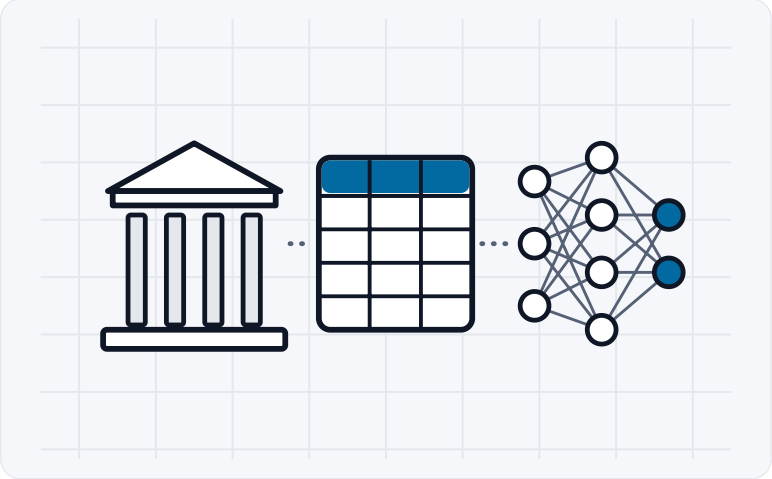
22. Februar 2023 Überblick · E-Government · KI · 11 Min. Lesezeit

Möglichkeiten zur Umsetzung eines österreichischen E-Government-Dienstes mittels offener Daten

Bianca Danczul



Autorinnen und Autoren	Bianca Danczul
Stand	2022-01-15
Quelle	https://technology2.a-sit.at/veroeffentlichungen/e-government-dienste-mit-offenen-daten/

Kernaussagen

- Auf data.gv.at sind 41154 Datensätze verfügbar, 35034 davon als CSV; 32669 stammen von offenerhaushalt.at (Gemeindefinanzen).
- Wirtschaft und Finanzen ist mit 33994 Datensätzen die größte Kategorie; Gesundheit sowie Bildung und Forschung haben je nur einen Datensatz.
- Mögliche KI-Anwendungsfälle sind prädiktive Modellierung, Optimierung, Analyse, Chatbots, Visualisierung und Überwachung.
- Offene Daten können Bias enthalten (Selektions-, Mess-, Bewertungs- und Kennzeichnungsfehler) und sollten vor der Nutzung geprüft werden.
- Auf data.gv.at sind 701 Anwendungen auf Basis offener Daten verzeichnet, die meisten davon Visualisierungen (276).

Die Nutzung offener Behördendaten wird immer wichtiger, um die Entwicklung von Anwendungen der Künstlichen Intelligenz (KI) im öffentlichen Sektor voranzutreiben. Offene Behördendaten sind Daten, die von Behörden gesammelt und in einem maschinenlesbaren Format öffentlich zugänglich gemacht werden, damit jeder sie nutzen kann.

Einer der Hauptvorteile offener Behördendaten besteht darin, dass sie die Entwicklung neuer Dienste und Erkenntnisse ermöglichen, die den Behörden helfen können, bessere Entscheidungen zu treffen und den Bürgern effektivere Dienstleistungen zu bieten. Auch Österreich stellt auf einer dedizierten Plattform offene Daten für jedermann und jederfrau zur Verfügung. Es ist jedoch unklar, welche Anwendungsfälle es dabei genau für AI-Applikationen, insbesondere Machine-Learning Applikationen, gibt und wofür die zur Verfügung gestellten Daten verwendet werden können.

Ziel des vorliegenden Überblickspapiers ist es daher, in einem ersten Schritt zu analysieren, welche Anwendungsmöglichkeiten offener Behördendaten in Österreich in Hinblick auf KI-Applikationen existieren. Im Zuge dessen sollen zudem etwaige Vorteile, Probleme und Forschungslücken aufgezeigt werden, um in einem Folgeprojekt tatsächlich diese offenen Daten zu nutzen und eine Applikation daraus zu erstellen.

