



SmartCity

Braunschweig

Wie passen Künstliche Intelligenz und offene Daten zusammen?

Gliederung

- I. Open Data in Braunschweig
- II. Zusammenspiel von KI und offenen Daten
- III. Anwendungsfälle
- IV. Fazit

Gliederung

I. Open Data in Braunschweig

II. Zusammenspiel von KI und offenen Daten

III. Anwendungsfälle

IV. Fazit

Open Data in Braunschweig



 BEVÖLKERUNG 10 DATENSÄTZE	 GEO 19 DATENSÄTZE	 VERWALTUNG, HAUSHALT UND STEUERN 5 DATENSÄTZE	 GESUNDHEIT 11 DATENSÄTZE	 INFRASTRUKTUR, BAUEN UND WOHNEN 5 DATENSÄTZE	 KULTUR, FREIZEIT, SPORT UND TOURISMUS 8 DATENSÄTZE	 BILDUNG UND WISSENSCHAFT 9 DATENSÄTZE
 POLITIK UND WAHLEN 2 DATENSÄTZE	 SOZIALES 8 DATENSÄTZE	 TRANSPORT UND VERKEHR 21 DATENSÄTZE	 UMWELT UND KLIMA 16 DATENSÄTZE	 WIRTSCHAFT UND ARBEIT 6 DATENSÄTZE		

- Seit Dezember 2022 im Einsatz
- Aktuell 58 Datensätze im Portal
- Kontinuierlicher Ausbau

Gliederung

I. Open Data in Braunschweig

II. Zusammenspiel von KI und
offenen Daten

III. Anwendungsfälle

IV. Fazit

Zusammenspiel von KI und offenen Daten

- KI gewinnt in der öffentlichen Verwaltung zunehmend an Bedeutung
- Neben der Ein- und Ausgabe von Texten rückt dabei die KI-basierte Analyse von Daten in den Fokus
- Anwendungsfälle wie kommunale Daten mit dem GPT Data Analyzer ausgewertet und aufbereitet werden können

Gliederung

I. Open Data in Braunschweig

II. Zusammenspiel von KI und
offenen Daten

III. Anwendungsfälle

IV. Fazit

Anwendungsfälle

- Geodaten → Fluglärm
- Haushaltsdaten → Ergebnishaushalt
- Sensordaten → Passantenfrequenzen

Anwendungsfälle

- Geodaten → Fluglärm
- Haushaltsdaten → Ergebnishaushalt
- Sensordaten → Passantenfrequenzen

Anwendungsfälle

Beispiel 1 | Geodaten → Fluglärm

Basisdaten aus dem Open Data Portal:

- Lärmkartierung Fluglärm in Braunschweig, 24 Stunden Tag (2022)
- Lärmkartierung Fluglärm in Braunschweig, Nacht (2022)
- Regionalkarte Braunschweig 1: 100.000

Anwendungsfälle

Beispiel 1 | Geodaten → Fluglärm



Lies bitte die folgende Datei ein. [Datei angefügt]



The file you've uploaded, [Dateiname], appears to be a shapefile. A shapefile is a common geospatial vector data format for geographic information system (GIS) software. It is developed and regulated by Esri, and it can be used to represent geographical features. We can use the geopandas library in Python to read and analyze this file. Let's first load the file and examine its structure. Would that be okay?



Ja, gerne.



Es scheint, dass beim Versuch, die Shapefile zu öffnen, ein Fehler aufgetreten ist. Shapefiles bestehen normalerweise aus mehreren zusammenhängenden Dateien mit den Erweiterungen .shp, .shx, .dbf und manchmal .prj, .sbn, .sbx, .shp.xml usw. Diese Dateien müssen alle zusammen vorhanden sein, damit die Shapefile korrekt gelesen werden kann. Der spezifische Fehler, den wir hier sehen, deutet darauf hin, dass die .shx-Datei fehlt oder nicht gefunden werden kann. Könnten Sie bitte überprüfen, ob diese Datei vorhanden ist und sie ggf. erneut hochladen?



