

Bäume im Bild: Wie ein transdisziplinärer Ansatz und KI zur Verbesserung der Datenqualität eines städtischen digitalen Baumkatasters beitragen können

Julia Mayer, Lena Hoff, Johannes Ruf, Martin Memmel, Sascha Henninger

(Julia Mayer, Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH, Trippstadter Straße 122, 67663 Kaiserslautern, DE, julia.mayer@dfki.de)

(Lena Hoff, RPTU Kaiserslautern-Landau, Fachbereich Raum- und Umweltplanung, Physische Geographie, Pfaffenbergstraße 95, 67663 Kaiserslautern, DE, lena.hoff@rptu.de)

(Johannes Ruf, Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH, Trippstadter Straße 122, 67663 Kaiserslautern, DE, johannes.ruf@dfki.de)

(Dr. Martin Memmel, Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH, Trippstadter Straße 122, 67663 Kaiserslautern, DE, martin.mommel@dfki.de)

(Prof. Dr. Sascha Henninger, RPTU Kaiserslautern-Landau, Fachbereich Raum- und Umweltplanung, Physische Geographie, Pfaffenbergstraße 95, 67663 Kaiserslautern, DE, sascha.henninger@rptu.de)

1 ABSTRACT

Städtische Baumkataster werden aus vielfältigen Gründen erstellt und gepflegt: Die Standorte und Arten der Bäume auf öffentlichem Grund spielen bei der Pflegeplanung - vor allem mit Blick auf die zu erwartenden Effekte durch den Klimawandel - eine zunehmende Rolle. Insbesondere der Zustand ist relevant für die Kommune, um ihrer Verkehrssicherungspflicht nachzukommen. Des Weiteren sind Bäume in der Bauplanung wesentlich und nicht zuletzt auch im Naturschutz von hoher Relevanz.

Die Digitalisierung städtischer Baumkataster hat in den letzten Jahrzehnten weitere Vorteile mit sich gebracht: Der Prozess der Datenerfassung und -bearbeitung kann einfacher und effizienter gestaltet werden. Durch die leichtere Zugänglichkeit ist eine bessere Planung, deren Koordination und eine Weitergabe relevanter Daten schneller realisierbar. Dies trägt auch zu einer transparenteren Kommunikation bei. Der Export von Daten ermöglicht zudem Analysen und automatisierte Anwendungen.

Allerdings stellt es die Kommunen vor eine große Herausforderung, das Baumkataster aktuell, vollständig und fehlerfrei zu halten, da es sich um eine sehr große Datenbasis handelt, die sich regelmäßig ändert.

Verschiedene (meist proprietäre) Anwendungen unterstützen die Mitarbeitenden von Kommunen durch intuitiv nutzbare Eingabemasken, um das städtische Baumkataster zu bearbeiten. Es bleibt jedoch immer die Herausforderung, dass die verschiedenen Merkmale einzelner Bäume oder Baumgruppen händisch angelegt oder bearbeitet werden müssen. Dies ist umso aufwändiger je mehr Merkmale ein Baumkataster enthält und birgt die Gefahr, dass durch mitunter mangelhafte fachliche Kenntnisse fehlerhafte Einträge vorgenommen werden. Es gibt zwar technische Hilfsmittel wie Bestimmungs-Apps, die bei nicht-trivialen Merkmalen wie bspw. Höhe herangezogen werden können, aber manche Eigenschaften, wie z.B. das Alter, können oft nur geschätzt werden.

Alle diese Gründe führen dazu, dass die Datenqualität städtischer Baumkataster oft unbefriedigend ist und die Daten nicht ohne Bedenken genutzt werden können. Im Rahmen des Klimawandels werden überdies Merkmale relevant, die bisher nicht erfasst wurden. Beispielsweise können ergänzende Angaben zum Standort (bspw. Baumscheibe oder Park) dazu dienen, einen potenziellen Trockenstress besser vorauszusagen. Eine Klassifizierung des Allergenpotenzials kann darüber hinaus Menschen mit Allergien eine adäquate raumrelevante Information zu „Allergen-Belastungsgebieten“ zur Verfügung stellen.

Um den großen personellen und zeitlichen Aufwand der Pflege des städtischen Baumkatasters in Zukunft zu minimieren und weitere für die Planung und Wissenschaft relevante Merkmale großflächig integrieren zu können, wird in Kaiserslautern ein transdisziplinärer Ansatz erprobt, der auf Grundlage verschiedener Datenbasen und mit Einbezug unterschiedlichster Akteursgruppen arbeitet.

Ausgangslage sind das vorhandene Baumkataster der Stadt, das lückenhaft ist und nur wenige Merkmale enthält, allerdings wegen der präzisen GPS-Standorte eine gute Datenbasis bildet. In Kombination mit Luftbildern kann ein Deep-Learning-Modell trainiert werden, das die fehlenden Bäume erkennt. Ein weiteres KI-Modell, das mit Panoramabildern und LiDAR-Daten aus Straßenbefahrungen arbeitet, soll zudem eine Artenbestimmung ermöglichen. Die Ergebnisse werden von Fachkundigen stichprobenartig evaluiert, sodass eine Feedback-Schleife entsteht, mit der die Modelle sukzessive verbessert werden.

Gleichzeitig wird eine Web-Anwendung entwickelt, die eine nutzerfreundliche Oberfläche bereitstellt, auf der Forschende und interessierte Bürgerinnen und Bürgern weitere Merkmale ergänzen können - auf diesen

Daten können weitere Modelle trainiert werden. Andere Merkmale, wie etwa Höhe des Baumes, können automatisiert aus den vorhandenen Datensätzen extrahiert werden. Im Rahmen des Prozesses sollen unterschiedliche Datenbasen unter Berücksichtigung der rechtlichen Rahmenbedingungen, der Kosten und der Verfügbarkeit als Trainingsdaten dienen und eine Aussage über die Qualität der jeweiligen Ergebnisse getroffen werden. Des Weiteren ist es ein erklärtes Ziel, die KI-Modelle übertragbar zu gestalten, so, dass sie in Regionen mit ähnlichem Baumbestand direkt angewendet werden können. Eine weitere Anwendung kann die Aktualisierung des Baumkatasters und das Monitoring der Datenqualität desselben sein, sobald neue Daten zur Verfügung stehen.

Keywords: artificial intelligence, machine learning, künstliche Intelligenz, Datenqualität, Baumkataster

2 EINLEITUNG

Bäume sind für die Aufenthaltsqualität im öffentlichen sowie privaten Raum sehr wichtig. Sie beeinflussen das lokale Klima und die Luftqualität, werten Straßen, Plätze und Parks optisch auf und bieten zudem einen Lebensraum für Tiere.

