
10. Winter service: Real time data from vehicles on environmental conditions, such as snowfall, for instance, could be used to optimize the deployment of winter service trucks.
11. Hazard warnings: Many workshop participants also pointed to the potential of identifying hazard such as black ice, pedestrians or obstacles on the road with sensor data and then communicating these hazards with approaching vehicles.

Based on the workshop results we then selected four use cases. The selection was meant to a) reflect the importance attributed by the workshop participants and b) to allow for an analysis of sensor data use in different contexts. Based on these criteria we selected four use cases:

1. Exposure: Who drives when and where?
2. Expanding accident statistics
3. Hazard warnings and real time prevention
4. Accident reconstruction and determining criminal and civil culpability and liability

We merged some of the use cases mentioned in the list above. For instance, we decided to merge the use of sensor data to determine reconstruction with criminal and civil responsibility.

We present the results of our use-case analysis in Chapters 8-11. During the analysis of the use cases, we have subsequently decided to divide some of the use cases in multiple sub-use cases.

5 Legal foundations

5.1 Introduction

In order to understand the legal rules that apply to sensor-collected data, we will first give an overview of the current legal framework in Switzerland (5.2.1) and some important acts and initiatives in the European Union (5.2.2). We will then present a selection of legal topics in more depth (5.3), which are relevant for the use cases presented in Chapters 8 to 11.

5.2 Current legal framework

5.2.1 Switzerland

Data protection

Data protection is regulated at the federal level by the Federal Act on Data Protection (FADP)[5]. This act aims to protect the personality and fundamental rights of natural persons whose personal data is processed (art. 1 FADP) by private persons or federal bodies (art. 2 al. 1 FADP). The FADP is supplemented by the OPDo[6]. The rules of the FADP which are relevant in the present context will be explained in detail below (5.3.1).

Road traffic

Several acts regulate road traffic in Switzerland.

The Road Traffic Act (loi fédérale sur la circulation routière, LCR)[7] contains a few provisions regarding the processing of personal data (see art. 89c ff LCR). In particular, art. 89i ff LCR contains rules about the information system relating to road accidents. Art. 89k lists categories of data collected in case of an accident, and art. 89l states which government entities are processing this data. These provisions are supplemented by the Ordinance on the Information System related to Road Accidents (Ordonnance sur le système d'information relatif aux accidents de la route, OSAR)[8]. Nevertheless, neither the LCR nor the OSAR expressly provide for the collection of data collected by sensors.

Some other provisions contained in road traffic legislation and related to data are more specific. For example, the Ordinance regarding the Technical Requirements for Road Vehicles (Ordonnance concernant les exigences techniques requises pour les véhicules routiers, OETV)[9] contains articles providing which cars need to be equipped with tachographs (art. 100 OETV; for example, cars driven by professional taxi drivers) or stating that emergency vehicles need to be equipped with a data recording device (art. 102 OETV).

[5] Federal Act on Data Protection of 25 September 2020, RS 235.1, available in English. The FADP has been fully revised and the new law into for on September 1, 2023, repealing the previous version adopted thirty years ago (Federal Act on Data Protection of June19, 1992. Although the revised FADP introduces changes, the main principles remain the same. Consequently, the case law issued in relation to the former FADP remains relevant.

[6] Ordonnance sur la protection des données du 31 août 2022, RS 235.11, not available in English.

[7] Loi fédérale sur la circulation routière du 19 décembre 1958, RS 741.01, not available in English.

[8] Ordonnance sur le système d'information relatif aux accidents de la route du 30 novembre 2018, RS 741.57, not available in English.

[9] Ordonnance concernant les exigences techniques requises pour les véhicules routiers du 19 juin 1995, RS 741.41, not available in English.

It should be noted that the LCR is currently in the process of being revised [2]. A modification of the law was adopted on 17 March 2023 but has not yet entered into force[10]. One of the aims of this revision is to include a set of new provisions relating to automated vehicles (levels 3 and 4) in the LCR[11]. Those provisions only concern automated vehicles and aim, in the first place, to allow the Federal counsel to regulate how a (human) driver can rely on the driving assistance of a car, and under which conditions cars equipped with an automation system can move around[12]. A new article 25g regulates access to data collected by the Data Storage System for Automated Driving (DSSAD). According to art. 25g al. 5 LCR, the authorities responsible for the registration of the vehicle have to send data collected by the driving mode recorder – in a form that does not allow the identification of the driver – to FEDRO[13]. FEDRO will in particular have the possibility to make this data available for research or analysis. This provision will not apply for non-automatized cars; the Message of the Federal Council specifies that the legislation does not regulate other types of data recorders and does not create a new legal basis for the processing of data by authorities[14]. The Federal Data Protection and Information Commissioner (FDPIC)[15] took a position on this revision in his 2021/2022 report [4]. In addition, a project of Ordinance on automated driving was presented in October 2023 [94]. The revised legislation is expected to come into force in the first half of 2025 [95].

In summary, there is currently no general provision regarding the processing of car-related personal data collected via sensors in the Swiss legislation. One or several provision(s) regulating this subject could however be added to the LCR in a future revision of this law.

Criminal procedure

When a road traffic accident occurs, a criminal investigation is usually opened. This procedure led by a cantonal prosecutor is subject to the Swiss Criminal Procedure Code (CrimPC)[16]. The FADP does not apply in this context (see art. 2 al. 3 FADP).

This investigation aims to identify if a criminal offense has been committed, and by whom. The outcome of the criminal procedure will, most of the time, also affect the outcome of a potential civil claim by a victim.

In the process, the accident is reconstructed, which requires using all available evidence. The data collected by connected cars can be very useful in this regard, because it allows them to establish facts that cannot be proved otherwise. But the use of this data raises questions in relation to the procedural rules contained in the CrimPC; the two following issues are particularly relevant and will be addressed in detail below (5.3.3.):

- Can car-collected data be used in a criminal proceeding without any formalities or is a search warrant needed?
- If the data has been previously collected in violation of the FADP, can it then be used in a criminal proceeding?

[10] See FF 2023 791.

[11] Message concernant la révision de la loi fédérale sur la circulation routière du 17 novembre 2021, FF 2017 3026 p. 3 ss.

[12] Message concernant la révision de la loi fédérale sur la circulation routière du 17 novembre 2021, FF 2017 3026 p. 33 ss.

[13] Federal Roads Office.

[14] Message concernant la révision de la loi fédérale sur la circulation routière du 17 novembre 2021, FF 2017 3026 p. 45.

[15] Federal Data Protection and Information Commissioner; the FDPIC is the Swiss authority in charge of monitoring the correct application of the federal provisions on data protection (art. 4 al. 1 FADP).

[16] Swiss Criminal Procedure Code of 5 October 2007, RS 312.0, available in English.

Telecommunications

In Switzerland, art. 29a OST[17] provides that, in the case of emergency calls to the European emergency number from specially equipped vehicles (eCall112), mobile radio dealers must extract the Minimum Set of Data (MSD) from the voice channel and transmit it to the location service. If Advanced Mobile Location (AML) is used, information collected by AML should also be transmitted to the emergency services.

The eCall system was first implemented in the EU (see Delegated Regulation (EU) No 305/2013, Decision No 585/2014/EU, Regulation (EU) 2015/758 mentioned below). It is used in vehicles and makes a free 112 emergency call if the vehicle is involved in a serious road accident [5]. It can also be triggered manually [5].

The system in Switzerland is the same as in the EU. When the eCall system is triggered (automatically or manually), a minimal set of data (MSD) is sent to emergency services (including vehicle location, time of the accident, travel direction, vehicle type, vehicle identification number) [6].

Intellectual property

Various acts protect intellectual property rights in Switzerland. They can be relevant as the technologies used by OEMs might be protected by intellectual property rights.

The most important intellectual property rights acts in the present context are the following:

- The PatA[18] regulates the grant of patents for new inventions applicable in the industry (see art. 1 PatA);
- The CopA[19] regulates (a) the protection of authors of literary and artistic works, (b) the protection of performers, producers of phonograms and audio-visual fixations and broadcasting organizations and (c) the federal supervision of the collective rights management organizations (art. 1 al. 1 CopA);
- The TmPA[20] regulates the protection of trade marks, defined as a sign allowing distinguishing the products or services of a company from those of other companies (art. 1 TmPA).

Manufacturing and trade secrets

Some aspects of the technologies used by OEMs that do not fall under the protections of rights may still be considered manufacturing and/or trade secrets. This could for example include information related to how the sensors are designed and work, how to access to collected data or some information contained in the data itself.

Manufacturing secrets and trade secrets are protected by various provisions of different acts, for example:

- SCC[21]: Under art. 162 SCC, the breach of manufacturing or trade secrecy can be punished. In addition, art. 273 SCC, which punishes industrial espionage, also protects manufacturing and trade secrets;

[17] Ordonnance sur les services de télécommunication du 9 mars 2007, RS 784.101.1, not available in English.

[18] Federal Act on Patents for Inventions of 25 June 1954, RS 232.14, available in English.

[19] Federal Act on Copyright and Related Rights of 9 October 1992, RS 231.1, available in English.

[20] Federal Act on the Protection of Trade Marks and Indications of Source of 28 August 1992, RS 232.11, available in English.

[21] Swiss Criminal Code of 21 December 1937, RS 311.0, available in English.

- UCA[22]: This act aims to ensure a fair and non-distorted competition (art. 1 LCD). Inciting someone to reveal or overhear manufacturing and trade secrets is deemed unfair (art. 4 let. c LCD) as well as exploiting or revealing manufacturing and trade secrets that were overheard or learned unduly (art. 6 LCD). Such behavior can lead to a civil claim (art. 9 LCD) and/or to a criminal sentence (art. 23 LCD);
- CO[23]: art. 321a al. 4 CO prohibits employees to exploit or reveal confidential information obtained while in the employer's service, such as manufacturing or trade secrets. This obligation of confidentiality remains after the end of the employment relationship to the extent required to safeguard the employer's legitimate interests.

Insurances

According to art. 63 ff. LCR, every motor vehicle must be insured. A lot of insurance companies are now proposing to their clients "pay as you drive" insurances. Some insurance companies ask their policyholders to install dongles in their windshields that collect data about the movements of the vehicle. The insurance company will collect data about how the insured car is driven and adapt the insurance premium accordingly [7].

The LCA[24] regulates insurance contracts. Its art. 3 al. 1 lit. g states that the insurance company shall inform the policyholder, before they enter into a contract, about the data processing, including the aim of the processing, the type of databank used, the data recipients and the data storage. Otherwise, this act does not contain specific provisions related to this type of premium calculation and the data processed for this purpose. The FADP is therefore applicable to all other issues regarding this data processing.

5.2.2 European Union

GDPR

The GDPR[25] lays down rules relating to the protection of natural persons concerning the processing of personal data and rules relating to the free movement of personal data (art. 1 par. 1 GDPR).

The GDPR applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not. (art. 3 al. 1). It also applies if the controller or processor is not established in the Union in two cases: (i) where the processing activities are related to the offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union (ii) where the processing activities are related to the monitoring of their behavior as far as their behavior takes place within the Union.

[22] Federal Act on Unfair Competition of 19 December 1986, RS 241, available in English.

[23] Code of Obligations of 30 March 1911, RS 220, available in English.

[24] Loi fédérale sur le contrat d'assurance du 2 avril 1908, RS 221.229.1, not available in English.

[25] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.

Directive 2010/40/UE

The Directive 2010/40/UE[26] establishes a framework in support of the coordinated and coherent deployment and use of Intelligent Transport Systems (ITS) within the Union, in particular across the borders between the Member States, and sets out the general conditions necessary for that purpose (art. 1 Directive 2010/40/UE).

Art. 2 of the Directive 2010/40/UE defines "priority areas" in which the specifications and standards shall be developed and used. The first priority area is titled "Optimal use of road, traffic and travel data".

Based on this directive, the following instruments have been adopted:

- Commission Delegated Regulation (EU) 2017/1926 of 31 May 2017 supplementing Directive 2010/40/EU of the European Parliament and of the Council with regard to the provision of EU-wide multimodal travel information services (OJ L 272, 21.10.2017, pp. 1-13);
- Commission Delegated Regulation (EU) 2015/962 of 18 December 2014 supplementing Directive 2010/40/EU of the European Parliament and of the Council with regard to the provision of EU-wide real-time traffic information services (OJ L 157, 23.6.2015, pp. 21-31);
- Regulation (EU) 2015/758 of the European Parliament and of the Council of 29 April 2015 concerning type-approval requirements for the deployment of the eCall in-vehicle system based on the 112 service and amending Directive 2007/46/EC (OJ L 123, 19.5.2015, pp. 77-89);
- Decision No 585/2014/EU of the European Parliament and of the Council of 15 May 2014 on the deployment of the interoperable EU-wide eCall service (OJ L 164, 3.6.2014, pp. 6-9);
- Commission Delegated Regulation (EU) No 886/2013 of 15 May 2013 supplementing Directive 2010/40/EU of the European Parliament and of the Council with regard to data and procedures for the provision, where possible, of road safety-related minimum universal traffic information free of charge to users (OJ L 247, 18.9.2013, pp. 6-10);
- Commission Delegated Regulation (EU) No 885/2013 of 15 May 2013 supplementing ITS Directive 2010/40/EU of the European Parliament and of the Council with regard to the provision of information services for safe and secure parking places for trucks and commercial vehicles (OJ L 247, 18.9.2013, pp. 1-5);
- Commission Delegated Regulation (EU) No 305/2013 of 26 November 2012 supplementing Directive 2010/40/EU of the European Parliament and of the Council with regard to the harmonized provision for an interoperable EU-wide eCall (OJ L 91, 3.4.2013, pp. 1-4).

These regulations are interesting because they provide some guidance about the legal requirements to collect and process data. Regulation No 886/2013[27] is of particular interest for this research and is presented in more detail in the following section. We will provide more information on this regulation in the context of use case 3 (Chapter 9).

The initiative "Access to vehicle data, functions and resources initiative"

The European Commission has launched an initiative called "Access to vehicle data, functions and resources". One of the aims of the initiative is to give public authorities access to data collected by car sensors, which would help them perform their tasks (for example

[26] Directive 2010/40/UE on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport.

[27] Commission delegated regulation (EU) No 886/2013 of 15 May 2013 supplementing Directive 2010/40/EU of the European Parliament and of the Council with regard to data and procedures for the provision, where possible, of road safety-related minimum universal traffic information free of charge to users.

monitoring CO2, ensuring compliance with pollutant emissions regulations or doing roadworthiness controls).

A public consultation has taken place from March 2022 to August 2022 on this topic[28]. This initiative may lead to a regulation, which, if adopted, will serve as an example of how law can compel vehicle manufacturers to give access to the data they collect.

5.3 Detailed presentation of relevant legal topics

5.3.1 Access to data

A large amount of data collected by car sensors is in the hands of private companies – OEMs (original equipment manufacturers), app developers, etc. They have, as GILL/METZGER put it, a "gatekeeper position" in this regard. Therefore, the question arises how the state can access data.

In this regard, several data sharing models exist: convince the companies to share their data voluntarily for free, buy the data or compel the companies to share their data.

The following considerations apply in case access to a large volume of data is desired, potentially in anonymous form. Access to data relating to a road traffic accident, after a criminal investigation has been opened, is a very different subject which will be discussed below in chapter 5.5.3.

Sharing on a voluntary basis for free

The exchange of car-collected data between private and public entities can benefit all parties and is therefore a possible solution, as can be seen is a position paper issued by the European Automobile Manufacturers Association (ACEA) in 2016[29]. In this paper, the ACEA stated that, for the purpose of road safety, vehicle manufacturers were prepared to make their data available to public authorities (or private operators entrusted with a public task such as road operators) on a reciprocal basis (which means that each party brings data and can access data brought by others in return)[30].

Data sharing for free on a voluntary basis is in fact the approach taken in the Data for Road Safety project, which is described in detail in section 10.3. Art. 3 of the Multi-Party Agreement signed by the parties participating in the project states that "Content is exchanged within the SRTI Ecosystem in-kind on the basis of reciprocity for the sole purpose of road safety [8][31]. Since this agreement was signed in November 2020, it seems that car manufacturers have stuck to the statement that ACEA made on their behalf in its Position Paper in 2016.

This solution is great for prevention and research purposes but can only work in case the private companies see an opportunity in sharing the data [9][32]. Sharing should therefore be reciprocal and allow private companies to benefit from the data as well.

[28] See European Commission website

[29] ACEA represents 15 car, van, truck or bus manufacturers, namely BMW Group, DAF Trucks, Daimler, Fiat Chrysler Automobiles, Ford of Europe, Hyundai Motor Europe, Iveco, Jaguar Land Rover, Opel Group, PSA Group, Renault Group, Toyota Motor Europe, Volkswagen Group, Volvo Cars, and Volvo Group (ACEA, p. 8).

[30] ACEA, p. 3.

[31] Art. 3 al. 1.

[32] Page 15.

Buying data available on the market

Another option to get access to data is simply for the State, or a State entity (such as FEDRO) to buy it from companies – OEMs or other companies whose business is to sell traffic data, such as Here [10] or Wejo [11][33]. The two companies mentioned now propose services targeting public services: Here proposes a service aiming at improving emergency services [12], whereas Wejo offers a service called "Road Health" which is designed to make road maintenance more efficient [13].

Paying is an easy way to access data. However, it does not give a lot of leeway, since the companies selling the data usually take most decisions regarding the processing of the data (what data is processed, how it is processed, how it is shared, etc.).

Compelling to share the data

The last option is to compel companies to share their data by adopting a new piece of legislation. An example of this is the eCall mentioned above (5.2.1 section Telecommunications) [14].

The advantage of this solution is that the state has a lot of margin to decide what data is required from private companies. Sometime it is difficult to compel companies established outside of Switzerland to share data. However, the revised Federal Act on Data Protection (FADP) now requires data controller established abroad to appoint a representative in Switzerland if they regularly process personal data on a large scale and such processing poses a high risk to the personality of the data subjects (which may be the case depending of the amount and granularity of data collected).

In Chapter 7.2.2 we will elaborate on the Swiss' policy with regards to accessing data and the role that the Swiss Federation envisages for government in this context.

5.3.2 Data protection

Personal, pseudonymized and anonymized data

The FADP only applies when *personal* data is processed (see art. 2 al. 1 FADP). This means that data that is not considered as personal data, in particular anonymous data, is not subject to the rules and constraints of this law.

Notion of personal data

Art. 5 lit. a nFADP provides a definition of personal data, that is "all information relating to an identified or identifiable natural person". This notion is construed broadly[34].

The person must be identified or *identifiable* – which means, according to the Swiss Federal Supreme Court, that information alone isn't sufficient to identify a person, but that it is possible to do so with additional information[35] or given the context[36]. To assess whether a

[33] and many others.

[34] Swiss Federal Supreme Court, 147 III 139, 10 December 2020, par. 3.1 and 3.4.1.

[35] Swiss Federal Supreme Court, 136 II 508, 8 September 2010, par. 3.2.

[36] Swiss Federal Supreme Court, 138 II 346, 31 May 2012, par. 6.1.

person is indeed identifiable, all circumstances of the case must be taken into account; in particular, the technical possibilities enabling identification should be considered[37].

Given the broad definition of personal data, a lot of information collected by cars should be considered as personal data [15][38]. As long as the information can be traced back to an individual, it will be qualified as personal data. Here are a few examples of personal data in this context:

- Data concerning the owner of the vehicle;
- Vehicle serial number (also called Vehicle Identification Number, VIN) [16][39];
- License plate number[40];
- Media Access Control (MAC) address [16];
- Location data[41];
- Distance covered[42];
- Wear and tear of the vehicle parts[43];
- Mileage[44];
- Driving style [16][45];
- Data collected by cameras[46].

In the data-science foundations chapter (6.2), we will provide further information on identifiers.

Anonymized and pseudonymized data

Personal data can be stripped from the elements allowing the identification of an individual by pseudonymizing or anonymizing data. These processes are, in themselves, data processing subject to the FADP[47].

Both processes consist in removing identifying information; the difference is that the process is reversible in case of pseudonymization [17]. Personal data can be pseudonymized by using a key (meaning an assignment rule [17], such as replacing identifying characteristics by a number). Pseudonymized data remains personal data for the persons who know the key[48] meaning that the FADP applies. By contrast, according to the majority opinion, if pseudonymized data is transferred to someone who does not know

[37] Swiss Federal Supreme Court, 138 II 346, 31 May 2012, par. 6.1; Swiss Federal Supreme Court, 136 II 508, 8 September 2010, par. 3.2.

[38] GILL/METZGER, p. 9; STÖRING, p. 9 f.

[39] CNIL, p. 6.

[40] CNIL, p. 6.

[41] CNIL, p. 6; EDPB, p. 5 ; FIALOVÀ, p. 2.

[42] EDPB, p. 5.

[43] CNIL, p. 5; EDPB, p. 5.

[44] CNIL, p. 5.

[45] CNIL, p. 5; EDPB, p. 5; FIALOVÀ, p. 2.

[46] EDPB, p. 5.

[47] See Swiss Federal Supreme Court, 4A_365/2017, 26 February 2018, par. 5.2.2; JOTTERAND, Personal Data or Anonymous Data: where to draw the lines (and why)?, Jusletter 15 August 2022, N 56.

[48] Handelsgericht Zürich, HG190107-O, 4 May 2021, par. 3.2.3; SHK DSG-RUDIN, art. 5 N 14; JOTTERAND, Personal Data or Anonymous Data: where to draw the lines (and why)?, Jusletter 15 August 2022, N 56.

the key (and have no other means to identify individuals), the data should be considered anonymized [17].

Anonymized data is not personal data anymore and is consequently not subject to the FADP [18, 19]. It is not always easy to determine if data is truly anonymous; the context has to be taken into account [17]. Data is still considered anonymized if a possibility of reidentification exists but would require significant means[49]. All means than can reasonably be used to identify a person should be taken into account[50]. In Chapter 6.35.2.3 we will elaborate on reidentification risks.

