

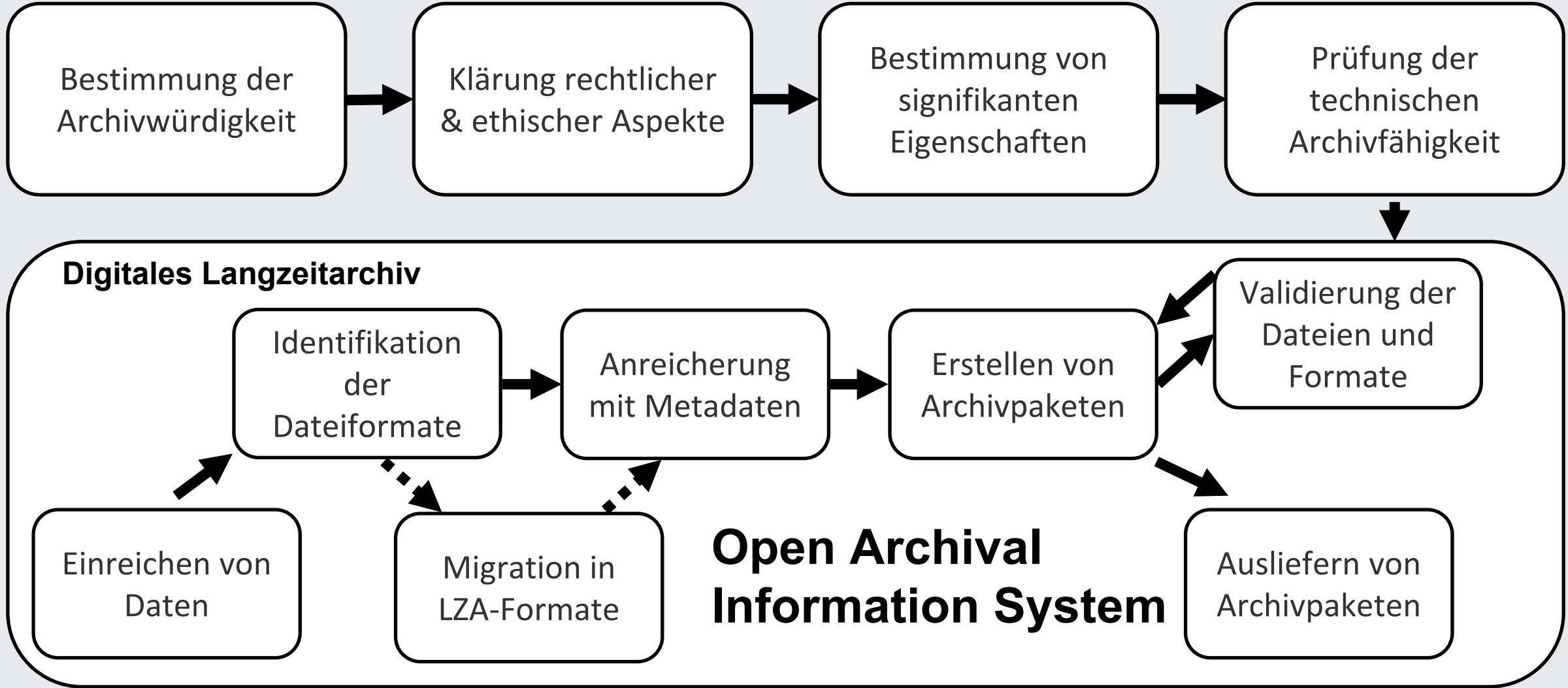


**Datenmodell und Implementierung eines OAI
Systems am Germanischen
Nationalmuseum Nürnberg**

GERMANISCHES
**NATIONAL
MUSEUM**

Robert Nasarek
Germanisches Nationalmuseum Nürnberg
r.nasarek@gnm.de
<https://fedihum.org/@rnsrk>

