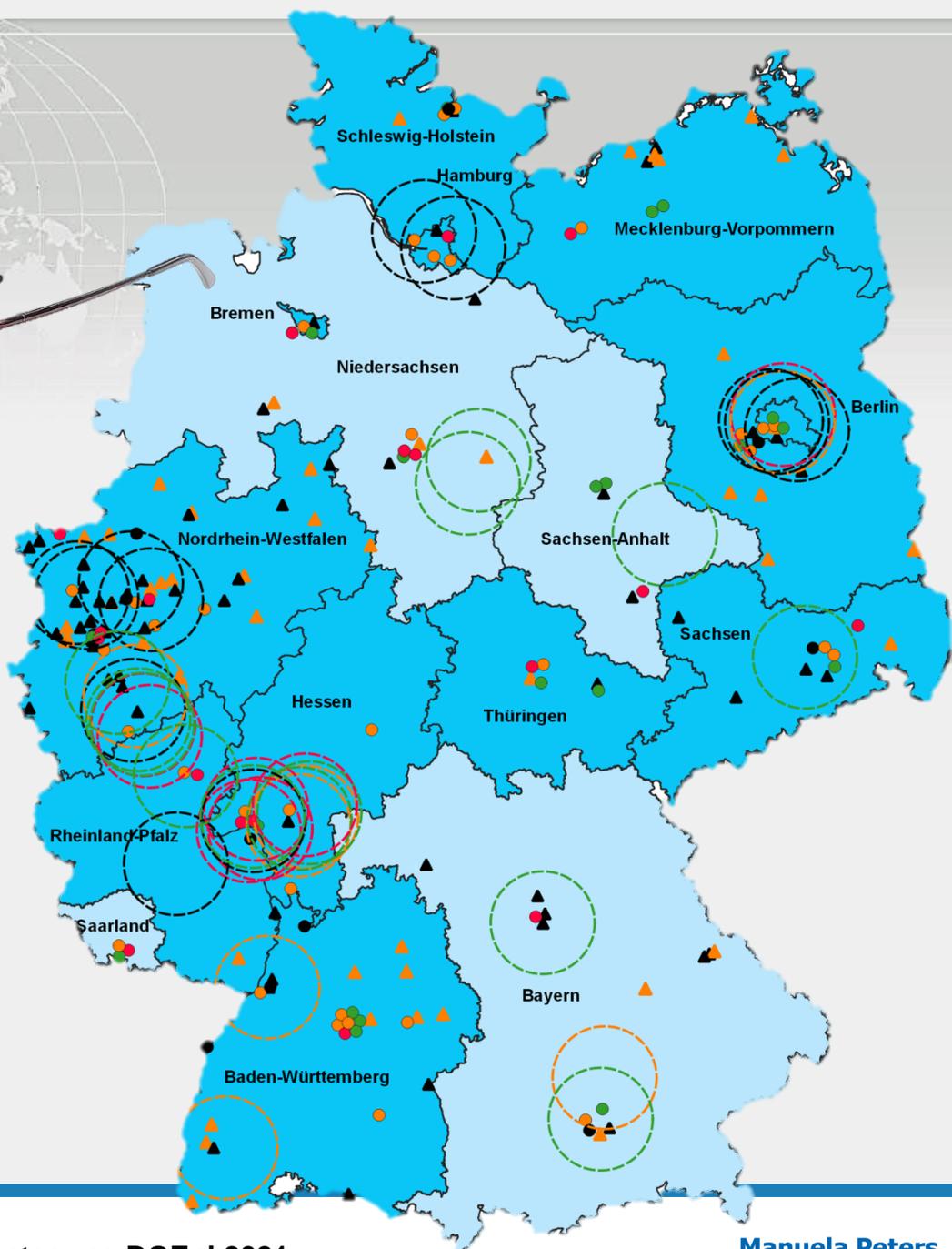




VERFÜGBARKEIT VON OPEN DATA IM KONTEXT VON RAUM UND GESUNDHEIT

Manuela Peters & Hajo Zeeb

*Leibniz-Institut für
Präventionsforschung und
Epidemiologie – BIPS*



Hintergrund: *Zusammenhang Raum & Gesundheit (& Daten)*

- Einflussfaktoren auf die menschliche Gesundheit = sozioökonomische Bedingungen, Umweltexpositionen, Qualität & Charakter d. natürlichen & bebauten Umgebung, Zugang zu/Erreichbarkeit von gesundheitsrelevanten Ressourcen, etc. → **geografische Variation**
- Ziel/ Maxime: Schaffung gleichwertiger, gesundheitsförderlicher Lebensverhältnisse vor Ort & Prävention gesundheitlicher Beeinträchtigungen (Health in all Policies-Ansatz, Ottawa Charta [Eriksson & Lindström 2008])
- Ein Monitoring des örtlichen Status Quo und spezifische Maßnahmen, im Sinne des „Public Health Action Cycle“, sind auf **datenbasierte Informationen** angewiesen