In the context of connected cars, anonymization can be viewed as a good measure to exploit the data while mitigating privacy risks[51]. However, it has been pointed out that there is a tendency, in the automotive industry, to consider that aggregated data is anonymous, which is not necessarily the case: given the increasing power of data analytics, cross-referencing data can lead to re-identification [20]. The process of de-identification should therefore be designed with particular care (see Chapter 3.2.3).

Lawfulness of the processing of data

The assessment of the lawfulness of a data processing differs depending on whether the data controller is a private person (meaning an individual or a company) or a federal body.

Processing by private persons

If the processing is carried out by a private person, the processing is lawful unless an unjustified breach of personality occurs (art. 30 ff FADP); the analysis is the following:

- The first step is to examine whether there is an infringement of personality; it is the case if (a) personal data is processed in violation of the principles of art. 6 or 8 FADP (for details about these principles, see below, «Data protection priniciple»), (b) personal data is processed in violation of the express wishes of the data subject or (c) sensitive personal data is disclosed to third parties (art. 30 al. 2 FADP);
- The second step is to assess whether this infringement of personality is justified by the consent of the data subject, an overriding private or public interest, or by law (art. 31 al. 1 FADP). Art. 31 al. 2 FADP lists cases in which the overriding interests of the data processor shall in particular be considered, for example when the processing is in direct connection with the conclusion or the performance of a contract and the personal data is that of a contractual party (let. a) or when data is used for purposes not related to persons, for example for research, planification or statistics (let. e).

This analysis must be carried out whenever the question arises as to whether personal data is being processed by a private person – in the context of this study, typically an Original Equipment Manufacturer (OEM).

[49] MEIER, Protection des données, N 440 ff.

[50] See Court of Justice of the European Union, C-582/14, Breyer v. Bundesrepublik Deutschland, 19 October 2016, par. 42.

[51] EDPB, Guidelines 01/2020 on processing personal data in the context of connected vehicles and mobility related applications, Version 2.0, 9 March 2021, N 79.

Processing by federal bodies

The FADP does not apply to cantonal bodies, but similar requirements apply to them based on cantonal laws. To be lawful, a data processing carried out[52] by a federal body, must, as a rule, have a legal basis (art. 34 al. 1 FADP). The legal basis has to be in a formal act – meaning an act adopted by the Swiss Parliament – if data is sensitive, if a profiling[53] is made or if the purpose or means of processing data could severely harm the data subject's fundamental rights (art. 34 al. 2 FADP). As an exception, a material act (like an Ordinance) is sufficient for data that is not sensitive, or in case of processing of sensitive data or a profiling, if the processing is essential for a task required by a formal act and the purpose of processing poses no particular risks to the data subject's fundamental rights.

There are a few limited exceptions allowing data to be processed without a legal basis, for example if the Federal Council has authorized the processing considering that the data subjects' rights are not threatened (art. 34 al. 4 let. a FADP) or if the data subject has consented to the processing (art. 34 al. 4 let. b FADP).

It should be emphasized that the International Working Group on Data Protection in Telecommunications, during its 63rd meeting – which was about connected cars – expressed an opinion in line with the FADP approach: they stated that "*Public authorities should mainly use data provided by connected vehicles for purposes that the data subjects have freely consented to or that are in the public interest, or to fulfill tasks that have been laid down in an appropriate law*" [16][54].

If a federal body – in particular FEDRO – processes data that is not (or cannot be) anonymized, the existence of a sufficient legal basis will need to be assessed.

Art. 39 FADP is a provision that applies specifically to the processing of personal data that was already legally collected by a federal body for a secondary purpose without requiring a legal basis for this secondary purpose, provided that the following conditions are cumulatively met:

- The purpose of processing is not related to specific persons (for example for research, planning or statistics purposes);
- The data is anonymized as soon as the purpose of processing permits;
- The results are only published in such a manner that the data subjects are not identifiable;
- The federal body only discloses sensitive personal data to private persons in such a manner that the data subjects are not identifiable; and
- The recipient only transmits the data to third parties with the consent of the federal body that disclosed the data.

Data protection principles

When processing (including collecting) data, the data controller has to comply with the following data protection principles:

- Lawfulness (art. 6 al. 1 FADP), meaning that personal data may only be processed lawfully and for federal bodies according to a legal basis;

[52] The legal basis requirement derives from art. 5 al. 1 of the Federal Constitution of the Swiss Confederation ("rule of law principle"), which states that "All activities of the state are based on and limited by law".

[53] According to art. 5 FADP, profiling means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location or movements.

[54] Par. 67 p. 12.

- Good faith (art. 6 al. 2 FADP). Among other things, it means that no processing should occur without the data subject's knowledge or against his or her consent [19];
- Proportionality (art. 6 al. 2 FADP). This principle implies that data can only be processed if it is suitable and necessary[55] to fulfill the purposes of the processing [18]. In addition, there must be a reasonable relationship between the purposes and the means used to fulfill them [18]. Data shall also not be retained longer than necessary;
- Purpose limitation (art. 6 al. 3 FADP). According to this principle, personal data can only be collected for specific and recognizable purposes. Purposes should be clearly defined [19]. In addition, art. 6 al. 3 FADP only allows data to be further processed in accordance with the initial purposes;
- Transparency (art. 6 al. 3 FADP). The data subject should be aware of the data collecting as well as its purposes. This is achieved if (i) the data subject has been informed about the processing, (ii) the processing is provided by the law or (iii) the processing is apparent from the circumstances [18]. This principle is supplemented by an obligation of information (art. 19 FADP);
- Accuracy (art. 6 al. 5 FADP). Anybody who processes personal data should make sure that the data is accurate. Consequently, data processors shall take any appropriate measure to correct, erase or delete data that is inaccurate or incomplete. The accuracy requirement is, however, not absolute, and rectification is only necessary in case the data subject is affected in his or her personality rights or if the data subject asks for rectification;
- Security (art. 8 FADP). The security of the data should be guaranteed by taking appropriate organizational and technical measures. The measures taken shall in particular safeguard data against any security violation – that is any violation which could lead to the loss, modification, erasure, deletion or divulgation of the data, as well as an unauthorized access to it [19]. Art. 8 FADP is supplemented by art. 1 ff nDPO, which explain how to assess the level of security and the measures needed (art. 1 nDPO) and define the aims of the measures (art. 2 et 3 nDPO).

Both private persons or federal bodies have the obligation to comply with the aforementioned principles when processing personal data.

5.3.3 Criminal procedure

As mentioned above, in case of a road accident, a criminal procedure is usually opened. This procedure led by a cantonal prosecutor is subject to the CrimPC and the FADP does not apply. Moreover, the outcome of the criminal procedure will, most of the time, also affect the outcome of a potential civil claim by a victim.

A criminal procedure takes place after an accident has occurred and does not constitute, as such, a preventive measure. However, repression of offenses is generally viewed as a having a deterrent effect and therefore serve the prevention of offenses. In addition, criminal procedures relating to road accidents generally involve a reconstruction of the events. Experts analyze all available evidence in order to understand the causes of the accident. These findings are then entered into accident statistics which, in turn, provide a better understanding of why, how and when accidents happen – which represents valuable information for determining what measures can be put in place to prevent road traffic accidents. Legal rules applying to criminal procedures are therefore relevant in the context of sensor data based accident research and prevention.

It should however be noted that the principles mentioned here only apply in the context of criminal proceedings. Their importance is thus limited compared to the rules regarding data protection explained above.

When an accident happens and a criminal procedure is opened, data collected by an Event Data Recorder (EDR) can be very useful. The advantage of the EDR is that, for most car brands, the data collected can be accessed in a readable format thanks to a system

[55] The principle of minimization derives from the proportionality principle.

developed by Bosch called “Bosch Crash Data Retrieval” (hereafter: “Bosch CDR”) [21, 80].

This raises the question of how the EDR data can be used in the procedure. In this regard, the legality of the data collection and processing by the EDR does not seem to be disputed. However, how EDR data can be obtained is debated: should it be considered as “forensic and other evidence” in the sense of art. 306 al. 2 let. a CrimPC (meaning that it can be secured and viewed by the police without any authorization), or should the EDR fall within the meaning of “data carriers and equipment for processing and storing information” of art. 246 CrimPC (which means that the public prosecutor, or, by delegation, the police has to issue a written search warrant [see art. 241 CrimPC] to authorize the consultation of the EDR data [22]) [23]? We discuss this question in Use Case 4 in Chapter 10.5.

Admissibility of unlawfully obtained evidence

Sometimes, behaviors that constitute criminal offenses are filmed by cameras attached to a vehicle (dashcams, GoPros, etc.). This raises the question of whether the event filmed can be used as evidence of the offense, in case the video was made in violation of the FADP.

The Swiss Federal Supreme Court takes the view that, when evidence is collected illegally by a private individual, it can only be used in a criminal procedure if (i) criminal authorities had been able to collect such evidence lawfully and (ii) a weighing of the interests involved pleads for its admissibility of said evidence[56]. When weighing interests, the severity of the offense must be taken into account – which means that if the offense is not serious enough, a piece of information illegally collected by a private individual cannot be used as evidence[57].

The systematic filming of traffic with a dashcam violates the FADP and is thus illegal (violation of art. 4 al. 4 former FADP – that states that the collection of personal data and the purpose of its processing must be evident to the data subject – which cannot be justified by any ground mentioned in art. 13 former FADP[58]; under the new FADP, the same conclusion would be reached[59]). If the criminal offense that was filmed or for which prosecution the video recording would be useful is only a violation of road traffic rule (art. 90 al. 1 and 2 LCR), the offense is, however, not serious enough to justify using the dashcam video as evidence[60].

We will return to questions of criminal procedure in use case 4 (Chapter 11).

[56] Swiss Federal Supreme Court, 147 IV 16, 13 November 2020, par. 1.1; Swiss Federal Supreme Court, 146 IV 226, 26 September 2019, par. 2.1.

[57] Swiss Federal Supreme Court, 147 IV 16, 13 November 2020, par. 1.1; Swiss Federal Supreme Court, 146 IV 226, 26 September 2019, par. 2.2.

[58] Swiss Federal Supreme Court, 147 IV 16, 13 November 2020, par. 2 ff; Swiss Federal Supreme Court, 146 IV 226, 26 September 2019, par. 3.

[59] The transparency principle has even been reinforced with an obligation to inform (art. 19 FADP). In addition – even if the Swiss Federal Supreme Court did not mention it – the processing appears to violate the principle of purpose limitation, because the scope of the processing is very broad (systematic filming and recording of the traffic) in comparison to its purpose (obtain evidence of a violation of a road traffic rule).

[60] Swiss Federal Supreme Court, 147 IV 16, 13 November 2020, par. 7.2; Swiss Federal Supreme Court, 146 IV 226, 26 September 2019, par. 4. As a general rule, an offense that is considered a felony (meaning that it carries a custodial sentence of more than three years, see art. 10 al. 2 SCC, such as a homicide) can be qualified as serious.

6 Data-science foundations

6.1 Introduction

This chapter outlines a systematic approach for managing the use of sensor data for accident research and prevention. The chapter delves into the privacy risks associated with sensor data. However, this is not a legal analysis but focuses of the data-science side of the issues discussed in the previous chapter. Through a focus on data layer-based measures and privacy-enhancing techniques, the Chapter outlines strategies for mitigating privacy risks, adjusting the level of applied protective measures to align with the unique requirements of different research objectives.

6.2 Understanding privacy challenge

This section elucidates privacy challenges inherent in sensor data analysis, emphasizing the importance of understanding identifiers and data structures for informed privacy risk management. It outlines the distinct implications of direct and indirect identifiers, explores the privacy risks associated with different data structures, and delves into the main types of privacy risks. The objective is to provide stakeholders with insights crucial for strategic privacy risk mitigation in sensor data analysis.

6.2.1 Distinguishing data categories with privacy implications

It is crucial for stakeholders to carefully distinguish between data categories acting as 'direct' and 'indirect' identifiers, as each requires tailored protection measures to safeguard privacy. Direct identifiers are pieces of information that uniquely identify individuals without the need for additional data, like a Vehicle Identification Number (VIN). A VIN, comprising a unique set of 17 characters, serves as a direct identifier, providing a direct means to identify a vehicle’s owner [24]. It is imperative to implement protection measures for such data since it can immediately pinpoint individuals.

On the other hand, indirect identifiers, while not identifying individuals outright, can do so when combined and analyzed together. For instance, many contemporary vehicles are equipped with GPS technology, continuously recording geographical coordinates (latitude and longitude) alongside timestamps (time and date). Though each piece of data in isolation may not reveal an identity, when combined and analyzed over a period, these data points can create unique behavioral patterns, potentially inferring or identifying individuals. Therefore, understanding these different categories of identifiers is essential for implementing the requisite protection measures, as even seemingly innocuous data can lead to identification when mishandled.

6.2.2 Analyzing privacy risks related to data structures

The complex landscape of data privacy is shaped by many factors, with one key element being the data’s structure. In accident research, observational studies and analytical approaches predominantly involve three principal data structures to study accidents’ causes, patterns, and implications: cross-sectional, time series, and panel data. Each of these structures carries unique privacy considerations to be carefully considered in mitigation strategies:

- Cross-sectional data: This type of data captures a snapshot of a particular population or phenomenon at a single point in time by recording various variables. For instance, a cross-sectional study on traffic accidents might collect data on the number of accidents occurring on a specific day, identifying the types of vehicles involved and the severity of the injuries sustained in each accident. Privacy risks arise when unique or rare attribute combinations make individuals identifiable or when combining cross-sectional data with publicly available data [25]. Additionally, sensitive information may

be unintentionally revealed. To mitigate these risks, anonymization techniques and stringent access controls are recommended [26].

- Time series data: Time series data comprises sequential data points for a single variable collected over time. This structure helps analyze trends and patterns. In the context of traffic accidents, a time series study might document the number of accidents occurring monthly, annually, or over different decades, allowing for a temporal analysis of accident frequencies and patterns. However, time series data poses privacy risks, as the sequences of events can be linked back to individuals, especially if the data reveals unique or consistent patterns over time [27]. Sensitive information may also be inferred with enough data points. To mitigate these risks, it is essential to aggregate data over time to prevent the detailed reconstruction of events and to apply techniques such as noise injection and obfuscation to protect individual privacy (see section 6.4.2).
- Panel data: Panel data integrates aspects of cross-sectional and time series data by collecting data on multiple variables over time for the same subjects. For example, a panel study on traffic accidents might compile data on various factors (like the number and severity of accidents and types of vehicles involved) for specific drivers or locations over successive years, offering a comprehensive view of accident dynamics. Nonetheless, panel data is inherently at a high risk of re-identification through potential record linkage attacks [28] and captures unique changes in data over time that can be distinguishable among other longitudinal patterns [29]. Continuous data collection over time can create detailed profiles of individuals, making them identifiable and exposing sensitive information. To mitigate these risks, it is essential to implement privacy-preserving measures such as k-anonymity or differential privacy (see section 6.4.2) and to adopt data minimization practices to collect only the necessary data.

Understanding the privacy risks of sensor-generated data in accident research requires considering the specific data structure underlying each data category. Commonly encountered data categories in this research field include accident event reports (which often start with event-based, cross-sectional data), geospatial data, survey and interview data, epidemiological data, simulation and modeling data, environmental and engineering data, historical records, and databases. Each category comes with its unique set of privacy considerations, dependent on its specific structure and application in the broader research context. Therefore, it is imperative to understand and effectively mitigate privacy risks, with each data structure necessitating tailored privacy-preserving approaches to safeguard individuals' confidentiality and prevent improper disclosures.

6.3 Understanding main types of privacy risks

Re-identification risk refers to the likelihood of accurately matching de-identified data back to a specific individual with high probability (see chapter 5.3.2 section Anonymized and pseudonymized data). Another concern involves the potential to link the identity of an individual to sensitive information, by establishing a logical association with other details about the individual.

Given the inherent potential for re-identification risk, access to vehicle sensor data necessitates the adoption of risk-based anonymization as a safeguarding measure. Hence, a comprehensive risk assessment becomes mandatory, serving as the compass to guide the necessary data transformations, ensuring dataset accessibility while minimizing risks. The risk assessment process helps estimate and manage the extent of re-identification risk effectively.

The evaluation of re-identification risk is intrinsically tied to the specific context in which the data is used, how the data is safeguarded, what external data sources could be used for matching, and the definition of the population under scrutiny. For instance, if external parties with potentially malicious intent [29] possess knowledge of the data subjects' origins in a particular region, that region becomes a defining factor in delineating the population.

Notably, the existing literature and legal frameworks outline various threats that warrant vigilant attention to mitigate privacy risks effectively [30, 31]. For instance:

- Singling out refers to the possibility of isolating some or all records about an individual in the dataset. Example: Just four spatio-temporal data points could be enough to identify 95% of individuals in a dataset [32].
- Linkability refers to the ability to link multiple records belonging to the same person or a group of persons, either in the same dataset or multiple different datasets. Example: Linking spatio-temporal data with publicly available datasets or datasets from direct sourcing can yield a comprehensive profile of an individual's life.
- Inference refers to the possibility of deducing, with a significant probability of correctness, the value of an attribute from the values of a set of other attributes. Example: Only a few spatio-temporal data points are enough to infer residential addresses and professional details [33].

Beyond privacy risks, several other considerations related to security should be addressed, such as data breaches where unauthorized access to sensor-related data can lead to privacy violations, data misuse, and potential harm to individuals. Consideration should be given to data subject consent, who may not always be fully informed about vehicle sensor data collection and usage. Finally, complying with data protection laws and regulations is essential to avoid privacy violations and legal consequences.

6.4 Developing a privacy-centric framework

Anonymization, alongside with other protective technologies and methods, is pivotal in facilitating data access within the accident research and prevention domain.

Effective reduction of privacy risks demands a holistic approach integrating technical, organizational, and legal measures. One comprehensive strategy for data access and sharing in different scenarios is the Five Safes framework [34], which provides a set of principles and best practices for managing the risk of re-identification when sharing and using sensitive data. Developed by the UK Office for National Statistics, it became widely adopted in academia, and government agencies. The Five Safes framework serves as a guiding principle for decision-making regarding data access and sharing. Furthermore, it continues to adapt in response to evolving technologies, data types, and privacy considerations.

The Five Safes framework encompasses five key dimensions:

1. **Safe projects**: This dimension emphasizes conducting meticulous reviews of research projects to assess their potential public benefits, including generating new knowledge, informing public policy, and benefiting specific societal groups. Every project must obtain ethical approval, ensuring that participants' rights and interests are respected and protected. The legal principles that need to be considered are described in Chapter 5.3.2.
2. **Safe people:** This dimension ensures that only qualified and authorized individuals have access to data, preventing misuse. This approach is underpinned by entrusting data handling to individuals proficient in ethical and responsible data management practices.
3. **Safe settings:** This dimension pertains to the creation and maintenance of secure environments that prevent unauthorized data access and use. It is imperative to establish settings where only authorized individuals have data access, providing a secure framework for data handling. Special attention should be given to potential re-identification risks in various environments, particularly when data copies are shared with recipients for use beyond the oversight of the data space.
4. **Safe data:** This dimension focuses on incorporating safeguards directly into the data to preserve privacy. Measures include the strategic removal or alteration of identifying features, application of privacy-preserving methods to prevent the deduction of confidential information, and enforcement of strict access limitations.
   Recognizing that even anonymized data may be employed inappropriately or harmfully is crucial. Equally important is the establishment of explicit data use agreements delineating precise data applications, required privacy safeguards, and protocols for sharing research findings.

5. **Safe outputs:** This dimension ensures that research results are non-disclosive and have undergone rigorous screening and approval processes. This scrutiny ensures that shared research outcomes do not violate privacy, maintaining a commitment to confidentiality and respect as insights from the data are disseminated to relevant stakeholders.

The operationalization of 'The Five Safes' framework encompasses several core processes that build upon the framework's dimensions:

- Initial data assessment: This foundational process involves examining the input data, focusing on its sensitivity, designated uses, and potential re-identification risks. This assessment should also account for the information accessible to potential external parties with malicious intent, evaluating how this might be amalgamated to achieve data re-identification.
- Data transformation techniques: The appropriate techniques are employed to render the data safe for analysis whilst retaining its utility. Implementing transformations that adhere to the ethical and legal parameters set by 'Safe Projects' is imperative, ensuring that the data's integrity is maintained for its intended purpose.
- Access control and authentication: Rooted in the 'Safe People' and 'Safe Settings' dimensions, this process is instrumental in fortifying data security. It ensures that data access and processing are stringently restricted to authorized individuals or teams, with each access request meticulously authenticated and logged. This requires the implementation of strict access controls and authentication mechanisms to ensure that only authorized individuals or teams can access and process the data. The framework stipulates policies specific to use cases regarding data access and processing, mandates regular compliance audits, and outlines requirements for secure data storage and processing environments. Furthermore, it advocates for the diligent review and sanitization of data analysis outputs to avert unintentional disclosure of sensitive information.

The Five Safes framework offers guidance for conducting sensor-based road safety research. The next section outlines a data management system that applies the Five Safes principles to address re-identification risks, adapt data to meet privacy requirements, and maintain optimal data utility.

6.4.1 Data layer-based measures for safe data

The Data Layer-Based Measures outlined in this section embody key principles of the Five Safes framework, primarily focusing on generating Safe Data while extending safeguards to other dimensions. This comprehensive, layered approach establishes a holistic and secure foundation for data handling and analysis in accident research.

The approach relies on the careful design of a cascade of layers, each aimed at progressively enhancing the privacy and security of the data (refer to Fig. 1 below). The foundational pseudonymized data layer relies on controlled pseudonymization techniques to protect data, setting the stage for subsequent transformations. Following this, data is further transformed in synthetic, anonymized, and aggregated layers through a series of robust measures. This makes it not only increasingly secure and anonymized, but also ensures the preservation of its utility.