Im Zuge der Verkehrssicherungspflicht, aber auch im Hinblick auf die Baumkontrollen, die Bauplanung und den Naturschutz, ist es für Kommunen von hoher Relevanz, die Lage und den Zustand der städtischen Bäume auf öffentlichem Grund zu kennen und zu erfassen. Hierbei spielen Baumkataster eine große Rolle. Dort werden Informationen zu den Bäumen gesammelt und gespeichert. Dazu gehören neben der Position des jeweiligen Baumes die Art und/oder die Gattung, die Höhe, der Kronendurchmesser, der Stammumfang sowie, falls bekannt, das Pflanzjahr und das Alter. Das Baumkataster stellt somit ein Hilfsmittel für die Baum- und Grünflächenpflege dar. Jedoch ist der Aufbau, die Ergänzung weiterer Merkmale und die Pflege eines solchen Katasters eine aufwändige Aufgabe, die viele Ressourcen beansprucht. Die Digitalisierung städtischer Baumkataster hat in den letzten Jahrzehnten vielerlei Veränderungen mit sich gebracht: Der Prozess der Datenerfassung und -bearbeitung kann einfacher und effizienter gestaltet werden. Durch die leichtere Zugänglichkeit ist eine bessere Planung, deren Koordination und eine Weitergabe relevanter Daten schneller realisierbar. Dies trägt auch zu einer transparenteren Kommunikation bei. Der Export von Daten ermöglicht zudem Analysen und automatisierte Anwendungen.

Jedoch haben nicht alle Kommunen digitale Baumkataster beziehungsweise sind diese oft unvollständig oder nicht aktuell. Weiterhin sind in den kommunalen Baumkatastern häufig nur die Bäume erfasst, die im öffentlichen oder halböffentlichen Raum stehen und in die Sorgfaltspflicht der Kommune fallen. Allerdings sind auch private Bäume für Aussagen über die Lufthygiene relevant und sollten für Anwendungsfälle in diesem Bereich aufgeführt sein [vgl. Albert et al. 2022]. Im Rahmen des Klimawandels werden überdies Merkmale von Bedeutung, die bisher nicht dokumentiert wurden. Unter anderem können ergänzende Angaben zum Standort (z.B. Baumscheibe oder Park) dazu dienen, einen potenziellen Trockenstress besser vorauszusagen. Eine Klassifizierung des Allergenpotenzials kann darüber hinaus Personen mit Allergien eine adäquate raumrelevante Information zu „Allergen-Belastungsgebieten“ zur Verfügung stellen.

Anhand dieser Problemfelder entstand die Idee, die Datenqualität städtischer Baumkataster zu verbessern. Um den großen personellen und zeitlichen Aufwand der Pflege des städtischen Baumkatasters in Zukunft zu minimieren und weitere, für die Planung und Wissenschaft relevante Merkmale, großflächig integrieren zu können, wird in Kaiserslautern ein transdisziplinärer Ansatz erprobt, der auf Grundlage verschiedener Datenbasen und mit Einbezug unterschiedlicher Akteursgruppen arbeitet. So soll es ermöglicht werden, dass die vorhandenen, aber unvollständigen Daten ergänzt und weitere Merkmale wie Baumart, Höhe und Kronendurchmesser sukzessive hinzugefügt werden.

Im Rahmen des Prozesses sollen unterschiedliche Datenbasen unter Berücksichtigung der rechtlichen Rahmenbedingungen, der Kosten und der Verfügbarkeit als Trainingsdaten dienen und eine Aussage über die Qualität der jeweiligen Ergebnisse getroffen werden. Des Weiteren ist es ein erklärtes Ziel, die KI-Modelle übertragbar zu gestalten, sodass sie in Regionen mit ähnlichem Baumbestand direkt angewendet werden können. Interessierten Bürgerinnen und Bürgern soll es durch die Entwicklung einer nutzerfreundlichen Web-Anwendung ermöglicht werden sich an der Evaluation der Ergebnisse zu beteiligen und ergänzende Merkmale zu erfassen. Sobald neue Daten zur Verfügung stehen, kann ein weiterer Einsatz darin bestehen, das Baumkatasters regelmäßig zu aktualisieren und die Qualität der Daten zu überwachen.

Die Vorteile eines solchen Vorgehens sind vielfältig. Durch Nutzung von Luftbildern können auch private Bäume erfasst werden, die für die Öffentlichkeit nicht ohne Weiteres zu sehen und häufig nicht kartiert sind. Dies ist u.a. für die potenzielle Bewertung von Bereichen mit hohen Allergiepotezialen durch (Baum-)Pollen sowie das Ozonbildungspotenzial durch biogene Kohlenwasserstoffemissionen von Interesse [Albert et al. 2022].

Es gibt einige Initiativen in deutschen Städten und Regionen, die sich bereits für eine Verbesserung des Baumkatasters einsetzen. Unter anderem in der Metropole Ruhr werden Baumstandorte auf Grundlage ähnlicher Daten und unter Anwendung von KI detektiert [Metz et al. 2023]. Da die Modelle Open-Source zur Verfügung gestellt werden sollen, könnten diese in den hier verfolgten Ansatz integriert werden. Es gibt des Weiteren bereits Deep-Learning-Modelle, die Bäume detektieren, bspw. Ventura et al. [2022] und Weinstein et al. [2019], auf die im nächsten Abschnitt näher eingegangen wird.

Ein erfolgreicher Anwendungsfall eines gut gepflegten Baumkatasters, kann der Einbezug der Bevölkerung bei der Bewässerung sein. Unter anderem in Berlin [Technologiestiftung Berlin] und Leipzig [OKF 2023] werden Plattformen genutzt, um das Gießen von Stadtbäumen zu koordinieren und nachzuvollziehen. In Magdeburg [Winter 2023] finden sich darüber hinaus noch Informationen zu gefälltten Bäumen.

Im Folgenden werden zunächst der Projektkontext und die beteiligten Akteurinnen und Akteure vorgestellt. Die Vorgehensweise wird im Allgemeinen erklärt und die vorhandenen Datengrundlagen erklärt und evaluiert. Es wird Bezug auf bereits existierende KI-Modelle zur Positionsbestimmung und Eignung in diesem Anwendungsfall genommen. Zusätzliche mögliche Merkmale werden priorisiert und Möglichkeiten bzw. Grenzen ihrer Bestimmung aufgezeigt. Zuletzt werden die Maßnahmen zur Einbeziehung der Zivilbevölkerung dargelegt.

