

"Wie erhöhen wir den Nutzen offener Daten in den nächsten 10 Jahren?"

In der Sitzung vom 30. September 2021 und 28. März 2022 hat die Cooperation OGD Österreich Themen erarbeitet, die für die Zukunft von Open in Data in Österreich wesentlich sind, weil Open Data ein wesentlicher Treiber für Verwaltungsmodernisierung, Innovation und Digitale Transformation ist.

Folgende Themen wurden in 4 Sub-Arbeitsgruppen (AG) vorrangig bearbeitet, die Ergebnisse liegen für die Sitzung der Cooperation OGD Österreich am 19. Juni 2023 vor und wurden im vorliegenden Dokument zusammengefasst:

Sub-AG 1 ... Datenqualitätsstandards und Feedbackschleifen (Seite 2-3)

Sub-AG 2 ... Identifier für Daten (Seiten 4-15)

Sub-AG 3 ... Data Spaces und wie sind diese bestmöglich zu nutzen (Seite 16-22)

Sub-AG 4 ... Zukunft Cooperation OGD Österreich (Seite 23-26)

Von jeder Arbeitsgruppe wurde eine zusätzlich eine Präsentation der Ergebnisse erstellt, die gemeinsam mit diesem Dokument diversen Gremien (z.B. BLSG, IKT-Bund, Städtebund – FIT-Tagung, Netzwerk der Open Data Beauftragten, etc.) zur Verfügung gestellt und von den Mitgliedern der Cooperation OGD Österreich zur Verbreitung der Ergebnisse genutzt werden kann.

Datenqualitätsstandards und Feedbackschleifen

Gerhard Hanzal Gerhard.Hanzal@bmdw.gv.at

Gerhard Hartmann gerhard.hartmann@wien.gv.at

Bernhard Krabina bernhard.krabina@ikt.linz.at

Michael Reichart michael.reichart@brz.gv.at

1.) Ausgangslage

Derzeit werden beim Import nach data.gv.at über die XML Metadaten-Schnittstelle überprüft, ob alle Felder ausgefüllt sind. Auch bei der Eingabe über das Webinterface werden die Pflichtfelder überprüft.

Beim JSON Harvesting und CKAN API finden keine Prüfungen statt. Es gibt auch keine Routine, die bestehenden Metadaten zu überprüfen.

2.) Vorschlag für Verbesserungen

- Linkchecker für Ressourcen und Links in den Metadaten - Report an Datenbereitsteller und an Fachgruppe data.gv.at
 - Auswahl, für welchen Bereich ich die Auswertung will
 - Möglichkeit, die Daten als csv zu exportieren
- Metadaten überprüfen - Report an Datenbereitsteller und an Fachgruppe data.gv.at
 - Auswahl, für welchen Bereich die Auswertung gemacht werden soll
 - Möglichkeit, die Daten als csv zu exportieren
 - Was soll überprüft werden:
 - Titel – Am Ende die entsprechende Organisation z.B. Kindergärten in Wien
 - Kategorie – Nur die Kategorien laut Metadatenstandard gültig
 - Schlagworte: Einheitlich erster Buchstabe groß Rest klein
 - Datensatz Format: Nur gültige Formate laut Metadatenstandard , bei PDF muss bei den Schlagworten OpenDocument angeführt sein
 - Datenverantwortliche Stelle: Muss ausgefüllt sein
 - Kontaktseite der datenverantwortlichen Stelle: Soll ausgefüllt sein
 - Geographische Abdeckung/Lage: Soll ausgefüllt sein
 - Geographische Ausdehnung: Soll ausgefüllt sein
 - Aktualisierungszyklus. Auswahl laut Metadatenstandard
 - Titel und Beschreibung Englisch: Soll ausgefüllt sein
 - Kontaktseite der veröffentlichenden Stelle: Soll ausgefüllt sein
 - Veröffentlichtende Stelle – E-Mailkontakt: Soll ausgefüllt sein, um ein Feedback geben zu können
- Linked Data - Identifier - Beispiele auf bestehende Vokabulare - URI aufs Vokabular
 - Use Case Umsetzungen
 - Infrastruktur zur Verfügung stellen
 - schema.org - Linz überlegt, das mitzuspeichern
- Musterdatenkatalog - URIs bei den eigenen Datensätzen verlinken - Weiterführende Metadaten eintragen
- Ansprechpartner – sollen mit Kontaktdaten erhoben und periodisch überprüft werden.
- Feedbackmöglichkeit:
 - Forumslösung für Datenbereitsteller
 - Art „Selbsthilfegruppe“

- Bewertung der Datensätze analog dem europäischen Datenportal
 - Im ersten Schritt nur für Datenbereitsteller und Fachgruppe data.gv.at sichtbar.
 - Später auch vorstellbar, analog dem europäischen Datenportal, das auch extern anzuzeigen.
- Alle Umsetzungen auch für opendataportal.at

3.) Umsetzung in der Fachgruppe data.gv.at im Mai2023 beauftragt

1.6.2023 – 31.12.2023

• Paket 1

- Qualitäts-Dashboard im Backend
- Möglichkeit zur Pflege Organisations-Daten (User, Ansprechpartner, ...)
- Auflistung von Fehlern
- Email-Information (nach Ausführung)
- Regelmäßige/manuelle Ausführung Überprüfungen

• Paket 2

- Linkchecker für Ressourcen-Files
- Überprüfung Metadatenstandard lt. Papier Sub-AG Cooperation OGD

Details der Überprüfung Metadatenstandard

- > Titel – Am Ende die entsprechende Organisation z.B. Kindergärten in Wien
- > Kategorie – Nur die Kategorien laut Metadatenstandard gültig
- > Schlagworte: Einheitlich erster Buchstabe groß Rest klein
- > Datensatz Format: Nur gültige Formate laut Metadatenstandard, bei PDF muss bei den Schlagworten „OpenDocument“ angeführt sein
- > Datenverantwortliche Stelle: Muss ausgefüllt sein
- > Kontaktseite der datenverantwortlichen Stelle: Soll ausgefüllt sein
- > Geographische Abdeckung/Lage: Soll ausgefüllt sein
- > Geographische Ausdehnung: Soll ausgefüllt sein
- > Aktualisierungszyklus: Auswahl laut Metadatenstandard
- > Titel und Beschreibung Englisch: Soll ausgefüllt sein
- > Kontaktseite der veröffentlichenden Stelle: Soll ausgefüllt sein
- > Veröffentlichende Stelle – E-Mailkontakt: Soll ausgefüllt sein, um ein Feedback geben zu können

1.1.2024 – 30.6.2024

• Paket 3

- Auswertung analog zu data.europa.eu (nur intern):
- Accessibility, Reusability, Contextuality, Findability, Interoperability, Distribution Quality

• Paket 4

- Darstellung der Datenqualität im Frontend („Adequate-Badge“, Sterne, analog zu data.europa.eu, Datum letzte Überprüfung, ...)

• Paket 5

- Vorschläge zur Verbesserung Datenqualität
- Anbindung Musterdatenkatalog
- Feedbackmöglichkeit/Community Bereitsteller
- Linked Data - Identifier
- Genauere Definition des Paketes notwendig

Identifizier für Daten

Autoren

- Gregor Eibl
- Bernhard Krabina
- Erik Obersteiner
- Michael Reichart
- Wolfgang Tinkl

Inhaltsverzeichnis

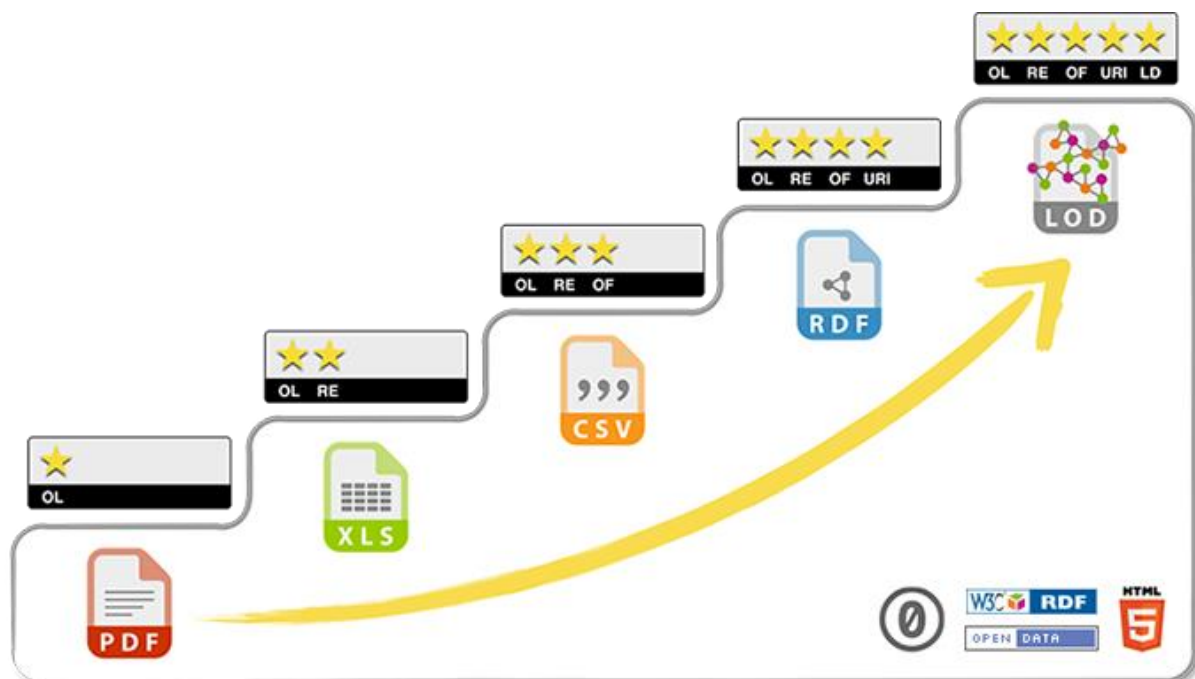
Motivation.....	5
Zielsetzung.....	6
Beschreibung der betrachteten Merkmale	7
Eindeutigkeit und Maschinenlesbarkeit	7
Persistenz.....	7
Einheitliches Schema.....	7
Menschenlesbarkeit.....	7
Auflösbarkeit.....	8
Betrachtungsebenen.....	8
Ebene 1: Metadaten	8
Ebene 2: Ressourcen	8
Ebene 3 Datenobjekt.....	9
Empfehlungen	10
Fazit: Lösungsvorschlag unter Verwendung von data.gv.at.....	11
ANHANG I: Hintergrundinformationen.....	12
DCAT.AP 2.0.....	12
Kontrollierte Vokabulare.....	12
URL-Shortener/Weiterleitungs-Services	12
URI-Schema aus Inspire.gv.at	13
DOI Verwendung am Beispiel BEV	14
ANHANG II: Vorgehensweise bei der Erstellung des vorliegenden Dokumentes.....	15