These structured, layer-wise measures collectively guarantee that the data is inherently safe, guarding against privacy risks at each stage of its transformation. Provisions embedded within each layer also implicitly address 'Safe Projects' by aligning data use with ethical standards and public interest goals. 'Safe People' is assured through stringent access controls and clearly articulated usage policies, which limit access to qualified and authorized personnel only. The secure environments, crafted meticulously through these layers, inherently establish 'Safe Settings' to access, process, and analyse data under tightly controlled and protected conditions. Finally, the framework ensures 'Safe Outputs' through provisions specific for each layer to manage the release of research outputs, thereby protecting data throughout its entire lifecycle.

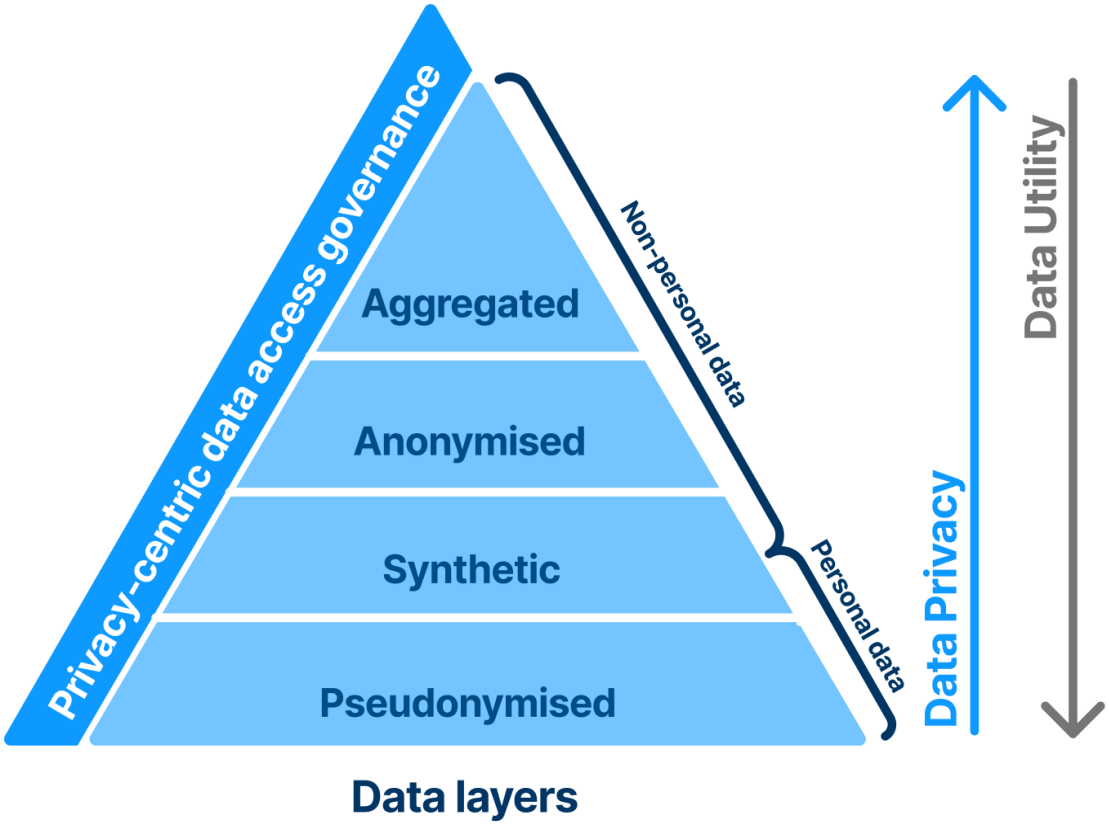

Fig. 7 *Data layer-based measures*

The Data Layer-Based Measures are structured as a pyramid consisting of four layers of data access, serving as a data transformation pipeline to minimize the risk of identification. The pseudonymized, synthetic, and anonymized layers constitute the most granular level, where each row represents a distinctive observation or entity within the dataset. The aggregate layer offers a comprehensive perspective on the datasets by presenting an aggregated view. Detailed descriptions of each layer and its role in the data transformation pipeline can be found in the following section.

1. Pseudonymized data layer

The pseudonymized data layer serves as the first line of defense against re-identification risks, handling the most sensitive data collected from providers (see section 4.3. 2 for legal definition). The sensitivity assessment of this data informs privacy goals and guides the application of privacy-preserving methods, ensuring the data is securely processed through subsequent layers or when accessed by recipients. This layer involves meticulous pseudonymization, which masks direct identifiers and transforms potential indirect identifiers, minimizing the risk of singling out individuals. Essential information required for re-identification is securely managed or destroyed. Robust data protection measures, including stringent access controls and clear usage policies, are crucial at this stage. Providing explicit details about the implemented pseudonymization techniques enhances transparency and fosters stakeholder trust.

2. Synthetic data layer

The synthetic data layer involves the creation of artificial data that closely mimics the structure and characteristics of the original dataset. Synthetic data can be sourced directly from data providers or generated based on pseudonymized data. To be effective, synthetic data must accurately represent the original dataset and align with its intended purpose.

This alignment can be confirmed by assessing utility, quality metrics. Utility metrics evaluate whether the synthetic data offers sufficient information for the intended analysis or research. Meanwhile, quality metrics assess whether synthetic data adequately replicates the statistical properties of the original dataset, ensuring reliability in analysis and research.

Despite being generated to protect individuals' privacy, synthetic data may still present privacy risks. Thus, it is crucial to evaluate the privacy implications of synthetic data to prevent any inadvertent disclosure of sensitive information. Privacy metrics are instrumental in determining if additional privacy-preserving methods are necessary before sharing or processing synthetic data.

Synthetic data has emerged as an innovative "data anonymization solution," aiming to address the limitations associated with traditional data sanitization techniques [35]. To evaluate the validity of this claim, which suggests that generating synthetic data is an effective anonymization method, it is essential to investigate whether synthetic data effectively mitigates the privacy risks outlined in the WP29 Guidelines on anonymization [30]. These risks include concerns related to linkability, inference, and singling out. In cases where there is no guarantee of adequate privacy protection, synthetic data is classified as personal data, necessitating the implementation of stringent access controls, individual consent, and comprehensive data protection measures.

3. Anonymized data layer

The anonymization layer employs privacy-enhancing technologies to ensure anonymity by permanently obscuring data, making it exceedingly challenging to re-identify individuals (see Chapter 5.3.2for a legal definition). This process includes indirect identifiers transformation to prevent data subjects from being distinguishable in the population, thus mitigating the risk of singling out individuals. Strong privacy measures, security protocols, and strict controls are applied to safeguard data and minimize the risk of re-identification, even in environments without access to additional data required for re-identification. Data collected from data providers may not have been sufficiently anonymized. Therefore, Data controllers should establish clear anonymization objectives and guidelines to ensure effectiveness and compliance. Furthermore, it is crucial to conduct periodic assessments that consider the unique context, intended data usage, and the probability and magnitude of associated risks to determine the degree of anonymization (see Chapter 5.3.2). While anonymization is a potent privacy safeguard, it can restrict data utility for certain types of analysis. Therefore, assessing data quality and utility is essential before deploying it for analysis.

4. Aggregated data layer

The aggregated layer is tailored for high-level data analysis, presenting insights about specific groups rather than individual data points. It leverages anonymized data to create aggregated categories, making it valuable when detailed data is not essential for research purposes.

While the aggregated layer offers the highest level of privacy protection through anonymized data and the application of aggregation rules, it is not entirely impervious to privacy risks. Hence, conducting regular assessments to evaluate the effectiveness of these aggregation rules becomes crucial in averting potential privacy breaches.

6.4.2 Privacy-enhancing technologies for safe and effective data use

Privacy-Enhancing Technologies (PETs) are essential tools for the secure, responsible, and effective use of sensor-related data that have the potential to offer strong privacy safeguards. These technologies, ranging from anonymization techniques like K-anonymity, differential privacy, and homomorphic encryption to cryptographic protocols offer robust

protection against identification and data linkage. Emerging PETs, as detailed in the UN [36] and OECD report [37], encompass data obfuscation tools, encrypted data processing tools, and federated and distributed analytics tools. Data obfuscation tools protect identity by altering data, encrypted data processing tools allow for secure data manipulation without decryption, and federated and distributed analytics tools enable task execution without direct data access. These technologies facilitate the alignment of practices with the Five Safes framework to promote secure and effective use of data, including data sharing, machine learning, and analytics. However, realizing their full potential and safeguarding privacy effectively requires further research, development, and careful application of these emerging technologies. In the following sections (see Tab. 1), we will explore different types of privacy-enhancing technologies, clarifying their impact on data integrity and confidentiality while addressing the challenges pertinent to their implementation.

Tab. 4 *Privacy Enhancing Technologies*

Types of PETs	**Key technologies**	**Current and potential applications**
Data obfuscation tools	K-anonymity	Ensuring that the location data of a user is indistinguishable from that of at least "k" other user.
	Differential Privacy	Analyzing and drawing insights from Big Data while preventing sensitive information from being traced.
	Zero-Knowledge proofs	Verifying information without requiring sensitive data disclosure.
	Pseudonimization techniques	Replacing identifiers with pseudonyms, to enable data analysis while preserving privacy.
	Anonymization techniques	Safe sharing of accident data between multiple organizations, researchers, or government agencies.
Secure computation with encrypted data processing	Homomorphic encryption	Performing privacy-preserving analytics on encrypted data to identify critical insights without the need to decrypt or expose individual information.
	Trusted execution Environment	Security boundary used for training and statistical modelling. Send the collected sensor data to a central public cloud component.
	Secure Multi-Party computation	Collaborative data analysis over a private dataset.
Distributed learning	Federated Learning	Collaborative machine learning over private data Ex: Collect knowledge from multiple drivers with privacy preserving.
	Split learning	

Data obfuscation techniques

Data obfuscation techniques are essential PETs playing a crucial role in processing data and ensuring privacy. These approaches protect identity by altering data in various ways, such as adding noise, removing identifying elements, or generating synthetic data.

Anonymization is a commonly used technique, promising to eliminate identifiable details effectively, but achieving true anonymization is elusive due to re-identification risks. Pseudonymization offering a reversible de-identification is classified as personal data due to the residual risk of re-identification. Synthetic data, an alternative approach, generates artificial datasets that mimic the statistical properties of original data, offering reduced privacy risks, but not without challenges and re-identification concerns.

Differential privacy, another well-developed and academically recognized obfuscation tool, introduces calibrated noise to query responses, protecting individual data while maintaining the dataset's value [38]. It operates with a privacy budget (parameter ε), balancing between

added noise and the resultant data utility. Although it doesn’t address direct identification, it provides an essential layer of protection against attribute inference. The practical application of differential privacy requires careful management of the privacy budget and understanding its varying effectiveness across different datasets.

The k-anonymity technique ensures each individual in a dataset is indistinguishable from at least k−1 others [39]. This approach, achieved through generalizing attributes or suppressing outlying records, hinders the re-identification of individuals within a dataset [40]. It is not suitable for scenarios where distinctive patterns or unusual data points are vital for statistical analysis because altering or concealing this information can result in inaccurate conclusions [41]. Lastly, Zero-Knowledge Proofs (ZKP) emerge as invaluable tools for validating the truthfulness of data without revealing additional information. These tools validate data truthfulness without revealing extra information, showing promise in the area of traffic management and similar applications.

Each tool and technique, while promising, comes with challenges and requires careful consideration and application to truly protect individual privacy. The reliability of anonymization techniques is often in question, as records from anonymized datasets are frequently re-identified post-release. Unintentional data leakage remains a concern, with the need for a balanced addition of 'noise' to the data, a currently non-existent standard. There is also a noticeable skills gap, with the implementation of obfuscation measures requiring knowledgeable data scientists to prevent unintentional data leaks.

6.4.3 Secure computation with encrypted data processing

Traditionally, the necessity for data decryption during processing has created vulnerabilities, exposing sensitive information to potential breaches. Recent advancements in encrypted data processing technologies have revolutionized this landscape, ensuring data remains encrypted while processed. This development is crucial in accident research, where there is a need for a balance between the accessibility of sensor-related data for analysis and preserving the privacy of the individuals and organizations involved. For example, these technologies facilitate secure computations on OEM sensor data without revealing confidential details to third parties, thereby safeguarding sensitive information while allowing for necessary data analysis in accident research.

Technologies such as Homomorphic Encryption (HE), Secure Multi-Party Computation (SMPC), and Trusted Execution Environments (TEEs) are at the forefront of this privacy-preserving initiative. HE allows for computations on encrypted data, offering a privacy-focused solution albeit at a higher computational cost. Despite being computationally intensive, it is a valuable tool for applications where the need for privacy is paramount. On the other hand, SMPC, a more mature technology, enables collaborative computation functions over private data by multiple parties. Furthermore, TEEs offer a secure enclave within a processor for data storage and processing, providing an additional layer of security.

The broader framework of Secure Computation integrates these technologies to form a robust system for collaborative computations among multiple parties without exposing individual data [42]. Secure computation can provide strong security guarantees, ensuring that even if some parties are malicious or untrustworthy, the privacy of the honest parties is maintained. It achieves this by using cryptographic protocols that guarantee privacy regardless of the actions of others [43]. One practical application of secure computation is evident in the privacy-preserving linkage of records from diverse data sources. When consolidating data from various vehicle sensors supplied by different OEMs, it's crucial to employ privacy-preserving record linkage techniques, including Private Set Intersection (PSI) and Homomorphic Encryption. SMPC is often used to implement PSI when parties need to securely find common elements in their data sets while preserving the privacy of the data. These techniques ensure records are linked securely and efficiently, minimizing the risk of unintentional data leakage and guaranteeing the privacy and confidentiality of sensitive identifiers.

Despite the substantial privacy and security benefits offered by encrypted data processing tools, they present specific challenges. Data encryption complicates the traditional data processing performed by analysts. Analysts cannot directly interact with the raw data, making it imperative to identify and rectify errors at the data submission stage. Moreover, while designed to secure processed data, these tools do not offer absolute guarantees against information leakage from the computational results. The risk of inadvertent information leakage through the results necessitates careful selection and structuring of the computational functions to be applied. Special care must be taken to ensure that the computational results do not reveal sensitive information about the input data. Finally, the higher computational costs associated with encrypted data processing tools can deter organizations from adopting these technologies unless there are explicit recommendations or mandates from governmental bodies advocating their use.

6.4.4 Distributed learning

Distributed learning represents an innovative approach to conducting analytical tasks, such as machine learning training, on data that remains invisible or inaccessible to those executing the tasks. These methodologies enable sensitive data analysis while it stays under the original data source's custody, with only summary statistics or results being transferred to the task executors. The European Commission's EU Data Strategy recognizes decentralized data processing as a viable method to enhance user control and compliance with data protection mandates [44]. And the Swiss Federation recognizes it in its Data Hub strategy [103].

Federated learning allows for raw data pre-processing at the data source level, which could be the data subjects themselves. This approach transfers only summary statistics and results to the data processor for amalgamation with other similar data. This mechanism significantly diminishes the necessity for sensitive data to exit the data subject’s device or be stored by data processors, enhancing privacy and data protection. In split learning, data is divided into parts, and computations are performed on these separate parts across different entities. This way, no single entity has access to the complete data or model, and the information is shared in a distributed and privacy-preserving manner, ultimately helping protect the privacy of individuals and organizations while still enabling the development of machine learning models. Despite their benefits, the application of these technologies requires careful consideration of legal and privacy concerns. Federated and split learning face the risk of potential information leakage, when the parameters sent back may inadvertently reveal sensitive information. Although researchers have proposed solutions like employing encrypted data processing techniques, including homomorphic encryption or secure multi-party computation, challenges still exist. Legal frameworks might need to address and adapt to the challenges and nuances introduced by federated and split learning to ensure that they do not become sources of data leakage. Finally, the effectiveness of federated and split learning heavily depends on stable connectivity. This reliance becomes a significant challenge for applications that require continuous availability of analytical results, as any disruption in connectivity could impede the analytics process.

6.4.5 Privacy risk assessment

Privacy risk assessment is a key component of the Five Safes to identify, assess, and mitigate potential privacy risks associated with sensitive data access. It will enable to identify any residual risks of re-identification or other privacy concerns for each data layer.

The privacy risk assessment establishes and enforces processes and procedures that provide controls and response mechanisms to manage risks associated with vehicle sensor data, both before and after it is made available to users. It may need to be more rigorous for the pseudonymized layer compared to the aggregated layer due to the different sensitivity levels. If the privacy risks are deemed acceptable and the anonymization objectives are met, researchers can access data according to the specific policies for each layer. For example, a researcher may only have access to the synthetic layer unless they can justify a need for accessing a more sensitive pseudonymized layer. If the anonymization objectives are not fully met, only authorized researchers with a significant

need can access data under restricted policies of the pseudonymized layer. ISO 27701 is an international standard that extends ISO 27001 and provides guidelines and best practices for privacy risk assessments and aligning them with an organization's privacy management system.

6.4.6 Enhancing consent and awareness in diverse data ecosystems

As data is coming from diverse sources, personal data can be collected unknowingly or unintentionally. This scenario can be particularly complex with sources such as vehicle sensor data. Given the mixed data ecosystem, individuals may not be aware that their data is being collected and used for such analysis. Therefore, it is crucial to establish clear consent mechanisms that cover all data sources involved.

- *Obtaining explicit consent*: Whenever possible and except if there is a legal provision allowing for the collection and sharing of data, get clear and informed consent from individuals whose data is being collected. This could mean working with data providers to include a consent clause for accident and near-accident research or asking vehicle owners to opt into accident and near-accident studies.
- *Broadening awareness*: Use public awareness campaigns to educate people about the nature and importance of their data collected from various sources for accident and near-accident research.
- *Implementing easy exercise of rights*: Provide individuals with easy ways to exercise their data subject rights, such as the right to access, rectify, oppose and delete their personal data.

6.4.7 Managing data sharing and third-party access through contractual agreements

A legally binding agreement is used to set up a relationship between the data controller, the data processor and the data recipient (third-party data controller). It specifies how input data (for example vehicle sensor data) and output data (result of the analysis performed on vehicle sensor data) may be collected, transformed, analyzed and shared. When the agreement ends, so does any processing of that personal data, as it needs to be destroyed by the processor. In instances of secondary analysis, it is critical to unambiguously define data usage purposes, limitations, and safeguards in a contract.

6.4.8 Access control mechanisms

Access control mechanisms guarantee that only authorized individuals have access to a privacy- based data layer. Researchers requesting access must undergo an approval process. Since the registry does not collect direct identifiers, researchers can only request indirect identifiers. Contractual obligations are established with the researcher's organization, and data provided are tailored for specific purposes within the designated environment.

6.4.9 Safeguarding data during retention and storage

Retaining sensor data beyond its designated retention period or implementing inadequate security measures can result in unauthorized access or data breaches, potentially compromising individuals' privacy and data protection. These risks can be effectively mitigated by:

- Establishing data retention and deletion policies: Define a precise data retention period aligned with analysis objectives and establish clear guidelines for secure data deletion or anonymization.
- Implementing secure storage practices: Adopt rigorous security measures, such as encryption at rest and robust access controls, to safeguard data. Regularly monitor and update security protocols to address emerging threats and vulnerabilities, ensuring the continuous protection of stored data.

6.4.10 Ethical considerations

Data analysis carries significant ethical implications that must be carefully addressed to ensure fairness, non-discrimination, and respect for individual's rights and dignity (See section 5.2.1 for ethical and legal principles). Trust in data analysis, particularly in the algorithms used, is a major concern in the realm of data ethics. Algorithms are trained on and applied to data, and they also generate data through inferences and predictions. While this generated data may not be identifiable, it can still be misused. It is important to note that this misuse may not always be intentional. Algorithms, which are designed to automate and derive insights, come with a range of inherent technical challenges when applied in practice.

Failing to consider ethical considerations can result in unjust treatment, stigmatization, and privacy and equity concerns. To address these challenges:

Ensure fair data processing: adopt unbiased data collection methods that reflect the diversity of the population. Perform analysis with rigorous methods, minimizing potential biases, and considering diverse perspectives.

promote transparency and accountability (See Section 5.3.1): provide clear information about data usage, maintain documented research practices, and establish mechanisms for individuals to access their personal data if it is not anonymized.

Conduct audits and ethics reviews: regularly review the ethical implications of the analysis, consult ethics committees or experts, and adapt to evolving ethical standards.

6.5 Checklist

We have developed a checklist that can help users apply the Five Safes framework. However, it may facilitate compliance, but it ensures alignment with privacy best practices. The inclusion of 'input' and 'factors' in this process indicates a comprehensive examination beyond legal and ethical boundaries, considering specific information, data, or conditions relevant to overall compliance. In Section 9.4 we will apply the checklist to the use case on exposure data.

Tab. 5 *Five Safes Checklist*

Factors	Input
Is the purpose defined? What is the benefit of using sensor data? (Usefulness, public interest) What are the risks of doing or not doing the project?	• Purpose of analysis (Data Usage Context)? • Legitimate interests include (non-exhaustive): public benefits, research, security and accident prevention, generating new knowledge, informing public policy, benefiting specific societal groups etc.?
Is it lawful and ethical?	• Does it respect the data protection principles (lawfulness, good faith, proportionality, purpose limitation, transparency, accuracy, and security)? • Does it respect the privacy by design and by default? • Does it respect the right to oppose of the data subject?
Who are the actors involved?	• Who is the data provider? • Who is the controller? • Who is the processor?
Do they have the motivation and capacity to re-identify?	• Motivations? • Conflicts of interest (Do they know people in the data set?) • Access to additional data sets that could potentially be used for re-identification? • Technical capabilities and resources necessary for a re-identification? • Reward versus cost of re-identification?

