



# LaVaH

Langzeitverfügbarkeit digitaler Inhalte an hessischen Hochschulen

## Welttag Digitale Erhaltung / World Digital Preservation Day 2024

### **Metadaten**

Dion Wagner | Frankfurt University of Applied Sciences

Wozu Metadaten?



## Wozu Metadaten?



- DNB: *„Strukturierte Daten, die Merkmale und Eigenschaften von Ressourcen, z. B. von Büchern, Webdokumenten, Videos, Bildern oder Daten, einheitlich beschreiben.“* (<https://www.dnb.de/glossar>)
- Metadaten sind „Daten über Daten“
  - beschreiben
  - strukturieren
  - geben Kontext
- dadurch erleichtern sie die Verwaltung, Auffindbarkeit und Nutzung von Daten

- **Datenorganisation und Strukturierung:** Übersichtliche Kategorien und gezielte Sortierung
- **Suchbarkeit und Auffindbarkeit:** Erhöhte Sichtbarkeit und Effizienz in Datenbanken
- **Rechteverwaltung:** Infos zu Urheberrechten und Zugriffsrechten
- **Langzeitarchivierung:** Sichern Kontext für künftige Nutzung

- Beispiele für Metadaten bei einem Buch:
  - Titel
  - Autor\*in
  - Erscheinungsjahr
  - Verlag
  - ISBN



- Ausgangslage für LaVaH:  
Daten aus unterschiedlichen Sammlungen
  - mit unterschiedlichen Metadaten(-standards)
  - in unterschiedlichen Formaten
  - in unterschiedlichen Katalogen nachgewiesen

- beschreibende Metadaten sollen für LaVaH-Archivierung in Dublin Core (DC) abgeliefert werden
- Metadatenschema zur Beschreibung von analogen und digitalen Ressourcen
- wird als Community-Projekt weiterentwickelt
- weltweit gebräuchlicher Standard
- DC kann z.B. in XML dargestellt werden
- dadurch sowohl von Menschen als auch von Maschinen lesbar

- Zwei Fassungen:
  - Simple Dublin Core: allgemeine Metadatenelemente
    - Beispiel: `<dc:date>` für ein allgemeines Datum
  - Qualified Dublin Core: spezifische inhaltliche Beschreibungen
    - Beispiel: `<dcterms:created>` für das Erstellungsdatum

- Projekt-AG erstellte Kernset:
  - Umfrage zu Metadatenlieferung mit allen beteiligten Einrichtungen
  - „Was soll?“ vs. „Was ist möglich?“
  - Kernset als „Kleinstes Gemeinsames Nenner“
  - bestehend aus sechs beschreibenden Metadaten

<b>Name</b>	<b>Beschreibung</b>
dc:date	Ein Zeitpunkt oder eine Zeitspanne im Zusammenhang mit einem Ereignis im Lebenszyklus der Ressource.
dc:identifizier	Ein eindeutiger Verweis auf die Ressource innerhalb eines bestimmten Kontexts.
dc:creator	Eine Person, die wesentlich für die Erstellung der Ressource verantwortlich ist.
dc:title	Name der Ressource.
dcterms:license	Ein Dokument, das die offizielle Erlaubnis erteilt, etwas mit der Ressource zu tun.
dcterms:accessRights	Informationen darüber, wer und unter welchen Bedingungen auf die Ressource zugreifen darf

- diese Felder müssen vorhanden sein
- Befüllung erfolgt durch Datenkurator\*innen
- keine Validierung durch Rosetta oder hebis VZ
- keine verbindlichen Vorgaben zur Befüllung, aber Empfehlungen
- weitere Metadaten selbstverständlich möglich

# Kernset: Datum (dc:date)

- Datumsangabe kann sich auf unterschiedliche Ereignisse/Zeitspannen beziehen, z.B.:
  - Jahreszahl
  - Datum
  - ermitteltes oder geschätztes Datum (in [ ])
- spezifizieren mit Qualifiern (DC terms), z.B.:
  - Date Available, Date Created, Date Accepted, Date Submitted, Date Copyrighted, Date Issued, Date Modified, Date Valid
- nur eine Variante sollte je Sammlung angewendet werden (dc:date ODER dcterms:xx), da unterschiedliche Prüfroutinen existieren
- bei unbekanntem Datum kann Platzhalter (YYYY-MM) oder (YYYY) angegeben werden