Motivation

Die Bereitstellung langfristig verfügbarer Daten erlangt zunehmend Bedeutung. Einerseits sollen oder müssen die Daten selbst, entweder aus gesetzlicher Notwendigkeit oder aus Gründen der Transparenz und Nachvollziehbarkeit abrufbar bleiben, andererseits sollen auch die dazugehörigen Metadaten auffindbar bleiben. Für Linked Open Data (LOD) ist das, neben den technischen Schnittstellen der dafür notwendigen Dienste oder APIs, eine Grundvoraussetzung der Vernetzung über das Internet.

Tim Berners-Lee, schlug ein **5-Sterne-Schema** für offene Daten vor. Für den ersten Stern muss das Material im Internet unter einer offenen Lizenz zur Verfügung gestellt werden. Für jeden weiteren Stern sind die Anforderungen der vorangegangenen Sterne plus die der neuen Stufe erforderlich. Stern 2 verlangt strukturierte Daten (z. B. Excel statt eines Bildscans einer Tabelle). Stern 3 erfordert das Öffnen in einem nicht-proprietären offenen Format (z. B. CSV statt Excel). Stern 4 umfasst URIs (webbasierte Identifikatoren) zur Identifizierung für physische oder digitale Dinge, damit andere auf den Inhalt verweisen können. Stern 5 schließlich fordert die Verknüpfung der eigenen Daten mit anderen Daten, um einen Kontext zu schaffen.¹

Das 5-Sterne Open-Data-Modell ist eine gute Methode, um zu messen, wie gut die Daten in das ‚Web‘ integriert sind. Daten, die im Internet zu finden sind, sind Daten, die von einer Webseite heruntergeladen werden können. Das **„Netz der Daten“** hingegen ist eine Reihe eng miteinander verknüpfter Datenressourcen, die sowohl von Menschen als auch von Maschinen gelesen und verstanden werden können. Um dieses Netz der Daten zu schaffen sind vorerst Daten notwendig, die untereinander verlinkt werden können.² Persistente Identifier (Star 4) sind daher Grundvoraussetzung für Linked Open Data (Star 5).



¹ <https://5stardata.info/en/>

² <https://data.europa.eu/elearning/de/module13/#/id/co-01>

Beim **Semantischen Web** geht es nicht nur darum, Daten ins Netz zu stellen. Es geht darum, Verknüpfungen herzustellen, so dass ein Mensch oder eine Maschine das Netz der Daten erkunden kann. Mit verknüpften Daten können Sie, wenn Sie einige davon haben, andere, verwandte Daten finden. Wie das Hypertext-Web wird auch das Datenweb aus Dokumenten im Web aufgebaut. Im Gegensatz zum Hypertext-Web, wo Links Beziehungsanker in HTML-Dokumenten sind, sind sie bei Daten Links zwischen beliebigen, durch RDF beschriebenen Dingen. Die URIs bezeichnen jede Art von Objekt oder Konzept.³

URIs sollten **langlebig** sein, d.h. URIs werden einer Ressource dauerhaft zugewiesen, sind also stabil, ändern sich nicht und verschwinden nicht im Laufe der Zeit. Die Langlebigkeit wird durch eine vertrauenswürdige zugrunde liegende Web-Infrastruktur gewährleistet. Data.gv.at stellt die langfristige Web-Infrastruktur zur Verfügung und stellt somit auch einen vertrauenswürdigen Dienst dar, der vom Datenersteller unabhängig ist. Alternativ könnte dieses Service auch von weiteren vertrauenswürdigen externen Diensten zur Verfügung gestellt werden. Dieser Dienst sollte bei Bedarf auch leicht auf andere Organisationen übertragbar sein. Es ist nicht notwendig, nur einen Dienst zu nutzen, aber ein Dienst ist in der Regel einfacher zu verwalten und kostengünstiger.⁴

Tim Berners-Lee empfiehlt für Linked Data die Verwendung von HTTP-URIs als Identifier. Seit den Anfängen des Web werden allerdings aus verschiedenen Gründen immer wieder neue URI-Schemata (und Unterschemata innerhalb des urn:-Schemas) wie LSIDs, Handles, XRIs, DOIs usw. erfunden. In der Regel geht es dabei darum, sich nicht an das etablierte Domain Name System (DNS) zu binden, sondern etwas unter eigener Kontrolle zu schaffen.⁵

Identifier spielen auch eine wichtige Rolle bei der Umsetzung der **FAIR (Findable, Accessible, Interoperable, Reusable)-Prinzipien**⁶, indem sie dazu beitragen, digitale Objekte eindeutig zu identifizieren, leichter auffindbar und zugänglich zu machen, die Interoperabilität zu verbessern und die Wiederverwendbarkeit von digitalen Objekten zu erhöhen.

Zielsetzung

Grundlegendes Ziel beim Einsatz von URIs in Offenen Daten ist die Erreichung des 4. Sterns gemäß Grafik oben, damit eine weitere Stufe in Richtung von Linked Open Data erreicht werden kann.

Da es bei der Gestaltung von URIs zwar eine vorgegebene Syntax⁷, jedoch – abgesehen von IANA-registrierten URI-Schemata⁸ - keine inhaltlichen Vorgaben für die Zusammensetzung der URI gibt, bestehen Freiheitsgrade für das, was mit dem Einsatz von URIs erreicht werden soll.

Aus Sicht der Cooperation OGD Österreich (Sub-AG URI) stellen die folgend aufgelisteten Merkmale die Voraussetzung für die Erreichung der Zielsetzung dar. Die Reihung von 1 bis 5 stellt die aus Sicht der Sub-AG URI gegebene Priorität und Relevanz dieser Merkmale:

- Eindeutigkeit und Maschinenlesbarkeit von URIs

³ <https://www.w3.org/DesignIssues/LinkedData.html>

⁴

https://data.europa.eu/sites/default/files/d2.1.2_training_module_2.3_persistent_uri_design_and_management_de_edp.pdf

⁵ <https://www.w3.org/DesignIssues/LinkedData.html>

⁶ Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

⁷ Zur URI Syntax siehe https://en.wikipedia.org/wiki/Uniform_Resource_Identifier

⁸ https://en.wikipedia.org/wiki/List_of_URI_schemes

- Persistenz der URIs
- Einheitliches Schema von URIs
- Menschenlesbarkeit von URIs
- Auflösbarkeit der URIs

Diese Begriffe werden im Folgenden zum gemeinsamen Verständnis kurz erläutert.

Beschreibung der betrachteten Merkmale

Eindeutigkeit und Maschinenlesbarkeit

Die Eindeutigkeit ist ein Merkmal, welches jedenfalls durch eine URI gewährleistet werden muss. Zur Herstellung der Eindeutigkeit kann z.B. die UUID als Teil des URI verwendet werden.

Beispiele:

- <https://www.data.gv.at/katalog/dataset/stadtischer-baumkataster> (Eindeutigkeit wird von data.gv.at sichergestellt)
- <https://www.data.gv.at/katalog/dataset/df4cc7ae-ead7-49aa-bcb8-b018f169162d> (Eindeutigkeit durch UUID)

Persistenz

Die Persistenz wird als wichtiges zu erreichendes Ziel gesehen. Erst durch Persistenz und somit Dauerhaftigkeit der URIs und in der Erreichbarkeit der Online Ressourcen (URLs), kann ein auf Dauer verlässlich funktionierendes System geschaffen werden. Die Persistenz kann unter anderem mit "Persistent Identifier" (PIDs)⁹ hergestellt werden. Sie stellt z.B. langfristig unveränderte/unveränderbare http(s)-URLs zu den Daten sicher. Persistenz kann mit Tools und Systemen sichergestellt werden. Zitate auf persistente Links stellen sicher, dass die Nutzer:innen auf die korrekte Quelle zugreifen können.

