

Information

zur Metadatenerfassung in der GDI-NW

bzgl. der

Kennzeichnung der Hochwertigen Datensätze (HVD) in den Metadaten der GDI-DE

17.07.2024



erstellt von: Gst. IMA GDI.NRW

1 Hintergrund

Die Durchführungsverordnung (EU) 2023/138 zur Festlegung bestimmter hochwertiger Datensätze und der Modalitäten ihrer Veröffentlichung und Weiterverwendung¹ („DVO-HVD“) regelt, dass öffentliche Stellen, die im Besitz der im Anhang der DVO-HVD aufgeführten hochwertigen Datensätze sind, diese als OpenData bereitstellen und dabei bestimmte Anforderungen erfüllen müssen. Zudem müssen die hochwertigen Datensätze in ihrer Metadatenbeschreibung gekennzeichnet werden (DVO-HVD Art. 3, Abs. 5).

Die GDI-DE hat hierzu eine entsprechende verpflichtende Konvention für die ISO-Metadaten eines Datenbestands, der unter die DVO-HVD fällt, formuliert (siehe Kap. 3: *Vorgaben für die Kennzeichnung*). Diese Vorgabe gilt dabei zusätzlich zu den bekannten Anforderungen an Metadaten zu Datenbeständen, die zu GovData übertragen und für das dortige OpenData-Portal berücksichtigt werden sollen (vergl. Kap. 6 der GDI-DE "Konventionen zu Metadaten"²). Erst mit der Weiterverwendung der entsprechend gekennzeichneten ISO-Metadaten aus der GDI-DE bei GovData gelangen auch die für HVD relevanten Geo-Datenbestände in die deutsche Gesamtmenge für die gem. DVO-HVD erforderliche Berichterstattung.

2 Geltungsbereich

Diese Unterlage betrifft ausschließlich die Fragestellung, WIE die erforderliche Kennzeichnung in den Metadaten der GDI-DE aussehen muss, um das Funktionieren der o.g. Prozesse zu gewährleisten. Die Frage, WELCHE Datenbestände überhaupt von der DVO-HVD betroffen sind (sofern die DVO-HVD im Einzelfall nicht eindeutig sein sollte), kann hier nicht beantwortet werden. Dies obliegt den jeweils fachlich verantwortlichen Stellen.

Bzgl. der explizit in der DVO-HVD genannten INSPIRE-Themen sei jedoch auf die Sicht der GDI-DE verwiesen, dass die seinerzeit für INSPIRE erarbeitete „Betroffenheit“ auch hier Anwendung findet („Für die HVD, die unter eines oder mehrere der referenzierten INSPIRE-Datenthemen fallen, können die Zuständigkeiten aus den INSPIRE-Zuständigkeiten abgeleitet werden.“; vergl. <https://wiki.gdi-de.org/pages/viewpage.action?pageId=1278214162>, Frage 2.4).

¹ <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32023R0138>

² https://www.gdi-de.org/download/AK_Metadaten_Konventionen_zu_Metadaten.pdf

3 Vorgaben für die Kennzeichnung

Die Anforderung aus der DVO-HVD, die hochwertigen Datensätze in ihrer Metadatenbeschreibung entsprechend zu kennzeichnen, wird in der GDI-DE in Form eines Schlüsselworts, das der jeweiligen HVD-Kategorie entspricht, und in Verbindung mit einer vorgegebenen Quellenangabe, die den HVD-Bezug erkennen lässt, erfüllt. Für den Fall, dass mehrere HVD-Kategorien zugeordnet werden sollen, ist dies durch entsprechend viele Schlüsselwörter zu realisieren.

Mit dieser Kennzeichnung wird erreicht, dass in der Ableitung der Metadaten für GovData die dort benötigten Inhalte gem. dem Profil DCAT-AP-HVD entsprechend belegt werden. Spezielle HVD-Angaben über die Kategorie(n) hinaus sind in den Metadaten der GDI-DE nicht erforderlich.

Für die Erfassung in den Metadaten bzw. für die Abbildung der Metadaten im Format XML stehen zwei Möglichkeiten zur Verfügung:

3.1 HVD-Kategorie und Quellengabe als Freitext:

Im Element *keyword* (GEOkatalog.NRW: „Schlüsselwörter“) wird die jeweils zutreffende HVD-Kategorie³ in deutscher Sprache und Schreibweise gem. nachstehender Aufzählung angegeben:

- **Georaum**
- **Erdbeobachtung und Umwelt**
- Meteorologie
- Statistik
- Unternehmen und Eigentümerschaft von Unternehmen
- **Mobilität**

Die zugehörige Quellengabe wird im Element *thesaurusName* wie folgt angegeben:

- *title*-Element (GEOkatalog.NRW: „Thesaurus“): „**High-value dataset categories**“
- *date*-Element (GEOkatalog.NRW: „Datum“): „**2023-09-27**“
- *dateType*-Element (GEOkatalog.NRW: „Datumstyp“): „**publication**“

Die Abbildung im XML-Metadatensatz (Auszug) sieht dann beispielhaft wie folgt aus:

```

...
<gmd:descriptiveKeywords>
  <gmd:MD_Keywords>
    <gmd:keyword>
      <gco:CharacterString>Georaum</gco:CharacterString>
    </gmd:keyword>
    ...
    <gmd:thesaurusName>
      <gmd:CI_Citation>
        <gmd:title>
          <gco:CharacterString>High-value dataset categories</gco:CharacterString>
        </gmd:title>
      </gmd:CI_Citation>
    </gmd:thesaurusName>
  </gmd:MD_Keywords>
</gmd:descriptiveKeywords>

