

# Kennzeichnung hochwertiger Datensätze (HVD) in den Metadaten

## 1 Hintergrund

Die Durchführungsverordnung (EU) 2023/138 zur Festlegung bestimmter hochwertiger Datensätze und der Modalitäten ihrer Veröffentlichung und Weiterverwendung („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"). Erst mit der Weiterverwendung der entsprechend gekennzeichneten ISO-Metadaten aus der GDI-DE gelangen die für HVD relevanten Geo-Datenbestände in das GovData-Portal und damit in die deutsche Gesamtmenge für die gem. DVO-HVD erforderliche Berichterstattung.

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

## 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.

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

#### 3.1 HVD-Kategorie und Quellenangabe als Freitext

Im Element *keyword* wird die jeweils zutreffende HVD-Kategorie<sup>1</sup> 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: „**High-value dataset categories**“
- *date*-Element: „**2023-09-27**“
- *dateType*-Element: „**publication**“

Die Abbildung im XML-Metadatenatz (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: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/re-
sources/gmxCodelists.xml#CI_DateTypeCode" codeListValue="publication"/>
            </gmd:dateType>
          </gmd:CI_Date>

```

<sup>1</sup> 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: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 EU<sup>2</sup>, der der jeweiligen HVD-Kategorie entspricht, als gmx:Anchor referenziert:

- [http://data.europa.eu/bna/c\\_ac64a52d](http://data.europa.eu/bna/c_ac64a52d) (Georaum)
- [http://data.europa.eu/bna/c\\_dd313021](http://data.europa.eu/bna/c_dd313021) (Erdbeobachtung und Umwelt)
- [http://data.europa.eu/bna/c\\_164e0bf5](http://data.europa.eu/bna/c_164e0bf5) (Meteorologie)
- [http://data.europa.eu/bna/c\\_e1da4e07](http://data.europa.eu/bna/c_e1da4e07) (Statistik)
- [http://data.europa.eu/bna/c\\_a9135398](http://data.europa.eu/bna/c_a9135398) (Unternehmen und Eigentümerschaft von Unternehmen)
- [http://data.europa.eu/bna/c\\_b79e35eb](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-Metadatenatz (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>
            <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>

```

<sup>2</sup> <https://op.europa.eu/de/web/eu-vocabularies/concept-scheme/-/resource?uri=http://data.europa.eu/bna/asd487ae75>

```
</gmd:thesaurusName>  
</gmd:MD_Keywords>  
</gmd:descriptiveKeywords>  
...
```

**Anmerkung:** die Kennzeichnung per Referenzierung ist mit einem Metadaten-Editor nicht zwangsläufig herstellbar! Das Eintragen der o.g. Referenzen auf das EU-Vokabular in die im Editor vorhandenen Freitext-Felder führt i.d.R. 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.

## 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/x/EgAwT>
- <https://www.govdata.de/web/guest/hochwertige-datensaetze>