Einheitliches Schema

Ein einheitliches Schema bei URIs auf der Seite der Datenprovider kann selten umgesetzt werden, in einzelnen Fällen aber sinnvoll sein. Bei Bedarf kann auf die Systematik von Inspire (Details siehe Anhang) zurückgegriffen werden.

Beispiel:

- [**https://data.inspire.gv.at/Geodatenstelle/Ressourcenbezeichner/Datenthema.FeatureType/localID \(/Version\)**](https://data.inspire.gv.at/Geodatenstelle/Ressourcenbezeichner/Datenthema.FeatureType/localID (/Version))

Menschenlesbarkeit

Als menschenlesbar bezeichnet man eine Informationsdarstellung, die ohne Hilfsmittel von Menschen gelesen werden kann. Somit sind auch URIs mit UUID grundsätzlich menschenlesbar (im Gegensatz zu einem Barcode oder QR-Code). Im Folgenden verstehen wir Menschenlesbarkeit aber im Sinne von "interpretierbaren" URIs im Gegensatz zu beliebigen Zeichenfolgen.

Beispiel:

- https://www.data.gv.at/katalog/dataset/stadt-linz_hotspotstandorte
- <https://doi.org/10.48677/6c316854-5453-4cf5-82cb-6586b2a2aa06>

⁹ https://de.wikipedia.org/wiki/Persistent_Identifier

Auflösbarkeit

Unter Auflösbarkeit versteht man die Eigenschaft einer URI, direkt im Web adressierbar zu sein (URL).

Betrachtungsebenen

Für die Zielsetzungen (Persistenz, Eindeutigkeit, Auflösbarkeit, Menschenlesbarkeit, Einheitliches Schema) ergeben sich drei unterschiedliche Betrachtungsebenen:

1. **Metadaten:** Metadaten beschreiben einen veröffentlichten Datensatz inklusive der Ressourcen. Metadatenportal ist data.gv.at, die Metadaten werden laut Metadatenstandard der Cooperation OGD Austria beschrieben.
2. **(Download)URL der Datenressource:** In der Regel erfolgt eine Verlinkung von den Metadaten auf eine Ressource, die üblicherweise nicht direkt auf data.gv.at liegt. Die Struktur und Qualität dieser Download bzw. Zugriffs-URLs liegt in der Verantwortung des Dateneinbringers.
3. **Datenobjekte:** Im Datenobjekt können persistente Identifier (z. B. Gemeindekennziffer, laufende interne Nummer) zur Identifikation der einzelnen (Daten-)Objekte als Schlüsselfelder oder auch URIs (z. B. zu kontrollierten Vokabularen, beispielsweise von SEMIC.eu) als Fremdschlüssel dienen, um Linked Open Data zu erzeugen. Bei eigenen Datenfeldern kann es sinnvoll sein, die eigene ID in Kombination mit einem Namespace als URIs oder URLs zur Identifikation zu verwenden.

Ebene 1: Metadaten

Eindeutigkeit und Persistenz ist auf data.gv.at für Datensätze sichergestellt (Ressourcen-URIs sind derzeit aus technischen Gründen nur teilweise persistent).

Standard sind auf data.gv.at maschinenlesbare, persistente UUIDs, z. B:

- <https://www.data.gv.at/katalog/dataset/f660cf3f-afa9-4816-aafb-0098a36ca57d>

Es gibt dazu auch eine menschenlesbare URL, in obigem Beispiel:

- <https://www.data.gv.at/katalog/dataset/baumkataster>

Beide URLs werden jeweils bei der Anlage automatisch erzeugt und auf Eindeutigkeit intern geprüft. Im Gegensatz zur UUID ist allerdings der menschenlesbare Link nicht persistent, weil sich der Name des Datensatzes jederzeit ändern kann. Dafür ist diese Angabe besser für die Kommunikation geeignet.

Auf data.gv.at gibt es sowohl für hochgeladene als auch für verlinkte Ressourcen eindeutige UUIDs. Beispiel:

- <https://www.data.gv.at/katalog/dataset/baumkataster/resource/d0db9035-0eb4-47d7-9d58-128bfe373063>

Ebene 2: Ressourcen

Der Aufruf der jeweiligen Ressource führt dann zu einem konkreten Link, z. B:

https://data.linz.gv.at/katalog/umwelt/baumkataster/2022/FME_BaumdatenBearbeitet_ODG_2022_0202.csv

Diese URL ist im Verantwortungsbereich des Dateneinbringers. In diesem Kapitel werden Beispiele und Empfehlungen zusammengefasst, die die Verwaltung dieser URLs der dateneinbringenden Stelle betreffen.

Es kann hier je nach Ressource unterschiedliche Arten von URLs geben:

1. URL zum Download eines einzelnen Files
2. URL zu einem auf data.gv.at hochgeladenen File (wird automatisch erzeugt und hier nicht weiter thematisiert)
3. URL zu einer API

Für einzelne Files wird eine zumindest lokal eindeutige URL in der Regel automatisch erzeugt, beispielsweise ist data.linz.gv.at eine Subdomain von linz.gv.at - der Domain die eindeutig für die dateneinbringende Stelle ist und von dieser verwaltet wird. Die weiteren Unterverzeichnisse sind zur freien Verwendung, hier ist nur wichtig, dass möglichst keine Änderungen im Laufe der Zeit vorgenommen werden, um die URL persistent zu halten. Am Ende des Verzeichnissystems (im Fall Linz über einen FTP-Server realisiert), befindet sich dann der Dateiname. Da sich Namensbezeichnungen etwa von Bundesministerien mitunter regelmäßig ändern, ist es empfehlenswert für allenfalls neu aufzubauende Datenrepositories gv.at-Domainnamen zu verwenden, die unabhängig von der Ressortbenennung sind. Beispielsweise ems.gv.at für Daten aus dem Epidemiologischen Meldesystem.

APIs bieten Daten über das Internet als Webservice an. Sie enthalten in der Regel eine fixe URL, an welche Parameter übergeben werden, die die Ausgabe des Webservices beeinflussen. Beispielsweise könnte als Parameter an eine URL ein ?format=json angehängt werden, um die Ausgabe im JSON-Format zu erhalten. Für die Erstellung von REST-Schnittstellen ist es empfehlenswert, der Spezifikation der OpenAPI Initiative zu folgen: <https://www.openapis.org>. (Andere Implementierungsmöglichkeiten sind [SOAP](#), [RPC](#) oder [WebDAV](#).)

Ebene 3 Datenobjekt

Verlinkung auf externe IDs

Die Verwendung von URIs innerhalb von Datenobjekten (z. B. Zeilen in einer Tabelle) ist ein weiterer Schritt in Richtung Linked Open Data und von den Dateneinbringern selbständig vorzunehmen. Die Weiterverwendung nationaler bzw. internationaler Vokabulare ist vom Thema des Datensatzes abhängig. Details siehe Anhang "kontrollierte Vokabulare".

Verwendung eigener IDs

Unabhängig von Veröffentlichungen ist es sinnvoll, in eigenen Datenbeständen möglichst eindeutige IDs zu verwenden. Beispielsweise könnte das im Baumkataster eine Baumnummer sein. Aus so einer internen ID kann dann relativ einfach eine URI aufgebaut werden:

<http://data.linz.gv.at/katalog/umwelt/baumkataster/BaumNr/00068>

Es muss sich hierbei nicht um eine URL handeln, die Daten ausliefert, sondern um eine URI, die ein Objekt in einem Datensatz bezeichnet. Auf diese Art und Weise könnte also der Eintrag mit der "Baumnummer" 68 im Baumkataster in Daten adressiert werden, der einen bestimmten Baum bezeichnet. So lange der Baumkataster als CSV veröffentlicht wird, deutet diese URI an, dass nach der Ausprägung "68" in der Spalte "BaumNr" gesucht werden muss, während beim Vorhandensein einer API ein Aufruf wie <http://data.linz.gv.at/katalog/umwelt/baumkataster/BaumNr=00068> (oder auch <http://data.linz.gv.at/katalog/umwelt/baumkataster/BaumNr/00068>) tatsächliche die Angaben zu dieser einen Zeile im CSV ausgeben könnte.

Empfehlungen

Die Empfehlungen der AG-URI basieren auf der Nutzung von data.gv.at. Die Empfehlungen sind nach Betrachtungsebene gegliedert.

Eindeutigkeit

URLs sind immer eindeutig, weshalb die Eindeutigkeit bei URLs gegeben sein muss. Andere Identifier auf die in Datenfeldern verlinkt wird wie ORCID, oder DOIs werden durch ihr Referenzschema eindeutig. Es können mehrere Identifier für eine Ressource existieren, wichtig ist aber, dass dabei immer jeweils dieselbe Ressource angesteuert wird.

	URI der Metadaten (data.gv.at)	URI des Metadaten-Ressourcenblattes (data.gv.at)	URI der Downloadressource (Datenprovider)	Datenfeld -Identifier im Datensatz selbst - (Datenprovider)
Eindeutigkeit	MUSS	MUSS	MUSS	SOLL

Persistenz

Die Persistenz (Langlebigkeit über viele Jahre) von Identifiern soll primär zu den Metadaten sichergestellt sein, an zweiter Stelle zu den Datenressourcen, und an dritter Stelle zu den Datenobjekten.