Kann ich dir auch einen ganzen Zip-Ordner hochladen, der die genannten Dateien enthält?

Anwendungsfälle

Beispiel 1 | Geodaten → Fluglärm



Ja, absolut! Sie können einen ZIP-Ordner mit allen Dateien hochladen, die zur Shapefile gehören. Sobald Sie das getan haben, kann ich versuchen, die Dateien zu entpacken und die Shapefile erneut zu laden.



[Zip-Ordner angefügt]



Die Dateien wurden erfolgreich entpackt. Wir haben jetzt alle Dateien, die zur Shapefile gehören, einschließlich der .shx-Datei, die vorher gefehlt hat. Lassen Sie uns nun versuchen, die Shapefile erneut zu laden. Die Shapefile wurde erfolgreich geladen. Hier sind die ersten paar Zeilen der Daten [...] Wie möchten Sie mit diesen Daten weiterfahren?



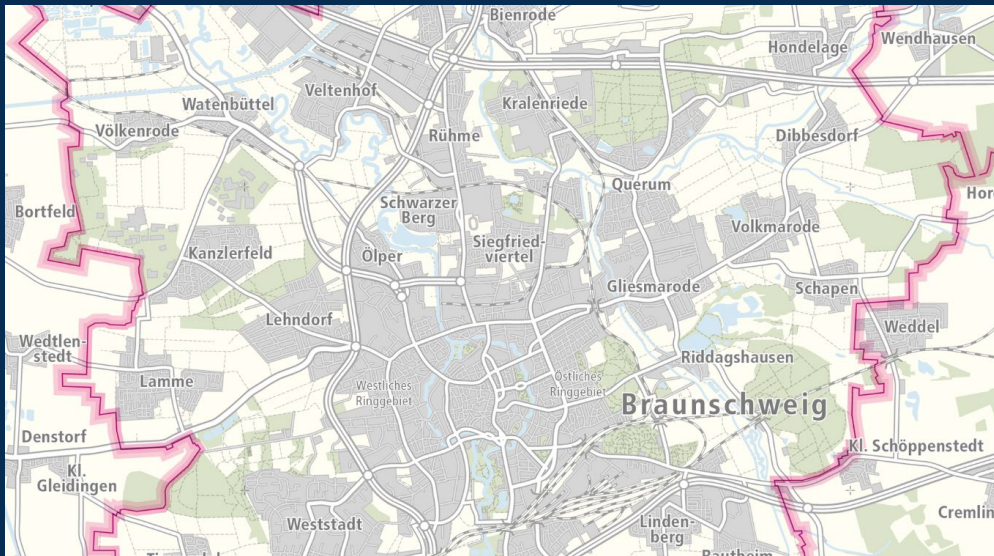
Kannst du mir die Daten bitte als Karte darstellen? Hinweis: Es handelt sich um eine Fluglärmkarte in der Einheit db(A).



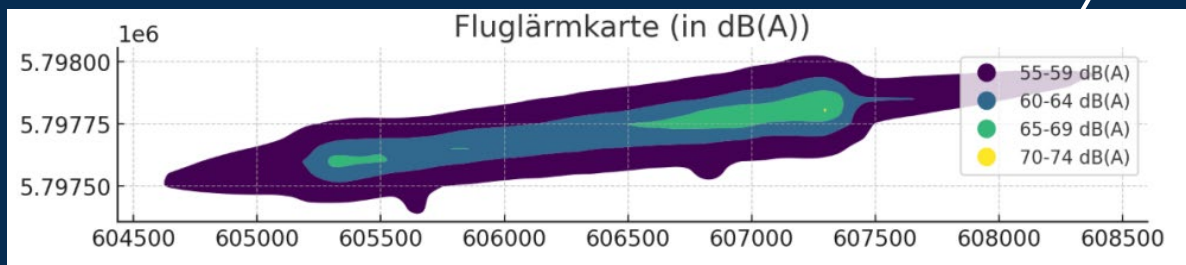
Hier ist die Fluglärmkarte, die auf den Daten in Ihrer Shapefile basiert. Jede Zone auf der Karte repräsentiert eine bestimmte Lärmkategorie (in dB(A)), die durch unterschiedliche Farben dargestellt wird. Die Farbskala recht von der Karte zeigt, welche Farben welche Lärmkategorien repräsentieren. Bitte beachten Sie, dass die Karte nur die räumliche Verteilung der Lärmkategorien zeigt und keine spezifischen geografischen Merkmale oder Bezugspunkte enthält. Wenn Sie zusätzliche Informationen auf der Karte darstellen möchten (z.B. Städte, Straßen, Flughäfen usw.), benötigen wir entsprechende Geodaten.

