

Modellierung urbaner Komplexität für datengestützte Anwendungen in der Smart City

Martin Memmel, Julia Mayer

Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH, Kaiserslautern

Die Qualität der Ergebnisse datengestützter Anwendungen hängt entscheidend von den verwendeten Daten ab. Diese Daten können unmittelbar oder in Form von aus Ihnen abgeleiteten Modellen vorliegen und sollten das in der Anwendung fokussierte Problem sowie den zugehörigen Kontext so repräsentieren und modellieren, dass eine algorithmische Nutzung möglichst aller für ein gutes oder unter Umständen sogar perfektes Ergebnis relevanten Informationen erfolgen kann (siehe Abbildung 1). Gleichzeitig sind aber auch Aspekte wie Datenschutz, Datenethik und Datensparsamkeit zu beachten.

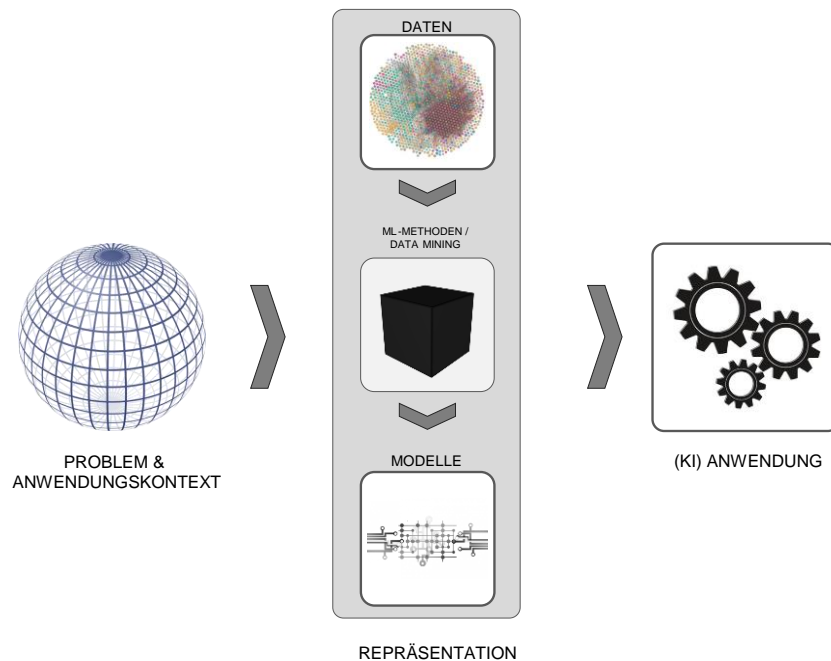


Abbildung 1: Problem und Anwendungskontext müssen für eine (KI)-Anwendung durch Daten oder daraus abgeleitete Modelle repräsentiert werden

Datengestützte Anwendungen in der Smart City sehen sich einer besonderen Herausforderung gegenüber, denn ihr Anwendungskontext ist die Stadt und betrifft damit auch die dort lebenden Menschen. Aber wie können Städte in ihrer Vielfalt, Komplexität und Dynamik adäquat repräsentiert und modelliert werden? Wie können Nachhaltigkeitsziele Eingang finden? Wie kann sichergestellt werden, dass Menschen in ihren individuellen Bedürfnissen nicht durch die Ergebnisse eines algorithmischen Verfahrens beeinträchtigt oder benachteiligt werden [1,2]? Und wie gehen die in den letzten Jahren vermehrt diskutierten oder im Einsatz befindlichen „Digitalen Zwillinge“ [3] von Städten mit diesen Herausforderungen um?

Zur Beantwortung dieser Punkte müssen zahlreiche Aspekte beleuchtet werden. Dies beginnt bei der grundsätzlichen Frage, welche Qualitätskriterien für eine adäquate Repräsentation und Modellierung existieren und wie, wann und von wem diese operativ angewendet werden können [4]. Weiterhin ist von Bedeutung, welche Daten im urbanen Umfeld überhaupt existieren, ob und wie auf diese Daten zugegriffen werden kann, oder ob Daten erhoben werden können.

