

|  |
| --- |
| data.gv.at  |
| Voraussetzungen für den automatischen Import von Metadatensätzen |
| V 1.5 |

****

18.5.2018

# Versionshistorie

|  |  |  |
| --- | --- | --- |
| **Version** | **Datum** | **Anmerkungen** |
| 1.0 | 2013 | Initiale Version |
| 1.1 | 30.04.2014 | Ergänzung zu „Löschen von Datensätzen“ |
| 1.2 | 05.09.2014 | Aktualisierung Link zu Metadatenspezifikation 2.1Beispiel API Land OÖ hinzugefügtHäufige Fragen und Antworten hinzugefügt |
| 1.3 | 3.4.2017 | Detailierung zu Tags (als Array nicht als String)  |
| 1.4 | 11.12.2017 | Hinweis zu Sonderzeichen in Tags hinzugefügt (in CKAN nicht möglich) |
| 1.5 | 18.5.2018 | Aktualisierung Verweis auf aktuelle Version des Metadatenstandards 2.4 |
| 1.6 | 16.12.2022 | Aktualisierung Verweis auf aktuelle Version des Metadatenstandards 2.6 |

# Allgemeines

Die Daten müssen grundsätzlich dem OGD Austria Metadatenstandard 2.6 entsprechen und im JSON-Format zur Verfügung stehen. Der automatische Import läuft in 3 Phasen ab:

1. Abfrage der Liste der geänderten Datensätze
2. Abfrage der Daten der jeweiligen Datensätze
3. Import in CKAN durch data.gv.at

Für 1. und 2. wird jeweils eine Schnittstelle benötigt, die die geforderten Daten per JSON in UTF8 zurückliefert.

Kontakt-Emailadresse bei Fragen: [data@brz.gv.at](https://brzoffice.portal.at/livelinkdav/nodes/43436513/data%40brz.gv.at)

# Abfrage der geänderten Datensätze

Für den Import ruft data.gv.at die URL für die Abfrage der Liste mit dem URL - Parameter (get) auf, beim Parameter „since\_time=“ übergeben wir als Wert einen Timestamp im ISO-Format. Als Rückantwort erwartet data.gv.at eine Liste von eindeutigen IDs der Datensätze, die seit dem Wert „since\_time“ geändert oder neu angelegt worden sind.

**Beispiel API – Land OÖ**

<http://data.ooe.gv.at/api/ogd/rest/dataset?since_time=2014-04-01T19:42:45.854533>

Beispiel-Antwort:

[

 "ooe-07B4E90B05184E268F2591C6A5505B1C",
 "ooe-730E3E2FE29E401C97BB2C05553D2F32",
 "ooe-8EFFF49E00E847C985B9C63E42851494",
 "ooe-C7872C3E78314D358A758E9983C5448D",
 "ooe-D3A6C0803D614A68AF45D1A809BC9EC9",
 "ooe-E108DCC3130440768619F2185C37EF81",
 "ooe-E8D2E799E4954929BBFA53672778F025"]

# Abfrage der Details der jeweiligen Datensätze

Für den Import der Daten ruft data.gv.at die URL mit der eindeutigen ID, die erhalten wurde, als Parameter auf. Als Rückantwort erwartet data.gv.at ein JSON-Objekt, das die Metadaten enthält und genau den Vorgaben im OGD Austria Metadaten-Standard 2.6 ([Veröffentlichte Informationen - Reference Server (wien.gv.at)](https://neu.ref.wien.gv.at/at.gv.wien.ref-live/web/reference-server/coopogdveroeffentlichteinformationen)) entspricht.

In der Spezifikation ist eine „Übersicht OGD Metadatenfelder“ zu finden, die Identifier sind in der Spalte „CKAN Feld“ zu finden, wobei Felder des „Metadatenkern“ verpflichtend zu befüllen sind – die Befüllung der „Optionale Metadatenfelder“ wird aber auch empfohlen.

**Beispiel API – Land OÖ**

<http://data.ooe.gv.at/api/ogd/rest/dataset/ooe-AF8B4B2105874E3880F0481544166CE1>

{

 "title": "Windkraftmasterplan",

 "maintainer": "Land Oberösterreich, Abteilung Umweltschutz",

 "tags": [

 "Windkraftmasterplan",

 "Windenergie",

 "Energie",

 "Windkraft"

 ],

 "metadata\_modified": "2013-07-17T14:36:54.00054",

 "license": "Creative Commons Namensnennung 3.0 Österreich",

 "resources": [

 {

 "format": "shp",

 "url": "http://data.ooe.gv.at/ /US\_Vorrangzone\_Windkraftmasterplan.zip",

 "name": "Vorrangzone Windkraftmasterplan"

 },

 {

 "format": "shp",

 "url": "http://data.ooe.gv.at/ US\_Ausschlusszone.zip",

 "name": "Ausschlusszone Windkraftmasterplan"

 }

 ],

 "extras": {

 "attribute\_description": "",

 "metadata\_identifier": "ooe-AF8B4B2105874E3880F0481544166CE1",

 "categorization": [

 "umwelt"

 ],

 "publisher": "Land Oberösterreich, Abteilung Presse",

 "begin\_datetime": "2013-05-24T00:00:00.00000",

 "metadata\_modified": "2013-07-17T14:36:54.00054"