Anwendungsfälle

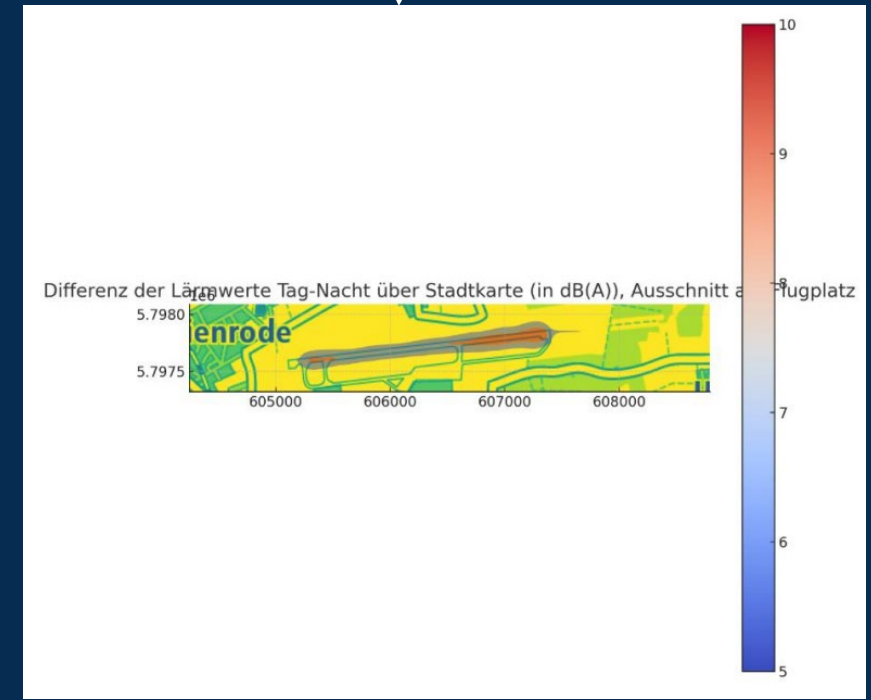
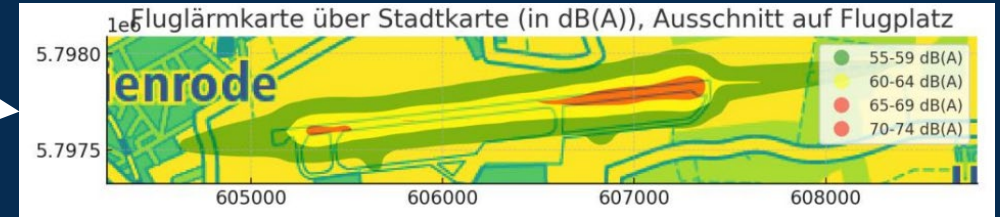
Beispiel 1 | Geodaten → Fluglärm



Regionalkarte Braunschweig



Lärmkartierung Flughafen



Anwendungsfälle

- Geodaten → Fluglärm
- **Haushaltsdaten → Ergebnishaushalt**
- Sensordaten → Passantenfrequenzen

Anwendungsfälle

Beispiel 2 | Haushaltsdaten → Ergebnishaushalt

Basisdaten aus dem Open Data Portal:

- Haushaltsplan der Stadt Braunschweig 2023-2024
(nach Umwandlung von PDF zu CSV)

Anwendungsfälle

Beispiel 2 | Haushaltsdaten → Ergebnishaushalt



Lies bitte die folgende Datei ein. Das Trennzeichen ist ein Semikolon.