1. Einleitung

Der Einsatz von Künstlicher Intelligenz (KI) im öffentlichen Sektor hat das Potenzial, eine Vielzahl von Dienstleistungen zu verbessern, von der Gesundheitsversorgung über das Bildungswesen bis hin zu Verkehr und öffentlicher Sicherheit. Ein wesentlicher Bereich, in dem KI besonders wirkungsvoll sein kann, ist der Bereich der elektronischen Behördendienste (E-Government), d. h. der Einsatz digitaler Technologien zur Verbesserung von Behördendiensten und zur Steigerung der Bürgerbeteiligung. Dabei können offene Daten (Open Government Data – OGD) verwendet werden, die seit Jahrzenten bereits in vielen Bereichen eingesetzt werden [1].

Die Verwendung offener Daten (Open Government Data – OGD) in speziell in KI-Anwendungen ermöglicht die Entwicklung neuer Dienste und Erkenntnisse, die Regierungen dabei helfen können, bessere Entscheidungen zu treffen und den Bürgern effektivere Dienstleistungen zu bieten. Diese offenen Daten können beispielsweise genutzt werden, um maschinelle Lernmodelle zu trainieren, die Verhaltensmuster bei Straftaten vorhersagen oder Bereiche identifizieren können, in denen Dienstleistungen am dringendsten benötigt werden.

OGD basiert auf verschiedenen Grundsätzen, die die Transparenz, den Austausch und die Nutzung von Regierungsdaten durch die Öffentlichkeit fördern. Zu den wichtigsten Grundsätzen von OGD [2] gehören:

- Zugänglichkeit: OGD-Daten müssen für die Öffentlichkeit leicht zugänglich sein, und zwar in einem Format, das leicht verstanden und genutzt werden kann.
- Wiederverwendbarkeit: OGD-Daten sollten für die Wiederverwendung und Weiterverbreitung durch die Öffentlichkeit ohne Einschränkungen zur Verfügung stehen.
- Interoperabilität: OGD-Daten sollten in einem Format vorliegen, das leicht mit anderen Datensätzen und Systemen integriert werden kann.
- Aktualität: OGD-Daten sollten so schnell wie möglich nach ihrer Erhebung zur Verfügung gestellt werden.
- Nicht-Diskriminierung: OGD-Daten sollten allen Mitgliedern der Öffentlichkeit zur Verfügung stehen, unabhängig von ihrem Hintergrund oder Status.
- Pflege: OGD-Daten sollten regelmäßig aktualisiert und gepflegt werden, um ihre Genauigkeit und Nützlichkeit zu gewährleisten.
- Lizenzfrei: OGD-Daten sollten ohne spezifische Lizenz verfügbar sein, damit die Daten frei verwendet, gemeinsam genutzt und weiterverteilt werden können.
- Maschinenlesbar: OGD sollten in einem Format vorliegen, das von Maschinen verarbeitet werden kann, zum Beispiel im CSV-, JSON- oder XML-Format.

Diese Grundsätze sind unerlässlich, um den öffentlichen Zugang und die sinnvolle Nutzung von Daten zu gewährleisten, die von öffentlichen Einrichtungen produziert und gesammelt werden, und um Zusammenarbeit und Innovation zu fördern.

Zur Förderung der Transparenz, dem Austausch von Informationen und der Nutzung offener Daten durch die Behörden sowie zur Festlegung von Regeln für die Weiterverwendung dieser Informationen gibt es in Österreich das Informationsweiterverwendungsgesetz (IWG) [3]. Dadurch wird sichergestellt, dass die von den Behörden gesammelten Daten in einem maschinenlesbaren Format öffentlich zugänglich gemacht werden, damit sie von Unternehmern, Forschern und anderen Gruppen zur Entwicklung neuer Dienste und Erkenntnisse genutzt werden können. Auf diese Weise kann der öffentliche Sektor die von ihm gesammelten Daten besser nutzen, was letztlich zu besseren Dienstleistungen für die Bürger*innen führen kann. Es muss dabei jedoch sichergestellt werden, dass gemäß der österreichischen KI Strategie [4] Künstliche Intelligenz und die verwendeten Daten im Einklang mit den europäischen Grundwerten, dem europäischen Rechtsrahmen und den Grund- und Menschenrechten, insbesondere dem Datenschutzrecht, eingesetzt werden.