Capacity to implement data protection according to 5 Safe?	• Training/capability? • Resources (time, money and technical resources)? • Existing training programs for employees on data ethics, privacy and security best practices? • Awareness about potential threats and the role of researchers in maintaining data security?
Do technical and organizational controls need to be implemented to deter intentional re-identification attempts and prevent potential data breaches?	• Access control, encryption, authentication and authorization, data agreements/ enforcement of the contracts? • Architecture?
Does the setting provide sufficient security?	Existing security controls and measures in place (aligning with privacy and security standard)?
Are personal details readily discernible?	Level of data identifiability (considering the people and settings of the data environment)?
What threats to the data need to be managed?	External and internal factors that could compromise data integrity, confidentiality, or availability? • Anonymization at the Source? • Anonymization Techniques? • Assessment of Anonymization Level? • Compliance with Privacy Standards?
Is data anonymization required?	Data Sensitivity Assessment?
Are Privacy Enhancing Technologies needed as a Privacy-Protective Measure?	No clear legal grounds for data analysis and the need to minimize the amount of data processed to help protect personal information?
Are analytical results and other outputs ensured to be non-disclosable in an inappropriate manner?	• Risk of identity disclosure analytical results and other outputs?
Is the data highly detailed, is it highly sensitive and personal in nature?	• Results of analytics? • Result of outputs?
Is there a potential injury to individuals from an inappropriate processing of the data?	• Potential injury to individuals inappropriate processing of the data?

6.6 Conclusion

Addressing privacy risks with a systematic approach is essential for ensuring responsible data access and processing, particularly when engaging with the intricate landscape of sensor-generated data. The strategic implementation of the Five Safes framework underpins the careful balancing between securing privacy and maintaining data's analytical utility in different analysis scenarios. It is essential to incorporate privacy considerations throughout the entire process, starting from the initial data collection stage to the eventual analysis. This requires the use of advanced privacy-enhancing techniques and governance mechanisms that not only safeguard individual privacy but also enable the extraction of valuable insights from the data. The deployment of robust data transformation pipelines optimized for data layers with different sensitivity is a key enabler in this regard, serving as a dynamic tool to minimize privacy risks while maximizing data utility for insightful analysis. Through a combination of privacy-centric governance, conscious data management, and the application of ethical principles, a secure and trustworthy environment for data analysis can be effectively established.

7 Governance Architecture

7.1 Introduction

In the previous chapters, we have focused on the legal requirements for using sensor data and the way that data science solutions can help meet these requirements. To unlock the full potential of sensor data for accident research and prevention, a comprehensive governance architecture is required. This involves bringing together diverse actors (OEMs and other data providers, drivers, passengers, citizens, researchers and prevention specialists, and public authorities) and addressing various concerns such as protection of personal data, protection of trade secrets, protection of public goods like road safety.

The concept of governance comes from political science. It arises from the recognition that technological change and economic globalization often lead to situations in which states (governments) can no longer solve political problems by exercising hierarchical control. Political issues like global warming, for instance, require a collaborative, horizontally coordinated approach involving multiple entities such as states, citizens, and private companies – referred to as 'governance'.

In our specific context, a similar need arises. Diverse actors must collaborate for the sharing and analysis of sensor data without resorting to coercion. To overcome these challenges, a governance architecture is needed.

7.2 Trustworthy data space

A trustworthy data space is a key component of the governance architecture. The concept of trustworthy data spaces draws inspiration from the IDS RAM [46] a reference model for a data space software architecture designed by the International Data Spaces Association (IDSA) as well as the study on “creating trustworthy data spaces based on digital self-determination" by DETEC and FDFA [41]. Data space, in this context, can be conceptualized as an organizational structure incorporating both technical and physical components, facilitating the connection between data consumers and providers with various data sources. Data spaces state rules governing access and how data is processed and used [45].

Accordingly, data space comprises two fundamental components [47]:

- **Technical infrastructure:** It provides both software and hardware components to facilitate the controlled, sovereign, and secure sharing of data. At this level, the infrastructure ensures the implementation of mechanisms that allow for meticulous control over data access, guarantee data sovereignty, and maintain a high level of security in the data-sharing processes. It also includes components that operate at the semantic level. This aspect focuses on preserving the format and meaning of shared data.
- **Governance architecture:** It enables collaboration and coordination among stakeholders. It is used to establish agreements that govern how data is shared, accessed, and used across diverse legal landscapes. By providing a common set of data sharing conditions, this level of governance facilitates secure and compliant data-sharing practices.

To sum up, the trustworthy data space provides the overarching framework for data governance and processing.

7.2.1 Technical infrastructure

Data space offers a comprehensive technical infrastructure, addressing both the technical and semantic aspects of data sharing and processing. The key components include:

- **Identity and Authentication**: This is used to identify and authenticate natural persons, organizations, or software components as legal entities. It is also used to manage and continuously verify the registration of identified and authenticated legal entities within a specific data space.
- **Authorization:** This involves defining access and usage control policies, registering these policies, and enforcing them. Access and usage control policies articulate the data provider's internal (business) data sharing policies and external (regulatory) policies.
- **Data catalog and Processing service:** This entails registering and managing metadata on data, processing, and service resources in individual data-sharing domains to make them searchable and available within and across data spaces.
- **The data repository is where data is stored**, organized, and managed. It serves as a digital warehouse where various datasets can be securely kept and easily accessed.
- **Interoperability standards and interface**s: they are essential for ensuring that different systems, applications, and data sources can work together seamlessly. They define common formats, protocols, and methods for data exchange. The clearing house and data repository need to adhere to these standards to enable smooth data transactions and integration.
- **Cloud storage:** this provides the physical infrastructure for housing data. It offers scalable and flexible storage solutions that can accommodate vast amounts of data.

7.2.2 Governance architecture

A governance architecture is required due to the following factors:

- **Implementation of the risk assessment**: If the Five Safes framework, as presented in the previous section, is to be implemented *someone* needs to assess whether the conditions defined in the framework are met. Some aspects of this process cannot be fully automated, raising questions about who holds the authority to conduct these assessments and ensure legitimacy.
- **Mobilization:** To encourage owners and custodians of sensor data to share it with researchers and prevention specialists, there is a need for someone to motivate and inspire trust by demonstrating the value of collaboration.
- **Mediation**: Aligning diverse interests requires the establishment of a set of collaboration rules that all parties can agree upon. This necessitates someone to oversee compliance with these rules, monitoring adherence, and having the ability to sanction those who fail to comply.

In the context of the last point, a series of governance questions need to be addressed by the data space:

1. Who is granted access and under what conditions? [48]
2. To what extent and how is the value created in the data space shared among its users, data providers (and its operator) [48]
3. To what extent may the data space operator intervene as curator?
4. How are conflicts resolved and how are individuals prevented from harming the ecosystem? [49, 50]
5. What incentives are offered to the different actor groups to make the data space a dynamic and innovative ecosystem? [51, 52]

All the aforementioned points raise an important question: Who should govern? This question can be answered in different ways. The answer also depends on political traditions. In a report on trusted data spaces by DETEC and FDFA to the Federal Council, the DETEC and FDFA point to different governance styles: In the US, governance is

typically provided by private enterprises, platform businesses to be specific. In China, the government tends to play a leading role. In the EU, decentralized governance architectures based on coordination between multiple stakeholders (public and private) tend to be more prevalent.

In its “Digital Strategy Switzerland” the Federal Council has declared that the role of government should be to enable and not to replace private initiatives [53]. In the “Digital Foreign Policy Strategy” the Federal Council has recommended following and closely coordinating with initiatives in the EU [54].

Therefore, we recommend that FEDRO monitors the developments in the data space domain.

7.3 Data space initiative

There are a number of initiatives that we recommend FEDRO to monitor:

- The EU has passed the Data Governance Act, which provides a framework for the common use of data. At the same time, the EU promotes the development and interoperability of European data spaces in all sectors (health, energy, mobility, finance, manufacturing, etc.) [55]
- The Franco-German project Gaia-X has become an important reference point for cross-sector and cross-country data spaces.
- DETEC and FDFA [47] speak of the need for a Swiss Data Hub that brings together various initiatives including the Data Science Competence Center, the “Koordinationsorgan für Geoinformation des Bundes”, and the national network for digital self-determination, etc. [103]
- The Mobility Data Space is a German project funded by the Federal Ministry of Digital Affairs and Transport (BMDV). It is part of the European cloud initiative Gaia-X. Numerous OEMs, mobility providers, insurers, research organizations, etc. have already joined the project, which is intended to develop into an ecosystem for mobility data.
- The "Data for Road Safety” project was launched by EU transport ministers in response to Regulation No 886/2013, which requires OEMs to share information about hazards detected by their vehicles. To do so, the project has created the Safety Related Traffic Information Ecosystem in which data is exchanged free of charge to create safety-related traffic information. In use case 3 (chapter 10), we will provide more information on this project.

7.4 Value provided by the data space

The data space provides value in two ways:

- It matches the suppliers and users of data.
- It provides economies of scope. When these are present, it becomes more efficient to produce multiple components together than producing them individually [56]. The data space promotes economies of scale by providing a stable technological core based on commonly accepted standards. Complementary extensions and applications are added around the core, which can be combined with each other in a modular way [57, 58].

In the best case, the interplay of these two functions can lead to the creation of data ecosystems [47].

The analogy of the "ecosystem" describes a community of heterogeneous actors who are only loosely connected with each other, but whose success and depend on each other [83]. In this respect, the ecosystem perspective differs from the pure transactional perspective, which focuses on the competition between the various platform participants.

Against this background, the concept of data spaces as ecosystems finds particularly strong resonance in the open innovation literature [59]. They allow the participants to share knowledge with the ecosystem (inside-out) and, at the same time, to absorb external knowledge (outside-in) [60]. This can be an important pull factor that could help motivate OEMs and data providers to share data with the data space.

8 Use case 1: Expanding accident statistics

8.1 The problem

The problem addressed in this use case is that when accidents happen, there is not enough data about what happened and what caused the accident.

Accident statistics are often considered incomplete and imprecise – many stakeholders and researchers who participated in our workshops confirmed this. This has two main reasons:

- Under-reporting: Minor accidents and self-caused accidents are commonly not reported to the police. A recent study based on survey results, for instance, shows that 86% of bicycle accidents in the city of Zurich might not be reported [61]. Also, near-accidents tend not to be reported either. Near-accidents are defined as situations in which there is an imminent crash danger that can be eliminated only through a successful avoidance maneuver [104].
- Misreporting: To date, the information on traffic accidents is mainly based on police reports. However, the information recorded by the police on site is sometimes incomplete and can only describe the situation found by the police. It does not include all the information that is uncovered during the investigations that follows. It does not include information retrieved from the EDR or witness testimonies.

How are the accident statistics compiled today? After the police have visited the site of an accident, they fill in the so-called accident reporting form (*Unfallaufnahmeprotokoll* UAP [62]), either in paper or via the online interface (radis). On radis, FEDRO also provides the opportunity for the police to attach files, which can be EDR readouts, videos, etc. To date, however, the police rarely use of this opportunity a FEDRO representative noted in an interview. For the future, FEDRO would like to incorporate more date sources, like dashcam data or pedestrian protection systems, the FEDRO representative stated.

It is important to note that the accident statistics are primarily based on this first impression of the police that visit the site of the accident. The statistics do not include any of the information about the accidents that is sometimes uncovered later, for instance, in potential legal proceedings or through a technical accident reconstruction analysis.

However, there is one important validation. Since 2011, the road accident statistics are compared with medical data of hospitals and accident insurance data. Once a year FEDRO matches the accident statistics with hospital data to determine the severity of injuries. The severity is ranked using the Naca score [63]. The matching is done by the Federal Statistical Office, who provides FEDRO with the hospital diagnoses of the persons involved in the registered accidents. The goal is to add further information to the accident statistics, such as type and severity of accidents, etc. (Interview with FEDRO representative [62]). Linkages with more sources of data are possible, as research projects have shown.[61]

To date, FEDRO manually studies around 500 accidents in depth to validate the data entries. However, FEDRO is trying to automate as much as possible the process and to make it easier for the police to enter data and be able to implement more validation tests.

[61] The research project VeSPA 64 consisting of several sub-projects has linked accident statistics to additional FEDRO statistics and, based on this, the project has been able to derive numerous recommendations for safety measures in the areas of education, enforcement and engineering, among others. At the same time, however, a substantial need for research was also identified, such as the need for more precise data on the human-machine interface (HMI) at higher levels of vehicle automation.

8.2 What sensor data could potentially be used?

To improve and complement accident statistics, the following sources of sensor data could potentially be used.

8.2.1 Near-accident data

Near-accidents are defined as situations in which there is an imminent crash danger that can be eliminated only through a successful avoidance maneuver – initiated by the driver or the vehicles emergency systems [104]. More and more OEMs are collecting near-accident data using multiple methods: Harsh braking (Wejo), g-force (Michelin), activation of emergency systems (ESP, etc.; Mercedes) and AEBS Warnings. The collected near-accident data typically includes a trigger event (harsh braking, emergency systems activated, g-force threshold exceeded, etc.), time and geolocation of the incident, and sometimes - but not always - the direction of travel.

Potential

Expanding the existing accident statistics with near-accident data carries great potential. It would enable researchers and prevention specialists to run more statistical analyses. Based on the existing accident statistics, this is not always possible because the number of accidents (observations) contained in the accident statistics is low. In 2022, for instance, there were 228 fatal accidents and 3'763 severe accidents in Switzerland. With so few observations, it generally is not possible to statistically analyze the impact of localized changes to the road infrastructure on road safety ([72] and assessment of our workshop participants). And before the accident statistics would show that a location is accident-prone, it may take a long time. More unfavorable, however, is that one would have to wait for an accident to happen, which would also raise ethical concerns.

With near-accident-data, however, one does not have to wait that long until statistical analyses become possible because the number of near-accidents considerably exceeds the number of accidents [65]. This makes it possible to conduct meaningful statistical analyses more often and on dedicated traffic spots.

Challenges

During our workshop participants with civil engineering offices (*Tiefbauämter*) of Swiss cities, they suggested that to date some of the near-accident data offered to them was not that relevant because it was sometimes missing important information like the direction of travel.

They also raised the concern that the quality of the data is difficult to assess from the outside because the companies that sell near-accident data provide little details about the way they detect and define near-accidents. It is unclear if and to what extent the data possibly violates scientific standards of validity, representativeness, and objectivity, as if it has been collected in a legally compliant way.

Providers of near-accident data explain what systems (i.e. activation of ESP etc.) they use to detect near-accidents. However, different systems and vehicle types may detect near-accidents differently. Detection methods are not standardized, nor are definitions of what qualifies as a near-accident. And while it might be true that harsh breaking, the sudden increase of g-forces or other incorporated measures correlate with accidents, there is little independent research into how well measures provide useful indicators for accident research and prevention. Moreover, the near-accident data may not be representative as it is not available for all vehicle types, population strata, and geography. It is obvious that as the data is based on sensors of comparatively new and possibly more expensive cars. This could lead, for example, to the neglect of low-income areas where a lot of near-

accidents might happen, but they are not recognized by any sensor equipped cars; or the data is based on a company that only integrates OEMs that sell comparatively expensive cars. All this information typically is not transparent hence it is difficult to assess the value of the data.

8.2.2 eCall data

The eCall system, that is described in Chapter 3, is intended to bring rapid assistance to motorists involved in an accident.

Potential

The use of eCall data provides two opportunities:

1. Add information to accidents that are already in the accident statistics database (i.e., the number of passengers or type of vehicle).
2. Collect data on accidents that may not be reported and are therefore not registered in the accident statistics.

Challenges

Information based on this system might be flawed due to unintended or purposefully wrong use by the people in the vehicle thus and in general the causation of the accident is still unclear.

More importantly, eCall data is used for the purpose of handling emergency situations and cannot be used for other purposes without violating the FADP. We will return to this challenge further below.

8.2.3 EDR data

Chapter 3 already provided a description of EDR.

Potential

A FEDRO representative who we interviewed expressed interest in expanding accident statistics with data from Event Data Recorders (EDR). The inclusion of EDR data would offer the opportunity to add more information about the causes of traffic accidents in the accident statistics.

EDR data is always stored in the vehicle and can only be read out via direct vehicle access. It should not be possible to transmit the data over the air.

A potential challenge of uploading the EDR readouts into the accident statistics database is that the EDR data alone are difficult to interpret. The EDR data always needs to be triangulated with traces on sight because the readouts are easily misinterpreted. For instance, a problem that can lead to misinterpretations of the EDR arises when the tires lose contact with the ground and spin freely. In these cases, the EDR data misrepresents the actual speed of the vehicle.

8.2.4 Floating car data

Description

There are various forms of floating car data, including image data, that we discussed in Chapter 3.5.

Potential

Floating car data in addition to the data mentioned above could potentially be used to complement and enrich accident statistics.

Road safety researchers that we have interviewed also suggest that it would be significant to eventually have as much information as possible:

- What happened before the accident? How did the traffic participants interact before the (near) accident? Time section data from the vehicles involved in the crash and from other traffic participants.
- How did the vehicle recording the near-accident interact with the other traffic participants before? Conversely: How did the traffic participants interact before the (near) accident? Ideally, researchers can have access to video data.
- Additional characteristics of the traffic participants could be relevant: Type of vehicle, age, gender, resident vs. non-resident, etc.

The more data can be provided on the events leading up to the (near-) accident the better the analysis of the relevant incident and the derivation of appropriate preventive measures, the interviewed researchers suggested.

There is great potential in using image data and other sensor data from the vehicles to study the causes of accidents or near-accidents. An interviewed OEM stated that in the case of an accident or near-accident the company transfers all data from the vehicles to their back offices, in order to study the incident in depth.

In particular, information on interactions with other traffic participants, are important, one interviewee suggested. To study such interactions, for instance, image data could potentially be an important source of information for researchers.

Challenges

To date, standards for the interpretation of the data is missing. This refers to both to the respective measures that are used and to their relevant thresholds. Time-to-collision (ttc), for example, might be meaningful to analyze at one traffic spot, at another it is useless – and the commonly used threshold of ttc ≥ 1 sec. might be significant at some points but at other ones – particularly on highways – it might be too low to assess safety (Steiner et al., 2023).

Again, as long as the full picture of a situation that led to an incident is not available, the presence of more data than today still might lead to a false sense of overview of the causes that might have contributed to the incident. In this respect, even such an additional piece of data is only a critical stopover on the presumably unattainable path to a complete understanding of what caused this event.

8.3 Can the data be accessed?

8.3.1 Near-accident data

This data is readily available and sold by various companies including Wejo, Michelin, Mercedes, etc. They provide near-accident data in aggregated and anonymized form, which means that the data is not personal (See Chapter 5.3.2). Therefore, data protection laws will not oppose to the use of that data.

8.3.2 eCall data

To organize emergency services, eCall data is readily available. Whether it could also be used for the purpose of expanding accident statistics is a legal question.

And the EU Regulation 2015/758 stipulates under Article 6 (2) that data generated via eCall "shall only be used for the purpose of handling the emergency situations." The purpose defined in the regulation does not include statistical purposes. Moreover Article 6(3) provides that eCall data „shall not be retained longer than necessary for the purpose of handling the emergency situations" and that the data "shall be fully deleted as soon as they are no longer necessary for that purpose."

In Switzerland, the legislation (in particular the OSE, which regulates the eCall system in Switzerland) does not contain a specific provision specifying that eCall data shall only be used for handling emergency situations. However, even without such a clause, another usage would not be allowed. Using eCall data for statistics would violate the purpose principle (art. 6 al. 3 FADP: personal data may only be collected for a specific purpose that the data subject can recognize). In addition, data processing by federal bodies requires in principle a legal basis, which does not exist in the present case. However, Art. 39 FADP (mentioned in Chapter 5.3.1 above) could allow for a processing for statistical purposes if the conditions mentioned in this provision are met.

Therefore, we recommend not to pursue the integration of eCall data (unless – after a thorough analysis – it appears that art. 39 FADP could allow such a use).

8.3.3 EDR data

In Chapter 11 (use case 4) we elaborate how EDR data can be accessed. We describe that the data needs to be read out manually and that often encryption keys need to be requested from OEMs and that a search warrant is required. In principle, however, the data should be available.

8.3.4 Floating car data

To date, more comprehensive floating car data that includes for instance image data cannot be accessed. OEM's policies vary. While some may, in select circumstances, share data, others categorically refuse to share such data, our interviews and workshops revealed. One OEM, for instance, who said that in the case of a (near-)accident they retrieve all sensor data from the vehicle and conduct a thorough analysis of the incident stated that they do not share the underlying data or the results of their internal analysis with third parties.

There appear to be two practical reasons for OEM's reluctance to share their data. First, they want to protect themselves from potential liability claims. Secondly, the data is essential for the training of automated driving systems. Strategically, it is therefore important to OEM's that the data is not shared with their competition.

Given OEM's reluctance, there are only two options for gaining access to the data. First, the governement could adopt a legal basis and force OEMs to share the data. In its "Digital Strategy Switzerland", however, the Federal Council has clarified that it would not force companies to share data arguing that the role of government should be to enable and not to replace private initiatives [53]. Therefore, we explore another option.

The second option is to motivate OEMs to share floating car data through a) incentives and b) an infrastructure that would allow them to safely share their data without having to fear liability claims or the loss of commercial secrets. A trustworthy data space, which we described in Section 7.2, could provide such a secure infrastructure.

The remaining question is then, how to motivate OEMs to join such a secure data space. One incentive could be that this could allow them an additional opportunity to monetize their data. An ecosystems approach might also provide them with the benefits by combining their data and services with others in the ecosystem.

8.4 How would the data need to be managed once access was secured?

8.4.1 Near-accident data

The data is sourced in an anonymized form from companies like Wejo, Michelin, or Mercedes, ensuring that the data does not contain personally identifiable information (See Chapter 5.3.2). Therefore, in principle, we do not identify any legal constraints. However, it shall be verified that data really is anonymous.