Am wichtigsten ist die Persistenz des Identifiers zu den Metadaten, damit der Datensatz immer wieder positiv identifiziert werden kann, und sie allenfalls ändernde Locations der dahinterliegenden Datenressourcen an einer zentralen Stelle (Metadaten) gewartet werden können. An zweiter Stelle soll die Persistenz zu den Datenressourcen sichergestellt sein, da bei Änderung dieser Identifier Probleme bei Anwendungen entstehen können. Es ist daher für die Nutzer:innen der Daten sinnvoll, dass die Datenanbieter ihre Download-URLs möglichst persistent zu halten.

Das Metadatenblatt kann auch noch persistent bleiben, wenn die Daten nicht mehr bestehen.

	URI der Metadaten (data.gv.at)	URI des Metadaten-Ressourcenblattes (data.gv.at)	URI der Downloadressource (Datenprovider)	Datenfeld -Identifier im Datensatz selbst - (Datenprovider)
Persistenz der URI/URL	MUSS	MUSS	SOLL	KANN

Einheitliches Schema

Die verwendeten Identifier auf data.gv.at folgen einem einheitlichen Schema, wie im Kapitel zu den Betrachtungsebenen auf Ebene 1 beschrieben.

Da die Ressourcen in unterschiedlichen Systemen gespeichert sind, die oftmals nicht in der vollständigen Kontrolle der Datenprovider stehen, ist ein einheitliches Schema bei der Downloadressource nicht praktikabel.

Wenn seitens umsetzender Stelle ein Schema für die URL benötigt wird, kann z.B. das Schema aus dem INSPIRE Kontext (siehe Anhang X) verwendet werden.

	URI der Metadaten (data.gv.at)	URI des Metadaten-Ressourcenblattes (data.gv.at)	URI der Downloadressource (Datenprovider)	Datenfeld - Identifier im Datensatz selbst - (Datenprovider)
Einheitliches Schema	MUSS	MUSS	Nicht praktikabel	Nicht relevant

Menschenlesbarkeit

Die Menscheninterpretierbarkeit von URIs ist optional und kann insbesondere hilfreich sein, wenn URIs in Medien geteilt oder auf Websites/in Publikationen angegeben werden. Data.gv.at bietet für die Metadaten und Metadaten-Ressourcenblätter optional menscheninterpretierbare URIs an.

	URI der Metadaten (data.gv.at)	URI des Metadaten-Ressourcenblattes (data.gv.at)	URI der Downloadressource (Datenprovider)	Datenfeld -Identifier im Datensatz selbst - (Datenprovider)
Menschenlesbarkeit	SOLL	KANN	KANN	KANN

Auflösbarkeit = URL

Die Auflösbarkeit im Netz als URL ist, bei Nutzung von Data.gv.at + Ressourcenblatt automatisch gegeben. Die im Ressourcenblatt angegebenen URLs müssen regelmäßig auf defekte Links geprüft werden.

	URI der Metadaten (data.gv.at)	URI des Metadaten-Ressourcenblattes (data.gv.at)	URI der Downloadressource (Datenprovider)	Datenfeld - Identifier im Datensatz selbst - (Datenprovider)
Auflösbarkeit als URL	MUSS	MUSS	MUSS	KANN (auch URI möglich)

Fazit: Lösungsvorschlag unter Verwendung von data.gv.at

- Der Datenprovider erfasst für seine Datensätze die entsprechenden Metadaten auf data.gv.at, gemäß Erfassungsleitfaden.
- Der Datenprovider legt je Datensatz bzw. Version des Datensatzes ein eigenes "Ressourcenblatt" auf data.gv.at an.
- Der Datenprovider stellt sicher, dass die URL zur Downloadressource aktuell gehalten wird und jedenfalls auflösbar ist, sowie zur gewünschten Ressource zeigt.
- Der Datenprovider hält sich an die in den Merkmalen der Lösung (siehe Tabellen oben) angeführten MUSS/SOLL/KANN Ausprägungen. Durch diese Lösung ist die Erreichung der Merkmale **Eindeutigkeit**, **Persistenz**, **Einheitliches Schema**, **Menschenlesbarkeit** und **Auflösbarkeit** gewährleistet.

ANHANG I: Hintergrundinformationen

DCAT.AP 2.0

DCAT.AP 2.0 sieht als ein optionales Attribut in der Klasse Datensatz 0 bis mehrere dct: identifier vor, das den Haupt-Identifizier für den Datensatz darstellt (als URI oder eindeutiger Identifizier im Kontext des Katalogs).¹⁰ Die derzeit zugewiesenen UUIDs auf Datensatz- und Distributionsebenen sollten für einen zukünftigen Umstieg auf DCAT.AP 2.0 ausreichend sein.

Kontrollierte Vokabulare

Data.gv.at verwendet derzeit beispielsweise für die Kategorien das vorhandene EU Vokabular. Über die API wird der Link auf dieses Vokabular ausgegeben, wie am folgenden Beispiel ersichtlich:

https://www.data.gv.at/katalog/api/3/action/package_show?id=f660cf3f-afa9-4816-aafb-0098a36ca57d

Hier wird die Angabe der Kategorie "Umwelt" in als URI

<http://publications.europa.eu/resource/authority/data-theme/ENVI> ausgegeben.

Darüber hinaus kann es sinnvoll sein, auch auf andere Vokabulare zu verweisen. Beispielsweise hat die Stadt Linz in den Metadaten die Angabe im Feld Geographische Abdeckung/Lage von früher "Linz, Oberösterreich, Österreich" auf die URI

http://publications.europa.eu/resource/authority/place/AUT_LNZ umgestellt, insbesondere, da es ohnehin menschenlesbare Angaben zur dateneinbringenden Stelle gibt. Eine Diskussion über zu verwendende Vokabulare sollte im Zuge der Weiterentwicklung des österreichischen Metadatenstandards Richtung DCAT-AP erfolgen.

Die Verwendung kontrollierter Vokabulare kann auch auf Datenobjektebene sinnvoll sein. Beispielsweise hat sich im Archiv und Bibliothekswesen die Verwendung der GND (Gemeinsame Normdatei) etabliert. Eine GND-Nummer auch in internen Systemen mitzuführen und dann auch im veröffentlichten Datensatz anzugeben, kann also viele Vorteile bringen. In anderen thematischen Bereichen (z. B. Umwelt, Geoinformation, Statistik) können ganz andere Vokabulare Verwendung finden, so dass Empfehlungen der Cooperation nur von allgemeiner Natur sind und die angegebenen Beispiele nur Hinweise darstellen.

Neben internationalen Identifiern wie [GND](#), [Wikidata ID](#) und anderen ist es insbesondere empfehlenswert, europäische Vokabulare zu verwenden: <https://op.europa.eu/en/web/eu-vocabularies> Beispielsweise kann es sinnvoll sein, Codes wie rail-s, tran-s, bus-s in internen Systemen zu verwenden, basierend auf dem [Transport service Vokabular der EU](#).

URL-Shortener/Weiterleitungs-Services

Unter einem Kurz-URL-Dienst versteht man einen Internetdienst, der für beliebige URLs, also zum Beispiel Links zu Webseiten, eine zusätzliche alternative URL zur Verfügung stellt. Diese Alias-URL führt beim Aufruf über eine Weiterleitung zur Zielwebseite. Der auch Short URL, URL-Alias oder Kurzlink genannte Alias besteht meist aus wenigen Buchstaben oder Zahlen. Dies dient dazu, für unhandliche URLs kurze, leicht eingebare oder leichter merkbare Alternativ-URLs bereitzustellen.

¹⁰ https://joinup.ec.europa.eu/sites/default/files/distribution/access_url/2019-12/12f0dc1d-50b6-43e4-90c2-0afe213ac2be/DCAT_AP_2.0.0.pdf

Durch Verwenden eines URL-Shorteners oder Weiterleitungsservices können die publizierten URLs persistent bleiben, auch wenn die dahinterliegende Ressource sich ändert. Es gibt allerdings ein paar Risiken bzw. Probleme bei der Verwendung von solchen Diensten:

- Missbräuchliche Verwendung durch Verstecken des eigentlichen Linkziels
- Zuverlässigkeit des Dienstes (Schutz vor Angriffen, Persistenz, langfristige Verfügbarkeit des Dienstes)
- Missbrauchspotenzial für User-Tracking
- Bei Ausfall des Dienstes keine Möglichkeit auf die eigentlichen Daten zuzugreifen

Aktuell wird unter go.gv.at eine Weiterleitungsservice für die Nutzung durch die Verwaltung angeboten, das von der Stadt Wien betrieben wird und dessen Backend über das Portal Austria allen Interessierten bei Bedarf zur Verfügung steht. Auch im Bundeskanzleramt gibt es Überlegungen ein entsprechendes Weiterleitungs-Service einzurichten oder eine österreichische Datenverwaltungsregistry analog zu DOI anzubieten.

URI-Schema aus Inspire.gv.at

Im Gesamtkonzept [INSPIRE-Konzept Umsetzung Netzdienste - Version 2.2](#) wird im Kapitel 10.1.2.2 (Seite 82 ff.) der Aufbau eines *persistenten Objektidentifikators* vorgeschlagen. Dieser setzt sich aus einem definierten, mehrteiligen Namespace sowie einer Objekt- und – optional – einer VersionsID zusammen. Der genannte mehrteilige Namespace ist zugleich der *INSPIRE-Ressourcenidentifikator (URI)* und somit Teil der Metadaten des Datensatzes (siehe [INSPIRE-Metadatenenerfassungsleitfaden Kap. 4.3.9, Seite 29](#)).