Digitale Langzeitarchivierung



Datenmodell: Ebenen

Content-
Management

Inhaltsbeschreibung

Archivpaket

Startseite

Intellektuelle Einheiten

Titel enthält

Anwenden

Abbildung

Schadenskartierung (statisch) UV-Abbildungen

Applikation

Datenmodell Drupal/WissKI-Instanz Graphdatenbank Pigmentdatenbank

Audio

Audiobeitrag

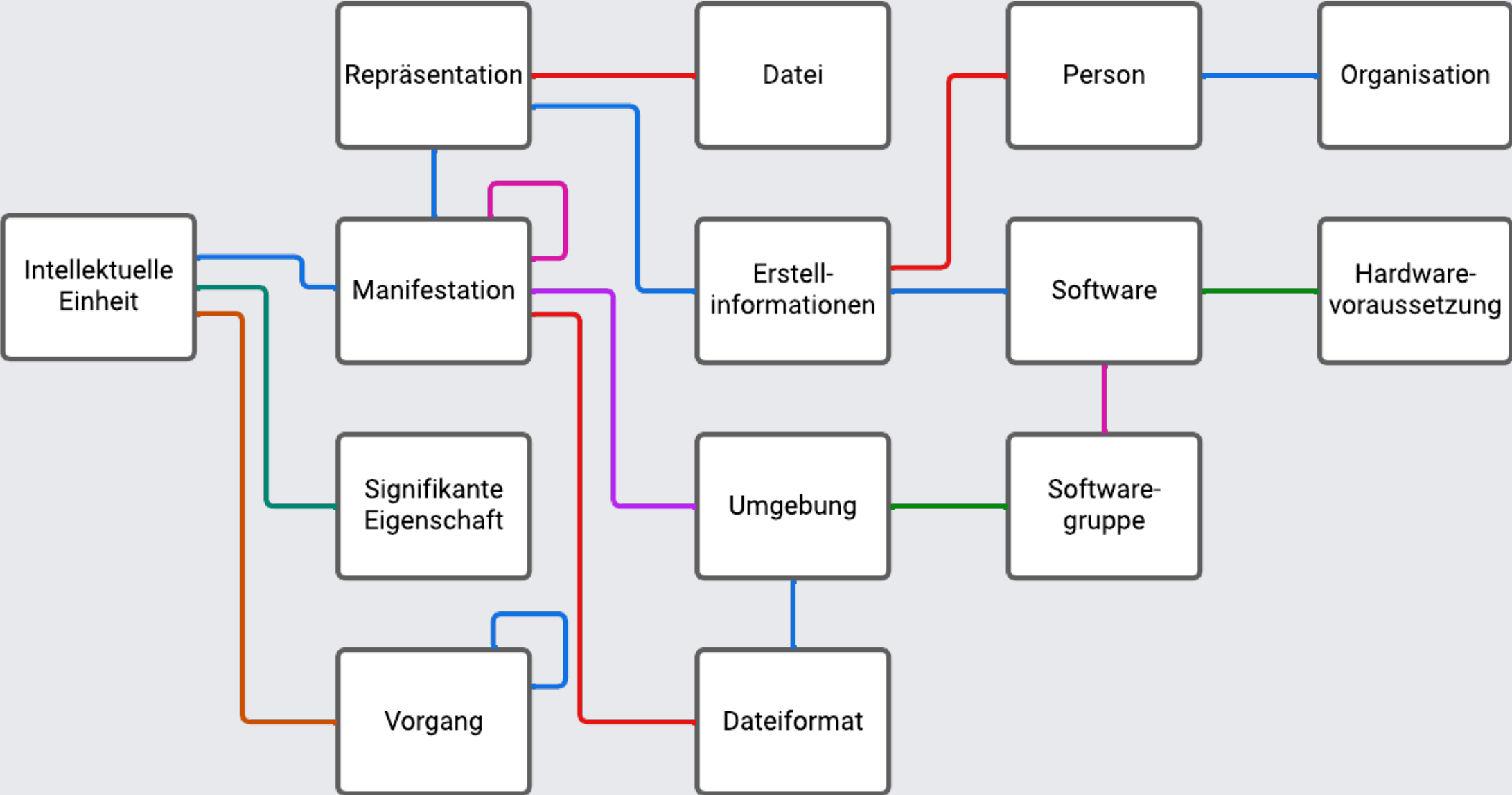
- Startseite
- Dokumentation ▶
- Richtlinien und Hilfe ▶
- Roadmap
- Werkzeuge ▶

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/
22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/
rdf-schema#"
  xmlns:crm="http://www.cidoc-crm.org/cidoc-crm/">