Das Ziel dieses Überblickspapiers ist somit, einen Überblick über die Anwendungsmöglichkeiten von OGD in Österreich in Hinblick auf KI-Applikationen zu geben und etwaige Vor- und Nachteile aufzuzeigen.

2. Offene Daten Österreich

Die Cooperation Open Government Data (OGD) Österreich[1] wurde im Jahre 2011 durch das Bundeskanzleramt in Zusammenarbeit mit den Städten Wien, Linz, Salzburg und Graz gegründet, um so die Basis für Open Government Data in Österreich zu bilden. Durch diese Kooperation soll die Nutzung und Verfügbarkeit von offenen Daten gefördert und ein Single Point of Contact zum Europäischen Datenportal dargestellt werden. Zu diesem Zwecke entstand aus dieser Kooperation im Jahre 2012 die Plattform data.gv.at[2]. Die Kooperation hat sich zum Ziel gesetzt, staatliche und öffentlich finanzierte Daten für die Öffentlichkeit zugänglicher zu machen, mit dem Ziel, Transparenz, Partizipation und Innovation zu erhöhen [2]. Erreicht wird dieses Ziel unter anderem durch folgende Maßnahmen:

- Lobbyarbeit für eine offene Datenpolitik auf nationaler und lokaler Ebene
- Bereitstellung von Ressourcen und Schulungen für Einzelpersonen, Organisationen und staatliche Stellen zur Nutzung offener Daten
- Organisation von Veranstaltungen und anderen Initiativen zur Förderung der Nutzung offener Daten
- Bereitstellung der Plattform data.gv.at zur Auflistung und gemeinsamen Nutzung offener Datensätze

Die Kooperation ist davon überzeugt, dass sie durch offenere und leichter zugängliche Daten mehr Transparenz und Rechenschaftspflicht in der Verwaltung fördern, die Öffentlichkeit besser informiert in die Entscheidungsfindung einbeziehen und Innovation und Wirtschaftswachstum unterstützen kann.

2.1. Verfügbare Datensätze

Die Cooperation OGD Österreich arbeitet daran, auf data.gv.at eine Vielzahl von Daten aus staatlichen und öffentlichen Quellen der Öffentlichkeit zugänglich zu machen. Dies umfasst Daten aus Bereichen, wie z.B.:

- Wirtschaftsdaten, z. B. Daten über Unternehmen und deren finanzielle Leistungsfähigkeit
- Umweltdaten, wie z. B. Messungen der Luft- und Wasserqualität
- Verwaltungsdaten, wie z. B. Daten über staatliche Haushalte und Ausgaben
- Statistische Daten, wie demografische Daten, Wirtschaftsdaten und Meinungsumfragen
- Verkehrsdaten, z. B. Informationen über Fahrpläne und Routen öffentlicher Verkehrsmittel
- Geodaten, wie z.B. Karten und geografische Informationen
- Sozial- und Wohlfahrtsdaten, z. B. Daten zu Gesundheitsversorgung, Bildung und Armut
- Daten aus der wissenschaftlichen Forschung.

Es ist erwähnenswert, dass sich die spezifischen Datensätze, die über Open Data Austria verfügbar sind, im Laufe der Zeit ändern können und von den Daten abhängen, die ihnen von Regierungen und anderen öffentlichen Organisationen zur Verfügung gestellt werden. Sie stellen auch sicher, dass die Daten in einem geeigneten Format und gut strukturiert sind, so dass sie leicht zugänglich und nutzbar sind.

