



Gewusst-wie:
Open Data bei der Stadt Bielefeld

Was ist Open Data?



Wie erreiche ich das Open Data-Portal und wer kann mir bei Fragen helfen?



In welchem Format können Daten für das Open Data-Portal abgegeben werden?



Glossar



Was ist das Open Data-Portal und welche Rolle spielt das Portal für die Stadt Bielefeld?



Wie werden diese Daten bereitgestellt?



Was muss zusätzlich bei Abgabe der Daten beachtet werden?



Kapitel 1

Was ist Open Data?

Als Open Data (aus dem Englischen für ‚offene Daten‘) werden Daten verstanden, die „von allen zu jedem Zweck frei verwendet, weiterverarbeitet und weitergegeben werden können“¹. Ausgeschlossen von Open Data werden Daten, bei denen ein Hindernisgrund² vorliegt.

Unsere Stadt sammelt unterschiedliche Daten in und über Bielefeld. Das können beispielsweise die Baumarten oder Standorte unserer Bielefelder Bäume sein, die Schulwegpläne der Bielefelder Grundschulen oder aber die Bielefelder Wahlergebnisse. Und wir alle können frei über sie verfügen, sie nutzen, weiterverarbeiten oder auch an andere weitergeben.

Gemeinsam und durch den Austausch der Daten und der Zusammenarbeit der gesamten Stadtgesellschaft bringen die offenen Daten Bielefeld näher an das Zielbild einer Smart City. Durch die Arbeit mit Open Data und dem Zusammenwirken der Stadt, wissenschaftlichen Einrichtungen, Bürgerinitiativen oder auch Wirtschaftsunternehmen, können große Projekte auf den Weg gebracht werden, um Bielefeld lebenswerter zu gestalten.

Ziel der Stadt Bielefeld ist die Bereitstellung von allen Verwaltungs-Daten auf dem Open Data-Portal der Stadt Bielefeld, bei denen kein Hindernisgrund vorliegt

¹ Definition der Richtlinie der EU über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors (Richtlinie (EU) 2019/1024).

