

Datenaufbewahrungs- und Löschkonzept

Universität Kassel

Projektkoordination: Fachbereich 16, Fachgebiet Kommunikationstechnik (ComTec)

Wilhelmshöher Allee 73

34121 Kassel

Deutschland

Bearbeitung: Prof. Dr.-Ing. Klaus David

Tel.: 0561 804 6314

Mail.: dynamo@uni-kassel.de

Titel des Projektes:

Sichere und nachhaltige Mobilität in der Stadt von Morgen –
Wie hilft Künstliche Intelligenz der Radverkehrssicherheit?
(DyNaMo)

Im Folgenden wird die Art der Erhebung personenbezogener Daten sowie das zugehörige Aufbewahrungs- und Löschkonzept im Rahmen des Genehmigungsverfahrens nach § 84 Abs. 1 HSchG für die Durchführung des Forschungsprojektes DyNaMo beschrieben.

Gesamtziel des Vorhabens

Das Ziel des Projekts ist die Entwicklung innovativer Technologien aus dem Bereich der künstlichen Intelligenz zur Verbesserung der Verkehrssicherheit für Radfahrende. Aufgrund der in Deutschland ansteigenden Zahlen der verunglückten Radfahrenden und der sich dabei als besonders gefährdet erweisenden Kinder und junge Erwachsene stellen Schüler und Schülerinnen eine besonders relevante Probandengruppe dar. Die im Projekt zu erhebenden Daten dienen insbesondere dazu Fahrfehler und gefährliches Fahrverhalten mittels künstlicher Intelligenz (KI) in realen Bedingungen vorbeugend erkennen zu können. Deshalb ist das Reallabor an Schulen für den Erfolg des Forschungsvorhabens maßgeblich, denn zuverlässige KI-Modelle können nur entstehen, wenn sie anhand von Daten aus authentischen Verkehrssituationen trainiert werden. Die Daten dienen einerseits zur (Weiter-)Entwicklung und Validierung solcher Technologien und fließen andererseits direkt in die Konzeption eines innovativen Verkehrssicherheitsprogramms für Schülerinnen und Schülern. Dahingegen kann im Fahrradsimulator das Radfahrverhalten der Schülerinnen und Schüler in verschiedenen Verkehrsszenarien unter kontrollierten Bedingungen untersucht werden. Aus den Daten wird also auch zwecks der Entwicklung von Verkehrssicherheitsmaßnahmen und Schulungen geschöpft.

Um die Projektziele zu erreichen, werden sowohl im Reallabor als auch im Fahrradsimulator Daten zum Fahrverhalten erhoben. Dies erfolgt mithilfe von Fragebögen, Kameras und LiDAR-Sensorik sowie optional durch den Einsatz einer Eye-Tracking-Brille. Im Fahrradsimulator werden zusätzlich Daten über die Sensorik am Fahrrad erfasst. Erhoben werden dabei folgende Datenkategorien: Kontaktdaten, demographische Daten von Schülern (u.a. Alter, Geschlecht, Größe, Gewicht), Daten über Radfahrerfahrung, Radfahrausrüstung und die persönliche Einstellung zum Radfahren, Augenbewegungsdaten, Radfahrverhalten (u.a. Arm-, Kopf-, Rumpfbewegungen, Blickrichtung), Standortdaten während spezifischer Fahrten, visuelle und akustische Daten über die Radfahrumgebung, sowie Daten des Fahrradsimulators (Daten aus einer simulierten Fahrt mit einem Fahrradsimulator, u.a. Geschwindigkeits-, Trittfrequenz-, Lenkwinkel- und Bremsaktivitätsdaten; Umfragedaten zur allg. Erfahrung mit dem Simulator, Einschätzung von Risiken im Straßenverkehr, Radweg zur Schule). Die Erhebung sämtlicher dieser Daten ist für die Gewährleistung der Leistungsfähigkeit der geplanten Maßnahmen und Technologieentwicklungen notwendig. Ohne die Daten kann das Radfahrverhalten nicht analysiert, erkannt oder vorhergesagt werden und somit auch keine Fahrfehler erkannt und vorgebeugt werden. Die Standortdaten sind erforderlich für die Positionierung der Radfahrenden. Ohne diese ist weder eine Fahrfehlererkennung noch eine Zusammenführung verschiedener Datenquellen möglich. Insbesondere ohne die Video- und LiDAR-Aufnahmen ist eine große Menge essenzieller Verhaltensdaten nicht erhebbar. Schließlich ist eine ausreichende Stichprobengröße für die Gewährleistung statistisch signifikanter Ergebnisse und möglichst geringer Fehlervarianz nötig. Über diese Projektzwecke hinaus können die Daten (optional) auch in der Lehre und Forschung anonymisiert oder pseudonymisiert eingesetzt werden.

Vorgehensweise zur Sicherstellung des Datenschutzes nach DSGVO, BDSG und HSchG

Insgesamt werden im Einklang mit den Grundsätzen in Art. 6 Abs. 1 DSGVO nur die Daten erhoben, die zwingend für die Zwecke des Forschungsprojektes erforderlich sind. Auch und insbesondere wird der Grundsatz der Speicherbegrenzung nach Art. 6 Abs. 1 lit. e) DSGVO berücksichtigt. Auch die Vorgaben von § 4 BDSG über den Einsatz von Videotechnik in öffentlich zugänglichen Bereichen werden eingehalten. Dabei sind jedoch einige Besonderheiten der Funktionsweise künstlicher Intelligenz zu berücksichtigen.

Die Videoaufnahmen werden aus zwei verschiedenen Perspektiven vorgenommen: ein mit Kamera und LiDAR ausgestattetes Auto und eine gleichfalls ausgestattete Road-Side-Unit (ein statisches Gehäuse). Dabei lässt es sich nicht vermeiden, dass die Schüler und Schülerinnen, die am Experiment teilnehmen erkennbar erfasst werden. Auch ist es möglich, dass sonstige Verkehrsteilnehmenden erkennbar erfasst werden. Dies gilt ebenfalls für die durchzuführenden Beobachtungen mit Eyetracking. Dabei wird die Fahrt eines Schülers bzw. einer Schülerin zur Schule aus der Sichtperspektive aufgezeichnet. Eine Anonymisierung oder Löschung personenbezogener Daten ist aufgrund der Funktionsweise der Technik und der Zwecke der Erhebung kaum möglich. Vielmehr wird, soweit möglich, mit Pseudonymisierung gearbeitet. KI-Modelle funktionieren nur dann zuverlässig, wenn sie unter Realbedingungen trainiert werden. An dem Personenbezug der Daten und insgesamt an der persönlichen Erkennbarkeit von Schülern und Schülerinnen sowie von Drittpersonen besteht allerdings kein Forschungsinteresse. Es lässt sich nicht vermeiden, dass solche Daten miterfasst werden, sie werden jedoch nicht zur Identifikation der Personen genutzt. Die Erkennbarkeit ist nur für die projektinterne Zuordnung von Datensätzen von Interesse. Modelle der künstlichen Intelligenz lernen aus (konkreten) Datensätzen verallgemeinerungsfähige Strukturen und Muster, ohne dabei vom spezifischem Personenbezug zu profitieren. Beispielsweise ist für ein Modell die Information, dass eine männliche radfahrende Person in einer spezifischen Altersgruppe und mit einer spezifischen Statur ein spezifisches Fahrverhalten aufweist, von Relevanz. Welche diese Person konkret ist, ist für die Modellentwicklung ohne Relevanz.

Alle im Rahmen der Erhebungen aufgezeichneten personenbezogenen Daten werden unverzüglich gelöscht, soweit sie nicht bzw. nicht mehr für die weitere Forschungsarbeit benötigt werden. Dieser Zeitraum umfasst jedenfalls die Projektdauer (bis Ende 2028). Dies ist erforderlich, da während der gesamten Dauer des Projekts damit an Anwendungen, Schulungskonzepte und Maßnahmen gearbeitet wird. Darüber hinaus würden die Daten benötigt, soweit das Projekt für eine weitere Zeitspanne finanziert wird, denn im Rahmen eines solchen Nachfolgeprojekts wird mit den Daten weiterhin nur zweckentsprechend gearbeitet. Auch könnten mit den Daten arbeitende wissenschaftliche Abschlussarbeiten länger als die Projektdauer brauchen. Insgesamt ist die unverfälschte Verfügbarkeit der Daten ein wichtiger Ausgangspunkt für künftige Forschung an der Universität Kassel. Nach Ablauf der Projektlaufzeit wird regelmäßig (mindestens einmal jährlich) überprüft, ob die Speicherung in nicht-anonymisierter bzw. pseudonymisierter Form noch notwendig ist. Fällt diese Prüfung negativ aus, werden die Daten gelöscht bzw. anonymisiert.

Im Rahmen der Lehre und der (über das Projekt hinausgehenden) wissenschaftlicher Publikationen wird nur mit anonymisierten bzw. pseudonymisierten Daten gearbeitet. Weiterhin werden alle nach Art. 32 DSGVO erforderlichen Technischen und Organisatorischen Maßnahmen (TOMs) getroffen. So wird u.a. sichergestellt, dass auf alle Daten nur befugte Personen zugreifen können. Die erhobenen Daten werden dazu nach Beendigung der Erhebungen von etwaigen mobilen Geräten (Kameras, Smartphones, etc.) gelöscht und auf separate Festplatten übertragen. Der Zugriff auf diese Festplatten erfolgt nur mit Kennwortschutz. Die Festplatten werden an einem Ort mit begrenztem Zutritt sicher aufbewahrt. Die Auswertung der Daten erfolgt ausschließlich von Computern und Netzwerken der Uni Kassel aus, ggf. wird noch die Rechenleistung eines Hochleistungscomputers der TU Darmstadt (HessianAI) für die Verarbeitung (nur) der pseudonymisierten Daten genutzt. Dabei erfolgt keine Übermittlung. Ein potenzieller (wenngleich unwahrscheinlicher) Zugriff von Administratoren der TU Darmstadt während des Verarbeitungsvorgangs kann dabei nicht ausgeschlossen werden. Die Daten werden bei den Verarbeitern wiederherstellbar gespeichert und es erfolgen entweder regelmäßige Backups über die IT-Services der Universität Kassel oder redundante Speicherprozesse, die vor Datenverlust schützen. Bei sämtlichen dieser Schritte werden die üblichen, dem Stand der Technik entsprechenden Sicherheitsverfahren der Universität Kassel eingehalten. Auf diese Weise wird der größtmögliche Schutz der Daten sichergestellt.

Quellen:

Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates v. 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung, DSGVO) (ABl. L 119 v. 4.5.2016, S. 1, ELI: <http://data.europa.eu/eli/reg/2016/679/oj>).

Bundesdatenschutzgesetz (BDSG) v. 30. Juni 2017 (BGBl. I S. 2097), zuletzt geändert durch Art. 7 G zur Durchführung der VO (EU) 2022/2065 v. 6.5.2024 (BGBl. 2024 I Nr. 149).

Hessisches Schulgesetz – in der Fassung v. 17. Dezember 2022 (GVBl. 2023 S. 150), zuletzt geändert durch Art. 2 G über die Ausgliederung der Stadt Hanau aus dem Main-Kinzig-Kreis und zur Änd. anderer Rechtsvorschriften v. 3.3.2025 (GVBl. Nr. 16).