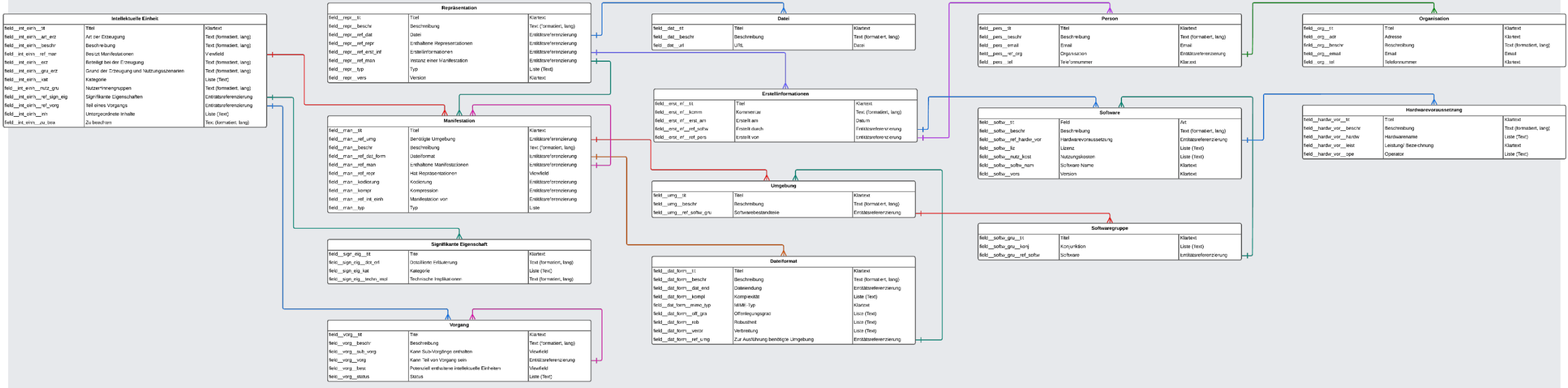
  <crm:E21_Person rdf:about="http://example.org/
person/123">
    <rdfs:label>Max Mustermann</rdfs:label>
    <crm:P1_is_identified_by>
      <crm:E41_Appellation>
        <rdfs:label>Mustermann, Max</rdfs:label>
      </crm:E41_Appellation>
    </crm:P1_is_identified_by>
  </crm:E21_Person>
</rdf:RDF>
```

- data ✓
- meta ✓
- bag-info.txt ✓
- bagit.txt ✓
- manifest-md5.txt ✓
- manifest-sha512.txt ✓
- tag-manifest-md5.txt ✓
- tag-manifest-sha512.txt ✓

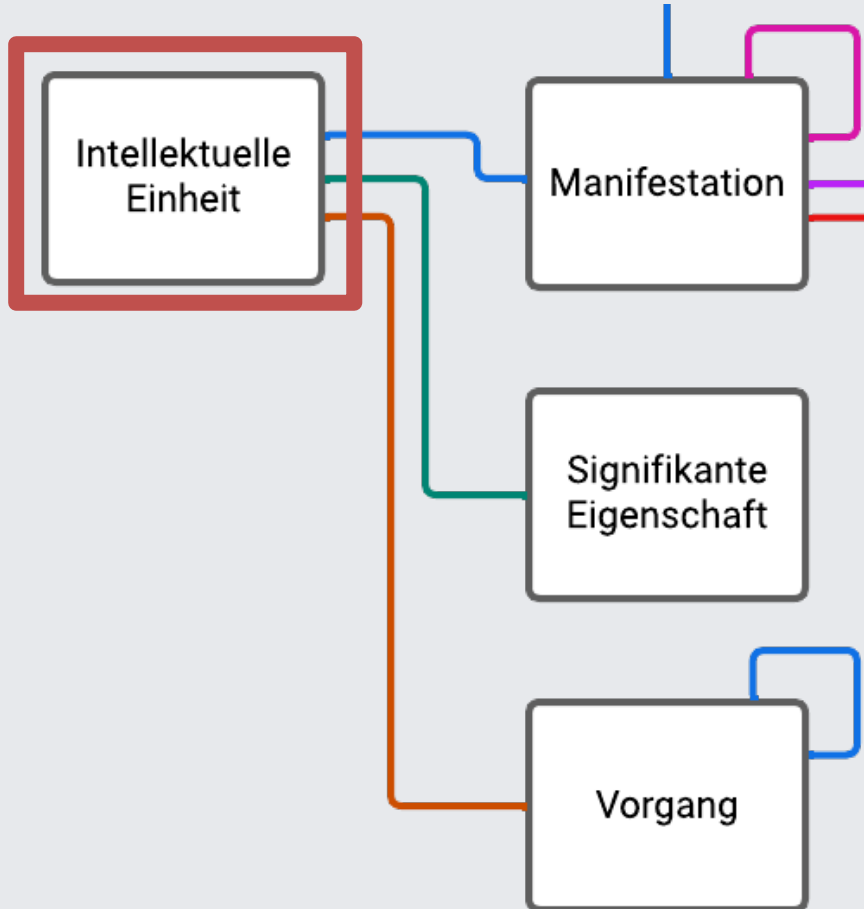
Content-Management



Content-Management



Content-Management



Drupal/WissKI-Instanz

Ansicht Bearbeiten Löschen Revisionen Devel-Modul

Gespeichert von **rnsrk** am Mi., 20.09.2023 - 14:21

[PDF anzeigen](#)

Kategorie

Applikation

Untergeordnete Inhalte

Abbildungen

Audio

Binaries

Text

Video

Beteiligt bei der Erzeugung

- Mitarbeitende des GNM

Art der Erzeugung

- Born digital

Grund der Erzeugung und Nutzungsszenarien

- Content Management
- Publikationsmedium

Nutzer*innengruppen

- Mitarbeitende des GNM
- Öffentlichkeit

Signifikante Eigenschaften