In letzten Jahren sind Portale für **offene Daten** (inkl. Verwaltungs- & Geodaten) entstanden. Auch Forschungseinrichtungen & (staatliche & private) Unternehmen & Bürgerinitiativen stellen Daten zur Verfügung.

= **Potenzial für die raumbezogene Public Health - Forschung**

Hintergrund: *Offene Daten als Ressource*

Offen sind Daten - in technischer und rechtlicher Hinsicht – vor allem dann, wenn sie **maschinenlesbar** und **strukturiert**, unter **Verwendung offener Nutzungsrechte** von **jedermann frei** über eine offene **Schnittstelle** verwendet, **nachgenutzt** und **verbreitet** werden können - und zwar zu **jeglichem Zweck**, ohne **Einschränkungen, Diskriminierungen und Gebühren** [Open Knowledge Foundation 2019].

Kriterien offener Daten [Sunlight Foundation 2010].

Kriterium	Beschreibung der Anforderungen
Vollständigkeit	so vollständig wie möglich, soweit dies die Regelungen zum Datenschutz zulassen; möglichst im Rohformat, inklusive beschreibender Metadaten
Primärquellen	aus Ursprungsquellen, mit Informationen zur Erhebung und Aufbereitung
Zeitliche Nähe	Veröffentlichung zeitnah zur Entstehung oder Aktualisierung der Daten
Leichter Zugang	Vermeidung/ Minimierung von Hindernissen, z.B. durch technische Probleme, Dateiformate, fehlende Metadaten, organisatorisch Barrieren (z.B. Pflichtregistrierung), rechtliche Vorgaben
Maschinenlesbarkeit	Verwendung einheitlicher, maschinenlesbarer Formate
Diskriminierungsfreiheit	Zugriff ohne zeitliche Restriktionen oder notwendige Registrierungen, keine Einschränkung auf ausgewählte Softwareanwendungen
Verwendung offener Standards	Verwendung von Formaten, die nicht nur von ausgewählten Programmen gelesen und verarbeitet werden können, keine Notwendigkeit der Abführung von Lizenzkosten an Software-Hersteller
Nutzungsbestimmungen/ Lizenzierung	klare und einheitliche Nutzungsbedingungen, eindeutig sichtbare Lizenz, die Nutzungsvoraussetzungen festlegt
Dauerhaftigkeit	permanente Verfügbarkeit, nachvollziehbare Dokumentation von Änderungen, Aktualisierungen und Löschungen mit Versionskontrolle und Archivierung
Nutzungskosten	keine Bereitstellung gegen Entgelt

→ effiziente Weiternutzung, wenn Daten in guter Qualität, als Rohdaten, bestenfalls inkl. Metadaten + klar definierten Nutzungsrechte vorliegen