Die Datei Ergebnis3.csv wurde erfolgreich eingelesen. Sie enthält folgende Spalten: [...] Was möchten Sie als Nächstes tun?



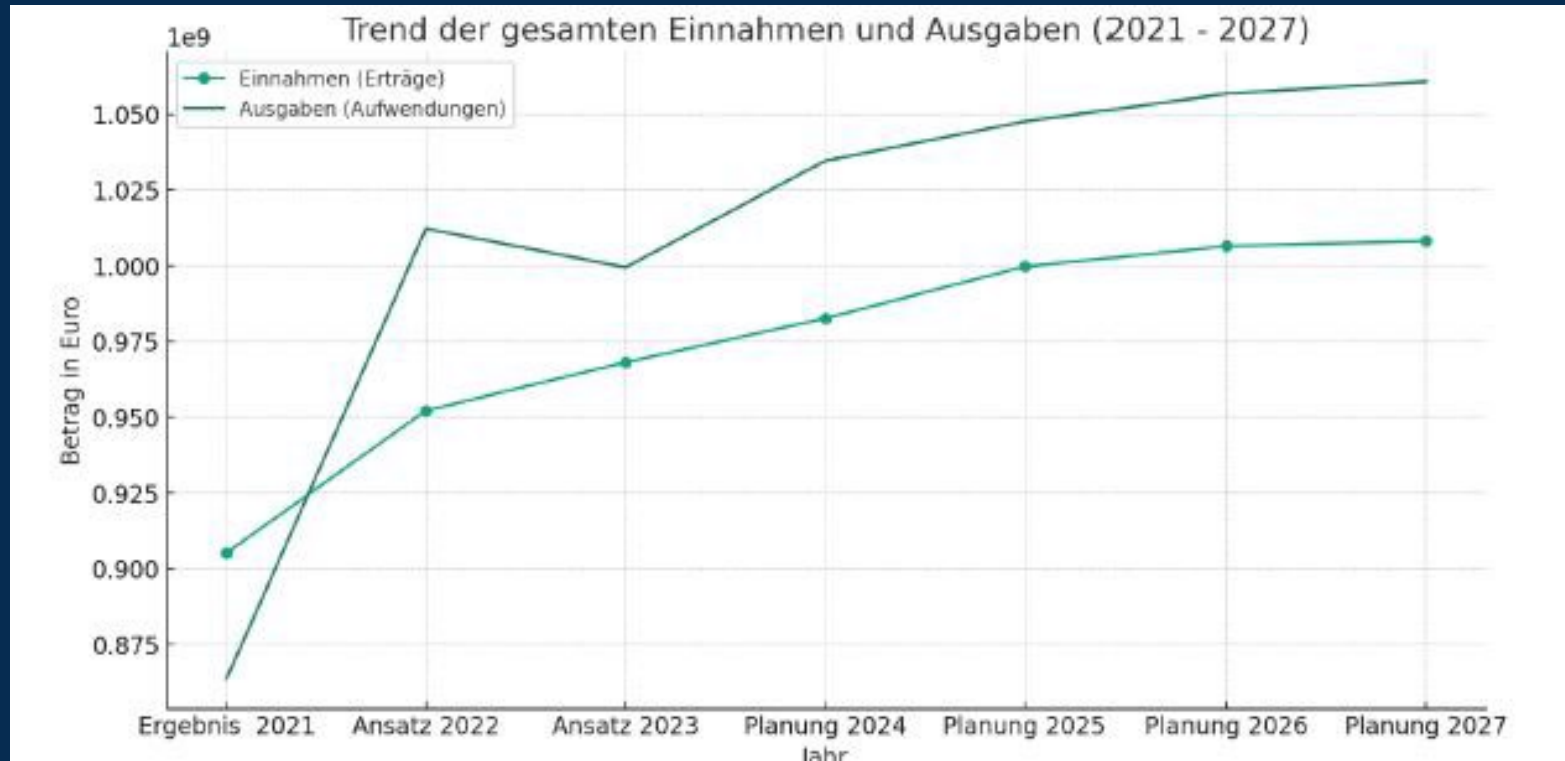
Es handelt sich um den Ergebnishaushalt (Ergebnis 2021 und Plan 2022-2027) einer deutschen Großstadt. Die ordentlichen und außerordentlichen Erträge sind die Einnahmen der Stadt und die ordentlichen und außerordentlichen Aufwendungen sind die Ausgaben der Stadt. Beide sind in der ersten Spalte kenntlich gemacht. Die zweite Spalte in der CSV-Datei zeigt die Unterkategorie an, die die Erträge/Aufwendungen verursachen. Die Spalte Sachkontobezeichnung gibt auf der Sachebene an, was die Erträge/Aufwendungen verursacht. Bitte stelle den Trend der gesamten Einnahmen (ordentliche + außerordentliche Einnahmen) und Ausgaben (ordentliche + außerordentliche Ausgaben) grafisch dar.