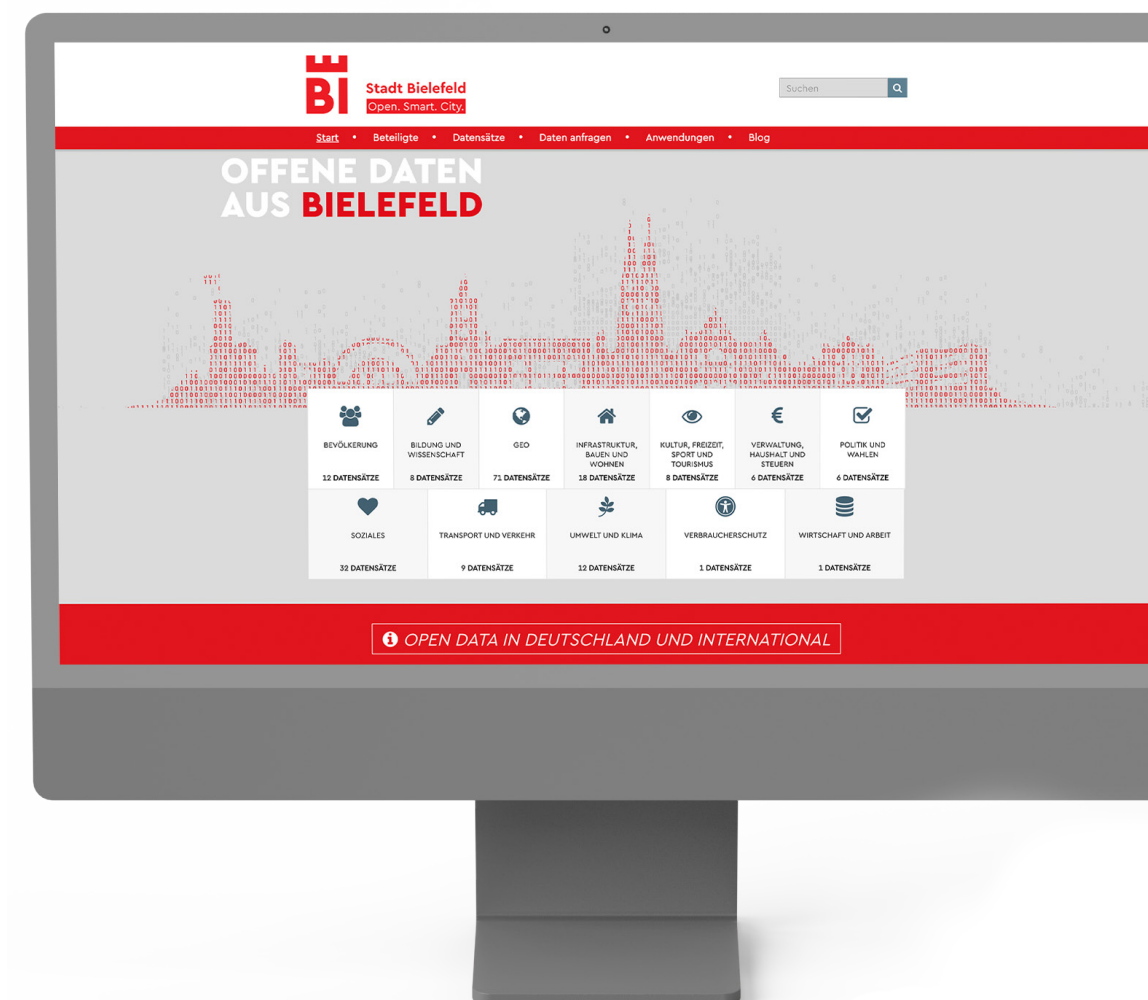
² Ein Hinderungsgrund liegt dann vor, wenn an den Daten kein oder nur ein eingeschränktes Zugangsrecht gemäß §§ 3 bis 6 Informationsfreiheitsgesetz besteht, beispielsweise wenn die Daten Betriebs- und Geschäftsgeheimnisse enthalten oder datenschutzrechtliche Regelungen entgegenstehen.

Kapitel 2

Was ist das Open Data-Portal und welche Rolle spielt das Portal für die Stadt Bielefeld?

Das Open Data-Portal ist ein „Katalog“ und eine „Sammelstelle“ für alle offenen Daten. Über das Open Data-Portal werden Verwaltungsdaten der Stadt Bielefeld bereitgestellt, die für die Öffentlichkeit zugänglich sind.

Die Bereitstellung und Nutzung von Open Data auf einem zentralen Portal bietet viele Mehrwerte für die Verwaltung und darüber hinaus auch für Einwohner*innen sowie Politik, Wirtschaft und Wissenschaft. Die Stadtgesellschaft kann auf alle bereitgestellten Daten an einem Ablageort zurückgreifen.



Für die Wirtschaft ergeben sich möglicherweise Innovationen und Wertschöpfungsgewinne, während die Wissenschaft von aktuellen und standortbezogenen Daten profitiert. Auch für die Zivilgesellschaft ergeben sich neue Möglichkeiten durch beispielsweise bürgerliche Initiativen, welche mit den Daten der Verwaltung neue Lösungen und Angebote im kulturellen oder sozialen Bereich erarbeiten.

Die Daten auf dem Open Data-Portal sind frei für die Besucher*innen der Webseite abrufbar und stehen zum Download bereit. Darüber hinaus bietet das Portal die Möglichkeit, Daten konkret über ein Online-Formular anzufragen.

 Mehr Informationen zum Open Data-Portal und Anwendungsbeispielen von Open Data unter: <https://open-data.bielefeld.de/>



Kapitel 3

Wie erreiche ich das Open Data-Portal und wer kann mir bei Fragen helfen?

Das Open Data-Portal ist online verfügbar unter:

<https://open-data.bielefeld.de/>.

Alle Fragen und Anmerkungen können auch an das zentrale Postfach opendata@bielefeld.de gesendet werden.