Nonetheless, we recommend conducting a privacy risk analysis based on the Five Safes framework (see Chapter 5 for more details). This assessment is particularly crucial when there is a risk of cross-referencing, as it helps evaluate the extent to which anonymized data can be linked or cross-referenced with other data sources, both internal and external. For example, the inclusion of geolocation data in anonymized near-accident data would depend on how the anonymization process is conducted and the specific details retained in the dataset. In a scenario where anonymized near-accident data retaining the geolocation of the incident is linked to a residential registry that contains information about individuals' residences, there is a risk of individual re-identification.

The significance of this assessment is underscored by historical incidents that exemplify the vulnerabilities inherent in data linkage. In 2002, Sweeney's research demonstrated that an anonymized medical dataset, when cross-referenced with voter registration information, could reveal the medical records of an individual [105].

Furthermore, the privacy risk analysis enables us to assess the effectiveness of the anonymization techniques used to obscure identifying information and evaluate their robustness against re-identification attempts.

8.4.2 eCall data

The police (operation centers) have access to this data. But as stated in the Regulation, they are only allowed to use it for the purpose of sending emergency services to the site of the accident. Under Swiss law, the purpose of the data processing is the same and this data could not be used for other purposes without violating the FADP, unless the conditions of art. 39 FADP are met.

8.4.3 EDR data

According to the new EU Directive (2019/2144), the last four digits of the VIN must be anonymised. The following EDR readout from a VW ID3 illustrates this. The VIN output without the last digits is in green.

CDR File Information

User Entered VIN	WVWZZZE1ZMP******
User	wws
Case Number	
EDR Data Imaging Date	02.08.2024
Crash Date	
Filename	WVWZZZE1ZMP1234_ACM.CDRX
Saved on	Donnerstag, Februar 8 2024 at 11:13:28
Imaged with CDR version	Crash Data Retrieval Tool 23.4
Imaged with Software Licensed to (Company Name)	DTC Dynamic Test Center AG
Reported with CDR version	Crash Data Retrieval Tool 23.4
Reported with Software Licensed to (Company Name)	DTC Dynamic Test Center AG
EDR Device Type	Airbag Control Module
Event(s) recovered	None

Fig. 8 *Anonymization: Example CDR file information from a VW ID3*

In that case, the data can be considered anonymized. No additional data protection requirements apply. Nonetheless, we recommend a privacy risk assessment to determine re-identification risks and to define appropriate mitigation measures.

However, for many older vehicles that do not fall under EU Directive (2019/2144) the full VIN is typically included in the readout. In that case, the readout is to be considered as personal data and legal restrictions apply. However, this can be easily addressed by omitting the VIN.

8.4.4 Floating car data more broadly

If such data were made available, complex data protection questions would arise. Moreover, OEMs appear to treat such information as trade secrets that they are not prepared to share. To comply with data protection requirements, we therefore recommend applying the Five Safes framework. To comply with data protection requirements and to accommodate OEM's commercial interests and liability concerns we recommend using a trusted data space as described in Chapter 6 would be required.

The secure data space could incorporate techniques like Federated Learning, Homomorphic Encryption and Secure Multi-Party Computation (SMPC) [50]. These advanced privacy-preserving technologies allow for collaborative data analysis without exposing sensitive information, safeguarding trade secrets while enabling the collaborative enhancement of AI models through insights derived from the data. This approach offers a win-win solution, benefiting third-party data users, third-party data providers and OEMs. Thereby the secure data space could potentially reconcile privacy concerns, address OEM reservations about trade secret exposure, and foster collaboration to collectively improve their autonomous driving algorithms (see Section 5.4.2).

However, implementation of a secure data space requires the set-up of a governance architecture that defines the roles, rights and responsibilities of the participants (OEMs, equipment and service providers, researchers, government actors, etc.). To date, no such governance architecture is available. And as described in Chapter 7, the Federal Council sees the setting up of such governance architectures essentially as a private sector activity. And it has also declared that the role of government should be to enable and not to replace private initiatives [53]and has recommended following and closely coordinating with initiatives in the EU. Therefore, we do not see any options for FEDRO except to follow ongoing initiatives in this domain.

8.5 Recommendations

1. Near-accident data: We recommend that FEDRO, Cantonal and city governments start to explore the use of near-accident data to identify risks in the road network. Several firms already provide near-accident data for sale. Given the strong public interests at stake we recommend that they explore legal ways requiring OEMs to share near-accident data free of charge. We also recommend that quality checks and validation tests are done to ensure the quality and comparability of the data. When accident data is used, a privacy risk assessment based on the Five Safes framework should be applied. We recommend using the checklist introduced in Chapter 6.5.
2. We recommend that FEDRO explores ways to integrate EDR data in accident statistics. On its own, EDR data can be difficult to interpret. Therefore, we recommend that FEDRO develops ways to validate and triangulate the EDR data with other sources. We assume that the data can be sufficiently anonymized so that no data protection concerns arise.
3. In the future, when policymakers develop new regulations that concerns sensor data, we recommend that they include in the law research and prevention as one of the purposes for which the data can be used. The example of eCall data shows that if this is not defined as the explicit purpose data protection rules may prohibit the research and prevention community from using the concerned data.

4. We recommend that FEDRO monitors and promotes the development of governance architectures to allow for the sharing of more sensor data with accident researchers and prevention specialists in academia, government and the private sector. We conclude that technically it is possible to set up systems that allow for the sharing of sensor data while both respecting data privacy requirements and accommodating the commercial interests of OEMs.

9 Use case 2: Exposure: Who drives when and where?

9.1 The problem

While use case 1 focuses on using sensor data to gain insights into accidents, use case 2 delves into scenarios where accidents do not occur. Specifically, it involves a counterfactual analysis, examining cases where accidents are averted. This approach aims to understand the factors and conditions contributing to accident prevention.

Such exposure data is critical to put the number of accidents happening in a given time and place into perspective and to draw meaningful conclusions (for instance [66, 67]). Exposure data is important because it allows comparing a location with lots of accidents but also high traffic intensity with a location that has fewer accidents but possibly also less traffic density.[62] Such comparisons are crucial for road authorities to be able to determine where to take action to improve safety. Objective criteria are required to determine whether a location is unsafe. Subjective perceptions of road safety often differ quite substantially from the facts [70]. Especially at places that are subjectively perceived as dangerous but have not been classified as accident hotspots so far, the derivation of preventive consequences is unclear. The problem is aggravated by the fact that places perceived as dangerous per se are sometimes particularly safe, for example due to more appropriate allocation of attention [71]. The dilemma of whether something needs to be changed at supposedly dangerous non-accident hotspots and - if so - what exactly, can only be clarified with the help of more detailed data.

To compare different states of safety, the road safety literature typically refers to the odds ratio [72] which is defined as the ratio of the probability of an accident happening to the probability of an accident not happening. With the odds ratio, it is possible to determine how safe or unsafe a location or situation is. Therefore, this use case is about accidents that do not happen.

Based just on data from the accident statistics, for instance, prevention specialists and the responsible civil engineering departments can only comparatively roughly determine the probability of an accident happening in a given location, period of time, and maybe which people (for instance, young vs. old, during daylight or night, residents or non-residents) are affected in the first place. However, particular causes are unknown.

Without taking exposure into account, civil engineering offices try to improve safety at locations whose "problem" merely might be that they are located in conurbations and are therefore heavily frequented. A particular hazard does not necessarily emanate from these locations. On the other hand, prevention measures would not be implemented at highly dangerous but low-frequented locations where only a few road users frequently have accidents.

Ignoring the risk exposure of individual road user groups would yield similar effects: Objectively non-existent needs of persons frequently present at certain locations might be overestimated, while those of others could be underestimated. Without exposure data, determining whether it is more dangerous to use a zebra crossing or to cross the street elsewhere becomes challenging [73]. Accident statistics might indicate more accidents on zebra crossings than when people cross the streets at unsignalized areas. Thus, to determine whether zebra crossings are genuinely more dangerous, it is necessary to know

[62] The Swiss VSS-Norm SN 641 724 [68] "Strassenverkehrssicherheit; Unfallschwerpunkt-Management / BSM", for instance, defines accident hotspots as areas where the number and severity of accidents over three years exceed a threshold value that is typical for comparable locations [69]. It thus refers directly to a counterfactual.

how often people use crosswalks versus crossing streets elsewhere and to what extent each behavior leads to accidents.

Another instance to consider involves exploring whether electric vehicles contribute to a higher rate of accidents compared to conventional vehicles. This inquiry was recently brought in Parliament[63]. However, accident researchers and prevention specialists cannot address this question due to a lack of essential exposure data. While we may have statistics on the sales of electric vehicles, we lack information about the time and the distance these vehicles spend on the road.

Currently, exposure data relies on two main sources. One is the microcensus on mobility by the Federal Office for Spatial Development (ARE) (Mikrozensus Mobilität und Verkehr) [89]. This survey, based on representative data from the Swiss population, occurs every five years. Despite its overall representativeness, it relies on a small sample size for subgroups and, as a result, cannot provide exposure data for specific locations in the road network.

Another source is traffic counters, which provide location-specific exposure data [106]. However, there are limitations. The amount of traffic counters that can be installed is limited, and they only provide data on the number of vehicles that cross a certain point. Most traffic counters lack the capability to differentiate between vehicle types and do not offer information about the origin and destination of the vehicles.

9.2 What data could potentially be used?

9.2.1 Description

As described in Chapter 3, mobility data is being generated from vehicles as well as satellite navigation devices and cell phones. Moreover, floating car data can provide information on more than mobility. Image data especially may provide information on the surroundings of vehicles and the events leading up to accidents.

9.2.2 Potential

The main potential is that, unlike the micro census, mobility data can provide location-specific exposure data. In principle, mobility data is available for any location. And in contrast to traffic counters, the mobility data can also reveal where vehicles are traveling from and to. From vehicles movement history it is even possible to deduce whether the driver is familiar with a certain location or not. This can be an important safety factor.

With floating car data, especially image data, one could also take the concept of exposure further. If for instance, researchers and prevention interests want to study a specific situation that often leads to accidents, floating car data could be used to measure how often these situations do not lead to accidents.

9.2.3 Challenges

The disadvantage of deriving exposure data from mobility data is that the mobility data is not representative.

That is because only certain vehicles can record and communicate location data and not everyone owns or carries a cell phone (e.g., children or the elderly) for instance. Therefore, we concluded that, today, there is no advantage to using sensor data. In the future, when

[63] Interpellation GRABER (19.3137) : Faut-il vraiment des voitures électriques qui produisent du bruit artificiellement? (2019).

mobility data for more traffic participants becomes available, this may change. Therefore, we do not see that mobility data from vehicles could replace the micro census, which, through systematic sampling, is more representative. However, we do see benefits in using sensor data to study very specific locations in the road network - “blind spots” - for which no exposure information exists.

In addition, mobility data generated from vehicles, satellite navigation devices and cell phones are personal data. Its processing shall respect the FADP.

9.3 Can the data be accessed?

Mobilty data is offered for sale by OEMs, mobile telecom operators and other data providers.

Floating car data by contrast is not readily available as mentioned in 8.3.4.

9.4 How would the data need to be managed once access is secured?

Data providers anonymize and aggregate mobility data they sell, potentially exempting it from data protection requirements (see Chapter 5.3.2). However, the risk of re-identification increases as more sources of mobility data and other datasets like the residency registry are combined. Also, in areas with lower population density or less traffic, individual movement patterns and locations may be more distinctive, making it easier to re-identify individuals even if the data has been anonymized or aggregated. Reduced anonymity can increase the privacy risks associated with the use of mobility data, emphasizing the need for careful consideration and safeguards when working with such datasets in these contexts.

Therefore, we recommend conducting a privacy risk assessment using the Five Safes framework described in Chapter 6. To illustrate the application of this framework, we provide a fictional example.

Considering a scenario where a city’s civil engineering department, responsible for road safety, receives complaints about a specific residential area as being deemed “dangerous”. In response, the department must act, but the question is how to proceed precisely. The first step involves collecting accurate data on the location and circumstances of any (near) accidents. To determine whether a genuine safety problem exists at one particular intersection, the traffic authority requires exposure data. This data helps understand the frequency of the reported incidents and to differentiate between situations that pose a safety concern and those that do not. Today, the responsible civil engineering department would request the installation of traffic counters. However, this may take some time and there is a limit to the number of traffic counters that can be installed. Moreover, traffic counters only count the number of passing vehicles. Sensor data may provide more information, for example, if primarily local drivers or non-local drivers are concerned. Additionally, considering vehicle-specific data could provide insights into the causes of accidents.

In this example, the civil engineering department could obtain the required data by from a commercial data provider, as some already offer such information. Additionally, they could opt for a more secure approach by leveraging a designated data space, as suggested in Chapter 6. This approach offers the advantage of securely combining data from various sources—both public and private, including mobility data from private providers—for a legitimate purpose.

Tab. 6 *Five Safes Checklist applied to the use case*

Factors	**Input**
Is the purpose defined? What is the benefit of using sensor data? (Usefulness, public interest) What are the risks of doing or not doing the project?	• The purpose is to contextualize the reported incidents by comparing them to the actual volume of traffic at the intersection. This comparison is important to determine whether safety measures need to be implemented. The intermediary assesses whether key legal principles are met. • The applicant has a legitimate goal since incidents have been reported, indicating a potential safety problem. To determine the need to put in place safety measures, the traffic authority needs to assess the actual risk and for this needs exposure data. Therefore, the intermediary determines that the application meets the usefulness criterion. • The data requested is proportionate to the purpose. For the analysis, aggregated and anonymized data is sufficient; pseudonymized data is not necessary and has not been requested.
Who are the actors involved?	• The data provider are OEMs or third party providers of mobility data. This are also the data controllers. • The civil engineering department is the data processor.
Do they have the motivation and capacity to re-identify?	• The public authority, serving as the applicant, demonstrates no apparent conflicting interests with the data protection objectives. Furthermore, the applicant has a history of acting in good faith, with no past instances of misconduct. As a result, the intermediary concludes that the applicant can be deemed as being in a "safe" category.
Capacity to implement data protection according to 5 Safe?	• The applicant also has the training to implement 5 Safe.
Do technical and organizational controls need to be implemented to deter intentional re-identification attempts and prevent potential data breaches?	• The intermediary requires the applicant to analyze the data within the secure data space. Access rights are allocated via the clearing house. Data analysis takes place in a secure data space [108], where the applicant can analyze the data but is restricted from moving the data outside that space. The result output is the only data allowed to be removed from the secure space, and this is permitted only after it has been assessed as safe from privacy risks (referred to as Safe Output). This limitation is important as it prevents the data from being matched with other sources, which could potentially lead to reidentification. In our example, the public authority might have access to the resident registration registry, which presents a reidentification risk. This is particularly significant in scenarios where, for instance, one of the streets concerned is a dead-end street with only a few residents. Reidentification risks may arise, especially when the applicant's objective is to differentiate between local and non-local drivers. For instance, a consistent commuting pattern could lead to re-identification. Within the secure data space, measures are in place to ensure that the provided data remains isolated and cannot be combined with external datasets. Moreover, access to the data within the secure space is restricted to authorized individuals or entities with a legitimate need for analysis. This limits the exposure of sensitive information to only those who have a valid reason to access it.
Does the setting provide sufficient security?	See above.

What threats to the data need to be managed?	Considering the purpose of the analysis, the intermediary assesses and determines the specific data layer to which the applicant is given access (refer to Chapter 6, section Data layer-based measures for safe data). The civil engineering office, having no need for microdata, is consequently granted access solely to anonymized and aggregated data. To ensure privacy, sensitive details are removed or obscured, preventing the direct association of data points with specific individuals. Aggregating data involves grouping information to a higher level (e.g., neighbourhood level) to prevent the identification of specific individuals. In this case, accident and near-accident data from the area are combined with data from nearby intersections and roads to create a larger dataset. This aggregated data allows for more meaningful analysis of accident trends without compromising individual privacy.
Are analytical results and other outputs ensured to be non-disclosable in an inappropriate manner?	Finally, the platform evaluates if privacy risks emerge if the results from the analysis are released from the secure data space (i.e., published). Considering that the data has been anonymized and aggregated, in this case, the intermediary allows the publication of the results without conditions.

9.5 Recommendations

5. Exposure data: Sensor data has great potential in the context of (risk) exposition. Sensor data, particularly mobility data, due to its current lack of representativeness may not be able to replace the micro census for mobility for now. And it may not yet replace data from traffic counters. However, mobility data may be used where census data or data from traffic counters is not available, for example, regarding specific neighborhoods or sections of road. We recommend FEDRO to support research that explores new ways of generating exposition data from mobility data.
6. Again, we recommend promoting the development of a secure data space. See recommendation 3 above.

10 Use case 3: Hazard warnings

10.1 The problem

Real time prevention is often difficult, the workshop participants and interviewees suggested, because relevant information about safety hazards is often lacking (weather conditions, accidents, potholes, people or animals on the road, road works, etc.).

Road authorities, for instance, often do not know where and when they need to clear roads from snow and ice.

10.2 What data could potentially be used?

10.2.1 Description

Modern cars are equipped with sensors which are designed to activate certain features of the car. For example, a car can detect rain and determine how heavy the rainfall is. It will then activate the wiper and set it to the right speed given the weather conditions. An automatic action taken by a car therefore gives indications about outside events.

Based on this premise, Regulation 886/2013 has defined eight types of hazards which can be detected by car sensors. In the context of the project Data for Road Safety, the hazards listed in the regulation have been put in relation with the corresponding sensor data – which gives a good overview of the data that could be used to detect hazards and therefore send warnings to road users:

Tab. 7 *Hazards defined in Regulation 886/2013*

Hazard	Corresponding sensor data
Temporary slippery road	Activation events of the electronic driving dynamic stabilization program of the vehicle (“lamp on”), absolute friction values as detected by the vehicle (“μ”)
Animal, people, obstacles, debris on the road	Object recognition from rich sensors for outside situations or emergency call / breakdown call from ego-vehicles, where ego-vehicles are vehicles equipped with sensor technology.
Unprotected accident area	Object recognition from rich sensors for outside situations or emergency call / breakdown call from ego-vehicle
Short-term road works	Sign recognition of road work signs
Reduced visibility	Activation events of the vehicle light (fog lights), rain sensor data, wiper activation
Wrong way driver	Object recognition from rich sensors for outside situations or ego-vehicle detection by sign-recognition
Unmanaged blockage of a road	Object recognition from rich sensors for outside situations
Exceptional weather conditions	Activation events of the vehicle light (fog lights), rain sensor data, wiper activation, activation events of the electronic driving dynamic stabilization program of the vehicle (“lamp on”), absolute friction values as detected by the vehicle (“μ”)

10.2.2 Potential

Vehicles can detect and communicate hazard warnings. In Europe, eight types of hazards are covered by the EU delegated regulation. There is already an ongoing project in Europe, called “Data for Road Safety”. It was started by companies and governments to figure out how Regulation 886/2013 can be implemented. Technical solutions are available. At this stage, Switzerland could join this project, which would make more sense than creating another hazard warning system from scratch.

The following paragraphs therefore explain how Data for Road Safety works and do not propose a new process to send hazard warnings to road users. Data for Road Safety made a live map of events in May 2023, which can be explored on https://data-intelligence.post.lu/dfrs/. This map shows in real time the hazards (called events on the website) registered by cars and helps understand the potential of such an instrument. As an illustration, the print screen below shows the hazards recorded on 8 January 2024, a day where it snowed in continental Europe. All the blue dots represent slippery road events. Altogether, 6546 events were recorded that day at 6 p.m. – which means that the number of events is significant enough to notice patterns and issue relevant warnings [107].

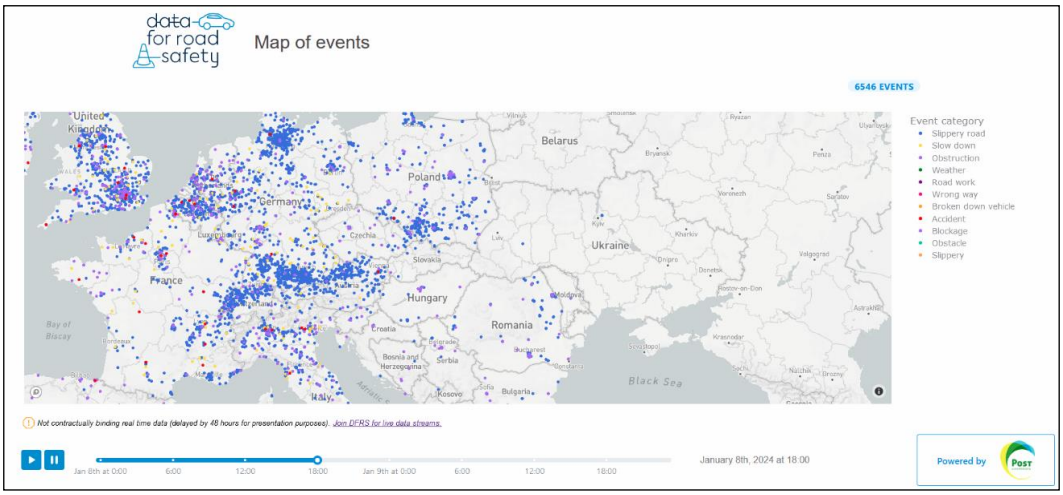

Fig. 9 *Map of events from the data for road safety project [86]*

The potential of the Data for Road Safety project is that it can effectively communicate warnings to approaching vehicles and the warnings can be used by the police to choose and prioritize appropriate responses to hazards.

While today, the Data for Road Safety project focuses on the eight hazards mentioned in the table above, the project could potentially be expanded to additional hazards. For example, in Scandinavia, road authorities use vehicle data to steer the allocation of winter service capacities [84]. Moreover, it also seems thinkable that the communication system set up by the Data for Road Safety is also used for the real-time communication of near-accidents, which we discuss in Chapter 8.