3 ANSATZ UND PROJEKTKONTEXT

Um das Vorhaben erfolgreich umzusetzen, bedarf es einer Zusammenarbeit verschiedener Akteursgruppen, die ihre jeweilige Expertise zur Verfügung stellen: es ist ein breites Fachwissen zu den Themengebieten Umwelt und Klima, Stadtplanung, Datenverarbeitung, Web-Design und KI gefragt, das durch die Forschenden abgedeckt wird. Mitarbeitende der öffentlichen Verwaltung kennen die vorhandenen Daten, stellen diese bereit und liefern Use-Cases aus der Praxis. Des Weiteren bestehen dort bereits Kommunikationskanäle, um mit der Zivilbevölkerung in Kontakt zu treten und diese für Crowdsourcing-Aktionen zu gewinnen. Hierbei sollen an Aktionstagen interessierte Bürgerinnen und Bürger mit Unterstützung der Forschenden Merkmale von Bäumen möglichst großflächig digital erfassen bzw. die Ergebnisse aus den KI-Modellen validieren.

Das Forschungsvorhaben ist aus dem durch die Carl-Zeiss-Stiftung geförderten Verbundprojekt „Ageing Smart – Räume intelligent gestalten“ entstanden. Im Fokus des Gesamtprojektes stehen die Babyboomer, also die geburtenstarken Jahrgänge von 1955 bis 1969, durch deren Eintritt in das Rentenalter Kommunen vor Herausforderungen gestellt werden. So sind diese oftmals damit konfrontiert, neben altersgerechten Wohnstandorten auch passende Versorgungs- und Freizeitstrukturen zu schaffen. Das Gesamtprojekt hat sich als Ziel gesetzt, ein datengestütztes Entscheidungsunterstützungssystem (Decision Support System) zu entwickeln, das Mitarbeitenden der öffentlichen Verwaltungen in ebendiesen Planungsprozessen unterstützt. Insgesamt zehn Teilprojekte aus den Fachbereichen Raum- und Umweltplanung, Informatik und Mathematik der RPTU am Standort Kaiserslautern bearbeiten interdisziplinär diese Fragestellung. Ebenso involviert ist das Fraunhofer-Institut für Experimentelles Software Engineering (IESE) sowie das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH (siehe dazu Abb.1).

Das Teilprojekt der Physischen Geographie untersucht bestehende, lokale, öffentliche Grünflächen im Hinblick auf die Frage, ob diese relevanten Rückzugsräume von allen Personengruppen gleichermaßen nutzbar sind. Ein besonderer Blick liegt hierbei auf den meteorologischen Rahmenbedingungen sowie lokalklimatischen und lufthygienischen Fragestellungen. Im Rahmen der Weiterentwicklung des Baumkatasters liefert das Teilgebiet mögliche Use-Cases und bietet die fachliche Expertise. Im weiteren Verlauf des Teilprojektes soll das methodische Vorgehen, das bereits im Jahr 2022 von Albert et al. vorgestellt wurde, auf Siedlungsräume übertragen werden. Hierbei ist ein (nahezu) vollständiges digitales Baumkataster erforderlich.

Bäume im Bild: Wie ein transdisziplinärer Ansatz und KI zur Verbesserung der Datenqualität eines städtischen digitalen Baumkatasters beitragen können

Das Team des SmartCity Living Lab des DFKI befasst sich im Rahmen des Forschungsprojekts, dessen Erkenntnisse und Umsetzungen auf vielfältigen Datenquellen fußt, mit der Data Governance und dem Daten-Management, um eine ausreichende Datenqualität zu gewährleisten. Diese ist ausschlaggebend für die Güte automatisierter Anwendungen und Methoden der Künstlichen Intelligenz. KI kann umgekehrt aber auch wie in diesem Fall genutzt werden, um eine gewünschte Datengrundlage zu schaffen oder zu verbessern.

Die Stadt Kaiserslautern wurde als eine von sieben Modellkommunen im Projekt ausgewählt und arbeitet zudem im Rahmen der MPSC-Förderung (Modellprojekte Smart Cities) des BMWSB schon seit einigen Jahren eng mit Forschenden des DFKI und der RPTU zusammen. Sie stellt Daten zur Verfügung und nutzt ihre Kommunikationskanäle mit der Zivilbevölkerung, um diese aktiv miteinzubeziehen.

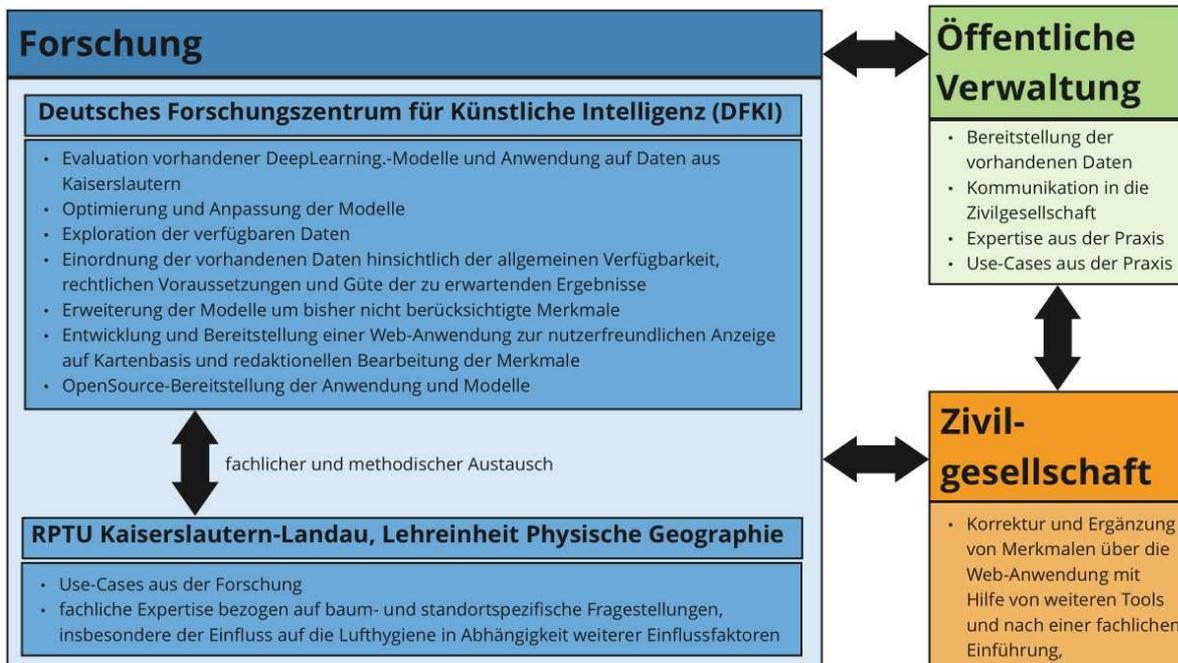


Abb. 1: Kooperationsnetzwerk zum Verbessern, Ergänzen und Nutzen des städtischen Baumkatasters

4 VORGEHENSWEISE

Zunächst werden die verfügbaren Daten auf ihre Qualität und Nutzbarkeit in diesem Anwendungsfall hin untersucht. Sie dienen als Grundlage, um das aktuelle Baumkataster mit den Merkmalen Position und Spezies unter Anwendung von vorhandenen und ggf. angepassten Deep-Learning-Modellen, um fehlende Positionen von Bäumen zu ergänzen. In einem weiteren Schritt wird ein weiteres Modell angewendet oder ggf. selbst trainiert, das die Gattung der neu detektierten Bäume ergänzen soll. Um die Qualität der Ergebnisse einordnen zu können, ist hier eine Kontrolle von fachkundigen Personen nötig. So können die Modelle sukzessive verbessert werden.