**Ansprechpartnerin für allgemeine Fragen
rund um das Thema Open Data:**

Bürge Uprak
opendata@bielefeld.de

Kapitel 4

Wie werden diese Daten bereitgestellt?

Sie entscheiden, welche Daten in das Open Data-Portal aufgenommen werden. Bei den Daten kann es sich sowohl um unverarbeitete **(Roh)daten**, als auch um weiterverarbeitete Daten oder Links zu Inhalten auf weiteren Seiten handeln, bei denen die Daten zur Verfügung stehen. Alle Daten werden mit **Metadaten**³ versehen. Durch die Weiterverarbeitung mittels Anwendung geeigneter Datenschutzverfahren können vorher **personenbezogene und personenbeziehbare Daten** so bereitgestellt werden, dass diese der **Datenschutzgrundverordnung (DSGVO)** entsprechen. Die Weiterverarbeitung bietet eine gute Möglichkeit der **aggregierten Datenbereitstellung**.

Es ist darauf zu achten, dass Dokumente nur einmal abgelegt werden. Von dem **Open Data-Portal** kann auf die Website der Stadt Bielefeld verlinkt werden, sofern das Dokument dort bereits zur öffentlichen Ansicht bereitgestellt worden ist. Zusätzlich ist darauf zu achten, dass ausschließlich eigene Inhalte bereitgestellt werden. Die Links, Daten und **Metadaten** werden durch das Amt für Geoinformation und Kataster, Abschnitt Geodatenmanagement (620.12, Florian Gehler) im Open Data-Portal eingepflegt.

³ Siehe Kapitel 7

Kapitel 5

In welchem Format können Daten für das Open Data-Portal abgegeben werden?

Generell können die Daten in verschiedenen Formaten eingereicht und als Open Data veröffentlicht werden. Einige Dateiformate wie z. B. PDF sind zwar digital, allerdings ist eine Weiterverarbeitung der darin enthaltenen Daten in dem Dateiformat nicht möglich.

Diese Daten sind in der Regel bereits ausgewertet und aufbereitet, sodass kein „**Rohdaten**“-Format vorliegt. Die Daten sind nicht **maschinenlesbar** und können nicht für die automatische Auswertung oder für Entwicklung von Mobilien Apps verwendet werden. Grundsätzlich ist daher ein **maschinenlesbares** Format zu wählen.

Es gelten folgende Format-Empfehlungen⁴:



Textformate



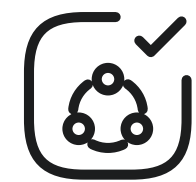
Tabellenformate



Bildformate



Geodatenformate



Schnittstellen

⁴ Weitere Formate können in Abstimmung mit den Open Data-Ansprechpartner*innen eingereicht werden.

Text-, Tabellen und Bildformate

- Comma Separated Value (.csv)
- Hypertext Markup Language (.html)
- JSON (JavaScript Object Notation) (.json)
- Extensible Markup Language (.xml)
- Klassische Textdateien (.txt)
- Weitere Datenformate: data (undefiniert)

Schnittstellen

- Web Map Service (.wms)
- Web feature Service (.wfs)
- Web Map Tile Service (.wmts)

Geodatenformate

- Geopgraphy Markup language (.gml)
- GPS Exchange Format (.gpx)
- Keyhole Markup Language (.kml)
- Shapefile (.shp)

Grundsätzlich sollten die identifizierten Open Data Datensätze in einem der genannten Formate aufbereitet werden. Dabei sind ein paar grundsätzliche Dinge zu beachten

- Im Dateinamen sollten keine Umlaute, Sonderzeichen oder Leerzeichen enthalten sein. Anstatt von Leerzeichen können bspw. Unterstriche _ verwendet werden.
- Werden Dateien regelmäßig aktualisiert, sollten die Daten stringente Namen haben (z. B. bodenrichtwerte_2021.csv / bodenrichtwerte_2022.csv)
- Link-Adressen auf Internetseiten müssen mit vollständigem Pfadnamen angegeben werden (z. B. <https://bielefeld.de>). Abkürzungen wie bielefeld.de oder die Nutzung von www. reicht nicht für eine automatische Link-Erstellung aus. Umlaute, Leer- und Sonderzeichen sind nicht zulässig.