Der Graph zeigt den Trend der gesamten Einnahmen (ordentliche + außerordentliche Erträge) und Ausgaben (ordentliche + außerordentliche Aufwendungen) der Stadt für die Jahre 2021 bis 2027. Was möchten Sie als Nächstes tun?

Anwendungsfälle

Beispiel 2 | Haushaltsdaten → Ergebnishaushalt



Anwendungsfälle

Beispiel 2 | Haushaltsdaten → Ergebnishaushalt



Welche Einspar- bzw. Ertragserhöhungsoptionen würdest du der Stadt vorschlagen? Bitte argumentiere anhand konkreter Sachkonten (Spalte Sachkontobezeichnung) und entwickle mögliche Szenarien. Bitte stelle für jedes Szenario jeweils die Vor- und Nachteile dar.



Basierend auf den höchsten Ausgaben und Einnahmen für das Jahr 2021 schlage ich zwei Szenarien für Einsparungen bzw. Ertragserhöhungen vor:

Szenario 1: Reduzierung der Personalkosten

Reduzierung der Ausgaben für "Tariflich Beschäftigte Entgelt" und "Beamtenbezüge" durch Effizienzsteigerung oder Personalabbau.

Vorteile:

1. Schnelle Kostensenkung: Personalkosten sind oft ein großer Posten im Budget, und Einsparungen hier können einen schnellen Einfluss haben.
2. Langfristige Einsparungen: Einsparungen in diesem Bereich könnten sich in den Folgejahren fortsetzen.

Nachteile:

1. Moral und Produktivität: Kostensenkungen in diesem Bereich könnten die Mitarbeitermoral und Produktivität negativ beeinflussen.
2. Qualität der Dienstleistungen: Eine Reduzierung des Personals könnte zu einer Verringerung der Qualität öffentlicher Dienstleistungen führen. [...]

Anwendungsfälle

- Geodaten → Fluglärm
- Haushaltsdaten → Ergebnishaushalt
- **Sensordaten → Passantenfrequenzen**

Anwendungsfälle

Beispiel 3 | Sensordaten (Passantenfrequenzen)

Basisdaten aus dem Open Data Portal:

- Tägliche Passantenfrequenzen für den Zeitraum 19.9.2016 bis 31.12.2022 für Braunschweig