## Beispiele

```
<dc:date>2017-01-31</dc:date>  
<dc:date>31-01-2017</dc:date>  
<dc:date>2017</dc:date>  
<dc:date>[2017-01-31]</dc:date> (ermitteltes Datum)  
<dc:date>[ca. 2017-01-31]</dc:date> (geschätztes Datum)  
<dc:date><dcterms:created>2017-01-31</dcterms:created></dc:date>  
<dc:date>YYYY</dc:date> (unbekanntes Datum)  
<dc:date>YYYY-MM</dc:date> (unbekanntes Datum)
```

# Kernset: Identifier (dc:identifier)

- Vielzahl unterschiedlicher Identifier, die von Einrichtungen genutzt oder vergeben werden können
- mehrere Identifier können angegeben werden, z.B. interne PPN und globaler DOI
- Empfehlung: Immer angeben, um welchen Identifier es sich handelt
- Besonderheit beim DC-Ingest beachten!
  - dc:identifier spezifiziert durch dcterms:URI lokalisiert die Datei in der Ordnerstruktur

## Beispiele

```
<dc:identifier>doi:10.22029/ko.2017.1004</dc:identifier>
```

```
<dc:identifier>isbn:978-3-902811-74-5</dc:identifier>
```

```
<dc:identifier type="doi">10.22029/ko.2017.1004</dc:identifier>
```

```
<dc:identifier>Signatur: 8° Ms. Hass. 122</dc:identifier>
```

- Personen, welche in Zusammenhang mit dem Dokument stehen, z.B.:
  - Verfasser\*in
  - Herausgeber\*in
  - Interpret\*in
  - Ersteller\*in
- falls keine Information vorhanden:  
„keine Angabe“ oder „unbekannt“ möglich

## Beispiele

```
<dc:creator>Schütz, Georg</dc:creator>
```

```
<dc:creator>Hochschule für Musik und Darstellende Kunst, Frankfurt am Main</dc:creator>
```

```
<dc:creator>keine Angabe</dc:creator>
```

- ein oder mehrere Titel
- falls keine Information vorhanden: Dateiname
- spezifizieren mit Qualifiern, z.B.:  
Sprache, Untertitel
- nur eine Variante sollte je Sammlung angewendet werden, da unterschiedliche Prüfroutinen existieren

## Beispiele

```
<dc:title>Klavierabend / Studierende der Klasse Prof. Florian Hölscher</dc:title>  
<dc:title>Wandertage im Harz</dc:title>  
<dc:title xml:lang="de">Wandertage im Harz</dc:title>
```

- Copyright Informationen, Creative Commons oder andere Angaben
- Klärt:
  - welche Aktionen können mit diesem Objekt durchgeführt werden?
  - welche Urheberrechte liegen vor?

## Beispiele

```
<dcterms:license>Creative Commons - No rights reserved (CC0)</dcterms:license>  
<dcterms:license>CC0</dcterms:license>  
<dcterms:license>http://creativecommons.org/licenses/by/3.0/</dcterms:license>  
<dcterms:license>https://rightsstatements.org/vocab/InC/1.0/</dcterms:license>  
<dcterms:license>openAccess</dcterms:license>  
<dcterms:license>Protected by German copyright law</dcterms:license>
```

- regeln, wer und unter welchen Umständen das digitale Objekt nutzen darf
  - räumliche Einschränkungen (z.B. „nur im Lesesaal“)
  - zeitliche Einschränkungen (Embargo)
  - weitere

## Beispiele

```
<dcterms:accessRights>Keine Beschränkung</dcterms:accessRights>  
<dcterms:accessRights>public -Frei abrufbar</dcterms:accessRights>  
<dcterms:accessRights>reader -Nur an Rechnern der Bibliothek  
abrufbar</dcterms:accessRights>
```



Vielen Dank für Ihre Aufmerksamkeit



Dion Wagner

Frankfurt University of Applied Sciences

[d.wagner@bibl.fra-uas.de](mailto:d.wagner@bibl.fra-uas.de)