Hintergrund: *technische & rechtliche Offenheit*

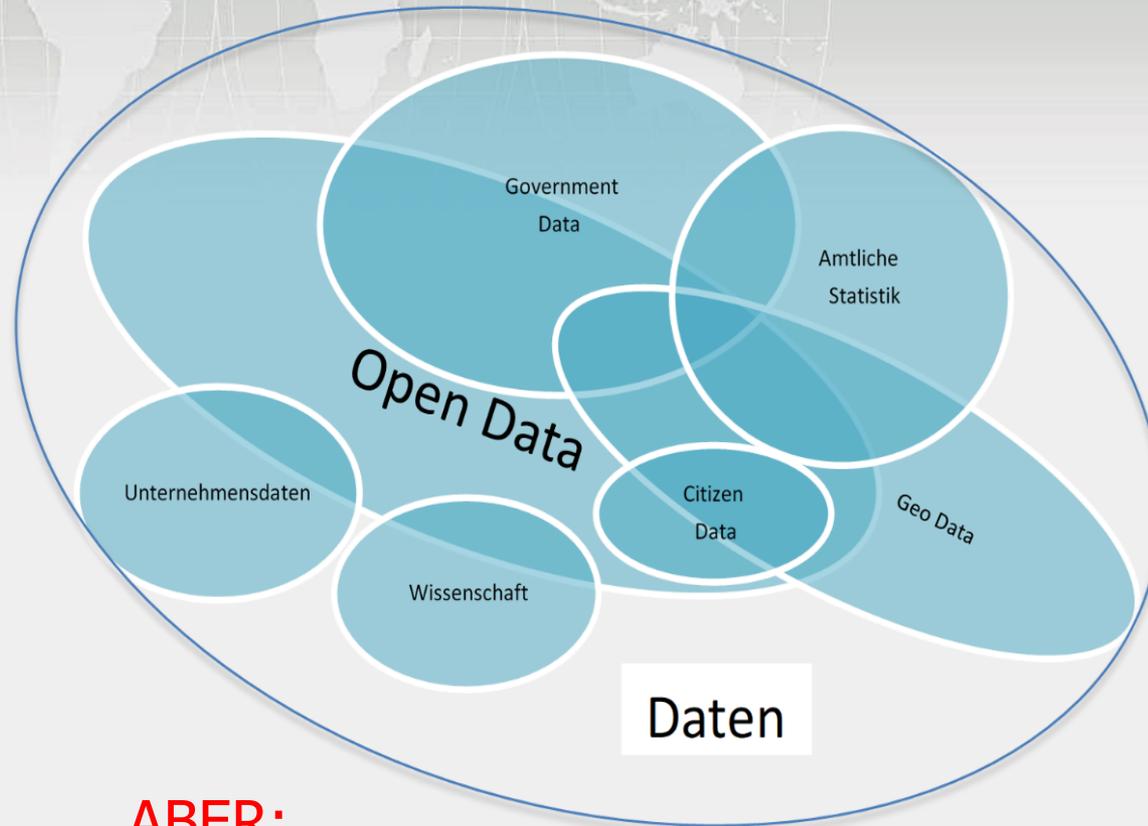
technisch: Formate, d. Weiternutzung mit proprietärer & freier Software erlauben

rechtlich: viele öffentl. Daten nur unter Beachtung ausgewiesener Bestimmungen nutzbar, z.B. jener, die kommerzielle Nutzung, Bearbeitung & Weitergabe erlauben → für gesetzl. geschützte personenbezog. Daten ist Nachnutzung überwiegend nicht gestattet

Rechtliche Offenheit: gängigste Lizenzen und Anwendungskontexte	
Creative Commons (CC)	Gilt nicht ausschließlich für Daten und erlaubt, Werke zu kopieren, zu verbreiten und anderweitig zu nutzen. CC ist weltweit einsetzbar, gilt für die Dauer des Urheberrechtes und wird häufig für Geodaten genutzt. Lizenzgeber können Erlaubnisse hinzufügen, um Bedingungen festzulegen. Demnach unterscheiden sich die Lizenzmodelle in den Versionen 1.0 bis 4.0, wobei Letztere für den internationalen Kontext gilt und z.B. von EU-Datenbereitstellern genutzt wird. <i>https://creativecommons.org/licenses</i>
Open Database License (ODC)	liegt in folgenden Versionen vor: <ul style="list-style-type: none">• Open Data Commons Public Domain Dedication and License (PDDL): Verzicht auf alle Urheberrechte an Daten• Open Data Commons Attribution License (ODC-BY): Teilen und Bearbeiten erlaubt; Namensnennung erforderlich• Open Data Commons Open Database License (ODC-ODbL): Teilen und Bearbeiten erlaubt; Namensnennung und Weitergabe unter gleichen Bedingungen <i>http://opendatacommons.org/licenses/odbl/</i>
Datenlizenz Deutschland	Hauptsächlich für OGD genutzt, liegt aktuell in folgenden Versionen vor: <ul style="list-style-type: none">• Datenlizenz Deutschland – Namensnennung – Version 2.0 – verpflichtet zur Nennung des Bereitstellers nach dessen Maßgabe und des Lizenznamens• Datenlizenz Deutschland – Zero – Version 2.0 - ermöglicht eine einschränkungslose Weiterverwendung <i>https://www.govdata.de/web/guest/lizenzen</i>
Technische Offenheit : offene und maschinenlesbare Dateiformate	
Formate für Texte, Tabellen & Bilder	<i>.txt, .csv, .html, .xml, .rdf, .odt, .ods, .rss, .pdf, .doc, docx, .xls, .xlsx, .rtf, .gif, .jpg, .png, .tif, .geotiff</i>
Formate für Geodaten	<i>.gml, .gpx, .kml, .dxf, .dwg, .shp, .shx, .dbf, .svg</i>