Der URI wird im Zuge einer vom LFRZ etablierten URL-Weiterleitung mit dem Metadatensatz selbst verlinkt und kann bei Bedarf zum erwähnten Objektidentifikator weiterentwickelt werden. Dieser kann dann – ebenfalls im Zuge einer URL-Weiterleitung – zu einem Objektaufruf im Rahmen einer API/eines Services werden.

Aufbau Objektidentifikator:

<https://data.inspire.gv.at/>**[\(/Geodatenstelle\)](#)****[/Ressourcenbezeichner](#)****[/Datenthema.FeatureType](#)****[/localID \(/Version\)](#)**

Beispiel – BORIS-Bodeninformationssystem:

<https://geometadatensuche.inspire.gv.at/metadatensuche/srv/ger/catalog.search.jsessionid=D7D708AC09EBEE96035A008B81D95460#/metadata/b0ab02b4-f172-41fb-ade8-cc3c46b322df>

<https://data.inspire.gv.at/b0ab02b4-f172-41fb-ade8-cc3c46b322df/so.ObservedSoilProfile/6428709/19981010>

- Metadaten - Unique Ressource Identifier (URI)
 - **Namespace:** <https://data.inspire.gv.at/b0ab02b4-f172-41fb-ade8-cc3c46b322df>
 - **INSPIRE-Basisnamespace:** <https://data.inspire.gv.at>
 - **Geodatenstelle:** keine nötig, da für den Ressourcenbezeichner eine UUID verwendet wurde
 - **Ressourcenbezeichner:** b0ab02b4-f172-41fb-ade8-cc3c46b322df
- Erweiterung zu Objektidentifikator (derzeit nicht umgesetzt)
 - **Datenthema.FeatureTyp:** so.ObservedSoilProfile (Featuretyp des INSPIRE-Schemas des SOIL)
 - **localID:** 6428709 (Standortnummer aus der Bodendatenbank)

- **Version:**19981010 (z.B. *gültigAb* als Versionsnummer)

DOI Verwendung am Beispiel BEV

Das BEV verwendet für die Identifikation und Beschreibung von Datensätzen DOIs (Digital Object Identifier). Dabei handelt es sich um eine Systematik für globale Daten, ähnlich der ISBN Nummer bei Büchern. Beispiel: <https://doi.org/10.48677/6c316854-5453-4cf5-82cb-6586b2a2aa06>

Das DOI System ist ein Beispiel eines externen Services zur Generierung eindeutiger und persistenter Identifier. Im Bund werden derzeit Überlegungen angestellt, ob ein Service für eindeutige IDs etabliert werden soll.

ANHANG II: Vorgehensweise bei der Erstellung des vorliegenden Dokumentes

In der ersten AG- Sitzung am 8.6.2022 wurde das Wissen der TeilnehmerInnen gesammelt und diskutiert, um im Anschluss erste Ziele für die Arbeitsgruppe zu definieren. Weiters wurden organisatorische Festlegungen wie die Frequenz der Treffen und die gemeinsame Ablage getroffen.

Dokumentation:

<https://app.mural.co/t/duk2947/m/duk2947/1654670723998/0ad01dfe5b6eac0c18647155af5ba45205975aca?sender=u3b1ec3b5f0ae00490bb44466>

In der zweiten AG-Sitzung vom 4.7.2022 wurden die Betrachtungsebenen diskutiert und eine Aufteilung wie folgt vorgenommen: Motivation (Eibl), Input zu BEV DOI (Tinkl), Inspire (Obersteiner), URL Shortener (Reichart), 3 Ebenenbeschreibung (Krabina)

In der dritten AG-Sitzung vom 27.7.2022 wurden die Ausarbeitungen und die sich daraus ergebenden Diskussionspunkte besprochen. Ein Hauptaugenmerk lag auf dem Identifier auf Datenebene. Außerdem wurde das Problem der unterschiedlichen Interpretationen von Ressourcen angesprochen. Die Textfragmente sollen bis zur nächsten Sitzung in Prosaform gebracht werden. Darüber hinaus werden wir überlegen, was die konkreten Ziele/Anforderungen für Identifier sein sollten und welche Empfehlungen wir auf verschiedenen Ebenen vorschlagen möchten.

In der vierten Arbeitsgruppensitzung am 31.8.2022 wurden die Ziele diskutiert und priorisiert, das Kapitel über die Beobachtungsebenen erörtert und der Unterschied zwischen Resource Sheet und Resource Link herausgearbeitet. Darüber hinaus wurde ein Brainstorming zu möglichen Empfehlungen aus Sicht der Arbeitsgruppe durchgeführt.

In der fünften AG-Sitzung vom 3.10.2022 wurden ein Erstentwurf zu den Empfehlungen erstellt, der für jede Zielsetzung in den unterschiedlichen Betrachtungsebenen MUSS/SOLL/KANN Empfehlungen vornimmt. Die Abgrenzung von URIs vs URLs wurde diskutiert und im Entwurf geschärft

In der sechsten AG-Sitzung vom 3.11.2022 wurde der Begriff Datenfeld diskutiert und geschärft, und ein Fahrplan zum Abschluss des Dokuments festgelegt.

In der Sub-AG Sitzung vom 17.11.2022 wurden die ersten Kapitel bis zu den Betrachtungsebenen konsolidiert und qualitätsgesichert.

In der Sub-AG Sitzung vom 22.11.2022 wurden die Kapitel Betrachtungsebenen, Empfehlungen und Anhänge qualitätsgesichert.

In der siebenten AG-Sitzung vom 6.12.2022 wurde die Version des Dokuments finalisiert und die Abschlusspräsentation erstellt.

In der achten AG-Sitzung vom 5.05.2023 wurde das Feedback der Sondersitzung der Cooperation OGD vom 13.03.2023 besprochen.

Sub-AG Data Spaces und wie sind diese bestmöglich zu nutzen

Martin Semberger (BMDW), martin.semberger@oesterreich.gv.at
Fritz Fahringer (Standortagentur Tirol), fritz.fahringer@standort-tirol.at
Christian Klingler (Tirol Werbung), christian.klingler@tirolwerbung.at
Michael Reichart (BRZ), michael.reichart@brz.gv.at
Robert Seyfriedsberger (BRZ), robert.seyfriedsberger@brz.gv.at
Günther Tschabuschnig (DIO), guenther.tschabuschnig@dataintelligence.at
Benjamin Witsch (AustriaTech), Benjamin.Witsch@austriatech.at

Ausgangslage

Die Handhabung datengesteuerter Prozesse war bislang häufig geprägt von Aktivitäten, die auf die Generierung, Sammlung und Kontrolle von teils großen Datenbeständen im eigenen Bereich bzw. Silo ausgerichtet war. Diese Tendenz führte zumeist zur Entstehung datenmächtiger Akteure, welche eine alleinige Kontroll- und Verfügungsgewalt über Datenbestände besaßen. Dabei wird der eigentliche Nutzen außer Acht gelassen, der durch den **Austausch und die gemeinsame Nutzung von Daten** entsteht. Gemäß den Prinzipien der Offenheit und Kollaboration gilt es diese Formen der Datenkonzentration, Dominanz und Ausschließlichkeit der Datenverwendung zu vermeiden und stattdessen die Zusammenarbeit und gemeinsame Datennutzung in den Vordergrund zu stellen.

Eine **offene Verwaltung** (Open Government) trägt in entscheidendem Maße dazu bei, positive gesellschaftliche und wirtschaftliche Entwicklungen auf der Grundlage einer veränderten Verwaltungskultur zu fördern. Von der Cooperation OGD Österreich wurden im vergangenen Jahrzehnt entscheidende Impulse und Grundlagen für die Standardisierung, Vernetzung und Bewusstseinsbildung im Datenbereich gesetzt. Nun ist es an der Zeit, die gemeinwohlorientierte Kooperation mit Unternehmen, der Forschung und Zivilgesellschaft weiter auszubauen. Diese Zusammenarbeit soll u.a. durch die Schaffung entsprechender Datenräume auf der Grundlage europäischer Konzepte verwirklicht werden.

Open Data als Grundlage für Datenräume

Strukturierte und für alle weiterverwendbare Daten sind Grundlage für die Entstehung von Daten-Ökosystemen und eine Notwendigkeit für eine datengestützte Wirtschaft und Gesellschaft. Offene Daten ermöglichen die Kollaboration und Teilhabe an vielfältigen Innovationsprozessen und verbinden unterschiedlichste Akteure im Bestreben Innovationen oder neue Geschäftsmodelle zu entwickeln oder um gesellschaftliche Entwicklungen zum Gemeinwohl voranzutreiben. Es ist diese offenste Form des Datenaustausches, die einen besonders hohen Grad an Kollaboration unter vielfältigsten Akteuren ermöglicht. Die übergreifende Datennutzung unterliegt jedoch häufig unterschiedlichsten Einschränkungen. Die Cooperation OGD Österreich wird sich auf der Grundlage

entsprechender Gesetzgebung für einen weitestgehenden **Abbau von Schranken für die Datenweiterverwendung** einsetzen.