[Bildqualität erhalten](#)

[Maschinenlesbarkeit](#)

[Menschenlesbarkeit](#)

[Struktur erhalten](#)

[Tonqualität erhalten](#)

Besitzt Manifestationen

Drupal 10 mit WissKI 3.14

Sicherung (Drupal

Datenbank Dump,

Triplesstore Dump, Sites-

Ordner, Pathbuilder)

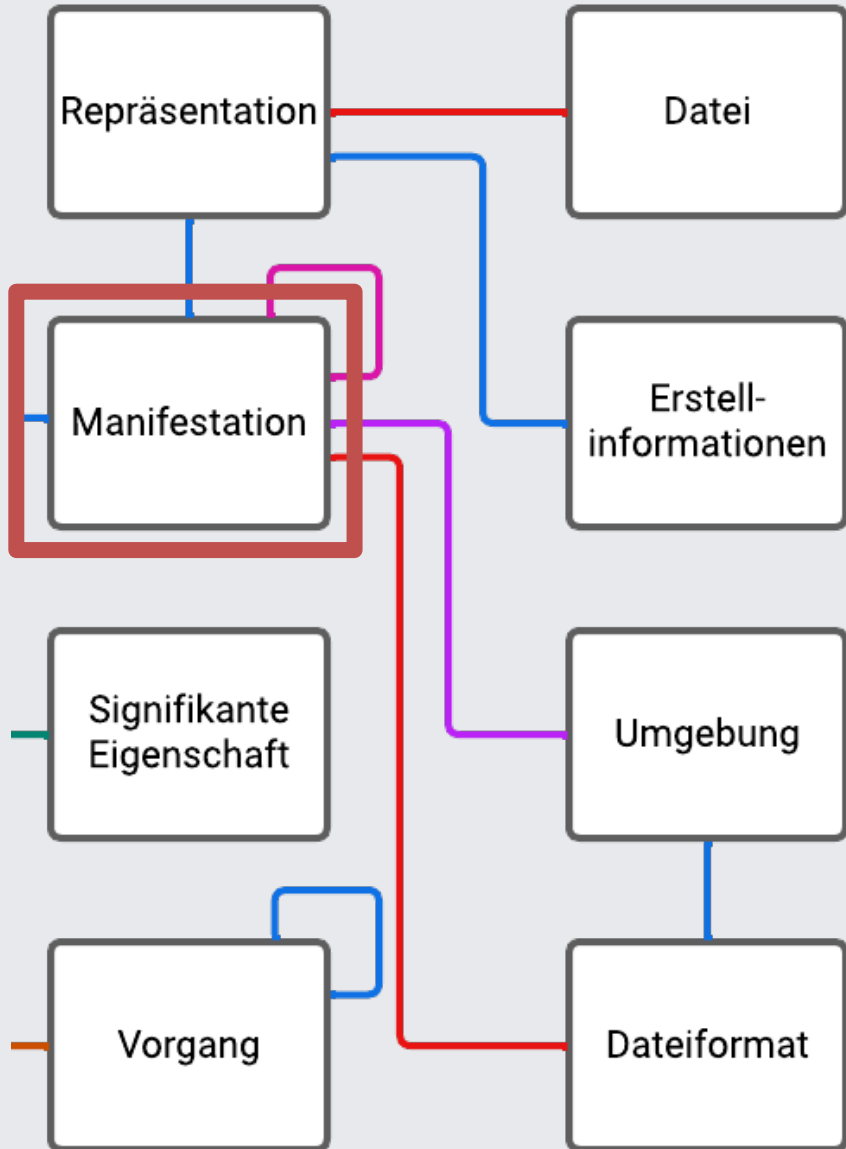
Intellektuelle Einheit

- Unabhängig von einer Implementierung
- Definiert durch signifikante Eigenschaften was erhalten werden soll

Manifestation

- Implementierungsform der Intellektuellen Einheit

Content-Management



Drupal 10 mit WissKI 3.14 Sicherung (Drupal Datenbank Dump, Triplesstore Dump, Sites-Ordner, Pathbuilder)

Ansicht Bearbeiten Löschen Revisionen Devel-Modul

Gespeichert von **rnsrk** am Fr., 13.10.2023 - 14:06

Benötigte Umgebung

LAMP+TS+IIP+Solr (2024)

Enthaltene Manifestationen

- Drupal 10 Webverzeichnis (tar-gepackt)
- MariaDB SQL Dump (sql)
- Triplestore Dump (n-quad)
- Pathbuilder Dump (xml)

Typ

Container

Manifestation von

Drupal/WissKI-Instanz

Hat Repräsentationen

- dlza.gnm.de
- Sicherung (CRON Script
- 28.02.24-16:00:00)