Insbesondere in kommunalen Verwaltungen liegt eine heterogene Datenlandschaft vor, die auf unterschiedlichste Arten generiert wird und in verschiedenen Bereichen zum Einsatz kommt. Diese Daten reichen von Bauplänen über Sensordaten bis zu Bevölkerungs- und Umfragedaten. Bei der Auswahl der Daten und bei ihrer Verwendung in einer Anwendung ist dann zu klären, in welcher Form die zu den Qualitätskriterien erfassten Kenntnisse Anwendung finden – sowohl bei der Selektion und algorithmischen Verarbeitung als auch in der Darstellung gegenüber Nutzerinnen und Nutzern oder in Schnittstellen für andere, ggf. weiterverarbeitende Verfahren. Hier spielen Konzepte und Forschungsansätze zur Realisierung von Transparenz und Erklärungsfähigkeit eine entscheidende Rolle.

Man kann grundsätzlich allgemeine, anwendungsunabhängige Aussagen über Datenqualität, etwa Korrektheit oder Aktualität, tätigen. Andere Qualitätsabschätzungen lassen sich jedoch oft erst im Zusammenhang mit der gewünschten Nutzung machen. Beispielsweise ist ein gut gepflegtes städtisches Baumkataster eine sinnvolle Grundlage, um Kontrollen zur Gewährleistung der Sicherheitspflicht zu planen. Es ist allerdings nicht geeignet, um eine Allergen-Belastung im gesamten Stadtgebiet zu ermitteln, da Bäume auf privaten Grundstücken fehlen und somit eine unvollständige Datenbasis vorliegt [5].

Kritisch werden solche Aspekte vor allem dann, wenn die Auswirkungen Menschen betreffen. Datensätze mit personenbezogenen Daten unterliegen besonderen Vorschriften und müssen der DSGVO entsprechend sensibel behandelt werden. Jedoch bergen auch Daten ohne direkten Personenbezug Risiken für eine Diskriminierung bestimmter Bevölkerungsgruppen, wenn sie ohne tieferes Verständnis algorithmisch verarbeitet werden. So können Daten etwa Informationen über Herkunft, Ethnie oder Geschlecht von Menschen enthalten und bei einer weiteren Verarbeitung zu ethisch und rechtlich problematischen Konsequenzen führen [6]. Sobald auch Anwendungen der KI hinzukommen, verstärken sich diese und ähnliche Effekte, da bei den dort verwendeten Modellen meist sogar für Expertinnen und Experten nur schwer oder gar nicht nachvollziehbar ist, wie und mit welchen Trainingsdaten die Modelle erstellt wurden.

Dieser Beitrag analysiert die Herausforderungen bei der Modellierung urbaner Komplexität für datengestützte Anwendungen in der Smart City und beschreibt darauf basierend beispielhaft das konkrete Vorgehen bei der Verwendung von Daten und Modellen im durch die Carl-Zeiss-Stiftung geförderten Forschungsprojekt „Ageing Smart - Räume intelligent gestalten“. Dort wird für ausgewählte Modellkommunen unter Verwendung von KI-Technologien ein datenbasiertes Entscheidungsunterstützungssystem entwickelt, welches kommunale Anwenderinnen und Anwendern bei einer nachhaltigen und generationengerechten Planung unterstützen soll.

Literatur

- [1] Williams, S. Data Action: Using Data for Public Good; The MIT Press, 2020. 13.
- [2] Jacobs, J. The death and life of great american cities; Jonathan Cape: London, 1962.
- [Charta] Smart City Charta - Digitale Transformation in den Kommunen nachhaltig gestalten; Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) / Bundesministerium des Innern, für Bau und Heimat (BMI), 2021.
- [3] Grieves, M., Virtually Intelligent Product Systems: Digital and Physical Twins. In Complex Systems Engineering: Theory and Practice; American Institute of Aeronautics and Astronautics, 2019; pp. 175–200. <https://doi.org/10.2514/5.9781624105654.0175.0200>.
- [4] Priestley, M.; O'Donnell, F.; Simperl, E. A Survey of Data Quality Requirements That Matter in ML Development Pipelines. *Journal of Data and Information Quality* 2023, 15, 1–39. <https://doi.org/10.1145/3592616>.
- [5] Julia Mayer, Martin Memmel, Johannes Ruf, Dhruv Patel and Sascha Henninger. Progressing towards Estimates of Local Emissions from Trees in Cities: A Transdisciplinary Framework Integrating Available Municipal Data, AI, and Citizen Science. In *Applied Sciences*, vol. 14, no. 1, 2024.
- [6] DSK. Hambacher Erklärung zur Künstlichen Intelligenz. Sieben datenschutzrechtliche Anforderungen.
Entschließung der 97. Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder. Technical report, DSK – Datenschutzkonferenz, 2019. Retrieved November 14, 2023, from https://www.datenschutzkonferenz-online.de/media/en/20190405_hambacher_erklaerung.pdf.