10.3 Can the data be accessed?

Regulation 886/2013, Data for Road Safety and the Multi-Party Agreement

As already mentioned in Chapter 4.2.2 EU Regulation No 886/2013 aims to provide road safety-related universal traffic information services. End users (meaning drivers; see art. 2 lit. k Regulation No 886/2013) can thus benefit from real-time traffic information relating to dangerous events or conditions such as temporary slippery road, animals, people, obstacles, debris on the road, or wrong-way drivers (see art. 3 lit. a, b and f Regulation No 886/2013).

Based on Regulation 886/2013 [74], European Transport Ministers, the European Commission and current industry partners have launched together the project “Data for Road Safety”, which aims to create a Safety Related Traffic Information (hereafter: SRTI) Ecosystem [74]. It was started because it was observed that the aforementioned Regulation was interpreted differently by public and private parties, and a clarification of the agreed extent and scope of the Regulation among the collaborating parties was needed [74].

To create the SRTI Ecosystem, the participating Member States and the industry first started a dedicated public-private task force called “the Data Task Force”. To extend their cooperation, they then signed a “Multi-Party Agreement” (hereafter: MPA). States and private companies (OEMs and service providers) can be parties to the MPA.

In this context, the data is made available on a reciprocal basis. As mentioned above, art. 3 MPA states that “Content is exchanged within the SRTI Ecosystem in-kind on the basis of reciprocity for the sole purpose of road safety”. Therefore, to access the data, each party to the MPA has to fulfill one of the following roles: Data Source, Aggregator, National Access Point (hereafter: NAP), Creator or Service Provider (art. 4 al. 1 MPA). Each party providing content (meaning data that is useful for the SRTI Ecosystem) grants to the others a license to use this content (art. 5 al. 1 MPA). The content can only be used to enable the provision of data free of charge to the end user (meaning drivers, as defined in art. 2 lit. k Regulation No 886/2013) and thereby improving road safety (art. 6 al. 1 MPA). State and public authorities are allowed to make a broader use of the content (art. 7 MPA).

To access the data, one must therefore first enter the MPA with one of the roles described above. Art. 13 MPA relates to the accession of new parties to the MPA. It provides that:

1. This Agreement will be open to the accession of new parties provided that this new party is able to fulfill at least one of the roles described in article 4.1.
2. Parties intend to promote this Agreement and the SRTI Ecosystem in order to expand this Agreement and the SRTI Ecosystem with third parties.
3. By signing this Agreement, all Parties authorize the incumbent Chair of the General Assembly to agree, on behalf of them, with the accession of an Entrant to this Agreement provided that this Entrant meets the following conditions:
   a) the Entrant formally declares it is willing and in a position to comply with all rights and obligations arising from this Agreement;
   b) the Entrant can provide evidence that it can contribute to the purpose of this Agreement as laid out in article 2 by fulfilling at least one of the roles (as laid out in article 4) within the SRTI Ecosystem;
   c) the Entrant declares which of the roles (see article 4) it intends to play within the SRTI Ecosystem and, thereby, what it is intending to contribute;
   d) the Entrant provides the information of Article 13.3 a, b and c in the self-declaration form (see Annex 2 [201005-IntakeFormSRTIEcosystem]) and distributes this to the Chair of the General Assembly.
4. The Chair of the General Assembly shall agree with the accession of an Entrant, provided that the conditions in article 13.3 are met, and shall inform each Party of the accession of a new Party.
5. The accession of new partners will be legally enacted only by a joint (digital) signature, as stated in article 19, of the following documents by the Legal Representatives respectively for A) the Chair of the GA at the time of signing, and B) the joining Party:
   a) The Multi-Party Agreement signature page;
   b) The self-declaration form.

The MPA will remain in force until October 31, 2025, and shall continue in force and effect from year to year with those parties that specifically approve such continuance at least annually (art. 15 al. 1 MPA).

So far, the MPA has been signed by 16 parties (Netherlands, Spain, Finland, Germany, Luxembourg, Mercedes, BMW, Ford, TomTom, HERE, Audi, Volvo, Belgium, Austria, Niradynamics, England).

Additionally, in May 2023, Data for Road Safety signed a cooperation agreement with National Access Point Coordination Organisation for Europe (NAPCORE), which is an organization that coordinates and harmonizes mobility data platforms across Europe [75]. The aim of this agreement is to “foster the exchange of intelligence on availability of data of relevance for road safety” [75]. Switzerland is part of NAPCORE [76].

It should be noted that, if Switzerland access this data, no further use of this data by the Swiss Government would be allowed. On the one hand, according to the purpose principle (art. 6 al. 3 FADP), personal data may only be collected for a specific purpose that the data subject can recognize; in addition, personal data may only be further processed in a manner that is compatible with this purpose. Using data from Data for Road Safety for other purposes than the project (meaning creating a system to send hazard warnings with the other stakeholders) would violate this principle. On the other hand, data processing by federal bodies requires in principle a legal basis, which does not exist in the present case.

10.4 How does the data need to be managed?

The functioning of the SRTI Ecosystem is, in summary, the following:

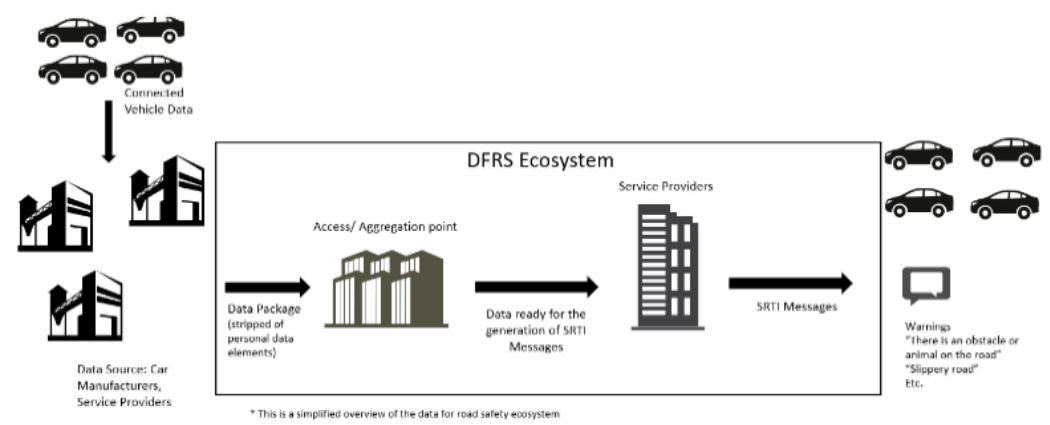

Fig. 10 *The SRTI Ecosystem [86]*

In this system, the events registered by a connected car are sent to a data source (generally an OEM). This data source then sends the data, stripped irreversibly of all identifiers, to an access point. This data consists of six data points: (1) event ID, (2) event type, (3) longitude, (4) latitude, (5) heading/direction of travel and (6) time stamp. The event ID contains a unique randomized rotating number within the ecosystem to solely identify the specific event. The data is then used to generate a SRTI message, which is sent by a service provider to different stakeholders (radio, navigation systems, etc.).

The data within the SRTI Ecosystem is anonymized and therefore not personal data. It is the responsibility of the data sources (OEMs or other) to make sure that the data entering the SRTI Ecosystem does not allow the identification of persons anymore. As explained in the Privacy Statement of Data for Road Safety,

> *“The source of the data exchanged within the Data for Road Safety ecosystem mostly comes from connected vehicles on the road in the European Union, produced by manufacturers which are part of Data for Road Safety. Car manufacturers collect various data points from such cars, depending on the services activated or consented to by the customer, agreements in place with customers, provided customer consent, as well as other factual elements in accordance with applicable law, differing from one manufacturer to another. Since this data is personal, each car manufacturer must have in place specific notices explaining the processing of personal data. Customers can obtain more information by contacting their respective*

> *manufacturer. It is important to note that this initial collection of car data by the manufacturer is not part of the Data for Road Safety ecosystem, and that data is fed into this ecosystem only after being stripped from all personal identifiers [...]". It is also specified that "All potential direct links to an individual and any other personal data have been irreversibly stripped from the data package before it enters the Data for Road Safety ecosystem in order to protect the privacy of individual vehicle owners and users".*

To sum up, the data used to send out a hazard warning is anonymous. It is therefore not personal data anymore and the constraints imposed by the data protection legislation do not apply.

10.5 Recommendations

To our knowledge, Switzerland is not part of Data for Road Safety. Therefore, two questions arise: (i) could Switzerland join the project and, if yes, (ii) should it?

Regarding the first question, the answer is yes: Switzerland has the possibility to join Data for Road Safety by signing the MPA. It seems that this option was considered in the MPA, in particular because art. 1 MPA defines "State" as „Any of the participating EU Member States, United Kingdom, Norway and Switzerland, including organizations working on behalf of the State or commissioned by the State". It should be underlined that, to sign the MPA, Switzerland will have to fulfill one of the roles mentioned in art. 4 MPA. It could in particular have the role of NAP, since Switzerland is already part of NAPCORE.

Regarding the second question, the answer is also yes: it would make sense for Switzerland to sign the MPA, because it could help contribute to improving road safety in Switzerland as well without having to create a new system. According to the Data Task Force's Final report [85], thanks to the system, 52% of incidents registered by vehicles were received by the NAP within 5 seconds, 85% within 1 minute and 96% within 5 minutes. It was also estimated that the average time gain for vehicle crashes is 11 minutes and 43 seconds, and for a broken-down vehicle 7 minutes and 30 seconds. It was therefore concluded that having Safety Related Traffic Information helps accelerate the deployment of emergency services and recovery. The reduction of the time necessary to detect an incident also helps traffic flow, because the average duration of the incident will be reduced as well. In addition, previous studies have shown that drivers, when receiving an in-car warning, adjust their speed. Finally, it was mentioned [85] that the data collected through the SRTI ecosystem could potentially be added to existing data sources.

Joining the project might also be an opportunity for knowledge transfer and this approach could potentially be expanded to other areas.

Moreover, it appears that signing the MPA was already contemplated by the Federal Department of the Environment, Transport, Energy and Communications (hereafter: DETEC): in a document titled "Provision and exchange of data for automated driving in road traffic" from December 7, 2018, the DETEC provides an overview of the measures to be taken.

We would therefore recommend that Switzerland join the MPA.

We emphasize that, since the data sent to the NAP is anonymized, there is no need for a legal basis from a data protection point of view. However, if Switzerland – via FEDRO or another entity – decides to join the project, it will be necessary to make sure that this entity has the capacity to sign the MPA.

7. We recommend that the Swiss Federation joins and supports the Data for Road Safety project. We also recommend that the project is used as a blueprint for other applications that require the collaboration of diverse actors, including OEMs and transport authorities.

11 Use case 4: Accident reconstruction

11.1 The problem

After a road traffic accident, the question arises how the accident happened and who is responsible. The public prosecutor, as the investigating authority in criminal proceedings, must get a picture of the course of the accident, while complying with the rules contained in the CrimPC (the FADP does not apply to criminal proceedings – see art. 2 al. 3 FADP – and the five safes framework is therefore not relevant in this context). The physical traces of the accident, however, only provide a limited picture of what happened. The fact that there are no break-marks, for instance, does not mean that the driver did not break. Moreover, witness reports may be unreliable and are inherently subjective. This is the main problem to be addressed in this use case.

It should be noted that an accident can also be followed by a civil procedure (if there is a victim claiming a damage) and/or by an administrative procedure (if the driver is at fault and his or her driving license needs to be suspended). The civil procedure is usually not conducted separately from the criminal one: the victim has the possibility to raise his or her civil claim during the criminal procedure and, if the driver is considered criminally responsible, he or she will be also considered civilly responsible. Even if the civil procedure is separate, conclusions will be drawn from the criminal procedure and there will be no additional accident reconstruction. The administrative procedure is always independent from the criminal procedure. However, the results of the criminal procedure regarding the cause of the accident will be used in the administrative procedure and – like in the civil procedure – no additional accident reconstruction will happen. In summary, the analysis of the accident and its causes will be conducted during the criminal procedure, which is why this use case focuses on this particular context.

Furthermore, there are two sub-problems in relation to accident reconstruction: The first problem is access sensor data (EDR). In most cases, reconstruction specialists have to ask the OEM to provide the keys to encrypt the EDR data. The second problem is that the police are concerned about the bureaucratic burden that is required to get the EDR readouts. We will further elaborate these problems below.

11.2 What data could potentially be used?

11.2.1 Description

Modern vehicles record a broad range of data that is relevant for accident reconstruction. In Chapter 3 we provide an overview of the relevant data.

11.2.2 Potential

A unique advantage of sensor data is that it can provide information on the events before and leading up to an accident.

In principle, all vehicle dynamics data from the EDR are relevant for accident reconstructions. In particular, physical parameters such as speed before the collision, accelerations, steering angles, brake activation and activation of the vehicle dynamics control system. The EDR records the vehicle dynamics 5 seconds before and up to 250 ms after irreversible safety systems (i. e. airbag deployment) are activated or the vehicle speed changes by more than 8 kph within 150 ms (delta v).

Until recently, the accident analyst had to concentrate exclusively on the trace situation. If there were no tracks in front of the collision point, it was unclear whether the vehicle had

not braked at all or had merely braked below the track drawing limit, i.e. with a correspondingly lower deceleration.

This data can be used to answer three questions:

1. How fast did the driver go and did he/she exceed the speed limit?
2. Did the driver respond appropriately and in time, for instance, by slowing the car down or by changing direction?
3. Were the available safety systems (ESP etc.) enabled?

In the future, camera data (from surveillance cameras but also from cameras build into the vehicles) might gain in importance for accident reconstruction. At the moment, however, camera data is not stored by the EDR.

11.2.3 Challenges

As argued above, the EDR data is sometimes difficult to interpret. It always needs to be contextualized and triangulated with the traces on site. If the car lifts off the ground, for instance, the rotation speed of the wheels does not provide an accurate representation of the real speed of the vehicle. Nevertheless, the EDR data is the best quality data stored in the vehicle for describing and classifying the accident.

Another challenge is that the vehicle sensors that send data to the EDR are not yet good at storing collisions between cars and vulnerable road users, such as pedestrians and cyclists. The problem is that such collisions often do neither lead to an airbag deployment, nor is the delta v threshold (more than 8 kph within 150 ms) exceeded. Therefore, the EDR is not activated and no data is stored.

First manufacturers have introduced sensors that can detect vulnerable road users. When a collision is detected, they then lift the engine hood of the car to dampen the impact for vulnerable road users or deploy special airbags (secondary VRU safety system).

11.3 Can the data be accessed?

Currently, the EDR data has to be red out after an accident on the vehicle. The data is not communicated outside the vehicle. It can only be obtained ex post, i.e. after the EDR has detected an accident. There are four options for practically accessing data:

1. Direct readout using Bosch CDR via the standardized OBD 2 interface (On Board Diagnostics 2). The EDR function (Event Data Reader) is integrated in the Airbag Control Module and the power supply to the control unit is intact.
2. Remove the control unit and read it out in the laboratory (also using the Bosch CDR tool). The EDR function (Event Data Reader) is integrated in the Airbag Control Module. If the power supply to the control unit is intact it can be read out in the vehicle. If the power supply is defective, the EDR has to be taken out of the vehicle and read out in the laboratory.
3. If the manufacturer has not provided access via Bosch CDR in the Airbag Control Module, the data can only be read out by the manufacturer of the control unit. That means that the EDR unit has to be physically removed and send to the manufacturer.
4. In older vehicles, the Airbag control Module often does not contain any usable data relevant to the accident.

All control unit manufacturers now require orders from the investigating authorities as well as a declaration of consent from the vehicle owner and, in some cases, the vehicle manufacturer. While some manufacturers provide the EDR readouts free of charge, others charge large fees or refuse to share the data.

Accident reconstruction specialists and the police are concerned about the bureaucratic burden that is required to get the EDR readouts. The problem is not obtaining the EDR or the EDR data. The unit can be accessed and the data downloaded. The problem is decrypting and reading it because the data is encrypted and OEMs may not provide decryption keys.

When the OEM does not collaborate and is only based abroad, it can only be compelled in compliance with the international rules, in particular the mutual legal assistance treaty (MLAT). This is not impossible, but it requires more time and effort and prosecutors will therefore often determine that it is not worth the investment of time and resources for the expected results.

Legally, however, this can be treated as a “technical” problem. It is a problem of finding an expert to make the data accessible (i. e. jailbreaking the device) or compelling the responsible companies (who in that sense do have the expertise) to provide the keys. At first, the authority will turn to the OEM but it could ask any other expert too.

When the manufacturer sits outside of Switzerland, it can be a good strategy to approach the importer or reseller in Switzerland. The importer normally should be able to read the EDR. The importer typically has to comply with a set of regulations before it can bring vehicles on the markets and the law could include an obligation to provide readable data. This approach, however, is time consuming. Again, prosecutors, in many cases, therefore decide not to invest the time and resources necessary.

Swiss policy makers could introduce a new legal basis (law or possibly an ordinance) through which the importer could be required to provide access to the EDR or the decryption keys. This is not something that needs to be regulated in the Criminal Procedure Code, but this could simply be a requirement as among the other conditions to be fulfilled to obtain a car homologation. For example, the law could mention that EDR data shall be easily accessible to law enforcement and car owners.

With the new EU regulation, the problem discussed above may eventually disappear. The regulation requires all manufacturers to allow for an EDR readout through standardized connectors. As new vehicles that have to comply with this regulation are brought into the market, the problem of access will gradually disappear. Until vehicles are still in the market that do not have to comply with the new EU regulation, however, the problem remains.

11.4 How does the data need to be managed once access is secured?

Access under criminal proceedings

In the course of a criminal procedure, the police can seize EDR data in accordance with art. 306 al. 2 let. a CrimPC. A police investigator mentioned legal uncertainties and conflicts about the analysis of this data, which is the step following the seizure of the data. Should EDR data be considered as “forensic and other evidence” in the sense of art. 306 al. 2 let. a CrimPC (meaning that it can be seized and examined by the police without any authorization), or should the examination of EDR data fall within the meaning of “data carriers and equipment for processing and storing information” of art. 246 CrimPC (which means that the public prosecutor, or, by delegation, the police has to issue a written search warrant [see art. 241 CrimPC] to authorize the examination of the EDR data [22]) [23]?

In 2015, Vuille and Arnold expressed the opinion that EDR data should be collected in accordance with art. 306 CrimPC, and that art. 246 CrimPC was not applicable to the analysis of the EDR content, because it does not contain the data subject’s thoughts [77]. Arnold took a slightly different approach in a 2017 article, where he indicated that EDR data should be qualified as evidence, but that given the broad scope of art. 246 CrimPC and the absence of an established jurisprudence, a search warrant should nevertheless be issued if the data needs to be analyzed [23]. In a later article, Vuille and Arnold mention that digital

data contained in vehicles are subject to a search within the meaning of art. 246 CrimPC [78]. Jeanneret also takes the view that collected directly by a connected vehicle should be searched in accordance with art. 246 CrimPC [79], meaning that a search warrant is necessary. On the other hand, Blanc/Zuber/Keusch/Liechti indicate that the Bosch CDR report has to be secured by the police according to art. 306 al. 1 let. a CrimPC and do not mention anything about the analysis of the data, which suggests that they consider EDR data as evidence available without a warrant [80].

Given the text of art. 246 CrimPC, which refers in particular to "equipment for processing and storing information"[64] we are of the opinion that EDR data falls under the scope of art. 246 CrimPC. This therefore means that this data is not readily available for analysis (unless the proprietor agrees to it; see art. 244 al. 1 CrimPC by analogy). A search warrant needs to be issued by the prosecutor– which requires that some conditions are fulfilled (see art. 197 CrimPC) [22], but this is in no way a complex process. This is slower (as a prosecutor shall be involved) but it also provides with more warranties and safeguards. The only action that the prosecutor, or the police (by delegation) is allowed to take without a warrant is to secure the data which is sufficient to save the evidence. The "proprietor" (meaning, in the present case, the owner of the car) has the right to comment before the search and can invoke its right to remain silent or to refuse to testify (art. 247 al. 1 and 248 al. 1 CrimPC). In this case, the data is sealed and the criminal authority needs to file a request for the removal of the seals within 20 days (art. 248 al. 2 CrimPC), otherwise the data may not be searched.

As for other data collected by connected cars, they are often not accessible without the help of the OEMs. Therefore, unless the OEMs are able and willing to help, this data will not be used [78]. It should be noted that art. 28 FADP provides for the right to portability, which allows data subject to request from data controllers their personal data in a conventional electronic format if, cumulatively, (i) the processing is automated and (ii) the data is processed with consent or in direct connection with the conclusion of the performance of a contract between the controller and the data subject. Even though this right cannot be relied on by criminal authorities to obtain data, it means that OEMs should have the means to provide data in a readable format.

To conclude, according to the law EDR data should be treated as "data carriers and equipment for processing and storing information" under art. 246 CrimPC and therefore a search warrant is required to examine the EDR data.

11.5 Recommendations

1. EDR: Because OEMs are often based in foreign countries, we recommend the introduction of a legal obligation (for example in traffic regulations) requiring importers of vehicles to provide access to EDR data to prosecutors in a readable format.
2. EDR: We consider it to be appropriate that a search warrant is needed to analyze EDR data and do not recommend a modification of the CrimPC. CrimPC shall remain technologically neutral and a specific regime for EDR data is not necessary.

[64] In German "*Datenträger sowie Anlagen zur Verarbeitung und Speicherung von Informationen*"; in French "*les supports informatiques ainsi que les installations destinées au traitement et à l'enregistrement d'informations*".

12 Discussion

12.1 Availability of sensor data

Modern vehicles are equipped with multiple sensors that record data with high relevance for accident research and prevention, a fact substantiated by our project. Often, OEMs provide limited information about the actual sensors they use. However, we find confirmation that these sensors exist in the EDR readouts, the hazard warnings that are communicated under Regulation 886/2013, the protocols developed by the Data for Road Safety project, or the data that is sometimes offered for sale.