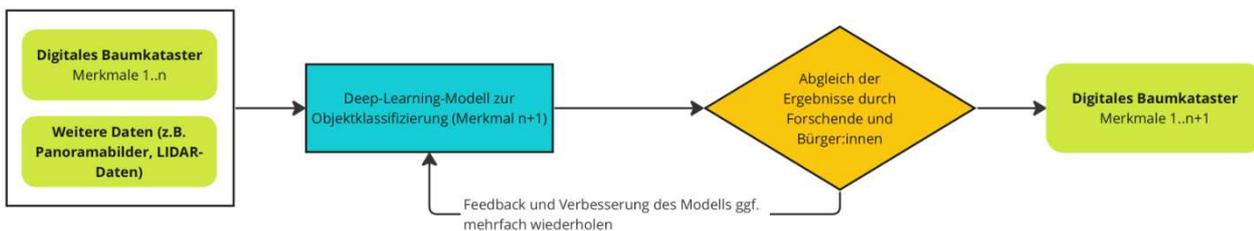


Abb. 2: Generalisierte Vorgehensweise zur Ergänzung des Digitalen Baumkatasters unter Einbezug der Bevölkerung

Des Weiteren soll das Baumkataster um weitere Merkmale, wie bspw. Höhe, Stammdurchmesser oder Standort ergänzt werden. Hierzu werden drei Ansätze verfolgt: Zum Ersten können vorhandene Deep-Learning-Modelle auf verfügbare Daten angewendet werden und wiederum durch eine Feedback-Schleife optimiert werden (vgl. Abb.2). Zum Zweiten können automatisierte Anwendungen (auch ohne Nutzung von KI) bestimmte Merkmale großflächig ergänzen. Zum Dritten gibt es die Option, ein eigenes Modell zu

trainieren. Hierfür werden ausreichend Trainingsdaten benötigt, die in einem Crowdsourcing-Ansatz gesammelt werden sollen. Interessierte Bürgerinnen und Bürger erfassen mit Unterstützung von Fachkundigen und unter Benutzung einer anwenderfreundlichen und intuitiven Software neue Merkmale von einigen Bäumen. So soll der Datensatz im Ganzen um die jeweiligen Attribute ergänzt werden.

4.1 Einordnung der Datengrundlagen

Es gibt einige Datenquellen, die Informationen zu den städtischen Bäumen in Kaiserslautern bereithalten. Diese müssen zunächst auf ihre Qualität geprüft werden. Dabei werden die verschiedenen Kriterien je nach Anwendungsfall unterschiedlich gewichtet und bewertet. In diesem Fall spielen insbesondere die Fehlerfreiheit, die Zugänglichkeit und die Verfügbarkeit der Daten eine übergeordnete Rolle. Unvollständige, aber verlässliche Daten können genutzt werden, um den Datensatz mithilfe von automatisierten Anwendungen zu vervollständigen.

Um Datenquellen als Trainingsdaten für KI-Methoden nutzen zu können, sollten diese möglichst fehlerfrei sein, da das Modell sonst eventuell diese Fehler reproduziert. Des Weiteren ist es sinnvoll, Datensätze zu nutzen, die den Kommunen ohnehin vorliegen oder ohne großen Aufwand zu beschaffen oder erzeugen sind, um ein Modell perspektivisch auch auf andere Kommunen anwenden zu können.

Die Aktualität der Daten ist in diesem Use-Case zunächst untergeordnet, da im Verhältnis zur Gesamtanzahl der Bäume nur ein kleiner Anteil gefällt bzw. neu gepflanzt wird. Andere Merkmale, wie die Höhe, ändern sich kontinuierlich. Hier muss generell mit Ungenauigkeiten gerechnet werden und diese mit angemessenen Methoden quantifiziert werden. Dies ist insbesondere in Use-Cases, die ebenjene Merkmale verwenden, zu berücksichtigen.

4.1.1 Das Baumkataster der Stadt Kaiserslautern

Das Baumkataster der Stadt Kaiserslautern enthält sehr präzise GPS-Koordinaten von ca. 16.600 Bäumen und der jeweiligen Bezeichnung der Gattung bzw. Art. Eine stichprobenartige Überprüfung der Daten durch Ortsbegehungen hat jedoch ergeben, dass viele Bäume hier noch nicht erfasst sind (s. Abb. 3). Es ist zudem entweder die Gattung oder die Art angegeben, wobei erstere aus letzterer hergeleitet werden könnte, umgekehrt aber nicht. Für die meisten Anwendungsfälle ist die Gattung als Information allerdings ausreichend.

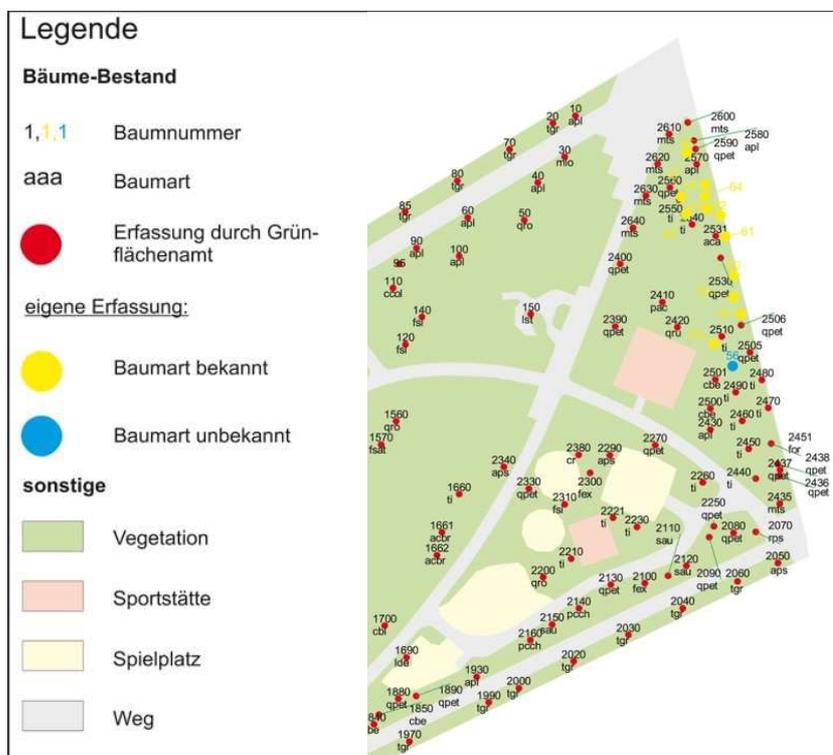


Abb. 3: Karte mit vorhandenen und fehlenden Bäumen im Stadtpark von Kaiserslautern (rot=durch Grünflächenamt erfasst, gelb & blau = eigene Erfassung) [Henninger 2014]