- Es ist eine Codierung zu wählen, die ggf. Umlaute und Sonderzeichen unterstützt. Empfohlen wird eine **UTF-8-Codierung**.

Für Tabellen gibt es weitere Besonderheiten:

- Für tabellarische Daten werden **CSV-Dateien** bevorzugt.
- **CSV-Dateien** sollten durch Semikolons getrennt sein. In den Spaltennamen oder bei Texten in der Tabelle dürfen keine Kommata enthalten sein.
- Spalten sollten verständlich bezeichnet werden, sodass Lesende verstehen, was bspw. die einzelnen Zahlen in einem Feld bedeuten. Können die Spalten innerhalb der Tabelle durch den Namen nicht eindeutig interpretiert werden, können diese in den **Metadaten** (siehe Kapitel 7) erläutert werden.
- In den Spalten sollte weder am Anfang noch am Ende ein Leerzeichen stehen.
- Spalten- und Zeilen müssen alle deutlich erkennbar sein.

- Die Spalten sollten „sauber“ formatiert sein. Für Zahlen sollte z. B. das jeweilige Zahlenformat gewählt werden und nicht „Text“ o. Ä.

Auf bestehende offene Datensätze, die schon im Internet verfügbar sind, kann im Open-Data-Portal über eine Adresse (URL) verwiesen werden. Es kann auch eine **Schnittstelle (API)** angegeben werden. So können auch Anwendungen (z. B. eine bielefeldKARTE) oder Verweise auf vorhandene Quellen oder Dienste wie z. B. WMS oder WFS eingebunden werden. Kann die **CSV-Datei** beispielsweise über einen WFS eingebunden werden, kann man auf diese **Schnittstelle** verweisen. Hier muss der vollständige Pfad angegeben werden (mit https://...).



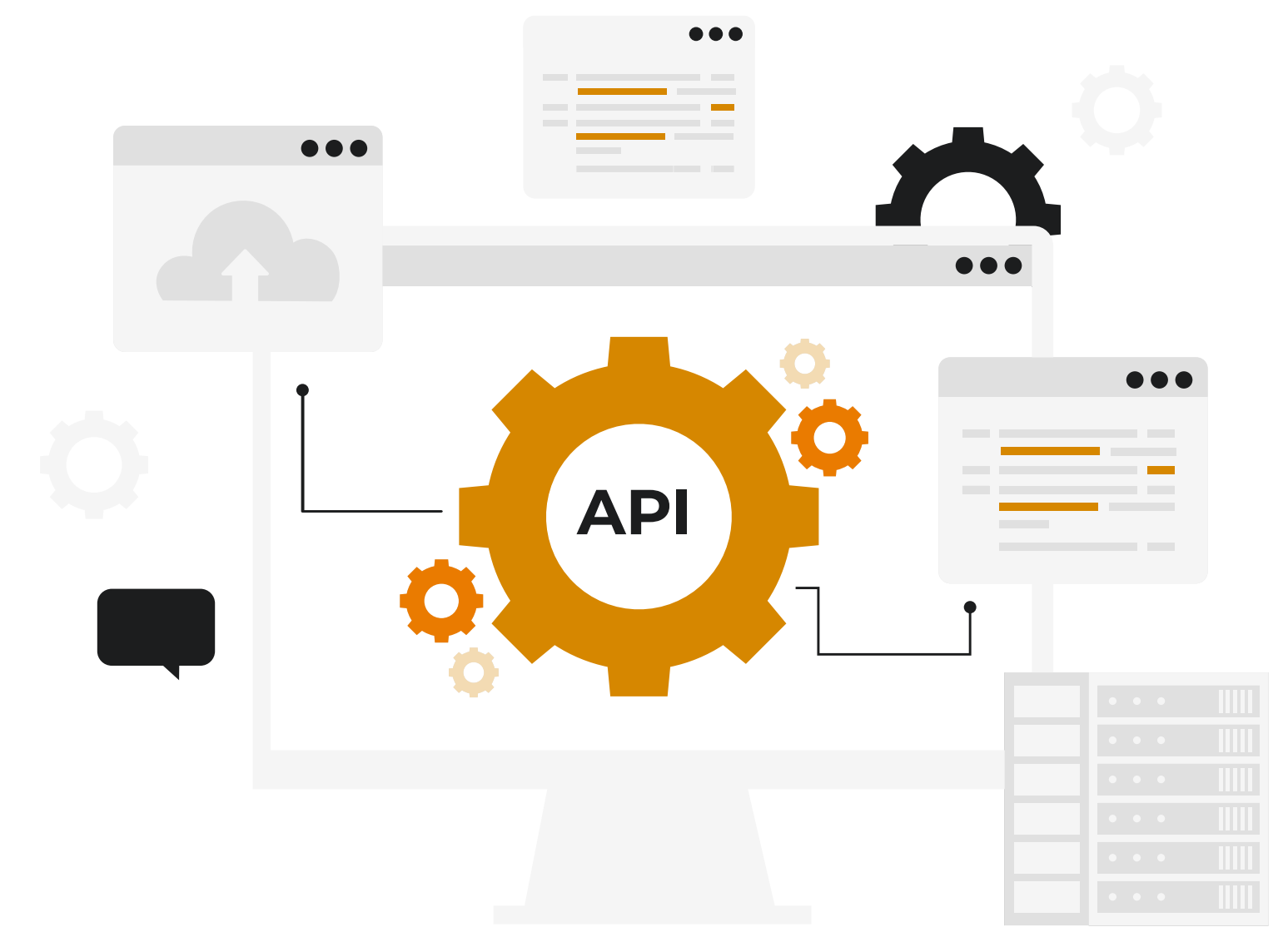
Bei Rückfragen wenden Sie sich an
Bürge Uprak opendata@bielefeld.de