Konkret sind aktuell 41154 Datensätze verfügbar, die größtenteils im csv Format (35034 Datensätze) vorliegen, aber auch andere Formate, wie pdf (4453 Datensätze), json (668 Datensätze) oder jpeg (268 Datensätze) werden zur Verfügung gestellt. Die meisten Daten (32669 Datensätze) werden von offenerhaushalt.at veröffentlicht, es handelt sich dabei um Gemeindefinanzen, doch auch Daten von Nationalparks Austria (3325), Statistik Austria (386) oder Daten der Bundeshauptstädte finden sich auf der Seite. Die größte Kategorie ist mit Abstand die der Wirtschaft und Finanzen mit 33994 Datensätzen, doch auch Umweltdaten (4438), Daten der Regierung und öffentlichen Verwaltung (1121 Datensätze), Regionen und Städte (1045 Datensätze) oder Bevölkerung und Gesellschaft (913 Datensätze) werden zur Verfügung gestellt. Vergleichsweise wenig Datensätze gibt es im Bereich Wissenschaft und Technologie (3), Energie (2), Justiz und Gesetze (2), Bildung und Forschung (1), Gesundheit (1) oder Gesellschaft und Soziales (1).

2.2. Anwendungsfälle für offene Daten

Data.gv.at ist das zentrale Datenportal der Republik Österreich, zur Verfügung gestellt von Offene Daten Österreich, auf dem, wie bereits erwähnt, eine Vielzahl von Datensätzen, die von den verschiedenen Regierungsebenen, Gemeinden und öffentlichen Einrichtungen bereitgestellt werden, der Öffentlichkeit zugänglich gemacht werden. Diese Datensätze könnten in mehrfacher Hinsicht für die Entwicklung wertvoller KI-Anwendungen für den öffentlichen Sektor nützlich sein:

- Prädiktive Modellierung: Mithilfe von Daten aus data.gv.at können maschinelle Lernmodelle trainiert werden, um genaue Vorhersagen über zukünftige Ereignisse wie Verkehrsmuster, Kriminalitätsraten und Naturkatastrophen zu treffen.
- Optimierung: KI kann zur Optimierung einer Vielzahl von öffentlichen Dienstleistungen eingesetzt werden, z. B. von Verkehrsnetzen, Notfallsystemen und Energieverteilungsnetzen.
- Analyse: KI kann auch zur Analyse von Daten aus data.gv.at verwendet werden, um Muster, Trends und Erkenntnisse zu erkennen, die in die öffentliche Politik und Entscheidungsfindung einfließen können.
- Chatbots und virtuelle Assistenten: Die Daten von data.gv.at können zur Entwicklung von Konversationssystemen wie Chatbots und virtuellen Assistenten verwendet werden, die der Öffentlichkeit einen schnellen und einfachen Zugang zu Behördendiensten und Informationen ermöglichen.
- Visualisierung: Die Daten können verwendet werden, um interaktive Visualisierungen, Dashboards und GIS-Tools zu erstellen, die einen Überblick über die Daten bieten und den Zugang und das Verständnis für die Bürger erleichtern.

- Überwachung: Die Daten können genutzt werden, um die Leistung der öffentlichen Dienste zu überwachen, Engpässe zu ermitteln und die Qualität der Dienste zu verbessern.

Dies sind nur einige Beispiele für mögliche Anwendungsfälle, deren Erfolg jedoch von der Genauigkeit der zur Verfügung gestellten Daten abhängt. Je genauer und verlässlicher die Daten, desto effizienter und effektiver ist die Applikation, die auf diesen Daten basiert.