4.1.2 OpenStreetMap

In der OpenStreetMap-Community gibt es das Bestreben, auch Bäume und Baumflächen mit diversen Merkmalen zu kartieren. Allerdings sind zumindest in der Stadt Kaiserslautern bisher wenige Eintragungen (nur rund 1000) vorhanden. Dazu kommt, dass mit dem Smartphone erfasste GPS-Standorte für bestimmte Anwendungsfälle zu ungenau sein können. Es ist darüber hinaus unklar, wie gut die Daten sind, die bisher erfasst wurden. In OpenStreetMap (OSM) sind einige Attribute vorgesehen, an denen man sich orientieren kann, um perspektivisch die neu generierten Daten in die Plattform zu integrieren. Erwähnenswert ist darüber hinaus jedoch, dass Merkmale enthalten sind, die ein anderes implizieren. Beispielsweise kann man durch Angabe der Gattung herleiten, ob es sich um einen Laub- oder Nadelbaum handelt. Umgekehrt ist das nicht möglich. Allerdings ist dieses Merkmal leicht von Personen zu erfassen, die weder Wissen über Baumarten noch technische Hilfsmittel zu deren Bestimmung haben. Es gibt auch eine Unterscheidung der Attribute `height` (Höhe) und `est_height` (geschätzte Höhe), die verhindern kann, dass falsch oder zu grob geschätzte Angaben als korrekt interpretiert werden [OSM Contributors 2023].

4.1.3 Luftbilder

Der Stadtverwaltung Kaiserslautern liegen Luftbilder aus Befliegungen zu unterschiedlichen Jahreszeiten und mit verschiedenen Auflösungen vor. Hier wird zunächst mit den Bildern der Frühjahrsbefliegung aus dem Jahr 2018 gearbeitet. Diese sind in einer Auflösung von 40cm frei über das Geoportal RLP [GDI-RP] verfügbar. Höher aufgelöste Luftbilder mit einer Bodenauflösung von 8cm sind aus Gründen des Datenschutzes nicht frei verfügbar, werden aber von der Stadtverwaltung zu Forschungszwecken zur Verfügung gestellt. Es soll überprüft werden, ob eine bessere Auflösung zwangsläufig auch eine bessere Kategorisierung nach sich zieht. Falls die Vorteile einer höheren Auflösung nicht signifikant sind, kann man die datenschutzrechtlich unbedenklichen und frei verfügbaren Daten nutzen. Dies macht eine Übertragung der Ergebnisse auf andere Kommunen und die Anwendung der Modelle auf neuere Luftbilder wesentlich unkomplizierter.

Des Weiteren wird auch ein Abgleich zu Befliegungen aus anderen Jahreszeiten und daher mit anderer Belaubung stattfinden. Es ist zu erwarten, dass je nach Use-Case unterschiedliche Jahreszeiten zu bevorzugen sind. Soll der Standort (Baumscheibe, Baumreihe oder Grünfläche) bestimmt werden, liefern die Luftbilder aus Jahreszeiten ohne Belaubung mutmaßlich bessere Ergebnisse.

4.1.4 Daten aus Straßenbefahrungen

Im Herbst 2021 hat die Stadt Kaiserslautern Befahrungen im gesamten Stadtgebiet ausführen lassen, bei der Panoramabilder und LiDaR-Daten erfasst wurden. Unter deren Nutzung können präzise Messungen (Höhe, Kronendurchmesser und Stammdurchmesser) am PC durchgeführt werden.

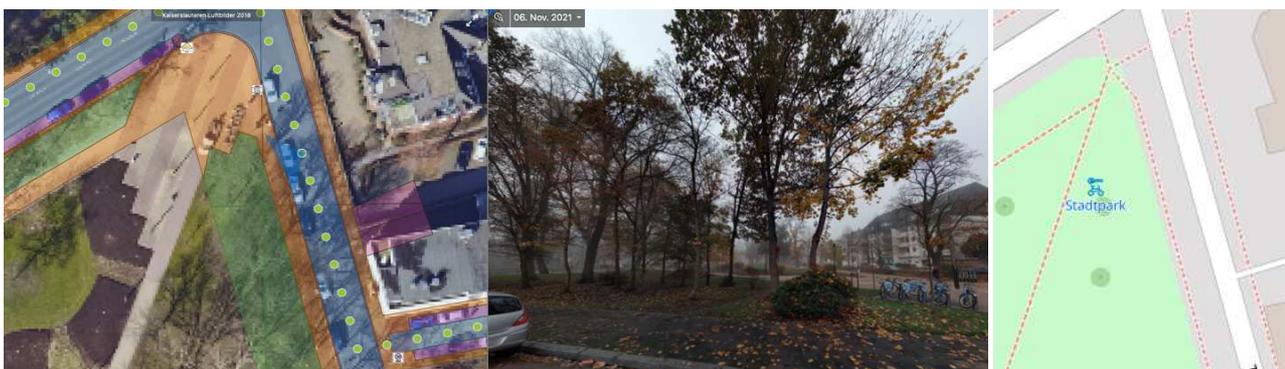


Abb. 4: Screenshots aus Cyclomedia (links) und OSM (rechts) auf denselben Bereich des Stadtparks von Kaiserslautern in Abb. 3

Ein direkter Vergleich der Bäume, die in OSM dargestellt werden, mit dem Bildausschnitt der Straßenbefahrung auf den entsprechenden Teil des Stadtparks, zeigt sofort, dass in OSM bisher nur ein Bruchteil der Bäume verzeichnet ist (vgl. Abb. 4).

4.2 Auswahl, Anwendung und Anpassung von KI-Methoden zur Bestimmung von Position und Spezies

Es gibt bereits einige Ansätze, die sich mit dem Kategorisieren von Bäumen beschäftigen. Es ist also nahe liegend, diese zu nutzen und darauf aufzubauen. Die Übertragbarkeit auf Kaiserslautern hängt unter anderem davon ab, mit welchen Daten die bestehenden Modelle trainiert wurden. Wurden Trainingsdaten aus Erdteilen mit einer im Vergleich sehr abweichenden Vegetation genutzt, ist es fraglich, ob das entsprechende Modell mit den hiesigen Bildern sinnvolle Ergebnisse liefert. Zudem muss die Jahreszeit berücksichtigt werden.