 },

 "id": "ooe-AF8B4B2105874E3880F0481544166CE1",

 "notes": "Vorrangzonen und Ausschlusszonen des \"Windenergiemasterplan Oberösterreich\"\r\nVorrangzone: für Windkraftnutzung potentiell geeignete konzentrierte Standorte, die eine für die Windkraftnutzung wirtschaftlich nutzbare Energiedichte aufweisen, gleichzeitig den ökologisch erforderlichen Rahmenbedingungen entsprechen und das Landschaftsbild mit überregionaler Bedeutung berücksichtigen. \r\nAusschlusszone: Negativdarstellung von Räumen in Oberösterreich, welche für Windkraftnutzung nicht geeignet sind, weil essentielle negative Auswirkungen auf einzelne Fachkriterien zu erwarten sind.",

 "state": "active"

}

# Löschen von Datensätzen

Soll ein Datensatz gelöscht werden, muss der Parameter **state** von **active** auf **deleted** gesetzt werden.

# Häufige Fragen und Antworten

1. Welche Parameter sind absolut minimal nötig und welche optional?

A: die "Metadatenkern-Felder" lt. 2.6er-Spezifikation (Seite 8) sind mindestens erforderlich, alle anderen sind optional, jedoch für ein besseres Verständnis empfohlen. Zusätzlich wird die Angabe des Parameters "state" empfohlen, welcher bei normalen Datensätzen auf "active" gesetzt sein muss (wird auch für Löschungen benötigt, mehr dazu bei der Antwort auf Frage 7)

2. Wie verhält sich der Import wenn Parameter fehlen, Null sind, die Feldlänge überschritten wird, oder zusätzliche Parameter im JSON enthalten sind?

A: Der Datensatz wird an sich immer importiert (solange Name und Titel vorhanden sind), Parameter die fehlen oder null sind, werden auch in CKAN als null gesetzt. Zusätzliche Parameter, die nicht dem Metadatenstandard oder einem zusätzlichen CKAN - Parameter entsprechen werden ignoriert.

3. Welche Felder sind nullable und welche nicht?

A: Im Prinzip sind alle nullable - entspricht dann aber nicht dem Metadatenstandard -> Verantwortlichkeit des Einbringers.

4. Ist das Nichtvorhandensein eines Parameters mit null value gleich zu setzen?

A: Im Prinzip ja

5. Werden Strings auf null oder Leerstring (““) initialisiert? In der Beispielantwort gibt es beides, es sind Null Strings und Leerstrings zu sehen.

A: Wird im Prinzip gleichbehandelt. Bevorzugt Leerstrings, weil beim Import von "null" teilweise Strings mit dem Wert "null" angelegt werden.

6. Welche Datumsformate (Pattern) können verarbeitet werden? Die Spezifikation besagt „YYYY-MM-DD.“ Im einem Beispiel sieht man Timestamps inkl. Zeit und Millisekunden mit den Pattern: „YYYY-MM-DD’T‘HH:mm:ss.SSSSSS“

A: bitte die Datumsformate lt. 2.1er-Spezifikation verwenden (zB metadata\_modified = YYYY-MM-DD, begin\_datetime = YYYY-MM-DDThh:mm:ss)

7. Den Parameter „state“ können wir in der Spezifikation nicht finden, dahingehend bitten wir um nähere Erläuterungen. Wir sind bislang davon ausgegangen, dass einmal veröffentlichte, Daten nicht mehr gelöscht werden sollten. Unter welchen Bedingungen soll/darf/muss „state=deleted“ eingesetzt werden?

A: der Parameter "state" ist in der Metadatenspezifikation nicht enthalten, da für die manuell Bearbeitung von Datensätzen über die Oberfläche nicht explizit angegeben werden muss. Dort können Datensätze auch gelöscht werden - in der Praxis wird dies aber z.B. nur gemacht, wenn z.B. ein Datensatz durch einen anderen ersetzt wird. Dies ist aber eine seltene Ausnahme (bis jetzt 1x vorgekommen) und prinzipiell stimmt ihre Annahme, dass einmal veröffentlichte Daten nicht mehr gelöscht werden.

8. resources:url und extras:metadata\_linkage sind jeweils als Multi-Value-Fields (N) spezifiziert, im Beispiel aber einfache Strings

A: das Beispiel war veraltet und wurde in v2.1 des Vorgabendokuments behoben (Link auf Datensatz des Landes OÖ mit 2 angefügten Ressourcen).

9. Sind Sonderzeichen Unicode (\uXXXX) zu codieren?

A: nein, bitte keine Kodierung von Sonderzeichen vornehmen.

10. Was bedeuten die Parameter „download\_url“ und „ckan\_url“? Worauf zeigt „download\_url“ wenn es mehrere resources Einträge (mehrere Files) gibt?

A: in der aktualisierten Version 2.4 des Vorgabendokuments wurde ein besser verständliches Beispiel eines Metadatenblattes hinzugefügt, der nur Ressourcen auf externen Servern verlinkt. Prinzipiell ist ressources:url für Links zu Ressourcen auf Ihren Servern zu verwenden. ckan\_url und download\_url sind CKAN-eigene Felder und müssen nicht angegeben werden.^

11. Wie sollen Tags angeliefert werden – als comma-separates values oder array in JSON?

A: Als Array in JSON – wichtiger Hinweis: CKAN erlaubt keine Sonderzeichen in Tags, entsprechende Datensätze können nicht geharvestet werden!