Anm.: kursiv: maschinenlesbar, fett: offenes Format, fett + kursiv: offen und maschinenlesbar

Hintergrund: *Potenziale für die Gesundheitsforschung*



Digitalisierung →

Datenangebot, Zahl d. Anbieter & Möglichkeiten zum Data Linkage stark gestiegen

ABER:

Deutschland liegt im internationalen Vergleich bei Nutzung dieses Potenzials für Public Health-Forschung zurück [Burgun et al. 2017].

Hintergrund: *Probleme*

- 1) Auffindbarkeit offener Daten:** unübersichtlicher Angebotsmarkt [Kubicek & Lippa 2015]: Zugänge & Angebote variieren nach Raumeinheit, Themengebiet, Zugangsart
 - nur einige Kommunen & Regionalverbände betreiben OD-Portale, zudem nur teilw. in übergreifende Kataloge eingebunden/ untereinander vernetzt
 - kommunale GBE ist länder-/kommunenspezifisch ausgerichtet – kein gemeinsamer Grundstandard
- 2) Verfügbarkeit offener Daten:** für kleinste räuml. Einheiten aus Datenschutzgründen keine VÖ von Rohdaten → Aufbereitung/ Aggregation je nach Verwendungszweck
- 3) Technische Barrieren:** offene Formate fehlen (z.B. FAIRe Daten → Auffindbarkeit, Zugänglichkeit, Interoperabilität & Wiederverwertbarkeit [Wilkinson et al. 2016]
- 4) Rechtliche Barrieren:** unterschiedliche Nutzungsbedingungen/ Lizenzmodelle
- 5) Probleme der Verknüpfung:** unterschiedl. Verwaltungsgrenzen und Aggregationen erschweren Vergleichbarkeit und Data Linkage mit Sachdaten
- 6) Föderale Strukturen:** heterogene Datenumfänge, -Zugänge und rechtliche Rahmen (teils eigene oder fehlende Informationsfreiheit- & Datenschutzgesetze)

Was war unser Ziel?

Überblick zu Datenangeboten für Betrachtung von Raum und Gesundheit auf verschiedenen adm. Ebenen geben, die **barrierefrei, ohne technische und rechtliche Einschränkung genutzt** und **weiterverbreitet** werden können

Konzeptioneller Rahmen

- zur Strukturierung von Daten → ganzheitliches humanökologische Modell [Barton & Grant 2006], das Einflussfaktoren aus sozialen, ökologischen und ökonomischen Systemen aufzeigt
- Dimensionen entsprechen Aspekten, die auch WHO [Neira & Prüss-Ustün 2016], neben individuellen (Lebensstil-) Faktoren, als unerlässlich für Wohlergehen identifiziert hat: strukturelle Bedingungen des Lebensumfelds wie:
 - (1) **soziodemografische und -ökonomische Faktoren,**
 - (2) **Faktoren des Zugangs und der Erreichbarkeit,**
 - (3) **die physische Umgebung** und
 - (4) **Klima- und Umweltfaktoren.**