Überblick über das Datenspektrum,

UK Open Data Institut (ODI) <https://theodi.org/about-the-odi/the-data-spectrum>

Data.gv.at als zentrale Drehscheibe

Über das zentrale Datenportal der österreichischen Verwaltung "data.gv.at" wird ein Überblick über das Vorhandensein offener Daten der österreichischen Verwaltung ermöglicht. Darüber hinaus fungiert das zentrale Datenportal als Anlaufstelle für Datensuchende. Auf der Grundlage dieser funktionierenden Infrastruktur und mit dem Ziel Datenräume voranzutreiben soll die Förderung des Austauschs unter Entwickler:innen und Dateninteressierten über die zentrale Informationsstelle "data.gv.at" u.a. durch technische, standardisierte Schnittstellen weiter ausgebaut werden. Überdies sind neue technologische Entwicklungen bei Datenräumen fortlaufend zu beobachten und anhand fairer Zugangsprinzipien mitzugestalten. Dies inkludiert den notwendigen Aufbau entsprechender personeller und technischer Kapazitäten (vgl. EU Data Governance Act).

Offenheit als Prinzip

Entsprechend der europäischen Vision zur Ausgestaltung eines dezentral organisierten europäischen **Binnenmarktes für Daten** steht die grenzüberschreitende Vernetzung von unterschiedlichsten Akteuren im Fokus der politischen, wirtschaftlichen und gesellschaftlichen Anstrengungen in ganz Europa. Ziel ist es, Impulse für eine andere Art der Wirtschaft zu setzen, als sie bislang von dominierenden Einzelakteuren vorgegeben wird. Beispielsweise bestehen derzeit hohe Abhängigkeiten von einzelnen Cloud-Anbietern. Ziel ist es, die Souveränität Europas basierend auf vielfältigen (anstelle monotoner) und untereinander vernetzten Ökosystemen zu stärken. Gemeinsam etablierte europäische Werte und Standards sollen weltweit aktiv verbreitet werden.

Für die Umsetzung dieser Vision entscheidend ist das **Prinzip der Offenheit** bzw. des Aufeinander-Zukommens. Um der Marktbeherrschung durch einige wenige Akteure entgegenzuwirken, braucht es Regeln zum fairen Zugang für alle Akteure und zur Sicherstellung der Fairness zwischen den Akteuren. Daher schöpfen innovative europäische Ansätze aus den Grundregeln des ursprünglichen, offenen Internets, offener Standards und der Nutzung quelloffener Software - allesamt Grundbegriffe, die von der Open Data (bzw. Open X) Community entscheidend geprägt wurden.

Als entscheidendes Leitprinzip für die Entwicklung von Datenräumen auf der Grundlage offener Daten erachtet wird (vgl. Open Data Richtlinie):

"as open as possible, as closed as necessary"
(*"so offen wie möglich, so geschlossen wie nötig"*)

Beitragende sollten demzufolge möglichst viele Datenbestände als offene Daten im Datenraum zur Verfügung stellen, damit größtmögliche Weiterverwendungsmöglichkeiten geschaffen werden. Mit der Verwendung **offener Lizenzen** (z.B. CC-BY) und offener technischer Standards wird dies unterstützt. Als Grundprinzip sollten Beitragende dazu motiviert werden, einen Grundbestand eigener Daten als Daten im Datenraum als **Open Data** bereitzustellen.

Das Konzept der Datenräume (Data Spaces)

Aufbauend auf der offenen Vernetzung zwischen unterschiedlichsten Akteuren beruht die europäische Vision darauf, eigene Räume für die Entfaltung der Kreativität und für die Stärkung des gegenseitigen Vertrauens bei der Datennutzung zu schaffen. Dies soll innerhalb gemeinsamer Datenräume geschehen, für welche dezidiert angepasste rechtliche und technische Rahmenbedingungen geschaffen werden.

Diese spezifisch angepassten Rahmenbedingungen sollen **Innovation** beflügeln und eine weitgehende Interoperabilität verschiedenster technischer Systeme ermöglichen. Die Rahmenbedingungen sind nicht nur auf eine Region oder einen Staat beschränkt, sondern ermöglichen aufgrund der notwendigen Skalierung Datenflüsse im gesamten europäischen Binnenmarkt. Ziel ist folglich die Verknüpfung bestehender Datenpools anhand gemeinsamer Zielsetzungen.

Die Etablierung unterschiedlichster Datenräume erfolgt entweder auf sektoraler Basis (Gesundheit, Landwirtschaft, Mobilität, Tourismus, Energie, etc.) oder anhand spezifischer Nutzungsszenarien (Use Cases). In Österreich wird die Kollaboration anhand konkreter Use Cases auch als "Datenkreis" bezeichnet. Im Rahmen der Initiative Gaia-X etwa werden konkrete Use Cases für die Wirtschaft im Einklang mit dieser europäischen Vision umgesetzt. Die European Open Science Cloud (EOSC) fördert den offenen Austausch unter Forschenden anhand der Prinzipien von Open Science und Open Access im Forschungsbereich.

Definition eines Datenraums

Der Begriff "Datenraum" bezieht sich auf eine Art von Datenbeziehung zwischen vertrauenswürdigen Partnern, die sich an dieselben hohen Standards und Richtlinien in Bezug auf die Speicherung und gemeinsame Nutzung von Daten innerhalb eines oder mehrerer vertikaler Ökosysteme halten. Ein entscheidender Aspekt des Datenraumkonzepts ist, dass die Daten nicht zentral gespeichert werden, sondern an der Quelle. Sie werden also nur bei Bedarf durch semantische Interoperabilität übertragen. Ein Datenraum ist die Summe aller seiner Teilnehmer, die Datenanbieter, Nutzer und Vermittler sein können. Datenräume können verschachtelt sein und sich überschneiden. Datenhoheit und Vertrauen sind für das Funktionieren von Datenräumen unerlässlich und unterstützen die Beziehungen zwischen den Teilnehmern.

Quelle: Gaia-X, <https://gaia-x.eu/what-is-gaia-x/core-elements/data-spaces/>

Laut Europäischer Kommission führen gemeinsame europäische Datenräume Dateninfrastrukturen und Governance-Vorschriften zusammen, um die gemeinsame Nutzung von Daten zu erleichtern.

Datenräume haben zum Ziel:

1. *Vernetzung von Infrastrukturen:*

Instrumente und Dienste für die gemeinsame Nutzung von Daten, mit denen Akteure Daten zusammenführen, verarbeiten und gemeinsam nutzen können.

2. *Rechtssicherheit und Datenhoheit:*

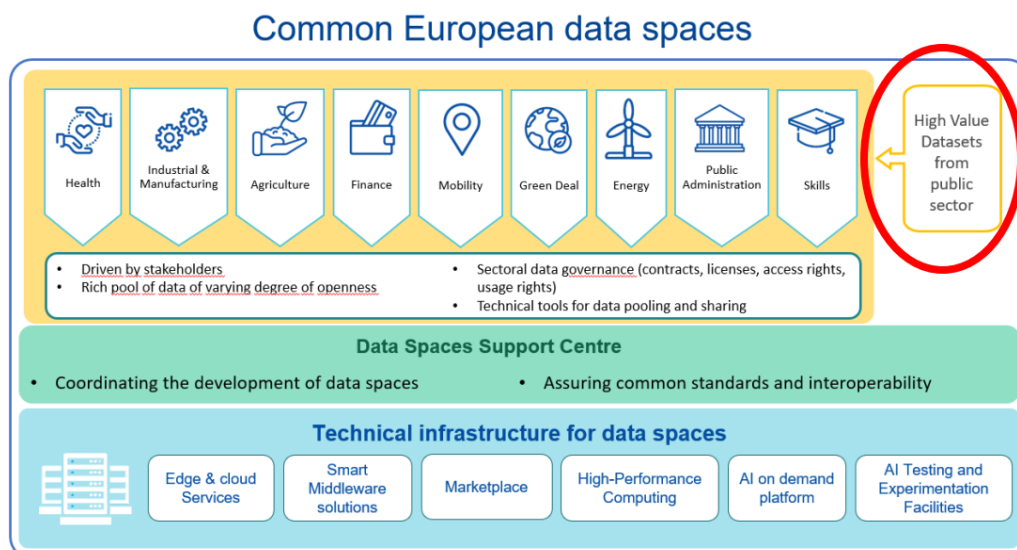
Kohärente Data-Governance-Strukturen, die auf transparente und faire Weise die Rechte in Bezug auf den Zugang zu den Daten und deren Verarbeitung festlegen.

3. *Verbesserung der Datenqualität und Interoperabilität:*

Verbesserung der Verfügbarkeit, Qualität und Interoperabilität von Daten, innerhalb bestimmter Sektoren als auch sektorübergreifend.

Quelle: Europäische Kommission,

<https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces>



Wie können Datenräume bestmöglich genutzt werden?

Zur besseren Ausschöpfung des Potentials offener Daten im Zusammenhang mit der Entstehung von Datenräumen werden von der Cooperation OGD Österreich die folgenden Maßnahmen angestrebt. Die Entwicklung von Datenräumen soll von allen Mitgliedern der Cooperation OGD Österreich in deren Wirkungsbereich aktiv mitgestaltet werden.

Maßnahme 1: Datensouveränität stärken

Die Cooperation OGD Österreich unterstützt die gemeinsame europäische Vision der **Stärkung der europäischen Datensouveränität** und wird sich für das Vorantreiben entsprechender Regulative auf EU-, nationaler und regionaler Ebene einsetzen.