4.2.1 Vorhandene Modelle zur Positionsbestimmung von Bäumen

Es gibt bereits Modelle, die Bäume detektieren. Erste Anwendungen auf Kaiserslautern haben ergeben, dass die Ergebnisse noch nicht zufrieden stellend bzw. nutzbar sind. Im Rahmen eines Masterprojekts mit dem Titel „Improving the Tree Cadastre of KL using Object Classification (Python)“ wird eines der Modelle angepasst und verbessert.

Ein Convolutional Neural Network (CNN) wurde auf Luftbildern aus Kalifornien trainiert und liefert so genannte Bounding Boxes, in denen Bäume stehen. Es detektierte in den Testdaten 73,3% der vorhandenen Bäume und 73,6% der detektierten Bäume waren tatsächlich existent [Ventura et al. 2022]. Diese Genauigkeit ist jedoch nicht ausreichend. Hinzu kommt, dass dem Modell vier Kanäle (RGBN) zur Verfügung standen, bei den Luftbildern in Kaiserslautern aber nur drei (RGB) vorliegen.

Ein weiteres vielversprechendes Modell wurde auf Bäume außerhalb von Städten trainiert [Weinstein et al. 2019] und dann mit Stadtbäumen in München feintrainiert.



Abb. 5: Ergebnisse der Modelle von Ventura et al. [2022] und Weinstein et al. [2019] angewendet auf die Luftbilder aus Kaiserslautern

4.2.2 Herausforderungen bei der Positionsbestimmung

Das Ziel der KI-Methoden besteht darin, eine hohe Präzision und Abdeckung zu erreichen. Das bedeutet, dass eine möglichst große Anzahl von Bäumen korrekt erkannt werden soll und dabei gleichzeitig möglichst wenige Fehler gemacht werden sollten. Es sollten also keine Bäume an Stellen gefunden werden, an denen tatsächlich keine stehen (False Positives, FP). Dies kann durch fehlerhafte Klassifizierung eines anderen Objekts oder durch ein Rauschen in den Daten verursacht werden. Umgekehrt sollten keine Bäume übersehen werden. Solche False Negatives (FN) treten auf, wenn die Bildqualität schlecht ist oder ungünstige Umgebungsbedingungen herrschen.

Da im städtischen Baumkataster schon viele Bäume fehlen, ist der Fokus auf eine hohe Abdeckung wünschenswert. Fälschlicherweise erkannte Bäume lassen sich schnell aus dem Datensatz entfernen, wohingegen nicht erkannte Bäume in Anbetracht der Angabe präziser GPS-Koordinaten etwas aufwändiger hinzuzufügen sind.

Da die vorhandenen Modelle Schwierigkeiten mit Baumflächen haben und es bereits Ansätze gibt, für Wälder auf Grundlage von Luftbildern das Verhältnis unterschiedlicher Baumarten zu bestimmen, wird diese Thematik zunächst hintenangestellt. Dennoch werden bei einigen Use-Cases die Baumgruppen und deren

Bäume im Bild: Wie ein transdisziplinärer Ansatz und KI zur Verbesserung der Datenqualität eines städtischen digitalen Baumkatasters beitragen können

Arten eine zentrale Rolle spielen. Daher muss langfristig auch für Baumgruppen eine Lösung gefunden werden.

In städtischen Baumkatastern sind häufig aus pragmatischen Gründen nur Bäume auf öffentlichem Grund verzeichnet, für die die Kommune auch eine Sicherungspflicht inne hat. Unter Anwendung von KI werden jedoch alle Bäume auf den gegebenen Bildern erkannt, also auch die auf halböffentlichem und privatem Grund. Diese sind insbesondere relevant für Anwendungsfälle, die sich mit Lufthygiene beschäftigen [vgl. Albert et al. 2022].

4.2.3 Artbestimmung von Stadtbäumen mit KI

Wesentlich komplexer ist die Bestimmung der Spezies der Bäume, die neu im Kataster ergänzt wurden. Hierfür werden zwei Ansätze verfolgt, die auf Grundlage des ursprünglichen Katasters fußen. Es soll erprobt werden, wie gut ein Modell mit hochaufgelösten Luftbildern die Gattung kategorisieren kann. Da hierfür erwartungsgemäß zu wenige Trainingsdaten vorliegen, wird ein zweiter Ansatz verfolgt, der die Panoramabilder nutzt. Diese und bereits existierende Modelle zur Baumbestimmung werden genutzt, um die fehlenden Angaben zur Spezies zu ergänzen.

4.3 Ergänzung weiterer Merkmale

Zusätzlich zu den Merkmalen Spezies und Position sind weitere Merkmale relevant, die soweit möglich, großflächig ergänzt werden sollen. Eine Priorisierung erfolgt in Abhängigkeit von den spezifischen Anwendungsfällen. Im Projektkontext haben Standort und Höhe eine hohe Priorität (vgl. Tabelle 1). Eine Orientierung geben unter anderen die in OSM verwendeten Merkmale.

Es ist unbedingt zu beachten, dass alle Attribute veränderlich sind. Die Höhe ändert sich stetig und es kommt bisweilen auch vor, dass ein Baum gefällt und an derselben Stelle ein Baum einer anderen Art gepflanzt wird. Das Baumkataster muss also regelmäßig aktualisiert oder mit entsprechenden Hinweisen versehen werden.

In OSM wird zusätzlich die Bedeutung eines Baumes erfasst. Diese umfasst zusätzliche Informationen z.B. ob ein Baum ein Naturdenkmal ist, eine besondere Landmarke darstellt oder historische Relevanz besitzt. Solche Attribute werden hier nicht berücksichtigt, da sie im lokalklimatischen Kontext keine Rolle spielen.

4.4 Automatisierte Ergänzungen

Messbare Merkmale wie Höhe, Stammdurchmesser und Kronendurchmesser können für Bäume, deren Position verfügbar ist, automatisiert aus den LiDaR-Daten errechnet werden (vgl. Abb. 6), die bei Straßenbefahrungen im Herbst 2021 erfasst wurden.