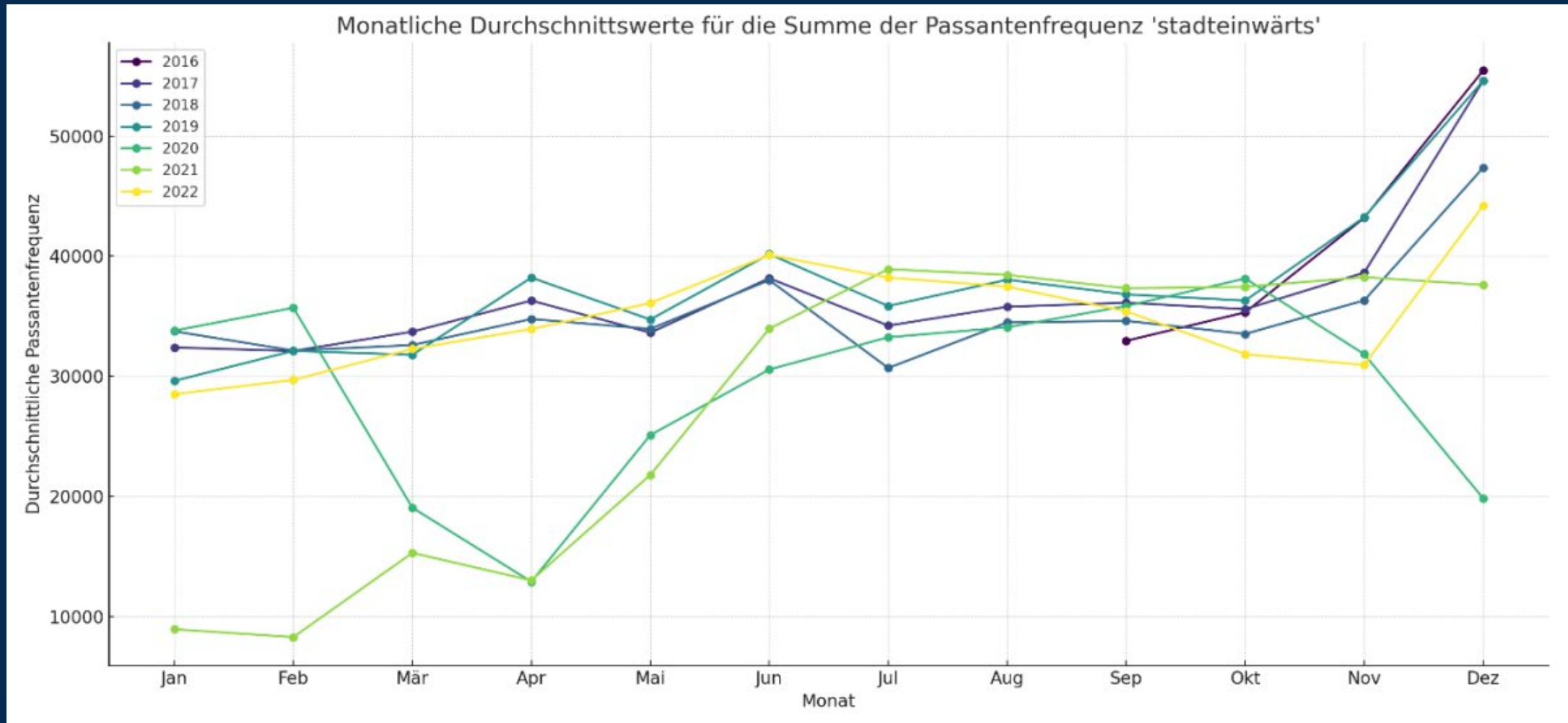
Anwendungsfälle

Beispiel 3 | Sensordaten (Passantenfrequenzen)

Datum	Damm (stadteinwärts)	Damm (stadtauswärts)	Damm (Gesamt)	Sack (stadteinwärts)	Sack (stadtauswärts)	...	max. Temperatur in °C	Wetter- Status	...
23.09.2016	8842	10623	19465	11029	10754	...	12	leichter Regen	...
24.09.2016	12720	16783	29503	17848	17475	...	8	Regenschauer	...
25.09.2016	9529	13539	23068	11978	11947	...	7	leichter Regen	...
26.09.2016	6692	7911	14603	8272	8841	...	9	wolkig	...
27.09.2016	7299	11598	18897	9094	9108	...	7	leichter Regen	...
28.09.2016	7586	8536	16122	9081	8999	...	6	leichter Regen	...
...