Abgesehen von einer gestärkten kollektiven Souveränität ist es für die Cooperation OGD Österreich essentiell, zur Stärkung der individuellen Souveränität und Teilhabe unterschiedlicher Akteure an Datenräumen beizutragen. Das Grundprinzip der **Datenhoheit** einzelner Akteure und verbundener Entscheidungen (Dezentralität) ist für die Cooperation OGD essentiell. Die Cooperation wird sich für souveräne Entscheidungsmöglichkeiten einsetzen, auch unter Berücksichtigung sicherer

technologischer Lösungen (z.B. Blockchain, Intermediäre, etc.) unter Einhaltung **europäischer Werte**. Ein Augenmerk dabei liegt auf der Vermeidung von Abhängigkeiten und unfairer Vertragsbestimmungen (z.B. Exklusivverträge).

Maßnahme 2: Fairen Zugang zu Datenräumen sicherstellen

Die Cooperation OGD Österreich setzt sich für die Einhaltung des Prinzips des **fairen Zugangs zu Datenräumen**, für offene Teilhabemöglichkeiten, die Wahrung der europäischen Vision grenzüberschreitender Datenräume und für das Konzept der offenen Innovation (Open Innovation) ein. Ausschließlichkeitsvereinbarungen und allfällige Datennutzungs-Einschränkungen im Bereich der Datenräume sind fortlaufend auf deren Notwendigkeit und ihren Nutzen für eine faire Datenökonomie zu überprüfen. Ebenso wird sich die Cooperation OGD für eine faire Rollenverteilung innerhalb von Datenräumen einsetzen.

Ungeachtet der Anwendungsszenarien zielen Datenräume darauf ab, Ökosysteme für Kollaboration zu schaffen. Vergangene Entwicklungen im Bereich der Datenwirtschaft zeigten häufig Tendenzen hin zur Entstehung exklusiver Räume für selektive Partner und für spezifische Eigenzwecke. Derartige Tendenzen zum offensichtlichen Nachteil der Allgemeinheit gilt es zu vermeiden. In Anlehnung an das Wettbewerbsrecht muss eine grundsätzliche Teilhabe (Bestreitbarkeit) an Datenräumen gewährleistet werden.

Maßnahme 3: Aufwertung von data.gv.at zur zentralen Daten-Drehscheibe

“Data.gv.at” ist mehr als nur eine Website. Data.gv.at soll als **Dreh- und Angelpunkt für Datensuchende und Innovatoren** etabliert werden und diese mit Akteuren der öffentlichen Verwaltung vernetzen. Das Team von data.gv.at soll hierbei die Funktion einer zentralen Anlaufstelle bilden, nicht nur zu Open Data, sondern zu vielfältigen Datenbelangen und so Auskunft geben können und Wegweiser sein zu Datenräumen. Dies umfasst folglich nicht nur technische Aspekte. Das Team von data.gv.at soll auch für weitergehende Fragestellungen im Datenbereich erreichbar sein und Anliegen gegebenenfalls auch weiterleiten können (vgl. EU Data Governance Act).

Entsprechend konzeptueller Vorstellungen von Gaia-X soll “data.gv.at” zu einem **Facilitator** für Datenräume ausgebaut werden. Dabei soll die Nutzung von standardisierten **Konnektoren** zu unterschiedlichsten Datenräumen und Sub-Netzwerken angestrebt werden. Ebenso sollen Unterstützungsdienstleistungen bezüglich der Anbindung an Datenräume und die Erbringung von Diensten etabliert werden. Data.gv.at könnte auch eine grundsätzliche förderierende Rolle (Federator) für offene Daten und viele weitere Datenbelange in Österreich zukommen.

Maßnahme 4: Datenverwendbarkeit erhöhen

Als Haupthemmnis für die unzureichende Nutzung von Daten wird wiederholt das mangelnde Vertrauen von Akteuren in den Datenaustausch (data sharing) identifiziert. Daher ist es entscheidend, dieses gegenseitige Vertrauen zu stärken. Die Cooperation OGD Österreich stellt zur **Stärkung des Vertrauens** die Prinzipien der Offenheit in den Vordergrund.

Rechtliche und technische Hürden gelten als wesentliche Hemmnisse zur Nutzung brachliegender Potenziale. Die Cooperation OGD Österreich setzt sich für möglichst einheitliche **rechtliche und technische Rahmenbedingungen** ein.

Datenräume sollen anhand einheitlicher rechtlicher und technischer Kriterien die Nutzbarkeit von Daten erhöhen und die Inklusion unterschiedlichster Akteure fördern. Dabei gilt es einerseits brachliegende Kapazitäten besser zu erschließen und andererseits die **Datenqualität** insgesamt zu erhöhen. Die Nutzbarkeit von Daten soll anhand von **Use Cases** (Datenkreise) besser veranschaulicht bzw. vorangetrieben werden.

Von der Cooperation OGD Österreich sollen Akzente gesetzt werden, welche laufend zur kontinuierlichen Erhöhung der Datenqualität führen. Dies umfasst auch einen verbesserten Überblick über **Bedingungen** für die Datenweiterverwendung (u.a. Lizenzen). Die Cooperation wird sich für den Einsatz etablierter, **offener Standardlizenzen** einsetzen.

Maßnahme 5: Interoperabilität und FAIR Data

Gemäß den **FAIR-Prinzipien** sollen Daten anhand folgender Charakteristika erstellt und aufbereitet werden: Diese sollen Findable (auffindbar), Accessible (zugänglich), Interoperable (interoperabel) und Reusable (weiterverwendbar) sein. Diese Grundprinzipien sollen wegweisend für die Etablierung von Datenräumen sein.

Die Cooperation OGD Österreich wird sich dafür einsetzen, dass für Datenräume **interoperable technische Standards, offene Formate** und **offene Anwendungsprogrammierschnittstellen (API)** möglichst als Standard ("per default" bzw. "per design") eingesetzt werden, zumal dies eine größtmögliche Interoperabilität und Weiterverwendbarkeit von Daten gewährleistet. Dies soll die erste und grundlegendste Stufe einer "Governance-Pyramide" für Datenräume darstellen. Ziel ist es, eine bestmögliche semantische Interoperabilität von Datensätzen zu gewährleisten.

Ebenso sollen offene Daten möglichst entsprechend gekennzeichnet werden, um hervorzuheben dass beispielsweise Anwendungen auf der Grundlage offener Daten entstanden sind (**Open Data Label**).

Die FAIR Prinzipien, FAIR Office Austria

<https://fair-office.at/ueber>

Maßnahme 6: Vision eines offenen Datenraums der Verwaltung

Entsprechend dem Konzept europäischer Datenräume bzw. der Initiative Gaia-X sollen von der Cooperation OGD Österreich Überlegungen angestellt werden zur Verwirklichung eines offenen Datenraums der öffentlichen Verwaltung (Public Data Space). Dabei sollen auch öffentliche Unternehmen zur Teilhabe ermutigt werden.

Hierbei soll das Potential von Daten der öffentlichen Verwaltung besser erschlossen und ausgeschöpft werden. Sämtliche Ebenen der öffentlichen Verwaltung sollen von der Expertise von Partnern profitieren und bei der digitalen Transformation unterstützt werden. Insbesondere hervorgehoben werden sollen anhand konkreter Beispiele Chancen für die öffentliche Verwaltung zum verbesserten Umgang mit eigenen Daten bzw. soll der übergreifende Nutzen des Datenaustausch unter Verwaltungseinheiten oder Partnern vorangetrieben werden. Dies umfasst ebenso Kapazitäten zur Stärkung der Datenanalyse, Anonymisierung, Aggregation, uvm. Eine wesentliche Rolle bei der Vernetzung könnten Open Data Beauftragte der betreffenden öffentlichen Stellen einnehmen (gemäß Informationsweiterwendungsgesetz IWG).

Als Ausgangspunkt hierfür könnte der “datahub.tirol” (www.datahub.tirol) herangezogen werden.

Mit dem [datahub.tirol](http://www.datahub.tirol) entsteht in Tirol ein regionaler Datenmarktplatz, auf dem Daten regional, national und international genützt und verknüpft werden können.

In den drei Datenräumen Energie, Mobilität und Tourismus werden Teilnehmende ihre Daten in Datenkreisen zur Verfügung stellen sowie bereits verfügbare Daten (z.B. Open Data) nutzen können. Zudem werden sie beim Übergang in die Datenwirtschaft, beim Einsatz und der Optimierung relevanter Technologien, bei der Marktbildung und beim Management von Daten unterstützt. Der datahub.tirol unterstützt Partner:innen dabei, den fortlaufenden Austausch und Handel ihrer Daten in einem Datenkreis, zu tätigen.

Der datahub.tirol soll NutzerInnen ermöglichen, neue Geschäftsmodelle zu entwickeln und neue Geschäftsideen umzusetzen.

Der datahub.tirol unterstützt als Initiator und Vernetzer die NutzerInnen dabei, Daten für ihre kreativen Lösungen regional, national und international zu nützen und zu verknüpfen - ohne dabei die Datenhoheit aufgeben zu müssen.

Der datahub.tirol soll der primäre Datenmarktplatz in den oben genannten Domänen werden und dabei auch auf bereits verfügbare Daten in regionalen wie in nationalen Datenräumen z.B. data.gv.at oder Data Spaces der Data Intelligence Offensive (DIO) zurückgreifen. Der datahub.tirol versteht sich somit als regionaler Facilitator von Daten und Standardservices und ist nicht nur regional, sondern national und überregional mit anderen Datenräumen verknüpft, indem er die europäische Vision von Gaia-X und deren Prinzipien mitträgt.