2.3. Mögliche Bias in den Daten

Die Daten von data.gv.at können, wie jede andere Datenquelle, Bias enthalten. Solche Bias können durch eine Reihe von Faktoren entstehen, wie z.B. die Art der Datenerhebung, die Art der Datenverarbeitung und bereinigung und die Art der Datenpräsentation. Es gibt verschiedene Arten von Bias, die in Daten vorkommen können, darunter:

- Selektionsfehler: Dies ist der Fall, wenn die verwendete Datenstichprobe nicht repräsentativ für die Grundgesamtheit ist. Wenn die Daten beispielsweise nur in bestimmten Gebieten oder von bestimmten Personengruppen erhoben wurden, repräsentieren sie möglicherweise nicht die Gesamtbevölkerung.
- Messfehler: Dies ist der Fall, wenn die Daten mit Methoden erhoben werden, die nicht genau oder zuverlässig sind. Wenn z. B. die Erhebungsfragen nicht gut konzipiert sind oder wenn Daten mit Instrumenten erhoben werden, die nicht richtig kalibriert sind, kann dies zu Verzerrungen in den Daten führen.
- Bewertungsfehler: Dies ist der Fall, wenn Daten in einer Weise interpretiert werden, die vorgefasste Meinungen oder Hypothesen bestätigt. Wenn ein Forscher zum Beispiel eine bestimmte Theorie beweisen will und nur Daten auswählt, die diese Theorie unterstützen, kann dies zu einer Verzerrung der Daten führen.
- Voreingenommenheit bei der Kennzeichnung: Dies ist der Fall, wenn Daten auf eine Weise gekennzeichnet werden, die nicht neutral oder objektiv ist. Wenn Daten beispielsweise mit subjektiven Begriffen gekennzeichnet werden, kann dies zu Verzerrungen in den Daten führen.

Auch die Daten von data.gv.at können Bias enthalten, weswegen klare Informationen darüber bereitstellt werden, wie die Daten gesammelt, verarbeitet und validiert wurden. Dennoch ist immer wichtig, die Daten vor deren Verwendung zu prüfen und zu verifizieren und mögliche Bias zu berücksichtigen.

3. Technische Details

Das offene Datenportal der österreichischen Bundesregierung, data.gv.at, bietet Zugang zu verschiedenen, frei verfügbaren oder unter Creative Commons Namensnennung 4.0 International lizenzierten Datensätzen über eine Programmierschnittstelle (API). Zur Nutzung der Daten von data.gv.at mit einer API sind folgende allgemeine Schritte notwendig:

- Durchsuchen der Website data.gv.at und finden des gewünschten Datensatzes.
- Auf der Seite des Datensatzes gibt es neben der Möglichkeit, die Daten im vorhandenen Format herunterzuladen auch ein API-Link zu allen Metadaten.
- Dieser API-Endpunkt kann dazu verwendet werden, um Anfragen zum Abrufen der Daten zu stellen. Möglicherweise sind für den Zugriff auf die Daten ein zusätzlicher API-Schlüssel oder andere Authentifizierungsdaten erforderlich.

Hier ist ein Beispiel für die Verwendung der Anfragenbibliothek in Python, um eine Anfrage an einen API-Endpunkt zu stellen und die Daten abzurufen:

```
import requests
api_endpoint = "https://data.gv.at/katalog/api/3/action/package_show?id=<dataset_id>"

response = requests.get(api_endpoint)
data = response.json()
print(data)
```

In diesem Beispiel sollte <dataset_id> durch den eindeutigen Identifikator des Datensatzes ersetzt werden, der abgerufen werden soll. Sobald die Antwortvariable die Daten enthält, ist es möglich, mit Hilfe des json-Formats auf die Datenfelder zuzugreifen und mit diesen zu arbeiten. Genauere Beschreibungen zur Verwendung der Daten finden sich in der Schnittstellendefinition [5].

Data.gv.at stellt neben der API-Schnittstelle auch einen Link auf Anwendungen zur Verfügung, die mittels der offenen Daten erstellt wurden. Insgesamt existieren derzeit 701 Anwendungen, die in folgende Bereiche untergliedert werden:

- Visualisierung (276)
- Web App (243)

- Mobile App (237)
- Datenaufbereitungstool (105)
- Konzept (37)
- Wissenschaftliche Publikation (21)
- Partizipation (19)
- Tutorial (12)
- Tool (10)

Die Registrierung einer neuen Applikation ist über ein dediziertes Formular[3] einfach und unkompliziert möglich.

4. Fazit

In dem vorliegenden Überblickspapier wurden Möglichkeiten aufgezeigt, offene Behördendaten in Österreich für konkrete, dem öffentlichen Sektor nutzenbringende Anwendungsfälle einzusetzen sowie technische Details dafür genannt. Solche Anwendungsfälle sind beispielsweise, dass offene Daten zum Trainieren von maschinellen Lernmodellen verwendet werden können, die Verbrechensmuster vorhersagen oder Bereiche identifizieren, in denen Dienstleistungen am dringendsten benötigt werden. Ein weiterer Vorteil offener Behördendaten besteht darin, dass sie Transparenz und Rechenschaftspflicht fördern, indem sie Bürgern und Organisationen den Zugang zu Informationen über staatliche Aktivitäten und Strategien erleichtern. Außerdem können Dritte auf bestehenden Diensten aufbauen und diese verbessern, wodurch sich möglicherweise neue Geschäftsmöglichkeiten ergeben.

Zusammenfassend lässt sich sagen, dass die Nutzung offener Behördendaten für KI-Anwendungen im öffentlichen Sektor von entscheidender Bedeutung ist, da sie eine große Menge wertvoller Daten liefert, die zum Trainieren von KI-Modellen, zur Verbesserung der Entscheidungsfindung, zur Förderung von Transparenz und Rechenschaftspflicht und zur Schaffung neuer Geschäftsmöglichkeiten genutzt werden können. Offene Behördendaten können auch Innovation, Unternehmertum und die Zusammenarbeit mit dem privaten Sektor fördern, es ist jedoch darauf zu achten, dass keine ungewollten Bias vorhanden sind.

Literatur

[1] C. P. Geiger und J. von Lucke, „Open Government and (Linked) (Open) (Government) (Data)“, JeDEM, Jg. 4, Nr. 2, S. 265–278, 2012, doi: 10.29379/jedem.v4i2.143.

[2] G. Eibl, B. Lutz, P. Parycek und S. Pawel, Rahmenbedingungen für Open Government Data Plattformen. Verfügbar unter: https://neu.ref.wien.gv.at/at.gv.wien.reflive/documents/20189/68315/Rahmenbedingungen_f%25C3%25BCr_Open_Government_Data_Portale_1.3_fin.pdf/51921bf6-cc0a-49c0-bc96-20fc23dcdfba.

[3] Bundesgesetz über die Weiterverwendung von Informationen öffentlicher Stellen, öffentlicher Unternehmen und von Forschungsdaten: Informationsweiterverwendungsgesetz 2022 – IWG 2022, 2022.

[4] BMK, Strategie der Bundesregierung für Künstliche Intelligenz. Artificial Intelligence Mission Austria. 2030.: (AIM AT 2030). [Online]. Verfügbar unter: https://www.bmdw.gv.at/Themen/Digitalisierung/Strategien/Kuenstliche-Intelligenz.html#:~:text=Um%20die%20Chancen%20von%20KI,(AIM%20AT%202030)%20entwickelt.

[5] David Zeilinger, Parlament OGD Schnittstellendokumentation. [Online]. Verfügbar unter: https://www.parlament.gv.at/ZUSD/PDF/Parlament_OGD_Schnittstellen-Dokumentation.pdf.

1. https://www.data.gv.at/infos/cooperation-ogd-oesterreich/ ↩
2. https://www.data.gv.at/ ↩
3. https://www.data.gv.at/applikationen/anwendung-einreichen/ ↩