```

³ Die DVO-HVD sieht sechs Kategorien vor. Lediglich in den drei fett gedruckten Kategorien sind erkennbar Geodaten zugeordnet, was jedoch keinen Ausschluss der drei übrigen Kategorien darstellt.

```

</gmd:title>
<gmd:date>
  <gmd:CI_Date>
    <gmd:date>
      <gco:Date>2023-09-27</gco:Date>
    </gmd:date>
    <gmd:dateType>
      <gmd:CI_DateTypeCode
codeList="https://standards.iso.org/iso/19139/resources/gmxCodelists.xml#CI_DateTypeCode"
codeListValue="publication"/>
    </gmd:dateType>
  </gmd:CI_Date>
</gmd:date>
  ...
</gmd:CI_Citation>
</gmd:thesaurusName>
</gmd:MD_Keywords>
</gmd:descriptiveKeywords>
...

```

3.2 Referenzierung des EU-Vokabulars für HVD-Kategorien als gmX:Anchor-Elemente

Im Element *keyword* wird einer der sechs Werte aus der Liste des Publications Office der EU⁴, der der jeweiligen HVD-Kategorie entspricht, als gmX:Anchor referenziert:

- http://data.europa.eu/bna/c_ac64a52d (Georaum)
- http://data.europa.eu/bna/c_dd313021 (Erdbeobachtung und Umwelt)
- http://data.europa.eu/bna/c_164e0bf5 (Meteorologie)
- http://data.europa.eu/bna/c_e1da4e07 (Statistik)
- http://data.europa.eu/bna/c_a9135398 (Unternehmen und Eigentümerschaft von Unternehmen)
- http://data.europa.eu/bna/c_b79e35eb (Mobilität)

Die Liste des Publications Office der EU selbst wird als zugehörige Quellengabe im Element *thesaurusName* wie folgt referenziert:

- *title*-Element: gmX:Anchor auf „<http://data.europa.eu/bna/asd487ae75>“
- *date*-Element: „2023-09-27“ mit *dateType*=“publication“

Die Abbildung im XML-Metadatensatz (Auszug) sieht dann beispielhaft wie folgt aus:

```

...
<gmd:descriptiveKeywords>
  <gmd:MD_Keywords>
    <gmd:keyword>
<gmX:Anchor xlink:href="http://data.europa.eu/bna/c_ac64a52d">Georaum</gmX:Anchor>
    </gmd:keyword>
    ...
    <gmd:thesaurusName>
      <gmd:CI_Citation>
        <gmd:title>
<gmX:Anchor xlink:href="http://data.europa.eu/bna/asd487ae75">High-value dataset
categories</gmX:Anchor>
        </gmd:title>
        <gmd:date>
          <gmd:CI_Date>
            <gmd:date>

```

⁴ <https://op.europa.eu/de/web/eu-vocabularies/concept-scheme/-/resource?uri=http://data.europa.eu/bna/asd487ae75>

```
<gco:Date>2023-09-27</gco:Date>
</gmd:date>
<gmd:dateType>
  <gmd:CI_DateTypeCode
codeList="https://standards.iso.org/iso/19139/resources/gmxCodelists.xml#CI_DateTypeCode"
codeListValue="publication"/>
  </gmd:dateType>
</gmd:CI_Date>
</gmd:date>
...
</gmd:CI_Citation>
</gmd:thesaurusName>
</gmd:MD_Keywords>
</gmd:descriptiveKeywords>
...
```

Anmerkung: die hier gezeigte Kennzeichnung per Referenzierung ist mit dem Editor zum GEOkatalog.NRW momentan nicht herstellbar. Auch das händische Eintragen der o.g. Referenzen auf das EU-Vokabular in die im Editor vorhandenen Freitext-Felder führt nicht zur gewünschten Abbildung im XML!

Bitte beachten: Beide Wege bedingen gleichermaßen, dass Schlüsselwörter, denen eine Quellenangabe oder ein Thesaurus zugeordnet ist, technisch (d.h. in der Abbildung in XML) von denjenigen Schlüsselwörtern aus anderen Quellen oder ohne Quellenangabe getrennt werden müssen (vergl. Kapitel 4.2.2.2.1 im NRW-Leitfaden zur Metadatenerfassung⁵). Bezogen auf die Erfassung im GEOkatalog.NRW bedeutet dies, dass die HVD-Kategorie unter den Schlüsselwörtern in einer eigenen Registerkarte zu erfassen ist und nicht mit anderen Schlüsselwörtern vermischt wird.

4 Weitere Informationen

Sowohl seitens GovData als auch durch die GDI-DE wurden umfangreiche Informationen zum Thema „Hochwertige Datensätze“ aufbereitet und in FAQ-Bereichen öffentlich bereitgestellt.

Die Fragensammlungen sind zu finden unter

- <https://wiki.gdi-de.org/pages/viewpage.action?pageId=1278214162>
- <https://www.govdata.de/web/guest/hochwertige-datensaetze>

Die hier dokumentierten Vorgaben zur HVD-Kennzeichnung werden im Rahmen der regelmäßigen Dokumentenpflege auch in die Metadatenkonventionen der GDI-DE sowie den NRW-Leitfaden zur Metadatenerfassung einfließen.

⁵ https://www.gdi.nrw/system/files/media/document/file/2024-04-23_leitfaden_metadaten_gdi-nw_v1_11.pdf