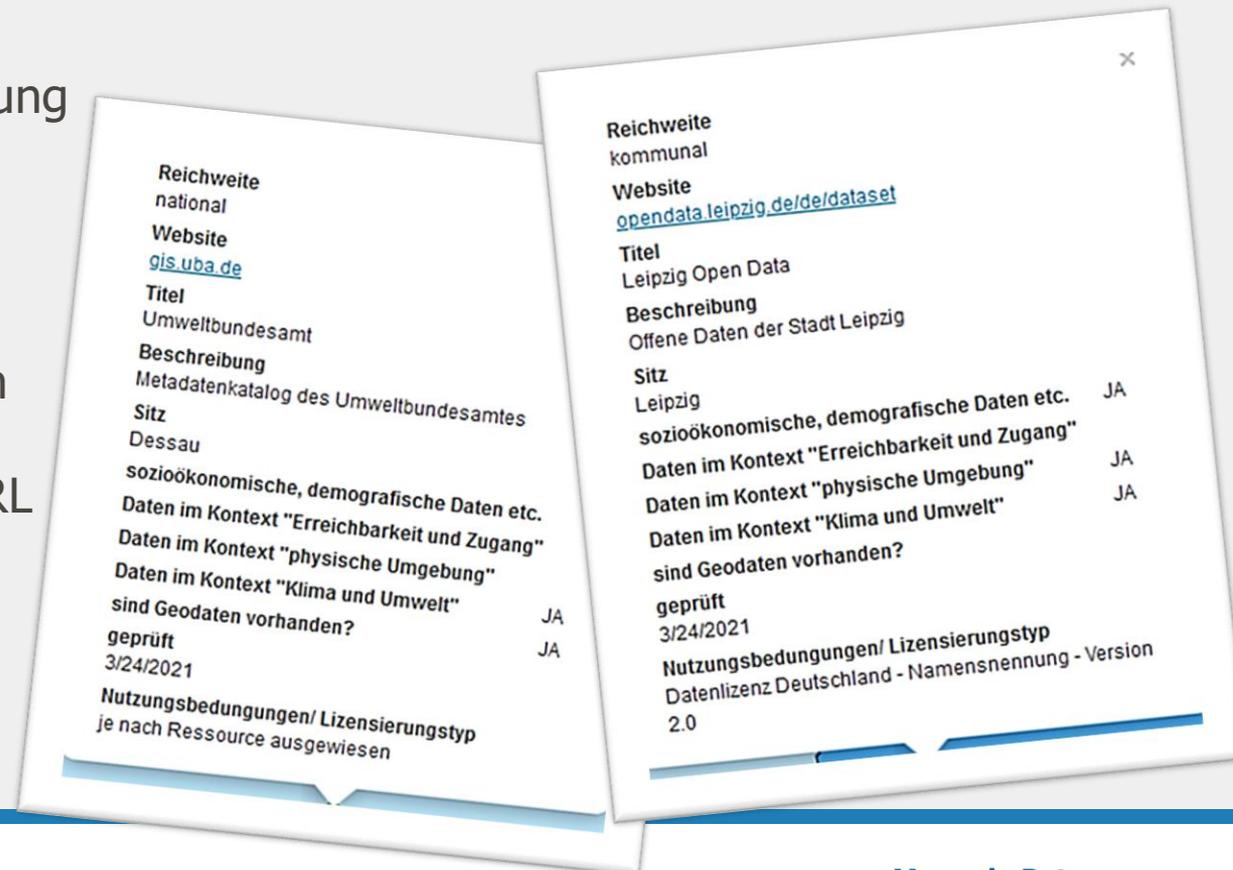
Wie sind wir vorgegangen?

Auswahl relevanter
Determinanten & Daten
orientiert an
theoretischen
Rahmenmodellen

Dimension/	Indikator
Faktoren der sozialen Umwelt <i>soziodemografische und sozioökonomische Faktoren, Urbanisierungsgrad soziale Infrastruktur, Gemeindewohl und Sicherheit</i>	EW-Dichte, Siedlungsdichte
	Durchschnittsalter der Bev., Anteil 65- 79-Jährige/ Bev., Altenquotient, Wanderungssaldo, Ausländeranteil, Bevölkerungsentwicklung
	Steuereinnahmen pro EW , kommunale Schulden, BIP
	Arbeitslosigkeit, Beschäftigtenquote, SGB II-Quote
	Bildung
	Bruttolohn & -gehalt, verfügb. HH-Eink.
	Öffentliche Bibliotheken
	Kulturstätten (Kinos, Theater, Museen , Clubs, etc.)
	Breitbandinternet
	Wohnfläche pro Person
	Verletzte/ Getötete im Straßenverkehr
	Verfügbarkeit Fußwege & Radwege
	Kriminalität/Straftaten
	ÖPNV-Infrastruktur/ ÖV Haltestellen
Verfügbarkeit, Zugang und Erreichbarkeit von Gesundheitsressourcen <i>medizinische Versorgung, Grün- und Blauflächen, Sporteinrichtungen</i>	PKW-Besatz
	Reisezeiten Mittelzentren, Oberzentren
	Erreichbarkeit IC/ ICE BH
	Erreichbarkeit von Supermärkten
	EW je Arzt
	Erreichbarkeit Ärzte, Apotheken
	Plätze in Pflegeheimen
	Anzahl & Erreichbarkeit von Krankenhäusern mit Regelversorgung
	Krankenhausbetten je 1000 EW
	öffentliche Grünflächen, Wasserflächen, Erholungsflächen, Gewerbe- und Industrieflächen, Verkehrsflächen
Physische Umgebung	Temperaturexposition (Hitze, Kälte)
	UV-Strahlung
	Bodenversiegelung
	Feinstaub, Ozon, CO ₂ , Stickoxide
	Lärm
	Trinkwasserqualität
	Radon
	Radon
Klima und Umwelt <i>Luftverschmutzung, Klimaexposition</i>	

Wie sind wir vorgegangen?