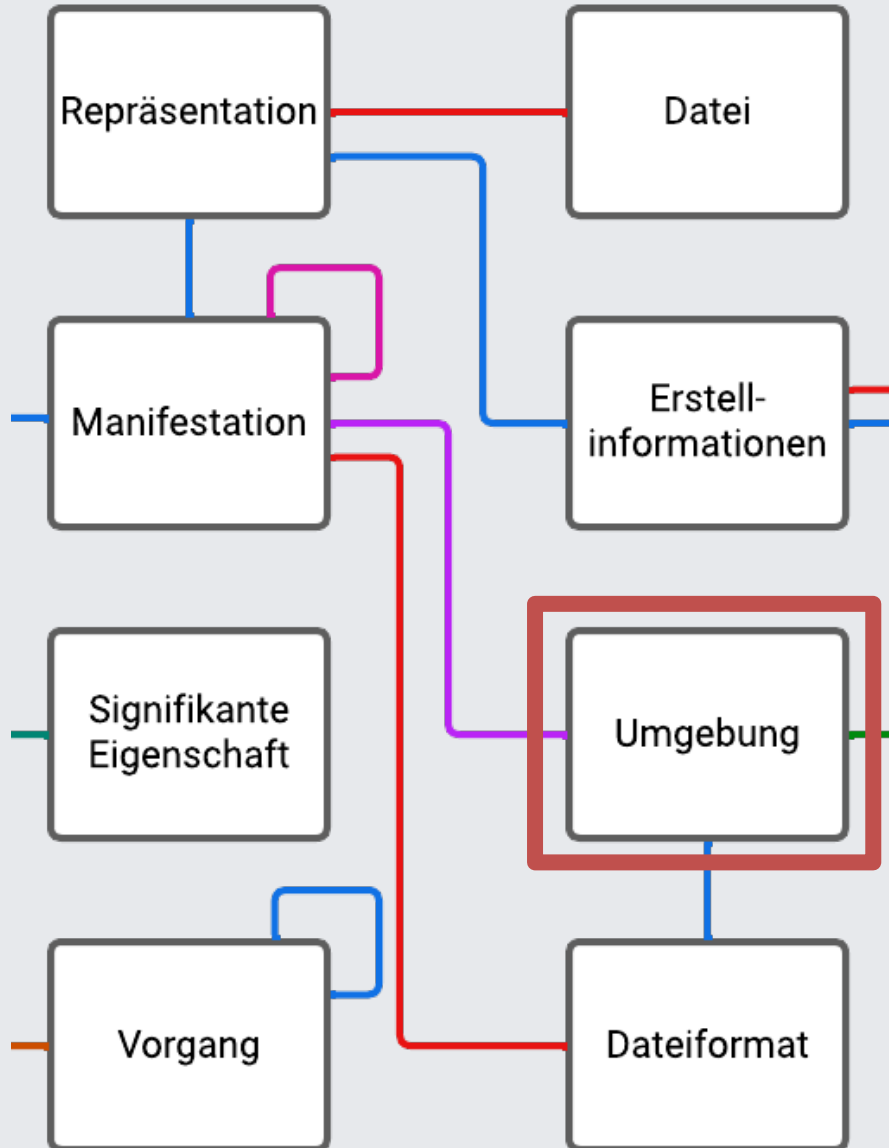
Umgebungen

- Abhängige Software oder Runtime zur Ausführung

Repräsentationen

- Tatsächlich existierende Instanzen

Content-Management



LAMP+TS+IIP+Solr (2024)

Ansicht Bearbeiten Löschen Revisionen Devel-Modul

Gespeichert von **rnsrk** am Mi., 28.02.2024 - 08:58

Beschreibung

Runtime Umgebung mit im Jahr 2024 unterstützten Versionen von Apache 2, MariaDB, PHP, Triplestore, IIP-Server und Solr-Server

Softwarebestandteile

Debian-basiertes Linux (Debian 11-12 oder Ubuntu 20.04-22.04)

Apache 2

SQL Datenbank (MariaDB 10.06/ 10.11 oder MySQL MySQL Community Server 8.0.36/ 8.3.0)

PHP (8.0/8.1/8.2)

PHP Extensions gd, mbstring, xml, mysql, zip, curl, apcu, opcache

Triplestore (Blazegraph 2.1.5 oder GraphDB free 9/10)

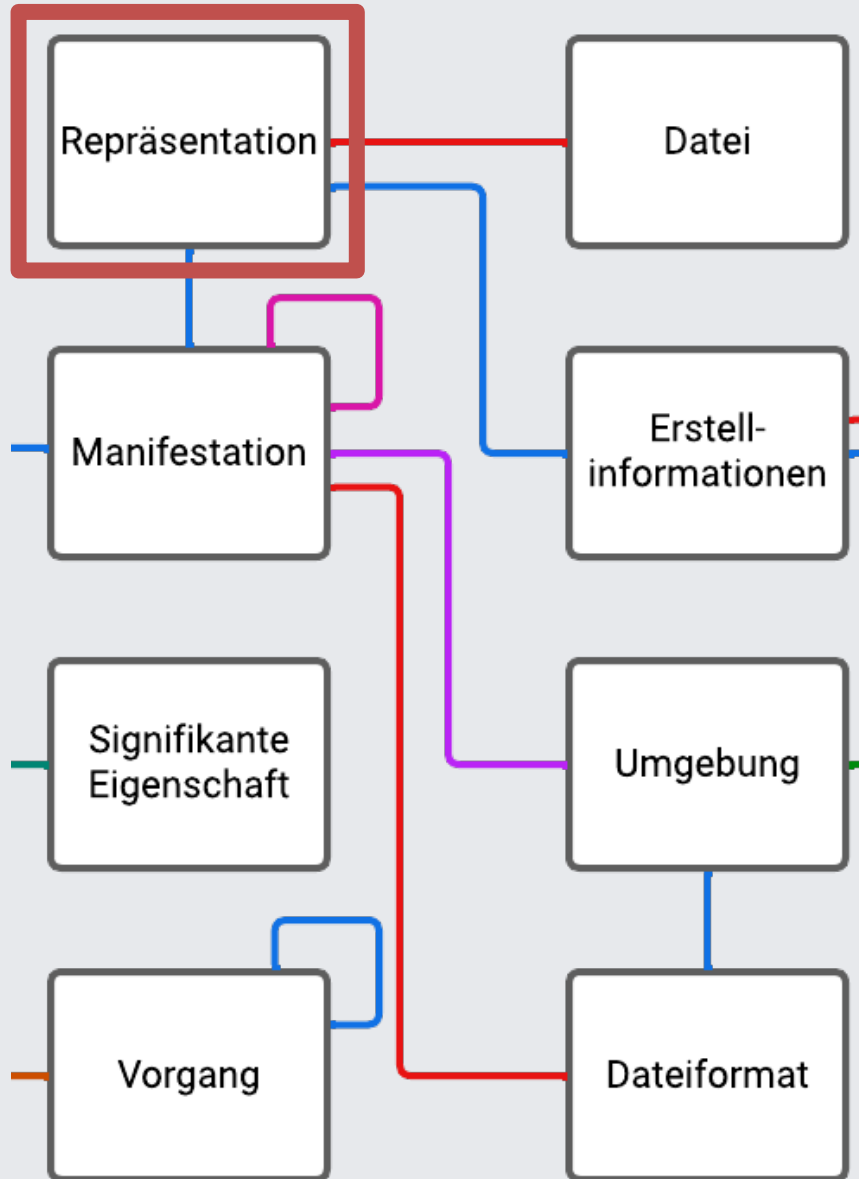
IIP-Server (IIPImage 1.1/1.2)

Solr Server (8/9)

Softwarebestandteile

- Austauschbare Software mit groben Versionsfaktor

Content-Management



dlza.gnm.de Sicherung (CRON Script 28.02.24-16:00:00)

Ansicht Bearbeiten Löschen Revisionen Devel-Modul

Gespeichert von **rnsrk** am Mi., 28.02.2024 - 16:00