Anwendungsfälle

Beispiel 3 | Sensordaten (Passantenfrequenzen)



Gliederung

- I. Open Data in Braunschweig
- II. Zusammenspiel von KI und offenen Daten
- III. Anwendungsfälle

IV. Fazit

Fazit

- Schnellere Analysen durch dialogbasierte Datenauswertung und -darstellung
- Potenziale = Analyse von nicht zusammenhängenden Daten in verschiedenen Datenformaten
- Qualität der Eingaben und Anfragen (sog. Prompts) bestimmt Qualität der Ausgaben der KI
- Menschliche Prüfung der Ergebnisse teilweise noch erforderlich
- Automatisierung immer stärker im Fokus: Einlesen und Analyse von (Roh-)Daten sowie daraus resultierende Handlungsempfehlungen oder Entscheidungen
- Qualität der Rohdaten bestimmt Qualität des Ergebnisses

Fazit

Um das volle Potenzial KI-basierter Datenanalyse in Zukunft auszuschöpfen, ist es erforderlich, den Umgang mit Daten zu standardisieren. Dazu müssen Verwaltungen...

... einen internen Prozess etablieren, um sich laufend einen Überblick über den vorhandenen Datenbestand zu verschaffen.

... den Datenbestand in geeigneter Form (z. B. über einen Metadatenkatalog) festhalten.

... innerhalb der Organisation Standards zur Datenaufbereitung einführen.

... die Mitarbeiterinnen und Mitarbeiter über den Umgang mit Daten aufklären.

... die Standardisierung des Umgangs mit Daten als Voraussetzung für die Entfaltung der Möglichkeiten von KI verstehen.

Arbeitspapier

KI-basierte Auswertung kommunaler Daten mit dem GPT Data Analyzer

Vollständiges
Arbeitspapier hier als
Download:



KI-basierte Auswertung kommunaler Daten mit dem GPT Data Analyzer

– Arbeitspapier –

Dr. Neven Josipovic
Braunschweig, 12.10.2023

Abstract

Der Einsatz von künstlicher Intelligenz (KI) gewinnt in der öffentlichen Verwaltung zunehmend an Bedeutung. Neben der Ein- und Ausgabe von Texten rückt dabei die KI-basierte Analyse von Daten in den Fokus. Das Arbeitspapier untersucht vor dem Hintergrund, wie kommunale Daten mit dem GPT Data Analyzer ausgewertet und aufbereitet werden können. Dazu wurden öffentlich verfügbare kommunale Daten in ChatGPT hochgeladen und dialogbasiert analysiert. Die ausgewählten Datensätze repräsentieren gängige offene Datenformate und umfassen sowohl Geodaten als auch strukturierte Tabellendaten. Im Ergebnis zeigt sich, dass sich GPT-4 insbesondere durch seine Fähigkeit zur Interaktion in natürlicher Sprache hervorhebt, was weniger Expertise im Hinblick auf die Analyse von Daten erfordert und den Analyseprozess beschleunigt. Als intelligentes System kann es außerdem Unregelmäßigkeiten in den Daten erkennen und beheben sowie Dokumentationen und Interpretationen durchführen. Die Ergebnisse der Datenanalysen mit GPT-4 verdeutlichen die Potenziale des Einsatzes von KI für die Automatisierung von Prozessen. Für eine optimale Nutzung sollten Verwaltungen den Umgang mit Daten standardisieren und ihre Mitarbeiterinnen und Mitarbeiter zur Kommunikation mit KI-Systemen schulen. Da KI-Modelle wahrscheinlich in externen Cloud-Umgebungen gehostet werden müssen, gewinnt der Datenschutz an Bedeutung. Um geeignete Lösungen für die öffentliche Verwaltung zu entwickeln, könnte eine enge Zusammenarbeit zwischen der öffentlichen Verwaltung und privaten IT-Dienstleistern von Vorteil sein.