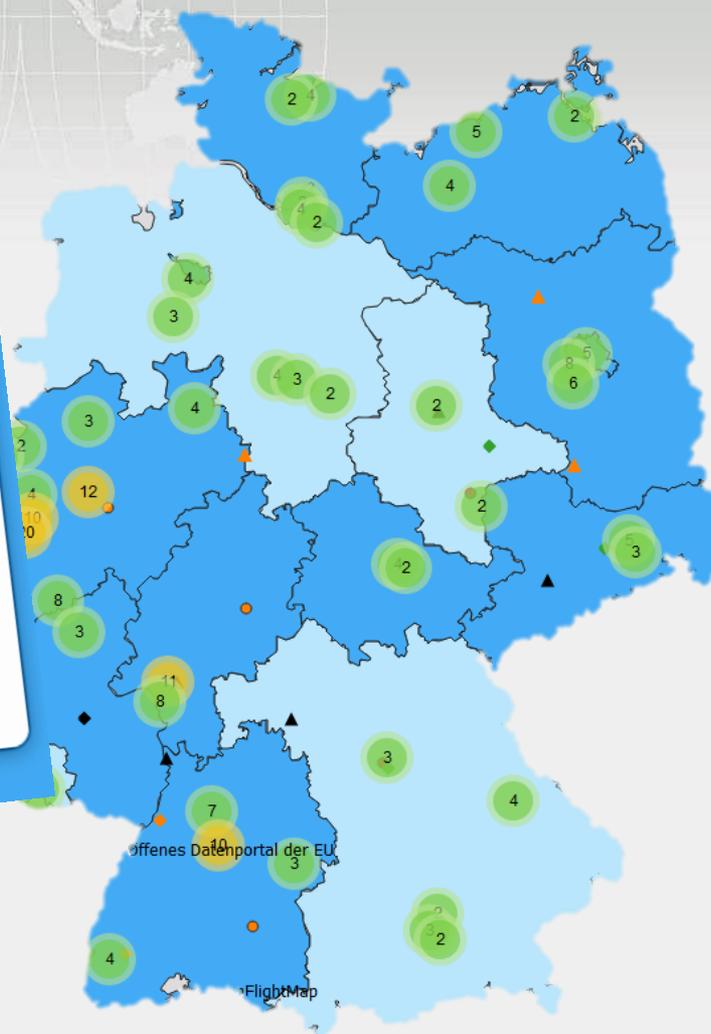
- einfache Webrecherchen, Nachverfolgen von Verweisen auf themenrelevanten Websites, Katalogen und Datenportalen, Anpassung und Erweiterung bereits vorhandener, offen zur Verfügung gestellter Datenübersichten & Webvisualisierungen (z.B. Open Data Atlas, OpenDataPortal) auf eigenen Themenschwerpunkt angepasst
- Entwicklung & Bereitstellung einer themenspezifischen Onlinekarte mit Informationen zum Themengebiet, zu rechtlichen & technischen Voraussetzungen, zur Reichweite der Daten, URL des Datenbereitstellers



Wie sieht unser Ergebnis aus?

Open Data for Health Map - offene Daten im Kontext von Raum und Gesundheit

Reichweite kommunal
Website geo.dortmund.de
Titel Geoportal Dortmund
Beschreibung Zugang zu raumbezogene Daten der Stadt Dortmund.
Sitz Dortmund
sozioökonomische, demografische Daten etc.
Daten im Kontext "Erreichbarkeit und Zugang" JA
Daten im Kontext "physische Umgebung" JA
Daten im Kontext "Klima und Umwelt" sind Geodaten vorhanden?
geprüft 3/24/2021
Nutzungsbedingungen/ Lizenzierungstyp Datenlizenz Deutschland - Namensnennung - Version 2.0