Kapitel 6

Was muss zusätzlich bei Abgabe der Daten beachtet werden?

Bei der Veröffentlichung von Datensätzen und Daten müssen zusätzlich **Metadaten** mitangegeben werden. **Metadaten** sind Informationen über die eigentlichen Datensätze und Daten. Sie beschreiben die Datensätze und Ressourcen und machen sie öffentlich auffindbar. Grundsätzlich gilt: Je detaillierter und vollständiger die Daten fachlich und technisch beschrieben werden, desto höher ist der Mehrwert für potenzielle Nutzer*Innen.

Grundsätzlich müssen Angaben zum **Datensatz** gemacht werden. Zusätzlich sind alle mit dem **Datensatz** verbundenen Daten / Ressourcen (z. B. **CSV-Dateien**, XML-Dokumente) zu beschreiben.



Die Tabellen in Anlage 1 (Hinweise zum Ausfüllen der Metadaten) geben klare Anweisungen zu den erforderlichen **Metadaten** und Mindestanforderungen für die Veröffentlichung von Datensätzen. Tabelle 1, „**Allgemeine Angaben für den gesamten Datensatz** (z. B. mehrere **CSV-Dateien**, **unterschiedliche Dateien oder Links zu einem Thema**)“, enthält Informationen, die den allgemeinen **Datensatz** betreffen, der zur Verfügung gestellt wird. Dies könnte relevant sein, wenn mehrere Ressourcen zu einem bestimmten Thema eingereicht werden und Tabelle 1 daher Angaben zu allen diesen Ressourcen umfasst.

Im Gegensatz dazu gibt Tabelle 2, „**Angaben für jede Datei im Datensatz** (z. B. eine **CSV-Datei**, **URL**, **Dokument**)“, detaillierte Informationen zu

den einzelnen Ressourcen. Diese Tabelle gibt spezifische Angaben zu jeder einzelnen Ressource, die möglicherweise mehrere Datensätze zu einem Thema umfasst. Die Mindestanforderungen in beiden Tabellen sind durch ein ***** als Pflichtfelder gekennzeichnet und müssen ausgefüllt werden, um eine erfolgreiche Veröffentlichung Ihrer Datensätze zu gewährleisten.