The advent of automated driving will lead to the production of even more sensor data in the future. OEMs need sensor data to train their algorithms for automated driving. And once the algorithms have matured, vehicles will depend on sensor data to operate autonomously. As a result, we can expect a surge in the number and sophistication of sensors in future cars, generating a massive volume of data.

12.2 Potential and limitations of sensor data for research and prevention

Gaining access to sensor data could open new avenues for advancing road safety research and prevention efforts. For instance, as we describe in use case 1, accessing near-accident data could allow for new statistical research not achievable with existing accident statistics. Having access to near-accident data will enable prevention specialists to swiftly identify safety risks in the road network, eliminating the need to wait for actual accidents to be reported. This way, injuries and damage could be prevented. From sensor data, researchers and prevention specialists could potentially also deduce exposure data that has not been available before (use case 2).

However, this potential comes with challenges. To date, there is a lack of sufficient threshold values for interpreting sensor data. For instance, there is no commonly accepted and used definition of what constitutes a near-accident. There are also issues related to data quality and comparability. The are no verifiable information, for instance, explaining OEMs detect near-accidents. Moreover, there is the risk of misinterpreting sensor data. As discussed in use case 4, for instance, EDR data is easily misinterpreted. This underscores the need for triangulation with physical evidence and traces on site to ensure accuracy.

12.3 Accessibility

In practice, researchers and prevention specialists frequently encounter challenges in obtaining the necessary sensor data. We have identified three main constraints that impede accessibility: First, the sensor data holds substantial commercial value for OEMs and other equipment providers, as they rely on data for tasks such as training their automated driving algorithms. Second, there are apprehensions among OEMs and equipment providers about potential liability claims arising from the use of sensor data. Third, data protection regulations impose limitations on the sharing of sensor data. There is an increasing risk of re-identification as more data sources become available.

12.4 Legal considerations

The study provides an overview of the various legal requirements that need to be considered when using sensor data. The analysis shows that, in many instances, sensor data falls under the category of personal data, regulated by the Federal Act on Data Protection (FADP). The FADP defines personal data as all information relating to an identified or identifiable natural person. This notion is construed broadly. According to the Swiss Federal Supreme Court, this interpretation implies that the sensor data alone may

not be sufficient to identify a person, but additional information or context could enable identification. Thus, given the broad definition of personal data, much of the information gathered by cars might be deemed as personal data. This includes details about the vehicle's owner, such as the vehicle serial number (VIN) or license plate number, as well as information on location and driving style.

12.5 Data science foundations

The study also provides an overview of the data science foundations. On the one hand, the analysis shows how technological advances create reidentification risks. On the other hand, the analysis shows that, in principle, privacy-enhancing technologies are available (e. g. Federated Learning, Homomorphic Encryption, and Secure Multi-Party Computation) to share sensor data with the research and prevention community while meeting data protection requirements and accommodating the commercial interests of the companies and people involved. However, the application of these technologies can be costly, and it often requires multiple actors to collaborate. For that purpose, governance architectures are required.

12.6 Need for governance architectures

Governance architectures define who gets is allowed to access data under what circumstance, and how costs and benefits are distributed. The Data for Road Safety project, for instance, shows how such a governance system could work. It has brought together all relevant OEMs and other stakeholders who have agreed on procedures and protocols for the sharing of hazard warnings (i.e. black ice, pedestrians on the road etc.) in real-time among vehicles in the vicinity of the hazard. Another example is provided by the eCall system. This is also supported by a governance system through which OEMs allow their vehicles, in the case of an accident, to share important information with the emergency services.

Both eCall and the Data for Road Safety project demonstrate that technically it is possible to share relevant safety information. If it is technically possible to share the location of an accident in real-time (in the case of eCall), it should also be possible to share the location of near accidents with the research and prevention community as we discussed in use case 1. And if it is possible to share hazard warnings with approaching vehicles in real-time (in the case of the Data for Road Safety), it should be possible to share exposure data as discussed in use case 2.

Apart from the eCall and Data for Road Safety governance frameworks, there is currently a lack of governance structures designed to support the sharing of sensor data with researchers and prevention specialists.

12.7 Role of state intervention

In the proposed use cases, state interventions have played an important role in making sensor data available. The EU has required OEMs to share hazard warnings (use case 3) and accident alerts (eCall in use cases 1 and 2). The EU regulation on EDR will also make it easier to access EDR data (use case 4). Today, it can be difficult to obtain the necessary encryption keys from OEMs. The EU regulation on EDR forces OEMs to make the EDR data accessible through standardized interfaces. Moreover, new legislation on DSSAD will also ensure access to vehicle data in the future. At the time of writing, however, it is in the public consultation phase still.

The example of the Data for Road Safety project suggests that without political pressure and a direct legal requirement, firms are unlikely to share sensor data even when there is a strong public interest. Already in 2013, the EU with Regulation 886/2013 legally required OEMs to share hazard warnings. It took OEMs around ten years to develop the technical systems, communication protocols, and governance necessary to make sharing hazard warnings possible.

The example of eCall is also interesting because here an opportunity was missed to define research and prevention as one of the purposes for which the data can be used. As a result, data protection rules prohibit the use of eCall data for other purposes. This is an important lesson that needs to be considered when new legislation is developed in the future, for instance, in the case of DSSAD.

12.8 Options for policymakers in Switzerland

As the accessibility of sensor data is concerned, FEDRO, the Swiss Federation more generally, but also Cantonal or city governments have several options. First, they can wait for the private sector to possibly make sensor data available to them. Once available they can purchase the data (as in the case of near-accident data) or they might explore ways to force OEMs to share sensor data with them where there is a legitimate public interest.

Secondly, through legislation, they can adopt laws that oblige firms to share sensor data. The EU regulations requiring OEMS to share accident and hazard information in real time, for instance, are examples of this.

Thirdly, the state can, through financial subsidies or coordination activities, facilitate the emergence of governance infrastructures.

12.9 International context

Swiss policymakers should consider the international context. Vehicles are mostly manufactured abroad, bought and sold in international markets, and easily cross borders with other countries. Moreover, research and development costs are spread across several national markets. Therefore, ideally, systems for sharing sensor data are developed internationally. The Federal Council recognizes this and declares that it would monitor relevant regulatory developments in the EU and the realm of industry-driven governance architectures. This leaves very little room for our initiative.

If FEDRO and policymakers more broadly do not want to wait for the EU or the private sector to develop new solutions for making sensor data available, they are well advised to identify areas with a high public interest, i. e. areas where there is a high potential for accident prevention. In these cases, they may, through legislation, order firms to share data. And they might also want to search for areas where they might influence international developments by developing best practice examples that might be adopted internationally.

12.10 Directions for future research

The projects' results point toward several new avenues of research. These include:

1. Our analysis shows that given the public interests at stake, it can be justified and may be legally possible to introduce laws requiring OEMs to share sensor data with the research and prevention community. However, this would represent a drastic intervention in the market. Therefore, we recommend that thorough regulatory impact assessments ("Regulierungsfolgenabschätzung") *are* conducted that weigh the costs and benefits of such interventions.
2. Once sensor data becomes available, we recommend that more research be done on the limitations of using sensor data for accident research and prevention. Our findings suggest a) that there may be quality limitations because the sensors are not calibrated or because the data is not standardized across OEMs and b) there are challenges regarding the interpretation of the data. Concerning the latter, our results suggest that EDR data is easily misinterpreted. Therefore, it always needs to be triangulated with the physical traces on site. If EDR data is incorporated in the official accident statistics, there would be a need for ways to triangulate and validate the EDR data. Concerning near-accidents, our findings show that the way that OEMs and other providers define near-accidents is not standardized. Research that defines

thresholds for near accidents could potentially help define a commonly accepted standard of what constitutes a near-accident.

3. The project focused on sensor data from cars. Mobile phones and other devices, however, may also provide information on pedestrians, cyclists, and other traffic participants. Potentially, this data could also be used by the research and prevention community to further improve road safety for these traffic participants.
4. Our data science analysis shows that the technical tools are available to accommodate OEMs' commercial interests, liability concerns, and privacy risks. Often, however, these tools are not used. In many cases, a broader governance architecture is necessary to coordinate the various actors that would need to be involved in the implementation of these tools. In our use case analysis, we have described what such governance architectures could look like. However, more research is necessary to determine the factors necessary to build the necessary governance architectures.

13 Recommendations

1. Near-accident data: We recommend that FEDRO, Cantonal and city governments start to explore the use of near-accident data to identify risks in the road network. Several firms already provide near-accident data for sale. Given the strong public interests at stake we recommend that they explore legal ways requiring OEMs to share near-accident data free of charge. We also recommend that quality checks and validation tests are done to ensure the quality and comparability of the data. When accident data is used, a privacy risk assessment based on the Five Safes framework should be applied. We recommend using the checklist introduced in Chapter 6.5.
2. We recommend that FEDRO explores ways to integrate EDR data in accident statistics. On its own, EDR data can be difficult to interpret. Therefore, we recommend that FEDRO develops ways to validate and triangulate the EDR data with other sources. We assume that the data can be sufficiently anonymized so that no data protection concerns arise.
3. In the future, when policymakers develop new regulations that concerns sensor data, we recommend that they include in the law research and prevention as one of the purposes for which the data can be used. The example of eCall data shows that if this is not defined as the explicit purpose data protection rules may prohibit the research and prevention community from using the concerned data.
4. We recommend that FEDRO monitors and promotes the development of governance architectures to allow for the sharing of more sensor data with accident researchers and prevention specialists in academia, government, and the private sector. We conclude that technically it is possible to set up systems that allow for the sharing of sensor data while both respecting data privacy requirements and accommodating the commercial interests of OEMs.
5. Exposure data: Sensor data has great potential in the context of (risk) exposition. Sensor data, particularly mobility data, due to its current lack of representativeness may not be able to replace the micro census for mobility for now. And it may not yet replace data from traffic counters. However, mobility data may be used where census data or data from traffic counters is not available, for example, regarding specific neighborhoods or sections of road. We recommend FEDRO to support research that explores new ways of generating exposition data from mobility data.
6. Hazard warnings: We recommend that the Swiss Federation joins and supports the Data for Road Safety project. We also recommend that the project is used as a blueprint for other applications that require the collaboration of diverse actors, including OEMs and transport authorities.
7. EDR: Because OEMs are often based in foreign countries, we recommend the introduction of a legal obligation (for example in traffic regulations) requiring importers of vehicles to provide access to EDR data to prosecutors in a readable format.
8. EDR: We consider it to be appropriate that a search warrant is needed to analyze EDR data and do not recommend a modification of the CrimPC. CrimPC shall remain technologically neutral and a specific regime for EDR data is not necessary.

Appendix

I Reaching out

In the following tables we document the stakeholders that participated in the workshops organized as part of the second work package (“reaching out”).

***Tab. 8** Workshops with policy-makers*

Invited people: 28		
Date	**Participants: 9**	**Projektteam: 5**
Sept 06 2022 (D)	City representative	Meyer Niclas
	Cantonale representative, civil engeineering office	Stoll Tanja
	Cantonale representative, civil engeineering office	Percassi Marie-Laure
	Cantonal police	Salem Maria
	Cantonal police	
	FEDRO representative	
Sept 22 2022 (D)	FEDRO representative	Meyer Niclas
	Cantonal civil engineering office	Stoll Tanja
	FEDRO representative	Reber Heinz
		Percassi Marie-Laure

***Tab. 9** Workshops with associations*

Date	**Participants: 4**	**Projektteam: 5**
Sept 06 2022 (E)	Representative of an EU level representation for pedestrians	Meyer Niclas
		Stoll Tanja
		Percassi Marie-Laure
		Salem Maria
Sept 22 2022 (D)	Representative of a foundation active in the field of road safety	Meyer Niclas
	Representative of an association representing cyclists	Stoll Tanja
	Representative of an association representing pedestrians	Reber Heinz
		Percassi Marie-Laure

***Tab. 10** Workshops with industry*

Date	**Participants: 5**	**Projektteam: 5**
Sept 21 2022 (E)	Representative of a Swiss data provider	Meyer Niclas
	Representative of a German OEM	Stoll Tanja
	Consultant working in the mobility domain, who previously worked for an equipment supplier	Percassi Marie-Laure
	Representative of the Data for Road Safety Project and employee of a ministry of an EU country	Reber Heinz
	Representative of a Scandinavian OEM	Ossey Sabrina

***Tab. 11** Workshops with legal experts*

Date	**Participants: 7**	**Projektteam: 3**
Sept 07 2022 (E)	Representative of the Federal Data Protection Commissionar	Sylvain Métille
	Representative of a data provider	Percassi Marie-Laure
	Representative of a German OEM	Meyer Niclas
	Representative of the Federal Data Protection Commissionar	
	Lawyer and representative of a regional association	
	Law professor	
	Lawyer specialized in data protection	

***Tab. 12** Workshops with researchers and prevention specialists*

Date	**Participants:**	**Projektteam:**
Oct 24 2022 (D)	Cantonal police	Heinz Reber
	Cantonal police	
	Cantonal police	
Oct 18 2022 (D)	Engineer	Heinz Reber
	Engineer	
	Insurance representative	
	Representative of a research institute	

***Tab. 13** Additional interviews with stakeholders*

16.3.2022 & 18.11.2022	Data provider
28.09.2022	Consultant spcialized in automated mobility
16.09.2022	Representative of a start-up in the domain of automated mobility
08.09.2022	Representative of a French equipment provider
26.07.2023	FEDRO representative
18.07.2023	Cantonal police
04.06.2023	Representative of the Data for Road Safety project

Glossary

In the following glossary we define the main technical terms used in the report.

Notion	Description
ABS	Anti-Lock Brake System (ABS) Prevents excessive brake slip through wheel-specific braking force reduction and enables greater drive stability and steerability under braking.
ACC	Active Cruise Control (ACC) A function within an electronic control system. A continuous steering process is triggered to assist the driver. Steering actuation can be performed by automatic evaluation of signals triggered on board the vehicle, possibly in conjunction with passive infrastructure features.
ACM	Airbag Control Module (ACM)
ACSF category A	Automatically Commanded Steering Function (ACSF) category A 'ACSF of Category A' means a function that operates at a speed no greater than 10 km/h to assist the driver, on demand, in low speed or parking manoeuvring.
ACSF category B1	Automatically Commanded Steering Function (ACSF) category B1 'ACSF of Category B1' means a function which assists the driver in keeping the vehicle within the chosen lane, by influencing the lateral movement of the vehicle.
ACSF category B2	Automatically Commanded Steering Function (ACSF) category B2 'ACSF of Category B2' means a function which is initiated/activated by the driver and which keeps the vehicle within its lane by influencing the lateral movement of the vehicle for extended periods without further driver command/confirmation
ACSF category C	Automatically Commanded Steering Function (ACSF) category C ‚ACSF of Category C' means, a function which is initiated/activated by the driver and which can perform a single lateral manoeuvre (e.g. lane change) when commanded by the driver.
ACSF category D	Automatically Commanded Steering Function (ACSF) category D 'ACSF of Category D' means a function which is initiated/activated by the driver and which can indicate the possibility of a single lateral manoeuvre (e.g. lane change) but performs that function only following a confirmation by the driver
ACSF category E	'Automatically Commanded Steering Function (ACSF) category E ACSF of Category E' means a function which is initiated/activated by the driver and which can continuously determine the possibility of a manoeuvre (e.g. lane change) and complete these manoeuvres for extended periods without further driver command/confirmation.
ADS	Automated Driving Systems (ADS)
AEBS	Automated Emergency Braking System (AEBS)
CC	Cruise Control (CC) Holds the speed on a driver defined value by an intervention in the motor control.
CDR	Crash Data Retrieval (CDR) A device which enables the readout of EDR data.
CSF	Corrective Steering Function (CSF)
DSSAD	Data Storage System for Automated Driving (DSSAD) Recording who had the control (driver or vehicle) over a vehicle at a certain time.
EDR	Event Data Recorder (EDR) Records and store technical data (information on driving dynamics and occupant safety systems) about the vehicle for a brief period of time before, during and after a crash (-5s to 2s). But it's no black box.
ESF	Emergency Steering Function (ESF) Detects automatically a potential collision on what it activates the vehicle steering system for a limited duration, to steer the vehicle with the purpose of avoiding or mitigating a collision.
GRVA	Working Party on Automated / Autonomous and Connected Vehicles

OBD	On Board Diagnostics (OBD)
OEM	Original Equipment Manufacturer (OEM)
PTI	Portable Test Inspection (PTI)
TPMS	Tire Pressure Monitoring System (TPMS)
TTC	Time to Collision (TTC) Duration until the collision. With unaccelerated movement = distance / relative speed.
VRU	Vulnerable Road User (VRU) Road users, who are easily injured and killed in a car-dominated road space. E.g. Pedestrians, bicyclists and motorcyclists.
Yaw Rate	Yaw rate is an indication of a vehicle's rotation from its vertical axis. The Yaw Rate is the angular velocity while rotating from its vertical axis.

References

Documentation

[1] Vargas, J., Alsweiss, S., Toker, O. Razdan, R. & Santos, J. (2021), "**An Overview of Autonomous Vehicles Sensors and Their Vulnerability to Weather Conditions Sensors"** (Basel) 21. https://doi.org/10.3390/s21165397

[2] Conseil fédéral (2021), "**Le Conseil fédéral adopte le message concernant la révision de la loi fédérale sur la circulation routière**".

[3] Porter, B. E. (ed) (2011), "**Handbook of Traffic Psychology**". Academic Press, San Diego.

[4] (2022) "**29e Rapport d'activités 2021/22"**, Berne.

[5] Your Europe (2023), "**eCall 112-based emergency assistance from your vehicle**". https://europa.eu/youreurope/citizens/travel/security-and-emergencies/emergency-assistance-vehicles-ecall/index_en.htm

[6] (2022) "**Prescriptions techniques et administratives concernant l'acheminement et la localisation des appels d'urgence**".

[7] Préposé fédéral à la protection des données et à la transparence (2015), "**Explications concernant les systèmes «Pay as you drive» (PAYD) et l'utilisation de «boîtes noires» dans les véhicules automobiles: Explications concernant les systèmes «Pay as you drive» (PAYD) et l'utilisation de «boîtes noires» dans les véhicules automobiles**". https://web.archive.org/web/20160227174557/http:/www.edoeb.admin.ch:80/datenschutz/00768/00774/00952/00953/index.html?lang=fr

[8] Data for Road Safety (2020), "**Multi Party Agreement: Deployment of the SRTI Ecosystem: Data for Road Safety**". https://www.dataforroadsafety.eu/images/Documenten/Multi_Party_Agreement-_SRTI_Ecosystem_Data_for_Road_Safety_final_bundled_PDF_signed_version.pdf

[9] Chitkara, A., Deloison, T., Kelkar, M., Pandey P. & Pankratz, D. (2020), "**Enabling data-sharing: Emerging principles for transforming urban mobility**". https://docs.wbcsd.org/2020/01/WBCSD_Enabling_data_sharing_Emerging_principles_for_transforming_urban_mobility.pdf

[10] HERE Global B.V. (2023), **Website Here**. https://www.here.com/

[11] Wejo (2023), **Website Wejo**. https://www.wejo.com/

[12] HERE (2023), "**Global B.V. Public Safety: Emergency Management**". https://www.here.com/solutions/public-safety

[13] Wejo (2023), "**Road Health: Products**". https://www.wejo.com/products/road-health

[14] Bird & Bird (2019), "**Big Data & Issues & Opportunities: Data Sharing Obligations**".

[15] Gill, D. (2022), "**The Data Act Proposal and the Problem of Access to In-Vehicle Data and Resources**". https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4115443

[16] International Working Groups on Data Protection in Telecommunications (2018), "**Connected Vehicles**", Budapest, Hungary.

[17] Jotterand, A. (2022), "**Personal data or anonymous data: where to draw the lines (and why)?**" Jusletter, 15 August 2022.

[18] Conseil fédéral suisse (2017), "**Message concernant la loi fédérale sur la révision totale de la loi fédérale sur la protection des données et sur la modification d'autres lois fédérales**". https://www.fedlex.admin.ch/eli/fga/2017/2057/fr

[19] Métille, S. (2021), "**La (nouvelle) Loi fédérale sur la protection des données du 25 septembre 2020 : des principes, des droits et des obligations**". https://smetille.ch/2023/02/01/la-nouvelle-lpd-principes-droits-obligations/

[20] BC Freedom of Information and Privacy Association (2015), "**The Connected Car: Who is in the Driver's Seat?**". https://fipa.bc.ca/connected-car/

[21] (2022) Ordonnance sur la protection des données: RS 235.11.

[22] Laurent, M. & Parein-Reymond, A. (2016), "**Petit commentaire–CPP Code de procédure pénale**", 2nd edn. Helbing Lichtenhahn Verlag, Basel.

[23] Arnold, J. (2017), "**Jahrbuch zum Strassenverkehrsrecht 2017: Verkehrstechnik und Unfallanalytik. Digitale Spuren im Strassenverkehr – die Zukunft hat begonnen! Digitale Spuren aus Fahrzeugen, Schliesssystemen, Aufzeichnungsgeräten, Navigationssystemen: Interdisziplinäre Arbeit und strafprozessuale Fragen.**" Dike Verlag.

[24] ISO (2009), "**Road vehicles-world manufacturer identifier (WMI) code: ISO 3780:2009 43.020**". https://www.iso.org/standard/45844.html

[25] Tomoaki, M., Masayuki, H., Shinsaku, K., Koji, K. & Atsuko, M. (2021), "**Privacy Risk of Document Data and a Countermeasure Framework**". *Journal of Information Processing* 29:778–786. https://doi.org/10.2197/ipsjjip.29.778

[26] Templ, M. (2017), "**Statistical Disclosure Control for Microdata: Methods and Applications**" in R, 1st edn. Springer.