Zukunft Cooperation OGD Österreich

Brigitte Barotanyi und Brigitte Lutz

§ Ausgangslage

Das Bundeskanzleramt, die Städte Wien, Linz, Salzburg und Graz gründeten am 13. Juli 2011 gemeinsam die „Cooperation Open Government Data Österreich“, kurz „Cooperation OGD Österreich“. Bund, Länder, Städte und Gemeinden wollen in Kooperation mit den Communities, Wissenschaft, Kultur und der Wirtschaft die Basis für die Zukunft von Open Government Data in Österreich legen. Durch die Einigung auf gemeinsame Standards sollen effektive Rahmenbedingungen geschaffen werden, die zum Nutzen aller Interessensgruppen sind. Das Zusammenwirken im Raum Deutschland – Österreich – Schweiz – Liechtenstein (D-A-CH-LI) wird forciert. In dieser Cooperation werden die Interessen von Ministerien, Ländern, Städten und Gemeinden vertreten, also all jener Akteure, die eine Open Government Data-Plattform betreiben, planen, erstellen oder sich daran beteiligen möchten.

Leitbild und Vision 2019 basierend auf der Salzburger Erklärung vom 29.10.2018

Wir sind die BegründerInnen und bedeutende Stakeholder für die Weiterentwicklung von OGD in Österreich. Wir nehmen aktiv an der Kooperation OGD D-A-CH-LI und anderen internationalen Netzwerken teil.

Durch unsere Aktivitäten soll Österreich Open-Data-Trendsetter bleiben und eine internationale Vorreiterrolle haben.

Daher ist es notwendig, dass die Bundesministerien, die Länder, der Städtebund und der Gemeindebund aktiv an der Cooperation OGD Österreich teilnehmen. Mittels politischem Schulterschluss gilt es, eine gemeinsame, zukunftsweisende Strategie zu entwickeln.

Was ist gelungen? Was wäre zu verbessern?

- ✓ Mit mehr als 130 Personen am Verteiler und mehr als 30 Teilnehmer*innen an den Treffen, die 2mal jährlich stattfinden, ist das Interesse sichtlich vorhanden.
- ✓ „Allianz der Willigen“, die Veränderung möchten
- ✓ Die Zusammensetzung aus Verwaltung, Wirtschaft und Wissenschaft ist gewinnbringend
- ✓ Arbeitet strategisch und operativ, unabhängig von politischen Entscheidungen
- ✓ Zur Erarbeitung von Papers / Mitarbeiten in Arbeitsgruppen finden sich immer wieder Expert:innen,
 - an der Ressourcenknappheit, oder weil von der jeweiligen Organisation keine Unterstützung gegeben wird, scheitern auch manche Erarbeitungen
 - Die Awareness für Open Data ist in den Ministerien, im Städtebund und Gemeindebund kaum vorhanden
 - Harmonisierung auf Datenebene ist nicht möglich, es fand zwar Definition von Kerndatensätzen statt, die mit High Value Datasets verknüpft worden, aber keine Umsetzung möglich
 - Durch die Freiwilligkeit gibt es keine fixen Mitgliedschaften – die Ministerien oder der Gemeindebund sind schwach vertreten
 - Die Cooperation OGD Österreich hat kein Budget und kann z.B. keine Dienstleistungen oder Studien selbst beauftragen
 - Die Cooperation OGD Österreich hat keine personellen Ressourcen zur Koordination und Unterstützung der Mitglieder

- Es ist keine rechtsverbindliche Organisation, kann weder Stellungnahmen bei Gesetzesentwürfen oder Umfragen zu Rankings (z.B. EU Portal) abgeben, noch Beschlüsse fassen oder verbindliche Empfehlungen aussprechen
- Beitritt zu Open Government Partnership nie möglich

Was ist für die Zukunft erforderlich?

- Das künftige IFG erfordert eine österreichweite Koordination
- Organisatorisch
 - Die Kontinuität ist derzeit nur der Initiative der langjährigen Sprecherin zu verdanken. Ein fixes Team, das die Agenden vorantreibt, ist erforderlich.
 - Stärkere Verbindlichkeit und Verankerung der Cooperation OGD Österreich mit einer Governance-Struktur
 - Fixe Mitgliedschaften
 - Netzwerk für die Open Data Beauftragten laut IWG
 - Marketing
 - Veranstaltungen
 - Community-Management
- Weiterentwicklung aktueller Themen mit Unterstützung der Wissenschaft und Forschung, z.B.:
 - DCAT-AP
 - Harmonisierung von Kerndatensätzen bzw. High Value Datasets
 - Data Spaces und wie sind diese bestmöglich zu nutzen
 - URI Koordinierungsstelle (Seiten 5-13)
 - Datenqualitätsstandards und Feedbackschleifen (Seite 14-15)
 - Linked Data
 - ...

§ Vorgehensweise

Informationen über die OGD-Organisationen in Deutschland und der Schweiz, sowie bestehende „Daten“-Organisationen in Österreich wurden eingeholt und die best practices zusammengefasst.

Deutschland:

Deutsches Dateninstitut geplant

- Positionspapier <https://github.com/Dateninstitut-de>
- Umfassender Ansatz
- Starke politische Abhängigkeit (Zuständigkeit? Digitalministerium / BMI(betreibt derzeit Open Data), BMWK),
- Form: Gemeinnützige GmbH?
- ev 2023 nächste Schritte

Kompetenzzentrum Open Data (Bundesverwaltungsamt) <https://opendata.bund.de>

- 4 Personen
- Beratung / Unterstützung der Behörden bei der Umsetzung von Open Data. Kontaktstelle für die Open Data - Stellen der Länder.
- Kompetenzaufbau
- Vernetzung

Bisher: Betrieb govdata.de mit Dienstleister, Standard weiterentwickeln

- Produktmanagement govdata.de

- Personelle Ausstattung: 4-5 Stellen empfohlen
- 800.000 – 1 Mio inkl. Personalkosten; + 620.000 EUR Projekt
- Von Fachgruppe govdata.de Bund-Ländern kontrolliert > formale Entscheidungsgremium
 - o Untergruppe für DCAT-Standard init AG

Ab 2023 FITKO (Föderale IT-Kooperation), 2022 gegründet, um Anwendungen des IT-Planungsrates zu betreiben, auch govdata.de

- Öffentlich-rechtliche Organisation durch Länder und Bund

Open Knowledge Foundation nicht sehr aktiv, eher Frag den Staat, Bitkom mit Industrie, Bertelsmann Stiftung, Adenauer Stiftung

Schweiz

Geschäftsstelle opendata.swiss bei Bundesamt für Statistik angesiedelt (Department des Innern)

- Mit 5 Personen besetzt (2 Datenspezialisten, 2 Kommunikation, 1 Person Migration Datenportal von externem Dienstleister)
- Umsetzung Strategie und Vorbereitung weitere Strategie)
- Organisationsverordnung – Gesetz open by default 2023 für Bundesverwaltung, Umsetzungsfrist 3 Jahre
- Geschäftsstelle OGD begleitet alle Ämter, statistischen Stellen, Kantone und Städter
- Datenmanagement
- Budget – digitale Verwaltung Schweiz – Idee Teil Kantone 1 Mio / Jahr
- Problem: swisstopo, Umwelt – ressortübergreifend
- „neutrale“ Bezeichnung

Zusammenspiel mit Verein Opendata.ch

- Regelmäßige Kontakte
- Datennutzer

Normen entwickelt <http://www.ech.ch/>

Österreich

Lenkungsgruppe + Fachgruppe data.gv.at (Bund + Länder)

- Finanzierung data.gv.at
- Weiterentwicklung data.gv.at
- Beauftragung BRZ

[Organisation INSPIRE in Österreich](#) - INSPIRE Assistenzstelle Österreich

- Rechtliche Grundlage
- Große Stärke: Harmonisierung

[gip.gv.at](#)

- Gespräch mit Irmgard Mandl-Mair
- Art 15 Vereinbarung mit Auftrag Verein zu gründen
- Eigener Finanzschlüssel - Mischung Verkehrslänge und Bevölkerung
- Mischsteuersatz

DIO - Data Intelligence Offensive (Verein):

- Mitglied der Cooperation OGD Österreich (= AG Open Data)

§ Ergebnisvorschlag

Einrichtung einer Koordinierungsstelle mit folgenden Aufgaben

- Impulsgeber:in
- Beauftragung von Studien
- Förderung Data Literacy, Data Excellence und Open Data in Bund, Ländern, Städten und Gemeinden, Begleitung der Open Data Beauftragten
- Internationale Kontakte
- Marketing
- Veranstaltungen
- Community-Management

Ressourcen:

- Empfehlung: Start mit ½ Vollzeitäquivalent im zuständigen Ministerium (derzeit BMF), Ausbau zu 3 Personen
- Budget: Empfehlung 70.000 EUR / Jahr für Sachkosten

Evaluierung nach 1,5 Jahren und Verbesserung der Vereinbarung und Geschäftsordnung

Erforderlich:

- Kooperationsvereinbarung (regelt Ressourcen, Rollen, Zuständigkeiten, Verantwortlichen, Anzahl Vertretungen, gemeinsame Beauftragungen über gemeinsames Budget,...)
- Geschäftsordnung für Cooperation OGD Österreich: Abwechselnder Vorsitz mit Ressourcen und eigener Organisation im Hintergrund,...
- Stellenbeschreibung Koordinationsstelle

Nächste Schritte:

- Erarbeitung von Kooperationsvereinbarung, Geschäftsordnung und Stellenbeschreibung
- Gespräch innerhalb des BMF zur Einrichtung der Koordinationsstelle