Beschreibung

Eine Webapplikation die es ermöglicht verschiedene im Germanischen Nationalmuseum anfallenden Daten digital langzeit zu archivieren. Sie ist ein vollständiges OAI-System und bietet Werkzeuge und Routinen zur Migration, Speicherung, Validierung und Reparatur von Daten an.

Instanz einer Manifestation

Drupal 10 mit WissKI 3.14 Sicherung (Drupal Datenbank Dump, Triplesstore Dump, Sites-Ordner, Pathbuilder)

Enthaltene Repräsentationen

- dlza.gnm.de Drupal-Web-Verzeichnis (28.02.24-15:20:30)
- dlza.gnm.de MySQL Dump (28.02.24-15:13:52)
- dlza.gnm.de Triplestore Dump (28.02.24-13:18:52)
- dlza.gnm.de Pathbuilder Dump (28.02.24-14:14:51)

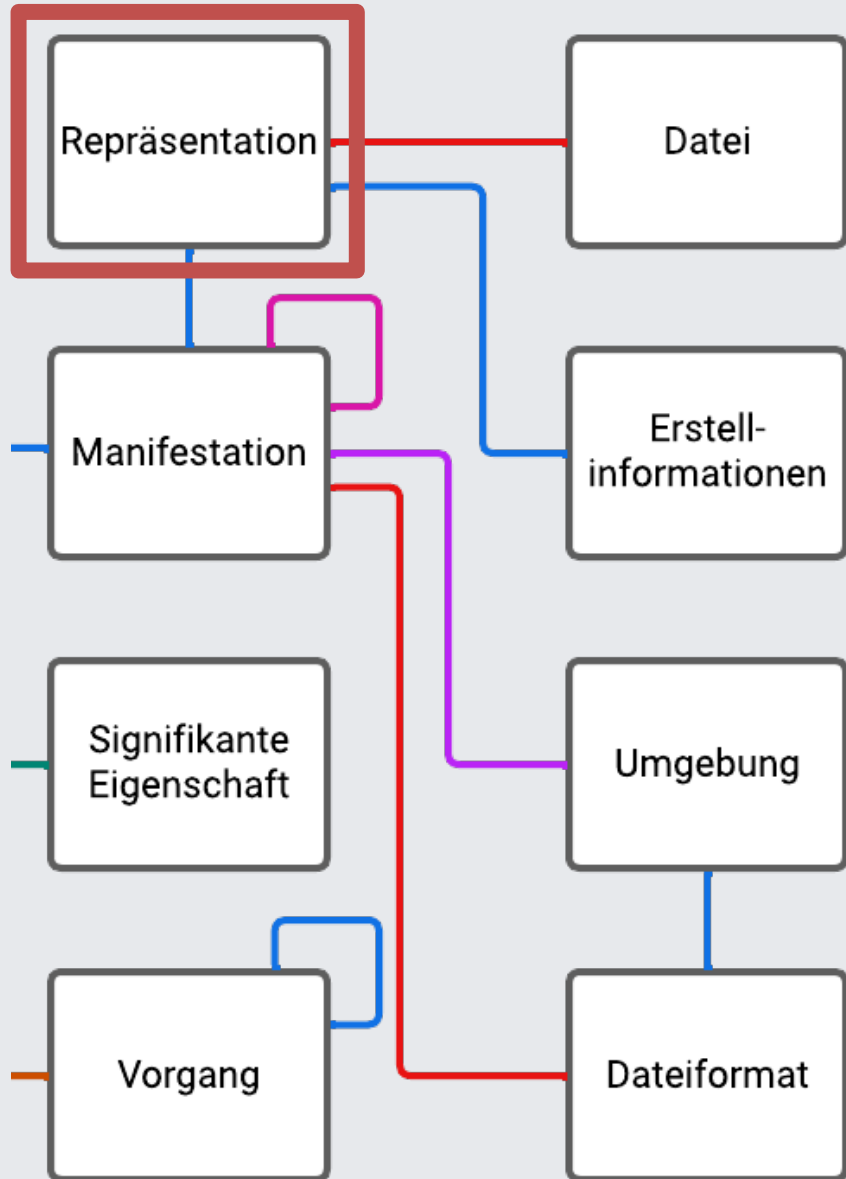
Erstellinformationen

GNM Website Archiving Script am 28.02.2024-16:00:00

Typ

Container

Content-Management



dlza.gnm.de MySQL Dump (28.02.24-15:13:52)

Ansicht Bearbeiten Löschen Revisionen Devel-Modul

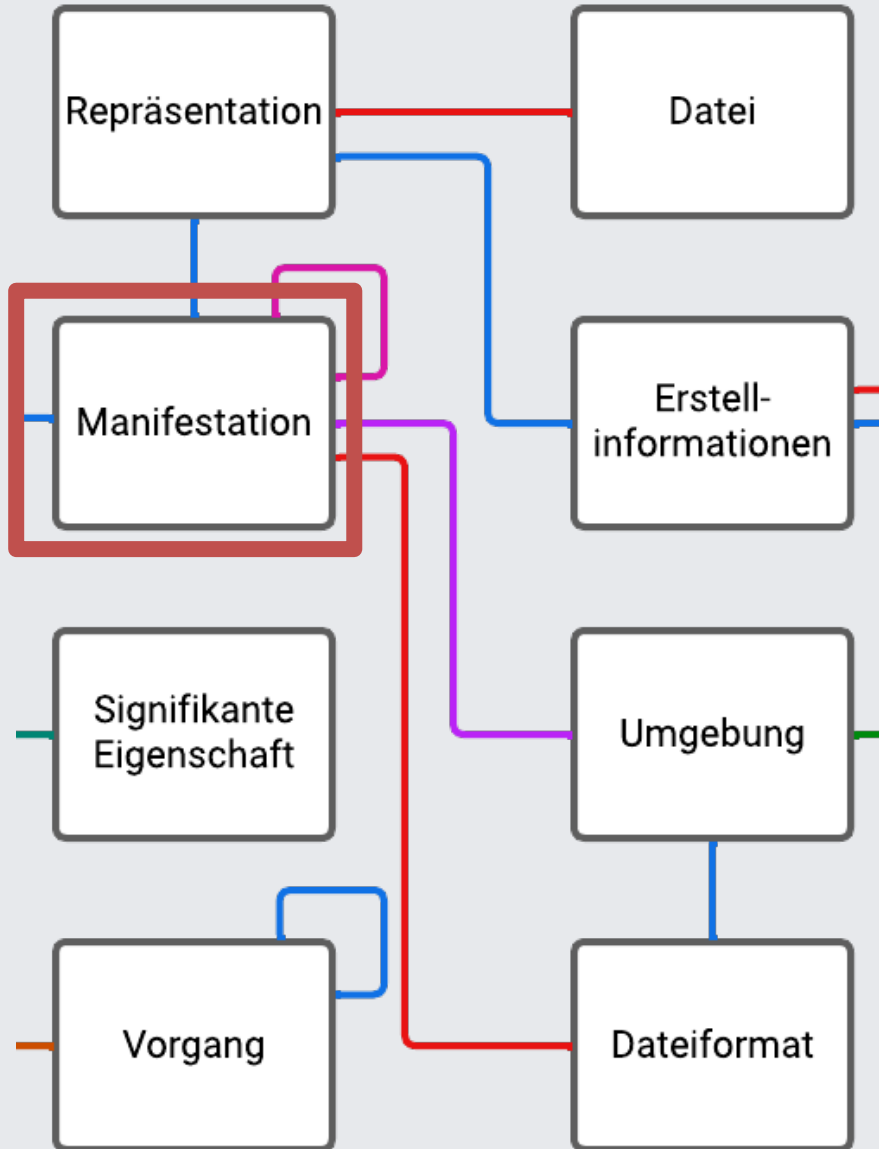
Gespeichert von **rnsrk** am Mi., 28.02.2024 - 16:04

Instanz einer Manifestation
MariaDB SQL Dump (sql)
Datei
dlza.gnm.de sql 28.02.25-15:13:52