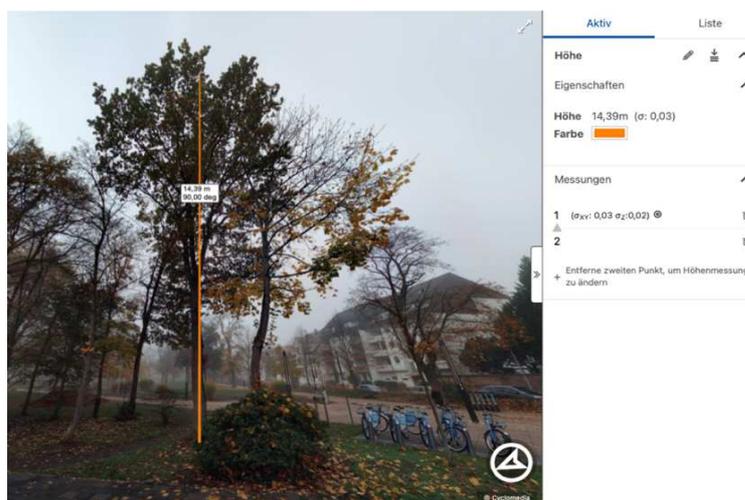


Abb. 6: Screenshot aus Cyclomedia mit der beispielhaften Messung der Höhe eines beliebigen Baumes

4.5 Bilden weiterer Deep-Learning-Modelle

Es sollen weitere Deep-Learning-Modelle gebildet werden, die eine Aussage über den Standort des Baumes treffen können: Steht dieser in einer Baumscheibe, einem Grünstreifen oder einer Grünanlage? Dies hat erheblichen Einfluss auf den Wasserbedarf des jeweiligen Baumes. In Baumscheiben sind arttypische

Wurzelausprägungen bis in die feuchteren Erdschichten hinein kaum möglich und werden durch Schutt im Untergrund sowie Medienträgern wie Rohre, Kabel und Leitungen zusätzlich erschwert.

		Merkmal	Beschreibung	Bemerkung	Priorität
In OSM	1	position	gegeben durch GPS-Koordinaten (Latitude, Longitude); in OSM auch Punkt oder Knoten	Die Präzision der Koordinaten ist für die Anwendung von KI relevant.	sehr hoch
	2	genus	Art; Sammelbegriff	kann aus (3) hergeleitet werden; in den meisten Use-Cases ausreichend	sehr hoch
	3	species	Gattung; Wissenschaftlicher Name	präziser, aber schwieriger zu bestimmen als (2)	mittel
	4	leaf_type	Laubbaum oder Nadelbaum	kann aus (2) hergeleitet werden	niedrig
	5	leaf_cycle	Beschreibt die Eigenschaft das Blattwerk abzuwerfen	kann aus (2) hergeleitet werden	niedrig
	6	est_height	geschätzte Höhe in Metern	falls keine genaue Angabe möglich	niedrig
	7	height	Höhe in Metern	relevant für Verschattung	hoch
	8	diameter_crown	Kronendurchmesser in Metern	relevant für Verschattung	mittel
	9	circumference	Stammumfang in Metern	kann aus (10) berechnet werden	niedrig
Weitere	10	diameter_trunk	Stammdurchmesser in Metern	kann aus (9) berechnet werden	niedrig
	11	standort	Umgebung des Baumes (Baumscheibe, Baumstreifen, Grünanlage)	relevant für Aussage über möglichen Trockenstress	hoch
	12	zustand	Gesundheitszustand des Baumes	kann nicht großflächig berechnet werden, da Trainingsdaten fehlen	keine
	13	sicherungsbedarf	Notwendigkeit Vorkehrungen zu treffen	kann nicht großflächig berechnet werden, da Trainingsdaten fehlen; unklare Rechtslage im Anwendungsfall	keine
	14	alter	Alter des Baumes	kann ggf. durch (2) und (7) geschätzt werden	niedrig
	15	standalter	Zeitpunkt der Pflanzung des Baumes am jetzigen Standort	zusätzliche Datenquellen sind nötig	keine
	16	besitz	Privatbesitz oder im Besitz der öffentlichen Hand	eine Verschneidung mit dem Liegenschaftskataster ist notwendig; relevant für die Sicherungspflicht	keine

Tabelle 1: Überblick Merkmale und ihre Priorisierung im Projektkontext

Die Erfassung dieses Merkmals kann also genutzt werden, um möglichen Trockenstress vorauszusagen und somit helfen, Priorisierungen beim Gießen von Stadtbäumen festzulegen.

Es gibt Merkmale, die in gut gepflegten Baumkatastern auftauchen und nicht leicht automatisiert auf Grundlage der bisher vorhandenen Datenquellen zu ergänzen sind. Dazu gehören Baumalter und Pflanzzeitpunkt. Ersteres könnte man anhand der Höhe und der Gattung grob abschätzen oder ein Intervall angeben. Zum Pflanzzeitpunkt könnten zusätzliche Datenquellen (z.B. Bebauungspläne) weitere Informationen bieten, wobei hierbei nur abgeschätzt werden kann, wann ein Baum frühestens gepflanzt wurde. Daher wird dieses Merkmal nicht berücksichtigt.

4.6 Einbezug der Zivilbevölkerung

Parallel zu den ersten Erprobungen und Verbesserungen der KI-Methoden wird eine nutzerfreundliche und intuitive Web-Anwendung entwickelt, die eine Einbeziehung interessierter Bürgerinnen und Bürger ermöglichen soll. Hierfür sollen diese unkompliziert weitere Merkmale ergänzen oder vorhandene editieren können. Das Nutzen weiterer Hilfsmittel wie Baumarten-Bestimmungsapps oder Messwerkzeuge bietet sich an. Fachkundige Personen sollen regelmäßig zur Einführung und Beratung zur Verfügung stehen.

In der Web-Anwendung werden ähnlich wie im Gieß-den-Kiez-Portal [Technologiestiftung Berlin] in einer Kartenansicht die vorhandenen Bäume angezeigt. Bei Auswahl einer dieser Punkte durch Anklicken öffnet sich ein Fenster mit zusätzlichen Informationen. Zudem soll es die Möglichkeit einer Registrierung geben, sodass nach einer Anmeldung die Möglichkeit besteht, Merkmale zu editieren oder zu ergänzen und Hinweise zu fehlenden oder gefällten Bäumen zu geben.

Initiiert und angeleitet von den Forschenden mit Unterstützung von Studierenden und Mitarbeitenden der Stadt, soll es regelmäßige Aktionstage geben, beispielsweise im Rahmen des bundesweiten Digitaltages, bei denen entweder vorhandene Merkmale kontrolliert oder neue Merkmale in der Web-Anwendung erfasst werden können. Jene könnten als Trainingsgrundlage für weitere KI-Modelle dienen.

4.7 Übertragbarkeit, Evaluation und Pflege

Der Programmcode der Deep-Learning-Modelle und auch die trainierten Weights der Modelle sollen offen zur Verfügung gestellt werden, sodass andere Kommunen unter Nutzung der benötigten Datenquellen mit erheblich weniger Aufwand in der Lage sind, eigene digitale Baumkataster zu erstellen oder zu verbessern.

