Handreichung

Standards für Künstliche Intelligenz

Vitako-Handreichung zur Entwicklung von Standards für Künstliche Intelligenz

Stand: 17. April 2020

VITAKO
Bundes-Arbeitsgemeinschaft der Kommunalen IT-Dienstleister e.V.

Vitako-Handreichung zur Entwicklung von Standards für Künstliche Intelligenz

Management Summary

Dieses Papier ist das Ergebnis der Vitako-Projektgruppe „Künstliche Intelligenz“. Es fasst die grundlegenden Ziele, Handlungsfelder und Probleme zusammen, mit denen die öffentliche Hand bzw. die Kommunen in Bezug auf Künstliche Intelligenz (KI) umgehen müssen. Unterstellt wird, dass KI im Sinne kommunaler Daseinsvorsorge gewinnbringend eingesetzt und Verwaltungsleistungen gegenüber Bürgern und Unternehmen verbessert werden können.

1. Einleitung

Die kommunalen IT-Dienstleister sehen es als eine ihrer Aufgaben, neue technologische Möglichkeiten im Sinne ihrer Eigner und Auftraggeber zu analysieren und ggf. zu nutzen. Für Vitako ist KI deshalb eine Schlüsseltechnologie für die effiziente Erbringung künftiger Verwaltungsleistungen. KI gewinnt bereits heute an Relevanz und wird laut verschiedener Studien[1] an Bedeutung noch deutlich zunehmen.

Im Mittelpunkt der vorliegenden Betrachtung stehen Aspekte, die sich vor allem um die spezifischen Belange eines KI-Einsatzes bei staatlichen und kommunalen Leistungen drehen. Es geht um Fragen wie Datenhoheit und -souveränität, Transparenz, Normen und Standards sowie um die Kontrolle von Algorithmen. Auch spielt die Frage eine Rolle, welche Daten bei Data-Analytics/Data-Mining genutzt werden dürfen.

2. Grundsätzliche Ziele und Annahmen

a. Worum es technologisch geht – schwache KI

Für kommunale bzw. öffentliche Belange spielt vor allem die „schwache Künstliche Intelligenz“[2] absehbar eine Rolle. Als schwache KI werden selbstlernende Systeme verstanden, die Menschen bei ihrer Arbeit assistieren, dabei Mehrwerte schaffen und zu einer Qualitätsverbesserung beitragen. Solche Systeme können Entscheidungen insbesondere dann antizipativ den Weg bereiten, wenn Ihnen sensorische Daten zur Auswertung zur Verfügung gestellt werden.

b. Zweck der Kommunen: Gemeinwohl durch Daseinsvorsorge

Die Daseinsvorsorge ist eine zentrale Aufgabe von öffentlichen Verwaltungen und öffentlichen IT-Dienstleistern. Ziel ist es, Aufgaben zu erfüllen, an denen ein besonderes allgemeines Interesse besteht. Vor allem im Bereich der Leistungsverwaltung bieten KI-Anwendungen gleichermaßen Chancen und Herausforderungen.

[1] Vgl. u. a. https://www.dfki.de/fileadmin/user_upload/import/9744_171012-KI-Gipfelpapier-online.pdf
[2] Vgl. etwa https://kiel.ai/kunstliche-intelligenz-definition/

c. Öffentliches Handeln muss nachvollziehbar sein

Ziel öffentlichen Handelns ist es, sowohl effizient und effektiv vorzugehen als auch legal und legitim – Vorgänge müssen diesbezüglich prüfbar sein. Künstliche Intelligenz fußt hingegen in weiten Teilen auf Lernprozessen, die bisher vielfach nur eingeschränkt nachzuvollziehen sind. „Explainable Artificial Intelligence“ (XAI)[3] kann solche „Black Boxes“ vermeiden und für Transparenz sorgen, sodass Verwaltungsleistungen von Bürgerinnen und Bürger akzeptiert und Fachleute in die Lage versetzt werden, Prozesse nachvollziehen und prüfen zu können.