Zusätzlich muss vor der Abgabe die datenschutzrechtliche Freigabe ausgefüllt werden. Diese finden Sie in der Anlage 2 (Freigabe-Erklärung).

Anlage 1

Anlage 2

Bitte reichen Sie Ihre Daten nach Klärung der datenschutzrechtlichen Freigabe in dem richtigen Dateiformat gemeinsam mit der ausgefüllten **Anlage 1** (Hinweise zum Ausfüllen der Metadaten) und **Anlage 2** (Freigabe-Erklärung) bei Bürge Uprak per E-Mail ein. Bei Rückfragen können Sie sich jederzeit bei Bürge Uprak wenden.



**Ansprechpartnerin für allgemeine Fragen
rund um das Thema Open Data:**

Bürge Uprak
opendata@bielefeld.de

Kapitel 7

Glossar

Aggregierte Datenbereitstellung

Die „aggregierte Datenbereitstellung“ bezieht sich auf den Prozess der Zusammenführung und Präsentation von Daten in einer kombinierten, übersichtlichen Form. Dabei werden oft mehrere einzelne Datensätze oder Datenelemente zu einem Gesamtbild zusammengefasst. Durch die Aggregation können große Mengen von Rohdaten in statistische Zusammenfassungen oder andere aggregierte Formen umgewandelt werden, um Trends, Muster oder allgemeine Erkenntnisse leichter erkennbar zu machen. Diese Praxis wird häufig in verschiedenen Branchen und Bereichen wie Wirtschaft, Wissenschaft, Technologie und

Finanzen angewendet, um Entscheidungsfindung und Analyse zu erleichtern. Aggregierte Daten bieten eine höhere Ebene der Abstraktion, was die Interpretation und Kommunikation von Informationen erleichtert.

CSV-Dateien

CSV steht für „Comma-Separated Values“ (komma-getrennte Werte) und bezeichnet ein einfaches Dateiformat zum Speichern von tabellarischen Daten. In CSV-Dateien werden Daten in Zeilen und Spalten organisiert, wobei jeder Dateneintrag durch ein Semikolon oder ein anderes definiertes Trennzeichen von den anderen

Einträgen getrennt ist. Jede Zeile in der Datei repräsentiert einen Datensatz, und die Spalten repräsentieren die verschiedenen Attribute oder Felder dieses Datensatzes. CSV-Dateien sind plattformunabhängig und einfach zu erstellen, zu lesen und zu bearbeiten, weshalb sie weit verbreitet für den Austausch von Daten zwischen verschiedenen Anwendungen und Systemen verwendet werden.

Datenhaltende Stelle (Amt / Institution / Person)

Die Datenhaltende Stelle ist die Organisation, Einrichtung oder Person, die für die Speicherung, Verwaltung und Sicherung von Daten verantwortlich ist. Diese Stelle trägt die Verantwortung für den sicheren und zuverlässigen Zugriff auf die Daten sowie die Einhaltung von Datenschutzrichtlinien.

Die Datenhaltende Stelle spielt eine entscheidende Rolle bei der Pflege der Datenintegrität und -verfügbarkeit.

Datenschutz Grundverordnung (DSGVO)

Die Datenschutz-Grundverordnung (DSGVO) ist eine europäische Verordnung, die am 25. Mai 2018 in Kraft getreten ist und den Schutz personenbezogener Daten innerhalb der Europäischen Union (EU) stärken soll. Die DSGVO legt Regeln für die Verarbeitung personenbezogener Daten fest und gewährt den Bürgern der EU mehr Kontrolle über ihre persönlichen Informationen. Zu den zentralen Grundsätzen gehören die Einwilligung der betroffenen Personen zur Datenverarbeitung, das Recht auf Information über die Verarbeitung, das Recht auf Zugang zu den eigenen Daten, das

Recht auf Berichtigung und Löschung von Daten, sowie Maßnahmen zur Sicherheit und zum Schutz der Privatsphäre. Die DSGVO gilt nicht nur für in der EU ansässige Unternehmen, sondern auch für Organisationen außerhalb der EU, die personenbezogene Daten von EU-Bürgern verarbeiten. Verstöße gegen die DSGVO können mit erheblichen Geldbußen geahndet werden.