Des Weiteren kann durch Nutzung der Modelle mit zukünftig erfassten Daten ein Monitoring stattfinden, bei dem abgeglichen wird, welche Bäume ohne Vermerk im Baumkataster gefällt, gepflanzt oder ersetzt wurden.

Langfristig soll auch eine Evaluation der genutzten Methoden stattfinden: Wenn eine Aktualisierung der Baumhöhen anhand von Gattung, Ursprungshöhe und vergangener Zeit seit dem Erfassen der Höhe stattgefunden hat, soll ein Abgleich mit der tatsächlichen neuen Höhe stattfinden, so dass auch hier wiederum eine Verbesserung der Vorgehensweise vorgenommen werden kann.

5 ZUSAMMENFASSUNG UND AUSBLICK

Zusammenfassend ist festzustellen, dass sowohl die Erstellung als auch die Pflege eines qualitativ guten Baumkatasters eine aufwändige Arbeit ist, die jedoch vor allem im Angesicht des Klimawandels immer dringlicher und bedeutsamer wird. Ein digitales Baumkataster kann Mitarbeitenden von Kommunen helfen, Umwelt- und Klimaaspekte effektiver zu berücksichtigen und nachhaltigere Entscheidungen in der Stadtplanung zu treffen, indem z.B. geeignete Baumarten für entsprechende Standorte gewählt werden. Darüber hinaus bietet es Unterstützung bei der Voraussage von potentiellm Trockenstress, Verschattung und möglichen lufthygienischen Einflüssen bei bestimmten Wetterlagen. Durch die Einbeziehung von privaten Bäumen und die Ergänzung weiterer Merkmale wird das Baumkataster umfassender und nutzbarer für weiter reichende Anwendungsfälle aus der Wissenschaft.

Die (Weiter-)Entwicklung von Deep-Learning-Modellen erlaubt nicht nur eine Übertragung auf andere Kommunen, sondern auch eine regelmäßige Aktualisierung unter Anwendung von erneut erfassten Luft- und Panoramabildern. Dies ermöglicht überdies ein Monitoring der Datenpflege. Unter Verwendung geeigneter Schnittstellen, Portale und Web-Anwendungen sollen die Daten öffentlich zugänglich und editierbar sein.

Die Zusammenarbeit verschiedener Akteursgruppen aus der Forschung, der Verwaltung und der Bevölkerung ist zugleich herausfordernd und entscheidend für den Erfolg des Projekts. Eine transparente Kommunikation, die einfache Erklärungen und nachvollziehbare Gründe für eine Mitwirkung beinhaltet, ist hier unabdingbar.

Es ist überdies empfehlenswert, dass nach der Erprobung verschiedener KI-Modelle unter Anwendung auf unterschiedliche Datenbasen eine Evaluierung stattfindet, bei der Kriterien festgelegt werden, nach denen die zu verwendenden Daten und Modelle auszuwählen sind. Des Weiteren sollen einheitliche und vereinfachte Prozesse erstellt werden, wie eine Erfassung bzw. wiederkehrende Aktualisierung des Baumkatasters ablaufen kann und insbesondere wie mit fehlerhaften Ergebnissen aus den KI-Anwendungen umgegangen wird, die sich nie gänzlich vermeiden lassen. Eine weiterer großer Schritt, der noch umzusetzen und zu integrieren ist, wird der Umgang mit Baumgruppen sein.

6 LITERATUR

ALBERT, L.; FRÖHLICH, N.; HAUSBRAND, N.; HENNINGER, S.; MAURER, T.; RUZIKA, S.: Human-biometeorologisch angepasste Routenführungen durch mathematische Optimierung. In: *Mobility, Knowledge and Innovation Hubs in Urban and Regional Development. Proceedings of REAL CORP 2022, 27th International Conference on Urban Development, Regional Planning and Information Society*. pp. 427-435. DOI: 10.48494/REALCORP2022.7061

BRUNS, L.; DITTWALD, B.; Meiners, F.: Leitfaden für qualitativ hochwertige Daten und Metadaten, <https://www.nqdm-projekt.de/de/downloads/leitfaden>, abgerufen am: 1.06.2023.

HENNINGER, S. (2014): The Impact of Biogenic Isoprene in Dependence on Meteorological Conditions within Urban Green. In: Rauch, S., Morrison, G., Norra, S. und N. Schleicher [Hrsg.]: *Urban Environment - Proceedings of the 11th Urban Environment Symposium*, Springer Verlag, pp. 153-162.

LENKUNGS AUSSCHUSS GEODATENINFRASTRUKTUR RP IM MINISTERIUM DES INNERNS FÜR SPORT DES LANDES RHEINLAND-PFALZ, GDI-RP, <https://www.geoportal.rlp.de>, abgerufen am: 01.05.2023.

- METZ, M; WEINMANN, A.; KRISZTIAN, L: Automatisierte Detektion von Baumstandorten in der Metropole Ruhr. In: FOSSGIS e.V. (2023). Tagungsband FOSSGIS-Konferenz 2023. p. 161. DOI: <https://doi.org/10.5281/zenodo.7573277>
- OPEN KNOWLEDGE FOUNDATION DEUTSCHLAND E.V., <https://codeforleipzig.github.io/giessdeibohm/>, abgerufen am:14.06.2023.
- OPEN STREET MAP WIKI CONTRIBUTORS, OpenStreetMap Wiki "DE:Tag:natural=tree"
<https://wiki.openstreetmap.org/w/index.php?title=DE:Tag:natural%3Dtree&oldid=2446549>, abgerufen am 02.06.2023.
- TECHNOLOGIESTIFTUNG BERLIN, <https://www.giessdenkiez.de>, abgerufen am: 14.06.2023.
- VENTURA, J; PAWLAK, C.; HONSBERGER, M.; GONSALVES, C.; RICE,J.; LOVE, N.L.R.; HAN, S.; NGUYEN, V.; SUGANO, K; DOREMUS, J.; FRICKER, G.A.; YOST, J.; RITTER, M.: "Individual Tree Detection in Large-Scale Urban Environments using High-Resolution Multispectral Imagery." arXiv:2208.10606 [cs], Oct. 2022.
- WEINSTEIN, B.G.; MARCONI, S.; BOHLMAN, S.; ZARE, A.; WHITE, E.: Individual Tree-Crown Detection in RGB Imagery Using Semi-Supervised Deep Learning Neural Networks. Remote Sens. 2019, 11, 1309.
- WINTER, J: Baumfreunde Magdeburg, <https://www.baumfreunde-md.de/home>, abgerufen am: 14.06.2023.