3. Handlungsempfehlungen zur Nutzung von KI

Architektur möglichst offen gestalten

Eine KI-Architektur ist für Veränderungen offen zu gestalten. Die Zusammenstellung ihrer Komponenten muss „klar“ aufgestellt werden. Ziel ist es, mit Blick auf die „Architektur des Trainings“ und die „Betriebsarchitektur“ den Algorithmus und die trainierte KI offen zu lassen, um „KI-Silos“ zu vermeiden. Dazu tragen Mindestqualitätsstandards sowie Transparenz über Quellen, Struktur und Zweck der Datenverarbeitung bei. Es muss die Möglichkeit einer "KI made in Germany" geschaffen werden, als Gegenstück zu KI-Systemen, die außerhalb des deutschen Rechtsrahmens entstanden sind.

Datenhoheit der Bürger sicherstellen

Digitale Souveränität und Datensouveränität[4] sind im Zweifel die Basis für die (digitale) Handlungsfähigkeit der öffentlichen Hand. Beides gilt es, fortlaufend zu evaluieren und bei Entscheidungen zu berücksichtigen. Damit Komplexität und Datensouveränität in KI-Systemen beherrschbar bleiben, sollten einige Grundsätze von vorne herein berücksichtigt werden. Dabei gilt grundsätzlich, dass Datenhoheit in jedem KI-Prozess als Standard zu etablieren ist:

- Bürgerinnen und Bürger müssen einwilligen, wenn ihre persönlichen Daten auch für KI-Zwecke genutzt werden sollen. Ihnen sollte zudem die Möglichkeit eingeräumt werden, die Entscheidung zu widerrufen bzw. personenbezogene Daten löschen zu lassen. Es ist hierbei darauf zu achten, dass ein KI-Modell (bestehend aus Daten und Algorithmus) nicht gänzlich verworfen werden muss, wenn einzelne Datensätze zurückgezogen werden.
- Es erscheint zudem angebracht, dass Bürgerinnen und Bürger der Nutzung ihrer persönlichen Daten später erneut zustimmen müssen. Hier gibt es etwa Regelungen, die Zeiträume von längstens zwei Jahre vorsehen, mit Option auf Verlängerung. Datensparsame Default-Einstellungen können diesen Aspekt ebenso unterstützen.
- Automatische Benachrichtigungen sorgen dafür, dass über Zugriffe auf Rohdaten mit Personenbezug informiert werden kann. Ein solches Element verursacht kaum Aufwand in der Verwaltung und fördert die Akzeptanz bei Bürgerinnen und Bürgern.

[3] Vgl. https://www.sciencedirect.com/science/article/pii/S1566253519308103

[4] Definition gem. Studie zum Thema „Digitale Souveränität“ der Kompetenzstelle Öffentliche IT (ÖFIT), aus Jan 2017, Gabriele Goldacker: „Die Fähigkeiten und Möglichkeiten von Individuen und Institutionen, ihre Rolle(n) in der digitalen Welt selbstständig, selbstbestimmt und sicher ausüben zu können“.

Klares Bekenntnis zu Open Source geben

KI-Systeme der öffentlichen Hand dürfen gegenüber internationalen Anbietern nicht in Abhängigkeit geraten. Es ist außerdem zu vermeiden, dass neu etablierte Standards zu „Lock-in-Effekten“ führen, die nur bestimmte Nutzungsweisen zulassen und so eine Fortentwicklung von Systemen sehr erschweren.

Im Gegensatz dazu könnte ein Standard im Sinne einer „KI DIN“ sowie ein klares Bekenntnis zu Open-Source helfen. In Kombination ließen sich offene freie Standards, Bibliotheken, Frameworks und Tools entwickeln. Private Unternehmen und öffentliche IT-Dienstleister wären so imstande, sich aktiv an Basistechnologien der betreffenden Communities zu beteiligen oder bei Bedarf auch neue Initiativen zu gründen. Dabei ist zu beachten, dass erfolgreiche Open Source-Projekte durch „Hyper-Scaler“ übernommen werden können. Es besteht die Gefahr, dass Internet-Konzerne diese dann eigenmächtig weiterentwickeln und ein Geschäftsmodell darauf aufbauen.

Hier ist eine politische Diskussion wünschenswert, ob „Source-Available“, welches nur im Kontext von Non-Profit Organisationen (NGOs) "verwertet" und bepreist werden kann, eine sinnvolle Open-Source-Variante darstellt. Dies wäre quasi eine Art NGO-Open-Source, die kostendeckend betrieben und gepflegt wird, aber keinen Gewinnabsichten obliegt. Öffentliches Geld wird damit in virtuelle, öffentliche und frei zugängliche Besitzgüter umgewandelt, die dem Gemeinwesen frei zur Verfügung stehen, aber keinen kommerziellen Interessen dienen.