Erstellinformationen
Robert Nasarek mit mysqldump am 28.02.2024-15:13:52
Typ
Datei

Datei

- Eigentliche Datei

Content-Management



MariaDB SQL Dump (sql)

Ansicht Bearbeiten Löschen Revisionen Devel-Modul

Gespeichert von **rnsrk** am Mi., 18.10.2023 - 11:06

Benötigte Umgebung

MariaDB Datenbank

Dateiformat

MariaDB SQL (sql)

Typ

Datei

Hat Repräsentationen

dlza.gnm.de MySQL

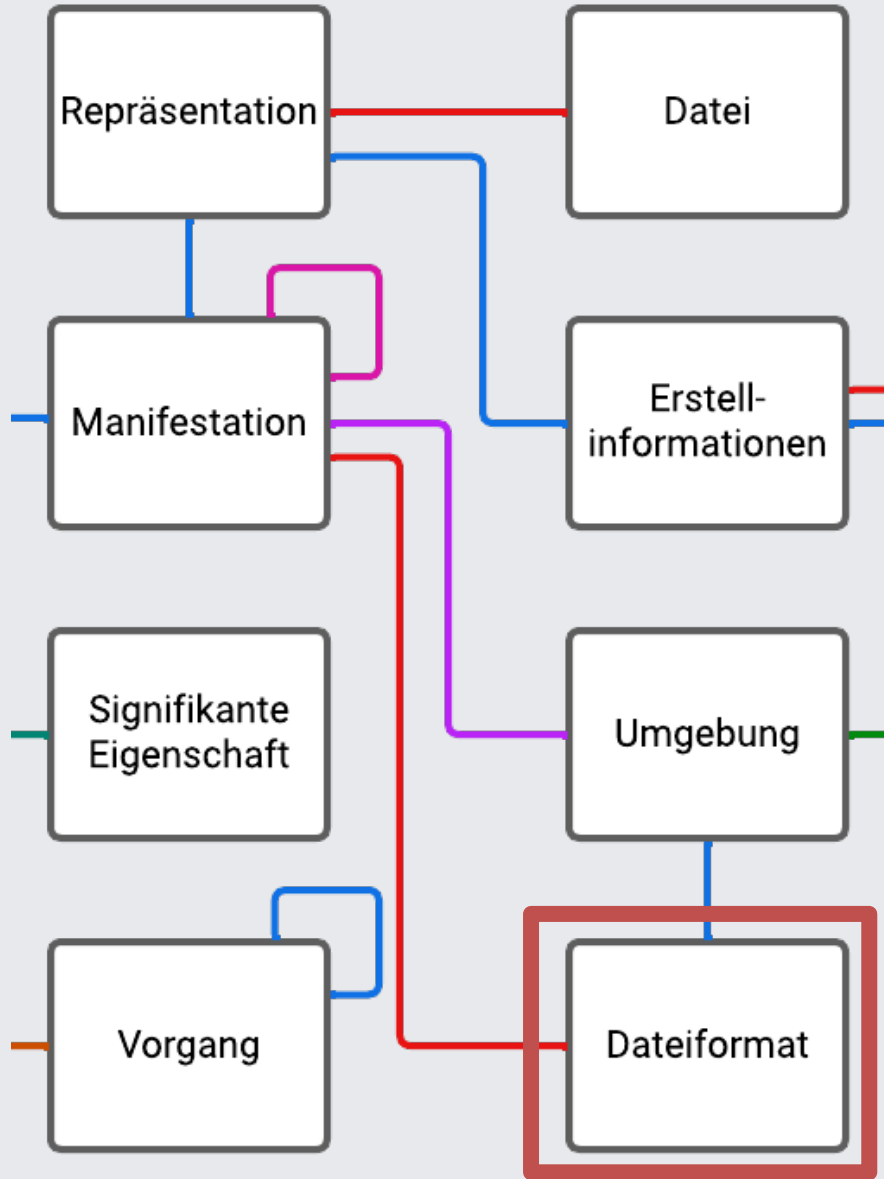
Dump

(28.02.24-15:13:52)

Dateiformat

- Ist Eigenschaft der Manifestation nicht der Datei

Content-Management



MariaDB SQL (sql)

Ansicht Bearbeiten Löschen Revisionen Devel-Modul

Gespeichert von **rnsrk** am Fr., 13.10.2023 - 14:30

Dateiendung

sql

MIME-Typ

application/sql

Offenlegungsgrad

offen und frei

Verbreitung

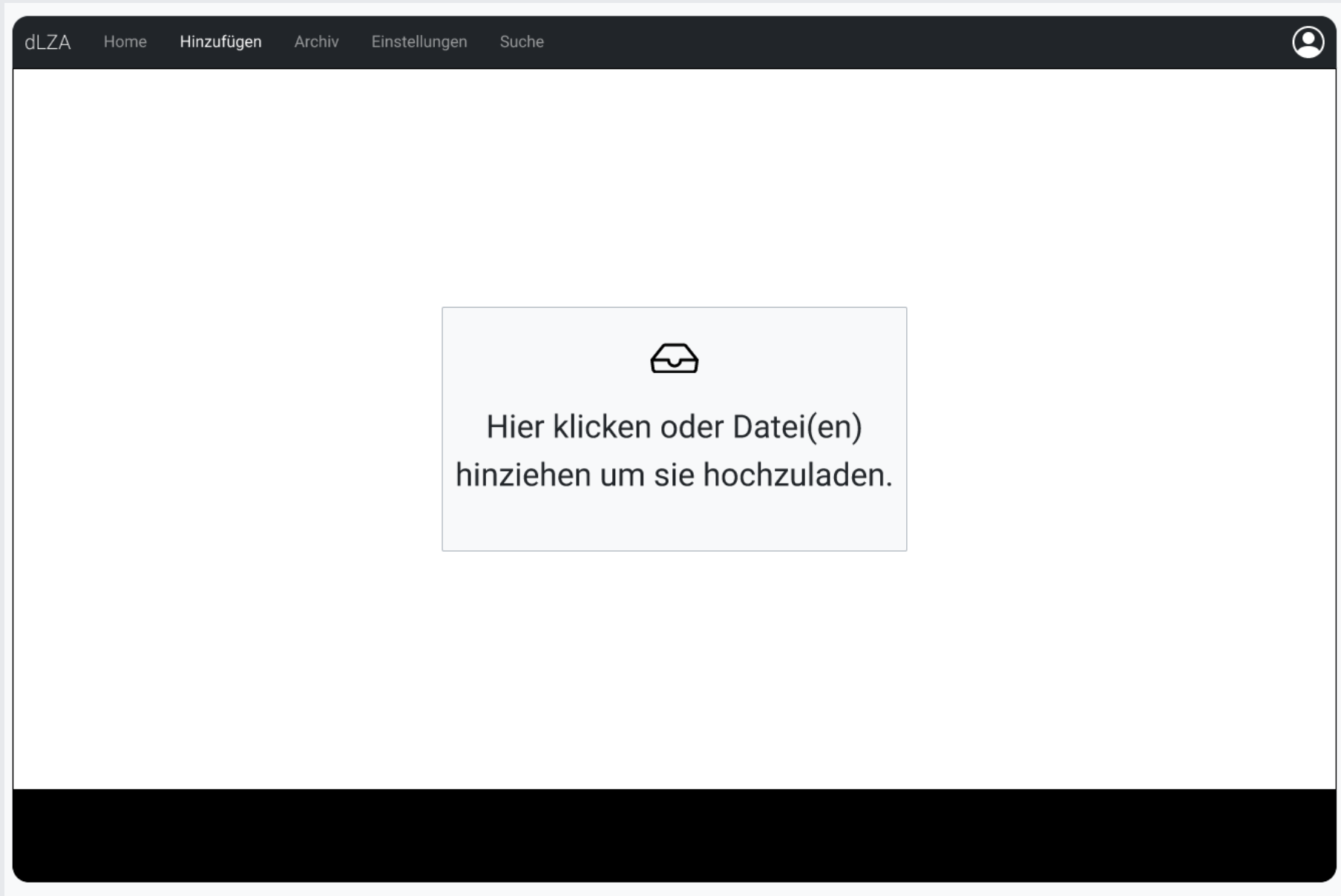
sektoral


Komplexität

einfach


Robustheit

sehr robust



dLZA Home Hinzufügen Archiv Einstellungen Suche 


Paket-Informationen ✕

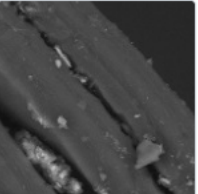
Abteilung: IKK Zeitraum: 03.11.2022 → 22.06.2023 Kommentar: ...  Titel: Spanischer Mantel des Stephan III. Pi