Datenurheber*in (Amt/Institution/Person)

Datenurheber*in ist die Person, Gruppe oder Organisation, die die ursprünglichen Daten erstellt oder generiert hat. Diese Rolle ist oft mit der Schöpfung oder Sammlung der Daten verbunden. Der/die Datenurheber*in kann auch als die Quelle der Informationen betrachtet werden und hat möglicherweise das Recht, Entscheidungen

über die Nutzung und Weitergabe der Daten zu treffen. Es ist wichtig zu beachten, dass der/die Datenurheber*in nicht unbedingt identisch mit der datenhaltenden Stelle sein muss; die Daten können von einer Stelle erstellt und von einer anderen verwaltet werden.

Datensatz

Ein Datensatz ist eine strukturierte Sammlung von Informationen, die zu einem bestimmten Thema oder Objekt gehören. Er besteht aus einzelnen Elementen, auch als Dateneinträge bezeichnet, die verschiedene Attribute oder Merkmale repräsentieren können. Diese Attribute enthalten Daten wie Zahlen, Texte oder andere Informationen. Datensätze werden häufig in Datenbanken gespeichert und können für verschiedene Zwe-

cke verwendet werden, wie beispielsweise für Analysen, Berichte oder um Informationen zu organisieren.

Rohdaten

Rohdaten sind ursprüngliche, unbehandelte und nicht verarbeitete Daten, die direkt aus den Quellen stammen. Diese Daten sind im Wesentlichen die grundlegenden, unveränderten Informationen, bevor sie einer weiteren Verarbeitung, Filterung oder Aggregation unterzogen werden. Rohdaten sind oft detailliert und bieten eine umfassende Sicht auf die Informationen, ohne dass sie in irgendeiner Weise verändert wurden. Die Veröffentlichung von Rohdaten im Rahmen von Open Data-Initiativen ermöglicht es Nutzer*innen auf den unmittelbaren Ursprung der Informa-

tionen zuzugreifen. Dies fördert Transparenz und erlaubt Anwender*innen, ihre eigenen Analysen, Interpretationen oder Visualisierungen durchzuführen. Die Offenlegung von Rohdaten ist entscheidend, um die Integrität und Glaubwürdigkeit von Open Data zu gewährleisten, da sie den Prozess der Datenherkunft und -sammlung transparent macht und gleichzeitig die Möglichkeit bietet, unterschiedliche Analyseansätze zu verfolgen.

Metadaten

Metadaten sind zusätzliche Informationen, die beschreiben, erläutern oder Kontext zu den veröffentlichten Daten bieten. Metadaten dienen dazu, die Verständlichkeit, Auffindbarkeit und Interpretation von Datensätzen zu verbessern. Sie bieten

Informationen über Aspekte wie Datenherkunft, Erstellungsdatum, Verwendungsbeschränkungen, Dateiformat, und andere relevante Eigenschaften, die es ermöglichen, die Daten effektiv zu nutzen. Metadaten spielen eine Schlüsselrolle, indem sie Nutzer*innen helfen, die Qualität und den Inhalt der bereitgestellten Daten zu verstehen, bevor sie sie herunterladen oder verwenden. Sie erleichtern auch das Auffinden spezifischer Datensätze innerhalb großer Datensammlungen. In Open Data-Initiativen sind klare, standardisierte Metadaten von entscheidender Bedeutung, um die Interoperabilität⁵ zwischen verschiedenen Datensätzen und Plattformen zu gewährleisten und somit die breite Nutzung und Wiederverwendbarkeit der Daten zu fördern.

Maschinenlesbar

Maschinenlesbar bedeutet, dass eine Datei automatisiert durch Software weiterverarbeitet werden kann. Das bedeutet nicht unbedingt, dass sie auch von Menschen lesbar ist. So ist ein PDF-Dokument zwar gut für den Menschen lesbar, doch handelt es sich hier um kein maschinenlesbares Format. Das PDF-Dokument müsste erst aufwendig konvertiert werden, um es lesbar für Software zu machen. Durch die Maschinenlesbarkeit kann eine Weiterverarbeitung der Daten ohne großen Aufwand erfolgen.

Personenbezogene und Personenbeziehbare Daten

„Personenbezogene Daten“ beziehen sich auf Informationen, die direkt einer identifizierten oder identifizierbaren natürlichen Person zugeordnet werden können, wie beispielsweise Name, Adres-

⁵ Interoperabilität bezeichnet die Fähigkeit von Systemen oder Software, nahtlos miteinander zu kommunizieren und Daten auszutauschen. Sie ermöglicht die effiziente Zusammenarbeit verschiedener Technologien oder Plattformen.

se oder Geburtsdatum. „Personenbeziehbare Daten“ hingegen beziehen sich auf Informationen, die nicht direkt mit einer bestimmten Person in Verbindung stehen, aber durch zusätzliche Maßnahmen oder Informationen mit einer konkreten Person verknüpft werden können. Personenbeziehbare Daten erfordern möglicherweise einen weiteren Kontext oder zusätzliche Daten, um eine Identifizierung zu ermöglichen. Bezogen auf das Open Data-Portal dürfen daher mehrere Datensätze, die auf dem Portal bereitgestellt werden, in Kombination nicht personenbeziehbar sein.

Schnittstelle / API

Der Begriff „Schnittstelle“ beschreibt einen definierten Punkt, über den externe Benutzer*innen oder Anwendungen auf die veröffentlichten Da-

ten zugreifen können. Diese Schnittstelle dient als Mittel zur Interaktion zwischen den Daten und den Anwendungen oder Nutzern*innen, wodurch der Austausch und die Nutzung von Informationen erleichtert werden.

Eine Open Data-Schnittstelle kann verschiedene Formen annehmen, darunter APIs (Application Programming Interfaces) oder andere Mechanismen, die den systematischen Zugriff auf die veröffentlichten Daten ermöglichen. Diese Schnittstellen stellen sicher, dass die Daten in standardisierter Form abgerufen werden können, erleichtern die Automatisierung von Prozessen und fördern die Integration in verschiedene Anwendungen und Plattformen. Eine gut gestaltete Open Data-Schnittstelle trägt zur Barrierefreiheit,

Benutzerfreundlichkeit und Wiederverwendbarkeit von Daten bei, indem sie einen einfachen und standardisierten Zugang zu den veröffentlichten Informationen ermöglicht.

UTF-8-Codierung

UTF-8 ist eine Methode, Texte auf Computern zu speichern und darzustellen. Es ermöglicht die Verwendung von Buchstaben und Zeichen aus vielen verschiedenen Sprachen. Besonders praktisch ist, dass UTF-8 sowohl für einfache Buchstaben als auch für komplexere Schriftzeichen genutzt werden kann, was es vielseitig einsetzbar macht. Es ist eine Art von „Übersetzung“, die es Computern ermöglicht, Texte in unterschiedlichen Sprachen richtig zu verstehen und anzuzeigen, was wichtig ist, wenn wir im Internet surfen oder digitale Dokumente erstellen.



**Stadt
Bielefeld**

Digitalisierungsbüro

Alter Markt 13

33602 Bielefeld

Telefon: 0521 51 – 5187

digitalisierungsbuero@bielefeld.de

Amt für Geoinformation und Kataster

August-Bebel-Str. 92 (3. Etage)

33602 Bielefeld

Telefon: 0521 51 – 3137, – 3139

geoinformationundkataster@bielefeld.de

Verantwortlich für den Inhalt:

Jens Edler, Dietmar Hermenau

Redaktion: Bürge Uprak, Florian Gehler

Stand: 02 / 2024