Open-Source ist zudem notwendig, um neben Datensouveränität die nötige Transparenz zu schaffen. Transparenz ermöglicht unabhängigen Experten in Konfliktfällen, jede Art von KI-Algorithmik besser auf den Prüfstand stellen zu können.

Metadaten ohne Personenbezug nutzen

Für die Anwendung und Entwicklung von KI muss eine klare, einheitliche und rechtsverbindliche Grundlage geschaffen werden. Dabei sind technologischer Fortschritt und gesetzliche Grundlagen gleichermaßen zu berücksichtigen. In Bezug zu Data-Analytics/Data-Mining sollte die Verwendung von Metadaten erlaubt sein, die keine expliziten Rückschlüsse mit individuellem Personenbezug ermöglichen. Dies würde die Entwicklung und die Anwendung von KI in einer rechtlichen Klammer ermöglichen.

Diskriminierung unterbinden

Diskriminierung muss bei durch KI getroffene Entscheidungen durch die Auswahl der Zielvariablen, der Testdaten und der zu berücksichtigen Attribute weitgehend ausgeschlossen werden. Jedoch ist zu beachten, dass auch ein diskriminierungsfreier Datensatz indirekt zu Diskriminierung führen kann. Folglich können die typischen diskriminierenden Merkmale gemessen und geprüft werden. Letztlich muss das Ergebnis diskriminierungsfrei sein.

Die Verwendung und Festlegung der Daten sind mit dem Trainingsergebnis zu dokumentieren. Andere Testdaten müssen geeignet sein, die Diskriminierungsfreiheit bis zu einem gewissen Grad nachzuweisen – dies ist in den Fällen besonders wichtig, wenn "fremde" Bibliotheken oder vortrainierte Modelle verwendet werden.

Kann dieser Nachweis bei "gravierenden" Entscheidungen nicht erbracht werden, ist ein abschließendes Urteil durch einen Menschen zu treffen. Die KI leistet in diesen Fällen trotzdem

wertvolle unterstützende Hilfe – das gilt gerade für komplexe Lagen bei einer Vielzahl relevanter Parameter. In einfacher gelagerten Fällen und bei automatisch getroffenen Entscheidungen sollten Nutzer einen Hinweis erhalten, dass an der Entscheidungsfindung ein KI-Algorithmus beteiligt war.

Klimaschutz berücksichtigen

Das Trainieren von Deep-Learning-Modellen mit enormen Datenmengen führt zu einer höheren Belastung der Speicher und Prozessoren. Folglich steigt der Rechenbedarf für KI-Systeme und damit auch der Energieverbrauch. Die Forschungsgruppe „Open AI“[5] ist zum Schluss gekommen, dass sich die benötigte Rechenleistung für größere KI-Modelle alle dreieinhalb Monate verdoppelt. Forscher der University of Massachusetts errechneten in ihrer Studie „Energy and Policy Considerations for Deep Learning in NLP“[6] , dass große bekannte KI-Modelle etwa fünf Mal so viel CO^2 emittieren wie die durchschnittliche Lebensemission eines Autos in den USA. Eine modern gedachte KI muss die eigenen Energiebedarfe daher regelmäßig hinterfragen. Es gilt, mithilfe energieeffizienter Server und Kühlungssysteme, KI längerfristig klimaneutral zu gestalten.

4. Fazit

Es ist zu erwarten, dass KI als Technologie ein wichtiger Bestandteil künftiger Verwaltungsleistungen und Services zur Daseinsvorsorge sein wird. Dabei muss der Begriff allerdings kulturell positiv geprägt werden, um gerade im öffentlichen Sektor die notwendige Akzeptanz der Bürgerinnen und Bürger zu erlangen. Hierbei sind insbesondere Maßnahmen zum Schutz persönlicher Daten und Rechte des Einzelnen zu berücksichtigen und Algorithmen entsprechend zu kontrollieren. Auf staatlicher/kommunaler Seite geht es zudem darum, eine souveräne Handlungsfähigkeit aufzubauen und langfristig zu gewährleisten. Wichtige Rollen dabei spielen die Architektur und Prozesse, Normen und Standards sowie eine grundsätzliche Ausrichtung an Open Source.

[5] Siehe: https://openai.com/
[6] Siehe: https://arxiv.org/abs/1906.02243