- International
- Portal für amtliche Daten
- Fachdatenportal (z.B. Umwelt, Klima)
- Geodatenportal
- Opendatenportal oder -Ressource
- national
- Portal für amtliche Daten
- Fachdatenportal (z.B. Umwelt, Klima)
- Geodatenportal
- Opendatenportal oder -Ressource
- regional (Bundesländer, Metropolregionen)
- Portal für amtliche Daten
- Fachdatenportal (z.B. Umwelt, Klima)
- Geodatenportal
- Opendatenportal oder -Ressource
- kommunal (Städte und Gemeinden)
- Portal für amtliche Daten
- Fachdatenportal (z.B. Umwelt, Klima)
- Geodatenportal
- Opendatenportal oder -Ressource
- Bundesländer
- BL in GOVdata.de enthalten
- BL, die GOVdata.de nicht begetreten sind

[Link zur Webvisualisierung](#)

Erkenntnisse/ Fazit

- auf versch. räuml. Ebenen stellen einige Akteure z.T. umfangreiche, andere A. thematisch nicht vergleichbare/ keine OD bereit → Menge & Qualität variieren
- Von >11.000 Kommunen in D. veröffentlichen 2021 nur ca. 146 Daten, davon mit teils weniger als 10 Datensätze *[Bertelsmann Stiftung 2021]*
- viele Daten prinzipiell verfügbar, setzen aber Wissen um Zugänge voraus
- aktuelles Angebot schwer bestimmbar - keine zentrale/offizielle Sammelstelle
- Vielzahl an D. grundsätzl. zugänglich, aber nicht maschinenlesb., offene Formate

Was tun?

- Für Suchmaschinenabfragen Indizierung d. Daten nötig (z.B. d. Metadaten)
- FAIRifizierung vorantreiben (z.B. Vereinbarung von Gemeinschaftsstandards, Ontologien & Fachvokabular) um Auffindbarkeit zu erhöhen
- Nutzung v. Open Citizen Projekten, wie OSM: Gewinn bezügl. Umfang & Zugang, **ABER** z.T. defizitär (Vollständigkeit, Qualität, Aktualität, Validität)
- (harmonisierte) Gesetzesgrundlagen auf Ebene der BL und Kommunen nötig
- **Anreize für offene Daten schaffen !!**

Referenzen

- Eriksson M, Lindström B (2008) A salutogenic interpretation of the Ottawa Charter. Health promot Int 23(2):190-199
- Open Knowledge Foundation (2019) Das ABC der Offenheit. <https://okfn.de/blog/2019/11/abc-der-offenheit/>. Zugriffen: 9. September 2021
- Sunlight Foundation (2010) Ten Principles for Opening Up Government Information. <https://sunlightfoundation.com/wp-content/uploads/sites/2/2016/11/Ten-Principles-for-Opening-Up-Government-Data.pdf> . Zugriffen: 9. September 2021
- Burgun A, Bernal-Delgado E, Kuchinke W et al. (2017) Health data for public health: Towards new ways of combining data sources to support research efforts in Europe. Yearbook of medical informatics 26(1):235-240
- Kubicek H, Lippa B (2015) Open Data: Appelle reichen nicht. Kommune 21(5):12-13
- Wilkinson MD, Dumontier M, Aalbersberg IJ et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific data 3(1):1-9
- Barton H, Grant M (2006) A health map for the local human habitat. Journal of the Royal Society for the Promotion of Health 126(6):252-253
- Neira M, Prüss-Ustün A (2016) Preventing disease through healthy environments: A global assessment of the environmental burden of disease. Toxology Letters 259:1
- Turcis T (2021) Open Data Atlas. <http://opendata.tursics.de/> . Zugriffen: 9. März 2021
- OpenGeoEdu (2020) OpenDataPortal. <https://portal.opengeoedu.de/> . Zugriffen: 29. März 2021
- Bertelsmann Stiftung (2021) Pressemitteilung. Musterdatenkatalog schafft erstmals Übersicht über offene Daten in Deutschlands Kommunen (03.03.2021). <https://www.bertelsmann-stiftung.de/de/themen/aktuelle-meldungen/2021/maerz/musterdatenkatalog-schafft-erstmal-uebersicht-ueber-offene-daten-in-deutschlands-kommunen> Zugriffen: 16. September 2021



Vielen Dank für Ihre Aufmerksamkeit!

Kontakt

Manuela Peters

Leibniz-Institut für Präventionsforschung und
Epidemiologie - BIPS
mpeters@leibniz-bips.de