[27] Erdemir, E., Dragotti, P. L. & Gündüz, D. (2021), "**Privacy-Aware Time-Series Data Sharing With Deep Reinforcement Learning**". IEEE Transactions on Information Forensics and Security 16:389–401. https://doi.org/10.1109/TIFS.2020.3013200

[28] Li, S., Schneider, M. J., Yu, Y. & Gupta, S. (2023), "**Reidentification Risk in Panel Data: Protecting for k-Anonymity**". Information Systems Research 34:1066–1088. https://doi.org/10.1287/isre.2022.1169

[29] (2022) ISO/IEC 27559:2022: "**Information security, cybersecurity and privacy protection – Privacy enhancing data de-identification framework 35.030**". https://www.iso.org/standard/71677.html

[30] WP29 **Guidelines on Anonymisation.**

[31] Recital 26, GENERAL DATA PROTECTION REGULATION (GDPR).

[32] de Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M. & Blondel, V. (2013), "**Unique in the Crowd: The privacy bounds of human mobility**". *Scientific Reports* 3:1376. https://doi.org/10.1038/srep01376

[33] Zang, H., Bolot, J. (2011), "**Anonymization of Location Data Does Not Work: A Large-Scale Measurement Study. In: Proceedings of the 17th Annual International Conference on Mobile Computing and Networking**". Association for Computing Machinery, New York, NY, USA, 145–156.

[34] Ritchie, F. (2017), "**The 'Five Safes': a framework for planning, designing and evaluating data access solutions**", London.

[35] Stadler, T., Oprisanu, B. & Troncoso, C. (2020), "**Synthetic Data – Anonymisation Groundhog Day**".

[36] Archer, D. W., de Balle Pigem, B., Bogdanov, D., Craddock, M., Gascon, A., Jansen, R., Jug, M., Laine, K., McLellan, R., Ohrimenko, O., Raykova, R., Trask, A. & Wardley, S. (2023), "**UN Handbook on Privacy-Preserving Computation Techniques**". https://unstats.un.org/bigdata/task-teams/privacy/UN%20Handbook%20for%20Privacy-Preserving%20Techniques.pdf

[37] OECD (2023), "**Emerging privacy-enhancing technologies**". https://doi.org/10.1787/bf121be4-en

[38] Dwork, C., McSherry, F., Nissim, K. & Smith, A. (2006), "**Calibrating Noise to Sensitivity in Private Data Analysis**". In: Halevi S, Rabin T (eds) *Theory of Cryptography*. Springer Berlin Heidelberg, Berlin, Heidelberg, 265–284.

[39] El Emam, K., Dankar, F. K. (2008), "**Protecting Privacy Using k-Anonymity**". *Journal of the American Medical Informatics Association* 15:627–637. https://doi.org/10.1197/jamia.M2716

[40] Samarati, P. & Sweeney, L. (1998), "**Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression**". https://epic.org/wp-content/uploads/privacy/reidentification/Samarati_Sweeney_paper.pdf

[41] Avraam, D., Jones, E. & Burton, P. (2022), "**A deterministic approach for protecting privacy in sensitive personal data**". *BMC Medical Informatics and Decision Making* 22:24.

[42] Li, X., Zhao, B., Yang, G., Xiang, T., Weng, J. & Deng R. (2023), "**A Survey of Secure Computation Using Trusted Execution Environments**".

[43] Lindell, Y. (2020), "**Secure Multiparty Computation**". *Commun. ACM* 64:86–96. https://doi.org/10.1145/3387108

[44] Janssen, H., Cobbe, J., Norval, C. & Singh, J. (2020), "**Decentralized data processing: personal data stores and the GDPR**". *International Data Privacy Law* 10:356–384. https://doi.org/10.1093/idpl/ipaa016

[45] DETEC (2022), "**Creation of trusted data spaces based on digital self-determination: Report from the DETEC and FDFA to the Federal Council**", Bern.

[46] Otto, B., ten Hompel, M. & Wrobel, S. (eds) (2022), "**Designing Data Spaces: The Ecosystem Approach to Competitive Advantage**". Springer International Publishing, Cham.

[47] DETEC (2022), "**Creating trustworthy data spaces based on digital self-determination: Creating trustworthy data spaces based on digital self-determination**". Report from the DETEC and FDFA to the Federal Council on 30 March 2022, Bern.

[48] Boudreau, K. (2010), "**Open Platform Strategies and Innovation: Granting Access vs. Devolving Control**". *Management Science* 56:1849–1872. https://doi.org/10.1287/mnsc.1100.1215

[49] Williamson, O. E. (1991), "**Comparative Economic Organization: The Analysis of Discrete Structural Alternatives**". *Administrative Science Quarterly* 36:269–296. https://doi.org/10.2307/2393356

[50] Hagiu, A. (2014), "**Strategic decisions for multisided platforms**". *MIT Sloan management review*:77.

[51] McIntyre, D. P. & Subramaniam, M. (2009), "**Strategy in Network Industries: A Review and Research Agenda**". *Journal of Management* 35:1494–1517. https://doi.org/10.1177/0149206309346734

[52] Evans, P. & Gawer, A. (2016), "**The Rise of the Platform Enterprise: A Global Survey**", The Center for Global Enterprise.

[53] Digitale Schweiz (2020), "**Digital Switzerland Strategy**". https://digital.swiss/en/

[54] (2020) "**Digital Foreign Policy Strategy 2021–24**".

[55] European Commission (2020), "**Towards a European strategy on business-to-government data sharing for the public interest**", *Final report prepared by the High-Level Expert Group on Business-to-Government Data Sharing*.

[56] Teece, D. J. (1980), "**Economies of scope and the scope of the enterprise**". *Journal of Economic Behavior & Organization* 1:223–247. https://doi.org/10.1016/0167-2681(80)90002-5

[57] Meyer, M. H. & Lehnerd, A. P. (1997), "**The power of product platforms: Creating and sustaining robust corporations**", Tuscon.

[58] Teece, D. J. (1986), "**Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy**". *Research Policy*, 15:285–305. https://doi.org/10.1016/0048-7333(86)90027-2

[59] Chesbrough, H. W. (2003), "**Open innovation: The new imperative for creating and profiting from technology**". Harvard Business Press.

[60] Enkel, E., Gassmann, O. & Chesbrough, H. (2009), "**Open R&D and open innovation: exploring the phenomenon**". *R&D Management,* 39:311–316. https://doi.org/10.1111/j.1467-9310.2009.00570.x

[61] Ringel, L., Kielhauser, C. & Adey, B. T. (2023), "**Wider view over bicycle crashes: Complementing and extending bicycle crash statistics in urban areas using surveys**", *Journal of Safety Research.* https://doi.org/10.1016/j.jsr.2023.09.018

[62] ASTRA (2023), "**Unfallerfassung**". https://www.astra.admin.ch/astra/de/home/dokumentation/daten-informationsprodukte/unfalldaten/grundlagen/unfallerfassung.html

[63] "**Wikipedia NACA-Score**" (20 Oct 2023), https://de.wikipedia.org/wiki/NACA-Score

[64] Buck, M., Grau, N. & Spacek, P. (2016), "**Forschungspaket VeSPA: Synthesebericht**".

[65] Heinrich, H. W. (1941), "**Industrial Accident Prevention. A Scientific Approach**", McGraw-Hill.

[66] Mckenna, F. P. (1983), "**Accident proneness: A conceptual analysis**". *Accident Analysis & Prevention* 15:65–71. https://doi.org/10.1016/0001-4575(83)90008-8

[67] Groeger, J. A. (2011), "**How Many E's in Road Safety?**" In: Porter BE (ed) *Handbook of Traffic Psychology*. Academic Press, San Diego, 3–12.

[68] Bähler, L., Baumann, D., Brucks, W. et al. (2015), "**Strassenverkehrssicherheit - Unfallschwerpunkt-Management**": *SN 641 724*, Zürich.

[69] Bundesamt für Strassen ASTRA (2018), "**Bericht: Ermittlung von Grenzwerten für das Unfallschwerpunkt-Management**".

[70] Olma, J., Bode, T., Ehlers, J. & Sutter, C. (2022), "**Road Users' Reports on Danger Spots: The Crowd as an Underestimated Expert?**" *Safety.* https://doi.org/10.3390/safety8040070

[71] Ghielmetti, M., Steiner, R., Leitner, J., Hackenfort, M., Diener, S. & Topp, H. (2017), "**Flächiges Queren in Ortszentren - langfristige Wirkung und Zweckmässigkeit**". *Forschungsberichte SVI*.

[72] Cerrelli, E. C. (1997), "**Fatal Crash Involvements -- What Are The Odds?**".

[73] Ghielmetti, M., Hackenfort, M., Scaramuzza, G. et al., "**Überprüfung der 50m-Regel bei Fussgängerstreifen**".

[74] Data for Road Safety (2021), "**Privacy Statement - Data for Road Safety**". https://www.dataforroadsafety.eu/images/Documenten/20210706_Privacy Statement._DFRS_ecosystem.pdf

[75] Data for Road Safety (2023), "**Cooperation Agreement between the DFRS SRTI Ecosystem and NAPCORE**", Lisbon.

[76] NAPCORE (2022), "**National Access Point National Access Point**". https://napcore.eu/description-naps/national-access-point/

[77] Vuille, J. & Arnold, J. (2015), "**Moyens de preuve techniques et appréciation des preuves lors de la reconstruction d'accidents de la route**". *Strassenverkehr/Circulation routière* 2015.

[78] Arnold, J. & Vuille, J. (2019), "**L'appréciation des preuves techniques enmatière de circulation routière – les traces numériques**". *Strassenverkehr/Circulation routière* 2019:62.

[79] Jeanneret, Y. (2019), "**La preuve en droit pénal de la circulation routière: questions choisies et nouvelles technologies**". Strassenverkehr/Circulation routière:56.

[80] Blanc, A., Zuber, S., Keusch, T. & Liechti, S. (2022), "**Digitale Unfallspuren im Event Data Recorder**". *Strassenverkehr/Circulation routière*:88.

[81] OFCOM (2019), "**Prescriptions techniques et administratives concernant l'acheminement et la localisation des appels d'urgence**», *Annexe 1.3 de l'ordonnance de l'Office fédéral de la communication du 9 décembre 1997 sur les services de télécommunication et les ressources d'adressage (RS 784.101.113 / 1.3)*. Biel-Bienne.

[82] Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation (2023), "**Vernehmlassungsverfahren: Verordnung über das automatisierte Fahren (AFV) und Verordnung über die Finanzhilfen zur Förderung neuartiger Lösungen für den Verkehr auf öffentlichen Strassen (ÖStFV)**". https://www.fedlex.admin.ch/eli/fga/2023/2420/de

[83] Iansiti and Levien (2004), "**Strategy as ecology**". *Harvard business review*, 82(3), 68–78.

[84] Klimator.se (2024), "**How connected cars improves winter operations**". https://www.klimator.se/klimator-articles/how-connected-cars-can-help-winter-maintenance-operations

[85] Data for Road Safety (2020), "**Data Task Force: Final report & recommendations, Data for Road Safety**". https://www.dataforroadsafety.eu/images/Documenten/DTF-REPORT-OCTOBER-2020-021020.pdf

[86] European Commission (2016), "**C-ITS Platform: Final report**". https://transport.ec.europa.eu/system/files/2016-09/c-its-platform-final-report-january-2016.pdf

[87] OICA (2019), "**Data Storage System for Automated Driving (DSSAD)**". https://unece.org/fileadmin/DAM/trans/doc/2019/wp29grva/GRVA-02-20e.pdf

[88] UNECE (2024), "**DSSAD Performance Elements GUIDANCE DOCUMENT**". https://wiki.unece.org/display/trans/SG-DSSAD-18

[89] Bundesamt für Statistik (2023), "**Mikrozensus Mobilität und Verkehr**". https://www.bfs.admin.ch/bfs/de/home/statistiken/mobilitaet-verkehr/erhebungen/mzmv.html

[90] Bastiaansen, H., Bramm, G., Ceballos, J., Gall, M. & Kolenstart, M. (2021), "**IDS Clearing House**", Zenodo.

[91] United Nations (2023), "**Agreement Concerning the Adoption of Harmonized Technical United Nations Regulations for Wheeled Vehicles, Equipment and Parts which can be Fitted and/or be Used on Wheeled Vehicles and the Conditions for Reciprocal Recognition of Approvals Granted on the Basis of these United Nations Regulations**". https://unece.org/sites/default/files/2023-10/R160E.pdf

[92] EU (2021), "**Acts adopted by bodies created by international agreements**". https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:42021X1215&qid=1707320224037

[93] UNECE (2023), "**DSSAD/EDR: Working Party on Automated/Autonomous and Connected Vehicles (GRVA)**". https://wiki.unece.org/pages/viewpage.action?pageId=87621709

[94] OFROU (2023), "**Consultation: Ordonnance sur la conduite automatisée**". https://www.astra.admin.ch/astra/fr/home/themes/intelligente-mobilitaet/rechtliche-situation/vernehmlassung-verordnung-automatisiertes-fahren.html

[95] DETEC (2023), "**Ordonnance sur la conduite automatisée: Rapport explicatif relatif à l'ouverture de la procédure de consultation**". https://www.fedlex.admin.ch/filestore/fedlex.data.admin.ch/eli/dl/proj/2023/75/cons_1/

doc_9/fr/pdf-a/fedlex-data-admin-ch-eli-dl-proj-2023-75-cons_1-doc_9-fr-pdf-a.pdf

[96] Otonomo (2024), "**Securely Access Mercedes-Benz Data**". https://otonomo.io/daimler/

[97] Swisscom (2024), "**Mobility Insights: Mobile ---- Bewegungs-daten wertstiftend nutzen**", Swisscom Ltd.

[98] Here (2024), "**HERE Fleet Telematics API - Developer Guide**". https://www.here.com/docs/bundle/fleet-telematics-api-developer-guide/page/topics/use-cases.html

[99] Here (2024), "**The HERE Technologies Privacy Charter**". https://www.here.com/here-privacy-charter

[100] BMW (2024), "**Wie funktioniert der BMW Remote 3D View?**". https://faq.bmw.ch/s/article/My-BMW-App-Remote-3D-View-Funktionsweise-qQLWT?language=de_CH

[101] Tesla (2024), "**Sentry Mode**". https://www.tesla.com/ownersmanual/model3/en_us/GUID-56703182-8191-4DAE-AF07-2FDC0EB64663.html

[102] Waze (2024), "**Traffic data**". https://www.waze.com/wiki/USA/Traffic_data

[103] Wüst, J. (2024), "**Digitale Transformation, IKT Lenkung: "Informationspapier DataHub – Einordnung**". https://www.digitale-verwaltung-schweiz.ch/application/files/7916/4606/7304/Informationspapier_DataHub_Einordnung_-_Maerz_2022.pdf

[104] Tarko, A. (2019), "**Measuring road safety with surrogate events**", Elsevier.

[105] Sweeney, L. (2002), "**k-anonymity: A model for protecting privacy**". *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2002;10:557–570.

[106] ASTRA (2016), "**Strassenverkehrszählung**". https://www.astra.admin.ch/astra/de/home/fachleute/weitere-bereiche/geoinformation/geobasisdaten/strassenverkehrszaehlung.html

[107] Erticonetwork (2023), "**ERTICO Activities: DFRS launches game-changing live map for road safety data**". https://erticonetwork.com/dfrs-launches-game-changing-live-map-for-road-safety-data/

[108] ONS (2024), "**Accessing secure research data as an accredited researcher**". https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme

Projektabschluss

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für
Umwelt, Verkehr, Energie und Kommunikation UVEK
Bundesamt für Strassen ASTRA

FORSCHUNG IM STRASSENWESEN DES UVEK Version vom 09.10.2013

Formular Nr. 3: Projektabschluss

erstellt / geändert am: 27.3.2024

Grunddaten

Projekt-Nr.:	MFZ_20_07A_02
Projekttitel:	Sensordatenbasierte Unfallforschung: Rechtliche und technologische Herausforderungen und Möglichkeiten
Enddatum:	31.05.2024

Texte

Zusammenfassung der Projektresultate:

Moderne Fahrzeuge sind mit zahlreichen Sensoren ausgestattet. Die Daten, die mit diesen Sensoren erhoben werden, haben das Potenzial neue Wege für die Forschung und Prävention im Bereich der Strassenverkehrssicherheit zu eröffnen. In der Praxis haben Unfallforscher und Präventionsspezialisten jedoch oft keinen Zugang zu den entsprechenden Sensordaten. Trotz des klaren öffentlichen Interesses, die Forschungs- und Präventionsbemühungen zu unterstützen, haben sie oft Schwierigkeiten, an die Sensordaten zu gelangen, wie z. B. detaillierte Informationen über das Fahrzeugverhalten oder anonymisierte Standortdaten. So entsteht ein Spannungsverhältnis zwischen dem öffentlichen Interesse, die Verkehrssicherheit zu erhöhen, an sichereren Strassen und dem begrenzten Zugang, der derzeit gewährt wird.
Das Projekt beschreibt die rechtlichen und technologischen Lösungen, mit denen der Zugang zu Sensordaten besser erschlossen werden kann.
Anhand von vier Anwendungsfällen zeigt das Projekt auf was konkret getan werden kann, um Sensordaten für die Unfallforschung und - prävention zugänglich zu machen.
Im ersten Anwendungsfall zeigt das Projekt, wie die bestehende Unfallstatistik um Sensordaten (Beinaheunfälle, EDR-Daten und Floating Car Data) ergänzt werden könnte.
Im zweiten Anwendungsfall wird diskutiert, wie mit Sensordaten Fragen der Exposition erklärt werden können.
Im dritten Anwendungsfall wird erläutert, wie Gefahrenwarnungen (wie bspw. Glatteis oder Gegenstände auf der Fahrbahn) in Echtzeit bereits heute geteilt werden.
Im vierten Anwendungsfall untersucht das Projekt, wie bei der Unfallrekonstruktion der Zugang zu Sensordaten erleichtert werden könnte.

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für
Umwelt, Verkehr, Energie und Kommunikation UVEK
Bundesamt für Strassen ASTRA

Zielerreichung:

Alle Projektmeilensteine wurden erreicht:
Erstens wurde ein umfangreiches Stock-Taking durchgeführt und erfasst, welche Sensordaten es bereits gibt und welche künftig verfügbar sein könnten.
Zweitens wurde ein intensives Stakeholder-Engagment durchgeführt. Es wurden fünf Workshops durchgeführt, an denen mehr als 30 Stakeholder teilgenommen haben und sieben Interviews mit weiteren Stakeholdern durchgeführt.

Folgerungen und Empfehlungen:

Auf Basis der vertieften Analysen der vier Anwendungsfälle wurden acht Empfehlungen formuliert, die sich auf die Nutzung von kommerziellen Mobilitätsdaten und Daten zu Beinahunfällen beziehen sowie die Erweiterung der Unallstatistik um EDR-Daten und die Verwendung von EDR-Daten im strafrechtlichen Kontext. Zudem macht das Projekt Empfehlungen zur Entwicklung notwendiger Governance-Architekturen.

Publikationen:

Bislang noch keine.

Der Projektleiter/die Projektleiterin:

Name: **Meyer** Vorname: **Niclas**

Amt, Firma, Institut: **BSS Volkswirtschaftliche Beratung AG**

Unterschrift des Projektleiters/der Projektleiterin:

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für
Umwelt, Verkehr, Energie und Kommunikation UVEK
Bundesamt für Strassen ASTRA

FORSCHUNG IM STRASSENWESEN DES UVEK

Formular Nr. 3: Projektabschluss

Beurteilung der Begleitkommission:

Beurteilung:

Das Ziel des Projekts war die Untersuchung der rechtlichen und technischen Lösungen, um Sensordaten aus Fahrzeugen für die Unfallforschung- und Prävention zur Verfügung zu stellen. Die Begleitkommission (BK) ist der Auffassung, dass das Projektteam wissenschaftlich fundiert und praxisnah aufzeigt, welche Möglichkeiten Sensordaten bieten, welche rechtlichen Probleme aktuell bestehen und wie die Daten von den OEMs den Behörden, Forschungsstellen und Privaten unter Berücksichtigung der rechtlichen Rahmenbedingungen und den kommerziellen Interessen der OEMs zur Verfügung gestellt werden können. Die detaillierten Empfehlungen an das ASTRA fokussieren auf die wichtigsten Anwendungsfälle wie die Unfallstatistik, Gefährdungswarnungen oder Expositionsdatenerfassung sowie Unfallanalysen und umfassen technische wie rechtliche Aspekte.

Umsetzung:

Auf Basis der Literatur hat das Projektteam Anwendungsfälle definiert und mit der BK abgestimmt. Mit einem transdisziplinären Ansatz - einerseits durch das Projektteam selbst und andererseits durch acht Workshops mit Vertreter/-innen aus der Praxis und potentiellen Nutzniessenden - wurde sichergestellt, dass die Anwendungsfälle ganzheitlich untersucht wurden. Die Anwendungsfälle wurden auf Basis der relevanten rechtlichen und datenwissenschaftlichen Grundlagen diskutiert und Möglichkeiten der Datennutzung heute und in Zukunft aufgezeigt.

weitergehender Forschungsbedarf:

Das Forschungsteam erkennt Forschungsbedarf bezüglich einer Regulierungsfolgenabschätzung für den Fall, dass OEMs dazu verpflichtet werden sollen, Fahrzeugdaten für die Forschung und Prävention zur Verfügung zu stellen. Weiter soll Forschung aufzeigen, wie Sensordaten interpretiert werden müssen und welche Limitationen

Einfluss auf Normenwerk:

Die Empfehlungen richten sich eher an die Gesetzgebung und weniger an die Normengremien. Im Fokus steht die Entwicklung einer Governance-Architektur, um die Nutzung von Sensordaten zu ermöglichen

Der Präsident/die Präsidentin der Begleitkommission:

Name: Zahnd Vorname: Bettina

Amt, Firma, Institut: EBP Schweiz AG

Unterschrift des Präsidenten/der Präsidentin der Begleitkommission: