Journal für Mobilität und Verkehr

ISSN 2628-4154

www.dvwg.de

Entwicklung eines konzeptionellen Frameworks zur Erfassung, Qualitätssicherung und Bereitstellung urbaner Umweltdaten als Grundlage für ein umweltsensitives Mobilitätsmanagement im Projekt AIAMO

Greta Schmalhorst, Timo Houben, Ema Vosgerau, Florian Berger, Michael Weger, Joost Hemmen, Alex Romanjuta, Paul Remmler, David Schäfer, Jan Bumberger, Thomas Trabert*

Siehe Autor*innenangaben

Abstract

Die ab 2030 geltenden, verschärften Grenzwerte der EU-Luftqualitätsrichtlinie stellen Kommunen vor die Herausforderung, verkehrsbedingte Immissionsspitzen gezielt und vorausschauend zu steuern. Herkömmliche Modellierungen und vereinzelte amtliche Messstationen bieten jedoch weder die räumlich-zeitliche Auflösung noch die operationelle Datenanbindung, die für ein dynamisches, umweltsensitives Verkehrsmanagement erforderlich sind. Dieser Beitrag stellt ein ganzheitliches methodisches Framework vor, das ein Netzwerk aus ca. 30 Sensoren im Stadtgebiet Leipzig mit Verkehrsfluss- und Wetterdaten verbindet. Die Systemarchitektur basiert auf dem Digital Ecosystem for FAIR Time Series Data Management und integriert ein Sensor Management System (SMS), die Plattform time.io sowie automatisierte Qualitätskontrollen mittels SaQC. Fehlende Daten in den Zeitreihen werden über ein vortrainiertes Foundation-Modell imputiert, wodurch ein konsistenter, FAIR-konformer und AI-ready Datenbestand entsteht. Auf dieser Grundlage ermöglicht ein KI-Vorhersagemodell eine 48-Stunden-Prognose der lokalen Stickstoffdioxid-Belastung (NO2).

Die verkehrswissenschaftliche Auswertung im Stadtgebiet Leipzig belegt eine ausgeprägte Koppelung zwischen Verkehrsfluss und Immissionen mit deutlichen Peaks zu Rush-Hour-Zeiten. Die Analyse differenziert zwischen strukturellen Dauerbelastungen im Leipziger Norden und temporären Spitzen durch Veranstaltungsverkehre. Im Rahmen des Projekts AIAMO (Artificial Intelligence and Mobility) fließen die Daten über eine interoperable OGC SensorThings API in automatisierte Wochenberichte für die Kommune ein und bilden das Fundament für umweltsensitive Routingkonzepte. Der Beitrag zeigt damit den konzeptionellen Weg von der reinen Datenerhebung bis zur Bereitstellung verlässlicher Entscheidungsgrundlagen für die Verkehrsplanung und umweltsensitive Verkehrssteuerung auf.

Schlagwörter:

Umweltsensitives Verkehrsmanagement, Sensorik, FAIR-Datenökosystem, Urbaner Raum, Monitoring

1 Einleitung / Problemstellung

Das anhaltend hohe Verkehrsaufkommen in verdichteten urbanen Räumen stellt Kommunen vor die komplexe Herausforderung, Mobilitätsbedürfnisse und Umwelt- sowie Gesundheitsschutz miteinander in Einklang zu bringen (Deng et al., 2021). Insbesondere im innerstädtischen Hauptverkehrsstraßennetz führt Verkehr zu Rush-Hour-Zeiten zu ausgeprägten Immissionsspitzen. Mit den ab 2030 geltenden Grenzwerten der neuen EU-Luftqualitätsrichtlinie werden die zulässigen Werte für verkehrsrelevante Schadstoffe wie Stickstoffdioxid (NO2) und Feinstaub weiter abgesenkt, sodass in zahlreichen

* korrespondierender Autor

europäischen Ballungsräumen ohne gezielte Interventionen mit signifikanten Grenzwertüberschreitungen zu rechnen ist (EU-Kommission, 2024). Um Kommunen zukünftig in die Lage zu versetzen, wirksame und dynamische Maßnahmen zu ergreifen, bedarf es einer hochaufgelösten und belastbaren Datenbasis sowie standardisierter Strukturen für deren Bereitstellung (Bibri et al., 2024).

Im Zuge der fortschreitenden Digitalisierung des Verkehrssektors eröffnen sich durch automatisierte Sensornetzwerke, vernetzte Fahrzeuge und Methoden der Künstlichen Intelligenz neue Möglichkeiten zur Erfassung und Analyse urbaner Mobilitätsdaten (Ismagilova et al., 2019; Minerva et al., 2020). Während datengetriebene Ansätze und digitale Zwillinge zunehmend als Schlüsseltechnologien für eine nachhaltige Mobilitätswende diskutiert werden (Weil et al., 2023; Supianto et al., 2024), stützt sich die Praxis häufig auf stark vereinfachte Modelle, ohne verkehrliche Simulations- und Messdaten mit realistischen meteorologischen Bedingungen zusammenzuführen (Trabert et al., 2026 und Weger et al., 2026). Zudem finden durchgängig validierte Systemarchitekturen, die Umweltsensornetze unter realen Betriebsbedingungen automatisiert an operative Verkehrssteuerungen koppeln, bisher kaum Anwendung (Weil et al., 2023). Ungeprüfte Rohdaten oder messnetzbedingte Datenlücken sind mit dem Risiko schwerwiegender Fehlsteuerungen im nachgelagerten Verkehrssystem verbunden.

Der vorliegende Beitrag begegnet diesen Herausforderungen, indem er den Aufbau und die praktische Implementierung eines durchgängigen methodischen Frameworks zur Erfassung, Qualitätssicherung, Imputation und verkehrlichen Integration von Umwelt- und Mobilitätsdaten vorstellt. Das Framework basiert auf den praktischen Erfahrungen des zweijährigen Aufbaus und Betriebs eines urbanen Sensornetzwerks im Rahmen des Forschungsprojekts AIAMO – Artificial Intelligence and MObility, in dem das Helmholtz-Zentrum für Umweltforschung (UFZ) das urbane Sensornetzwerk in der Pilotregion Leipzig aufbaut und betreibt. Die Sicherung einer belastbaren Datenqualität hinsichtlich Vollständigkeit, Konsistenz und Vergleichbarkeit sowie die Kalibrierung der Hardware bilden dabei die zentrale Voraussetzung, um Umweltdaten verlässlich in operative Verkehrssysteme einzuspeisen.

Am konkreten Anwendungsfall der AIAMO Pilotregion Leipzig wird demonstriert, wie die heterogenen Zeitreihen (Luftqualitätsdaten) von knapp 30 stationären Sensoren über offene Schnittstellen kontinuierlich aggregiert, qualitätsgesichert und für zielgerichtete Maßnahmen des umweltsensitiven Verkehrsmanagements operationalisiert werden können.

Durch die Bereitstellung über standardisierte Schnittstellen (OGC, STA-SensorThings API) und die Anbindung an städtische Verkehrsdaten zeigt der Beitrag exemplarisch auf, wie datenbasierte Ansätze von der reinen Messwerterhebung bis hin zur praktischen Umsetzung umweltsensitiver Steuerungskonzepte im urbanen Raum beitragen können. Abbildung 1 demonstriert an dieser Stelle das aufgebaute Umweltmessnetz.

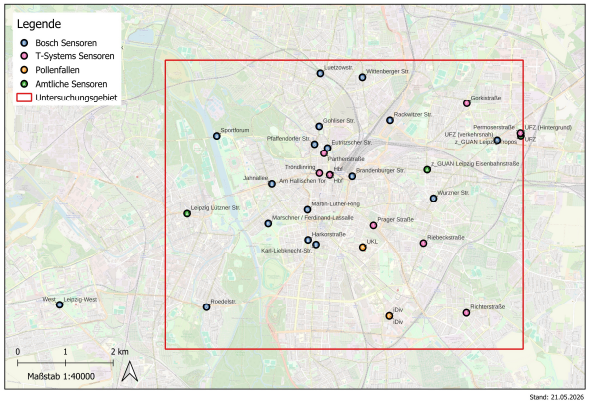


Abbildung 1: Sensornetzwerk in Leipzig (Eigene Darstellung UFZ, 2026)

2 Methodisches Framework

Zur Beschreibung des entwickelten Frameworks (Abbildung 2) gliedert sich der methodische Aufbau in drei Schritte: (1) Identifikation und Schaffung der Datenbasis, (2) Datenmanagement, (3) Datenbereitstellung und -integration in KI-Lösungen.

2.1 Identifikation und Schaffung der Datenbasis

Um eine valide Grundlage für das umweltsensitive Verkehrsmanagement zu schaffen, erfordert das Framework eine synchrone Erfassung von Luftqualitäts- und Verkehrsdaten sowie meteorologischen Parametern. Die Datenbasis setzt sich aus drei komplementären Teilbereichen zusammen.

Den Kern bildet ein installiertes Messnetzwerk aus ca. 30 stationären Luftqualitätssensoren, welche an straßenseitiger Infrastruktur wie Laternenmasten und Trägern des Parkleitsystems montiert wurden.

Die Festlegung der Sensorstandorte erfolgte in enger Abstimmung mit der Kommune anhand verkehrsplanerischer Kriterien. Dabei wurden gezielt Hauptverkehrsstraßen im Stadtzentrum und Verkehrsknotenpunkte, an denen häufige Halte- und Anfahrvorgänge zu ausgeprägten Emissionsspitzen führen, gewählt. Ergänzend wurden Hintergrundmessungen in verkehrsberuhigten Bereichen eingerichtet. Das Sensornetz erfasst primär die Schadstoffe Stickstoffdioxid (NO2) und Feinstaub (PM10 und PM2.5) sowie Ozon, begleitet von Temperatur und relativer Luftfeuchte in minütlicher zeitlicher Auflösung.

Um den bekannten Messunsicherheiten der Sensoren, insbesondere Sensordrifts sowie Querempfindlichkeiten gegenüber Temperatur- und Feuchtigkeitsschwankungen, wirksam zu begegnen, durchlaufen alle Geräte ein Kalibrierungsverfahren. In einem ersten Schritt werden die Sensoren unter kontrollierten Bedingungen an der Referenzmessstation des Helmholtz-Zentrums für Umweltforschung vorab kalibriert, um sensorindividuelle Parameter für den Nullpunkt-Offset und die Sensitivität zu bestimmen. Daran schließt sich eine feldbezogene Co-Location-Kalibrierung im realen Straßenraum an. Feste Re-Kalibrierungszyklen im März 2025 sowie im Sommer 2026 stellen zusammen mit einer bedarfsgesteuerten Wartung bei auffälligen Messabweichungen die langfristige Datenqualität sicher. Dass eine fundierte Kalibrierung für Low-Cost-Sensorik zwingend erforderlich ist, belegen Zhivkov et al. (2025), die zeigen, dass sich das Bestimmtheitsmaß R2 durch Kalibrierungsverfahren von 0,29 auf bis zu 0,95 steigern lässt (Zhivkov et al., 2025).

Ihre verkehrswissenschaftliche Relevanz erhalten diese Luftqualitätsdaten durch die synchrone Verknüpfung mit detaillierten Verkehrsflussinformationen. Hierfür werden Detektordaten aus dem städtischen Straßennetz herangezogen, die über Induktionsschleifen oder Radardetektoren fahrstreifen- und richtungsgetrennte Zähldaten liefern.

Da die Schadstoffausbreitung im Straßenraum maßgeblich durch Witterungsbedingungen bestimmt wird, werden parallel amtliche Wetterdaten eingebunden. Parameter wie Windgeschwindigkeit, Windrichtung, Luftdruck und Niederschlag ermöglichen es, meteorologisch bedingte Konzentrationsschwankungen und Inversionslagen rechnerisch von den rein verkehrsbedingten Effekten zu isolieren.

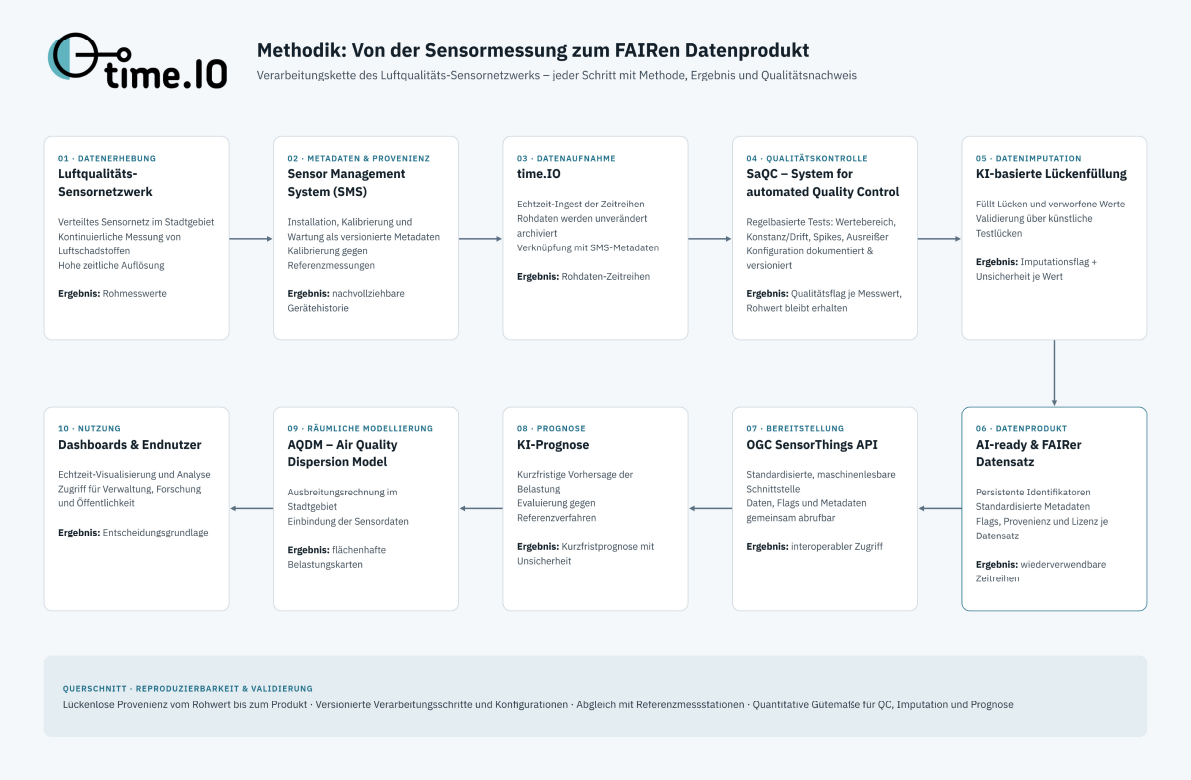


Abbildung 2: Konzeptionelles Framework (eigene Darstellung UFZ, 2025)

2.2 Datenmanagement

Die kontinuierliche Verarbeitung und Bereitstellung der erhobenen Messwerte stützt sich auf eine modular aufgebaute Datenpipeline, die eine automatisierte Übernahme, Qualitätsprüfung und Veredelung der heterogenen Datenströme nach den FAIR-Prinzipien (Findable, Accessible, Interoperable, Reusable) gewährleistet. Die grundlegende Dateninfrastruktur basiert dabei auf dem Framework eines „Digital Ecosystem for FAIR Time Series Data Management in Environmental System Science“ (Bumberger et al., 2024), welches aus drei eng miteinander gekoppelten Kernkomponenten besteht.

Die grundlegende Administration der Messinfrastruktur erfolgt über das Sensor Management System (SMS), welches die Metadaten der eingesetzten Hardware (Sensoren), deren physische Standorte im städtischen Straßennetz sowie historische Änderungen an Konfigurationen und Kalibrierungen im Zeitverlauf dokumentiert (Lorenz et al.2026). Außerdem wird hier die Zuordnung der individuellen Datenströme verwaltet. Durch die Umsetzung internationaler Standards wie OGC, SensorML und der JSON:API-Spezifikation sowie eindeutiger Persistent Identifiers (EUDAT B2INST) gewährleistet das SMS ein einheitliches Metadatenmanagement sowie die nahtlose, interoperable Auffindbarkeit und Integration aller Messstellen in übergeordnete Dateninfrastrukturen (Brinkmann et al. 2024, Bumberger et al. 2024).

Als zentrale Datenintegrationsplattform fungiert das vom Helmholtz-Zentrum für Umweltforschung entwickelte System time.io. Die Plattform übernimmt die Steuerung und Konsolidierung aller Daten, indem sie heterogene Datenströme, darunter die über MQTT und (S-)FTP eingehenden Rohmesswerte sowie vorverarbeitete Referenzdaten des Umweltbundesamtes und Wetterzeitreihen des Deutschen Wetterdienstes, automatisiert aggregiert und persistent speichert. Als technisches Bindeglied synchronisiert time.io die Metadaten aus dem SMS und leitet die Zeitreihen direkt an die nachgelagerte Qualitätskontrolle (Bumberger et al. 2024, Schäfer et al. 2025).

Auf den minütlich eingehenden Zeitreihendaten setzt unmittelbar eine automatisierte Qualitätskontrolle auf, die auf dem Framework SaQC (System for automated Quality Control, Schmidt et al., 2023) basiert. Das System greift auf statistische Verfahren zurück, um Ausreißer und unplausible Signalsprünge in Echtzeit zu detektieren. Anstatt die Rohdatenbestände irreversibel zu überschreiben, markiert das Framework auffällige Messwerte über ein standardisiertes Flagging-Schema (Data Annotation). Auf diese Weise wird eine lückenlose Nachvollziehbarkeit und Reproduzierbarkeit gemäß den FAIR-Prinzipien für alle nachgelagerten Analyse- und Steuerungsprozesse garantiert (Schäfer et al. 2025, Schmidt et al. 2023).

Auf dieser Datenbasis setzt die weiterführende Veredelung auf. Um sensorspezifische Datenlücken zu schließen, durchlaufen die bereinigten Zeitreihen eine nachgelagerte KI-basierte Imputation auf Basis des Foundation Modells TSPulse (Ekambaram et al., 2025). Foundation Modelle für Zeitreihen werden auf großen Datenmengen vortrainiert und lassen sich effizient auf Downstream-Aufgaben adaptieren (Wang et al., 2025). Sie erlernen allgemeine temporale Strukturen wie Saisonalität, Periodizität und Trends (Kottapalli et al., 2025). TSPulse ist auf Aufgaben wie Imputation, Anomalieerkennung und Klassifikation spezialisiert und kombiniert zeitliche mit spektraler Rekonstruktion (Ekambaram et al., 2025). Dies ist beispielsweise für Luftqualitätsdaten relevant, die tägliche, wöchentliche und saisonale Muster aufweisen. Das feinabgestimmte Modell rekonstruiert fehlende Messwerte unter Einbezug zyklischer Zeitvariablen, räumlicher Abhängigkeiten und meteorologischer Parameter. Auf diese Weise entsteht ein lückenloser, verlässlicher und konsistenter AI-ready Datensatz, der als fundierte Datenbasis für weiterführende Verkehrsmodelle fungiert.

Die operative Überwachung des Messnetzes und die Visualisierung der aufbereiteten (qualitätsgesicherten) Daten erfolgen in Echtzeit über webbasierte Dashboards. Diese ermöglichen die Darstellung der Immissionslagen im Straßennetz mittels Karten, die vergleichende Analyse von Zeitreihen einzelner Sensorstandorte sowie die direkte Sichtbarkeit der vergebenen Qualitätsflags. Dass die zeitreihenbasierte Datenhaltung in spezialisierten Datenbanken in Kombination mit Dashboards und KI-Modellen eine hochpräzise Echtzeit-Überwachung sowie die Vorhersage von Grenzwertüberschreitungen im urbanen Raum ermöglicht, demonstrieren Naveed et al. (2025).

Die konsolidierte Datenbasis bildet zudem das Fundament für ein nachgelagertes KI-Modell zur Vorhersage stationsspezifischer Schadstoffimmissionen mit einem Prognosehorizont von 48 Stunden. In der gegenwärtigen Ausbaustufe konzentriert sich das Modell auf den verkehrsrelevanten Leitschadstoff Stickstoffdioxid NO2. Als Trainingsgrundlage dient ein statischer Datensatz (ein Jahr) aus historischen Vor-Ort-Messreihen und meteorologischen Parametern.

2.3 Datenbereitstellung und Anbindung von Umweltdaten

Den abschließenden Schritt des Frameworks bildet die Etablierung eines Datenmanagementsystems, welches die aufbereiteten Umwelt- und Verkehrsdaten bündelt und über die zentrale AIAMO Integrationszone bereitstellt. Über die Integrationszone kann eine Anbindung an das Ausbreitungsmodell Air Quality Dispersion Model (AQDM) erfolgen. Durch die synchrone Bereitstellung von Echtzeit-Messwerten, Wetter- und Verkehrsdaten erlaubt das Modell eine räumlich hochaufgelöste Simulation der Schadstoffverteilung im gesamten Stadtgebiet und schließt damit die Lücke zwischen den verteilten Sensorstandorten und einer flächendeckenden Luftqualitätsbewertung.

3 Das Framework im verkehrlichen Kontext

Die Anwendbarkeit und der verkehrssystemische Mehrwert des vorgestellten konzeptionellen Frameworks werden nachfolgend am Beispiel des Stadtgebiets Leipzig demonstriert. Ziel dieses Kapitels ist es zu zeigen, wie aus der Verknüpfung von Umweltsensorik, Verkehrsdaten und Wetter fundierte Erkenntnisse über verkehrsbedingte Belastungsdynamiken gewonnen und in operative sowie planerische Anwendungsfälle des umweltsensitiven Verkehrsmanagements überführt werden können.

Die Bereitstellung über die AIAMO Integrationszone ermöglicht die Einbindung der Umweltmonitoring-Pipeline in die umweltsensitive Verkehrssteuerung. Während das vorgestellte methodische Framework die messtechnischen und datenseitigen Voraussetzungen schafft, entfaltet das Gesamtsystem seinen praktischen Nutzen mit der konkreten Einbindung in den Stadtverkehr.

3.1 Korrelation von Verkehrsfluss und Schadstoffdynamik

Die kontinuierliche Verknüpfung von Verkehrs- und Luftqualitätsdaten ermöglicht eine differenzierte räumliche Bewertung des Straßennetzes. Insbesondere an stark belasteten Verkehrspunkten treten während der morgendlichen und nachmittäglichen Verkehrsspitzen charakteristische Schadstoffkonzentrationen auf. Die Auswertung der hochaufgelösten Messdaten zeigt deutlich, wie eng Schadstoffbelastung und Verkehrsaufkommen im Tagesverlauf zusammenhängen.

Durch den Abgleich der Immissionsverläufe mit den Geschwindigkeitsdaten und Zählschleifenwerten lässt sich der direkte Einfluss von Stauerscheinungen und Stop-and-Go-Mustern auf die lokale Stickstoffdioxid- und Feinstaubexposition empirisch nachweisen und von Hintergrundeffekten isolieren.

Im Vergleich zu den städtischen Hintergrundmesspunkten, die ein weitgehend homogenes Tagesprofil ohne verkehrstypische Spitzen aufweisen, belegt dieser Kontrast die direkte Kausalität zwischen den gemessenen Immissionsspitzen und dem Verkehrsgeschehen im Straßenkorridor.

3.2 Identifikation von Hotspots

Die Identifikation räumlicher Belastungsschwerpunkte erfolgte in enger Abstimmung mit den Fachämtern der Kommune. Die Auswertung der kombinierten Messreihen offenbart dabei grundlegend verschiedene Typen verkehrsbedingter Hotspots im Stadtgebiet.

Einerseits lassen sich strukturell bedingte Dauerbelastungszonen lokalisieren, wie sie insbesondere im Leipziger Norden auftreten. Dort führen ein kontinuierlich hohes Verkehrsaufkommen und zentrale Straßenknotenpunkte sowie die Autobahnanbindungen zu einer dauerhaft erhöhten Schadstoffgrundlast.

Andererseits können temporäre, ereignisbezogene Belastungsspitzen identifiziert werden (z. B. Großveranstaltungen). An Veranstaltungstagen überlagert der An- und Abreiseverkehr des motorisierten Individualverkehrs das reguläre Verkehrsgeschehen, was kurzzeitig zu lokalen Überlastungen des Straßennetzes führen kann.

Durch den synchronen Abgleich von Geschwindigkeitsabfällen mit den gemessenen Immissionsanstiegen lassen sich gezielt Knotenpunkte im Stadtgebiet lokalisieren, an denen Stauerscheinungen und Stop-and-Go-Zustände die Grenzwerte der EU-Luftqualitätsrichtlinie akut gefährden. Diese differenzierte und evidenzbasierte Erfassung bildet die Voraussetzung, um Steuerungsmaßnahmen räumlich zu adressieren.

3.3 AI-ready Daten für das umweltsensitive Verkehrsmanagement

Der direkte Mehrwert für die Verkehrsplanung und die städtischen Akteure zeigt sich in zielgruppenspezifischen Anwendungen für die kommunale Praxis. Im Rahmen von AIAMO werden die Umweltdaten der Stadt Leipzig für das strategische und operative Mobilitätsmanagement kontinuierlich mit den Verkehrsfluss- und Detektordaten des Deutschen Zentrums für Luft- und Raumfahrt fusioniert und in Form automatisierter Wochenberichte zur Verfügung gestellt. Diese Berichte dienen als evidenzbasierte Entscheidungsgrundlage, um kommunale, verkehrssteuernde Maßnahmen einzuleiten oder verkehrsplanerische Eingriffe an hochbelasteten Achsen zu evaluieren.

Auf operativer Ebene ermöglichen die aktuellen Messwerte und 48-Stunden-NO2-Prognosen die Konzeption umweltsensitiver Navigationsstrategien. Das Ziel besteht darin, den motorisierten Individualverkehr frühzeitig um prognostizierte Hotspots herumzuleiten und den Fuß- sowie Radverkehr über emissionsarme bzw. gemeinwohlkonforme Routen zu führen, bei denen anstelle einer reinen individuellen Fahrzeitminimierung primär die Reduktion von Schadstoffexposition und lokaler Umweltbelastung im Vordergrund steht.

4 Diskussion

Die Ergebnisse aus dem zweijährigen Betrieb des Messnetzes in Leipzig belegen, dass die automatisierte Zusammenführung von Sensorik, amtlichen Wetterdaten und Verkehrsflussparametern eine praxistaugliche Grundlage für ein umweltsensitives Verkehrsmanagement schafft. Insbesondere vor dem Hintergrund der ab 2030 verschärften Grenzwerte aus der EU-Luftqualitätsrichtlinie bietet das konzeptionelle Framework Kommunen die notwendige Datendichte, um drohende Grenzwertüberschreitungen nicht mehr nur reaktiv festzustellen, sondern ihnen durch proaktive Netzbeeinflussung und umweltsensitive Verkehrssteuerung gezielt entgegenzuwirken. Da sich dieses Framework derzeit noch in der Erprobung befindet, werden weitere notwendige Schritte getestet. Das konzeptionelle Framework kann hierfür AI-ready Daten bereitstellen, um ein umweltsensitives Mobilitätsmanagement zu ermöglichen. Dahingehend werden mögliche Empfehlungen für Verkehrsmanagementmaßnahmen bereitgestellt, die sowohl situativ als auch perspektivisch Einfluss auf das Verkehrsgeschehen nehmen können.

5 Fazit und Ausblick

Zusammenfassend zeigt das hier vorgestellte Framework eine mögliche Vorgehensweise zur Bereitstellung qualitätsgeprüfter Daten auf – von der Erfassung und Verknüpfung von Umwelt-, Verkehrs- und Wetterdaten über Qualitätssicherung und KI-basierte Imputation bis zur standardisierten Bereitstellung und Integration in KI-Lösungen für das umweltsensitive Mobilitätsmanagement.

Im Gesamtkontext von AIAMO schafft diese Datenbasis die Voraussetzung, hochaufgelöste Umweltdaten gemeinsam mit Verkehrs- und Wetterdaten in ein umfassendes KI-gestütztes Mobilitätsmanagement zu integrieren und für die kommunale Praxis nutzbar zu machen. Damit erhalten Kommunen eine gemeinsame Datengrundlage für fachbereichsübergreifende Entscheidungen: Verkehrliche Maßnahmen können bereits vor ihrer Umsetzung hinsichtlich ihrer zu erwartenden verkehrlichen und ökologischen Wirkungen bewertet werden. Nach der Umsetzung lassen sich die tatsächlichen Wirkungen auf Basis der integrierten Daten umfassend analysieren und die Maßnahmen darauf aufbauend weiter optimieren.

Zugleich ist zu berücksichtigen, dass Messunsicherheiten und Sensordrifts auch das Training der KI-Anwendungen beeinflussen können. Eine vollumfängliche Validierung dieses Ansatzes ist dahingehend noch ausstehend.

Der strukturierte Aufbau dieses Frameworks zeigt einen Leitfaden zur Bereitstellung und Nutzung dieser Datensätze in verschiedenen Anwendungen. Besonderer Dank gilt der Pilotregion Leipzig, in der das Framework im Rahmen des vom Bundesministerium für Digitales und Staatsmodernisierung (BMDS)

geförderten Projekts AIAMO (Artificial Intelligence and Mobility) erprobt wird.

Literaturverzeichnis

Bumberger, J., Abbrent, M., Brinckmann, N., Hemmen, J., Kunkel, R., Lorenz, C., Lünenschloss, P., Palm, B., Schnicke, T., Schulz, C., van der Schaaf, H., Schäfer, D. (2025). Digital ecosystem for FAIR time series data management in environmental system science. SoftwareX, 29, 102038. https://doi.org/10.1016/j.softx.2025.102038.

Deng, Tianhu, Keren Zhang und Zuo-Jun (Max) Shen (2021). A Systematic Review of a Digital Twin City: A New Pattern of Urban Governance toward Smart Cities. Journal of Management Science and Engineering, 6(2), 125–134. https://doi.org/10.1016/j.jmse.2021.03.003.

Ekambaram, Vijay, Subodh Kumar, Arindam Jati, u. a. (2026). TSPulse: Tiny Pre-Trained Models with Disentangled Representations for Rapid Time-Series Analysis. arXiv:2505.13033. Preprint, arXiv, 4. März 2026. https://doi.org/10.48550/arXiv.2505.13033.

EU-Kommission (2024). EU air quality standards. https://environment.ec.europa.eu/topics/air/air-quality_en?prefLang=de. Abgerufen im Dezember 2024.

Ismagilova, Elvira, Laurie Hughes, Yogesh K. Dwivedi und K. Ravi Raman (2019). Smart Cities: Advances in Research—An Information Systems Perspective. International Journal of Information Management, 47, 88–100. https://doi.org/10.1016/j.ijinfomgt.2019.01.004.

Kaffashzadeh, Najmeh (2023). A Data-Driven Persistence Test for Robust (Probabilistic) Quality Control of Measured Environmental Time Series: Constant Value Episodes. Atmospheric Measurement Techniques, 16(12), 3085–3100. https://doi.org/10.5194/amt-16-3085-2023.

Kottapalli, Siva Rama Krishna, Karthik Hubli, Sandeep Chandrashekhara, u. a. (2025). Foundation Models for Time Series: A Survey. arXiv:2504.04011. Preprint, arXiv, 5. April 2025. https://doi.org/10.48550/arXiv.2504.04011.

Lorenz, C., Brinckmann, N., Bumberger, J., Hanisch, M., Kuhnert, T., Loup, U., Moorthy, R., Obersteiner, F., Schäfer, D. und Schnicke, T. (2026). Sensor management system (SMS): Open-source software for FAIR sensor metadata management in earth system

sciences. SoftwareX, 35, 102858. https://doi.org/10.1016/j.softx.2026.102858.

Minerva, Roberto, Gyu Myoung Lee und Noel Crespi (2020). Digital Twin in the IoT Context: A Survey on Technical Features, Scenarios, and Architectural Models. Proceedings of the IEEE, 108(10), 1785–1824. https://doi.org/10.1109/JPROC.2020.2998530.

Naveed, Khazina, Tariq Umer, Aamer Bilal Asghar, u. a. (2025). Machine Learning Assisted Predictive Urban Digital Twin for Intelligent Monitoring of Air Quality Index for Smart City Environment. Environmental Modelling & Software, 192, 106559. https://doi.org/10.1016/j.envsoft.2025.106559.

Schäfer, David, Martin Abbrent, Florian Gransee, u. a. (2025). time.IO – A Digital ecosystem for FAIR time series data management. Version 2025-09-12. Zenodo, veröffentlicht am 12. September 2025. https://doi.org/10.5281/ZENODO.8354839.

Schmidt, L., Schäfer, D., Geller, J., Lünenschloss, P., Palm, B., Rinke, K., Rebmann, C., Rode, M. und Bumberger, J. (2023). System for automated Quality Control (SaQC) to enable traceable and reproducible data streams in environmental science. Environmental Modelling & Software, 169, 105809. https://doi.org/10.1016/j.envsoft.2023.105809.

Trabert, Thomas, Elmar Brockfeld, Alexander Sohr, u. a. (2026). Urban Digital Twin for Environmentally Sensitive Mobility Planning: Conceptual Framework and Application in Pilot Region Leipzig. International Journal of Applied Earth Observation and Geoinformation, 152, 105477. https://doi.org/10.1016/j.jag.2026.105477.

Wang, Jun, Wenjie Du, Yiyuan Yang, u. a. (2025). Deep Learning for Multivariate Time Series Imputation: A Survey. arXiv:2402.04059. Preprint, arXiv, 20. Mai 2025. https://doi.org/10.48550/arXiv.2402.04059.

Weger, Michael, Thomas Trabert, Timo Houben, u. a. (2026). Hyperlocal Urban NO_2 Hotspot Modeling Driven by Microscopic Traffic Data. npj Climate and Atmospheric Science, 9(1), 150. https://doi.org/10.1038/s41612-026-01471-z.

Weil, Charlotte, Simon Elias Bibri, Régis Longchamp, François Golay und Alexandre Alahi (2023). Urban Digital Twin Challenges: A Systematic Review and Perspectives for Sustainable Smart Cities. Sustainable Cities and Society, 99, 104862. https://doi.org/10.1016/j.scs.2023.104862.

Zhivkov, Peter, Stefka Fidanova und Ivan Dimov (2025). Digital Twin-Based Framework for Real-Time Monitoring and Analysis of Urban Mobile-Source Emissions. Atmosphere, 16(6), 731. https://doi.org/10.3390/atmos16060731.

Autor*innenangaben

Greta Schmalhorst
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: greta.schmalhorst@ufz.de

Timo Houben
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: timo.houben@ufz.de

Ema Vosgerau
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: ema.vosgerau@ufz.de

Florian Berger
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: florian.berger@ufz.de

Dr. Michael Weger
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: michael.weger@ufz.de

Joost Hemmen
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: joost.hemmen@ufz.de

Alex Romanjuta
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: alex.romanjuta@ufz.de

Paul Remmler
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: paul.remmler@ufz.de

David Schäfer
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: david.schaefer@ufz.de

Dr. Jan Bumberger
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig
German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig
Puschstraße 4
04103 Leipzig

E-Mail: jan.bumberger@ufz.de

Dr. Thomas Trabert
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: thomas.trabert@ufz.de