Geschäftsgang: Restaurierung Verknüpft mit Objekt(en): T550 Sammlung: Textilien-Schmuck
Nummer: 4 Version: 1 Sichtbar für: IKK Mitarbeiter*innen

< T550 >


Paket-Dateien

Nicht zugeordnet Berichte Analysen

<input type="checkbox"/>	#	Format ⁱ	Dokument-Typ	gehört zu Dokument	Zuletzt geändert	Dateigröße	Software	<input type="button" value="🗑"/>
	1	docx	Stoffprobenbericht <input type="text"/>	- <input type="text"/>	23.06.2023-16:43:04	120KB	Word 2013	
		Verfasser	Dateiname					
		Lora Ipsum <input type="text"/>	T550__Spanischer Mantel__R_4__SP_1.docx					

<input type="checkbox"/>	#	Format ⁱ	Dokument-Typ	gehört zu Dokument	Zuletzt geändert	Dateigröße	Höhe	Breite	<input type="button" value="🗑"/>
	2	png	Stoffprobe <input type="text"/>	Stoffprobenbericht 1 <input type="text"/>	23.06.2023-16:43:04	1,2MB	1200px	1800px	
		Lizenz	Verfasser	Dateiname					
		CC BY-NC-ND <input type="text"/>	Lora Ipsum <input type="text"/>	550_Spanischer_Mantel_R_4__SP_1__Abb_1.png					

Hier klicken oder Datei(en) hinziehen um sie hochzuladen. [Weiter](#)

dLZA Home Hinzufügen Archiv Einstellungen Suche 

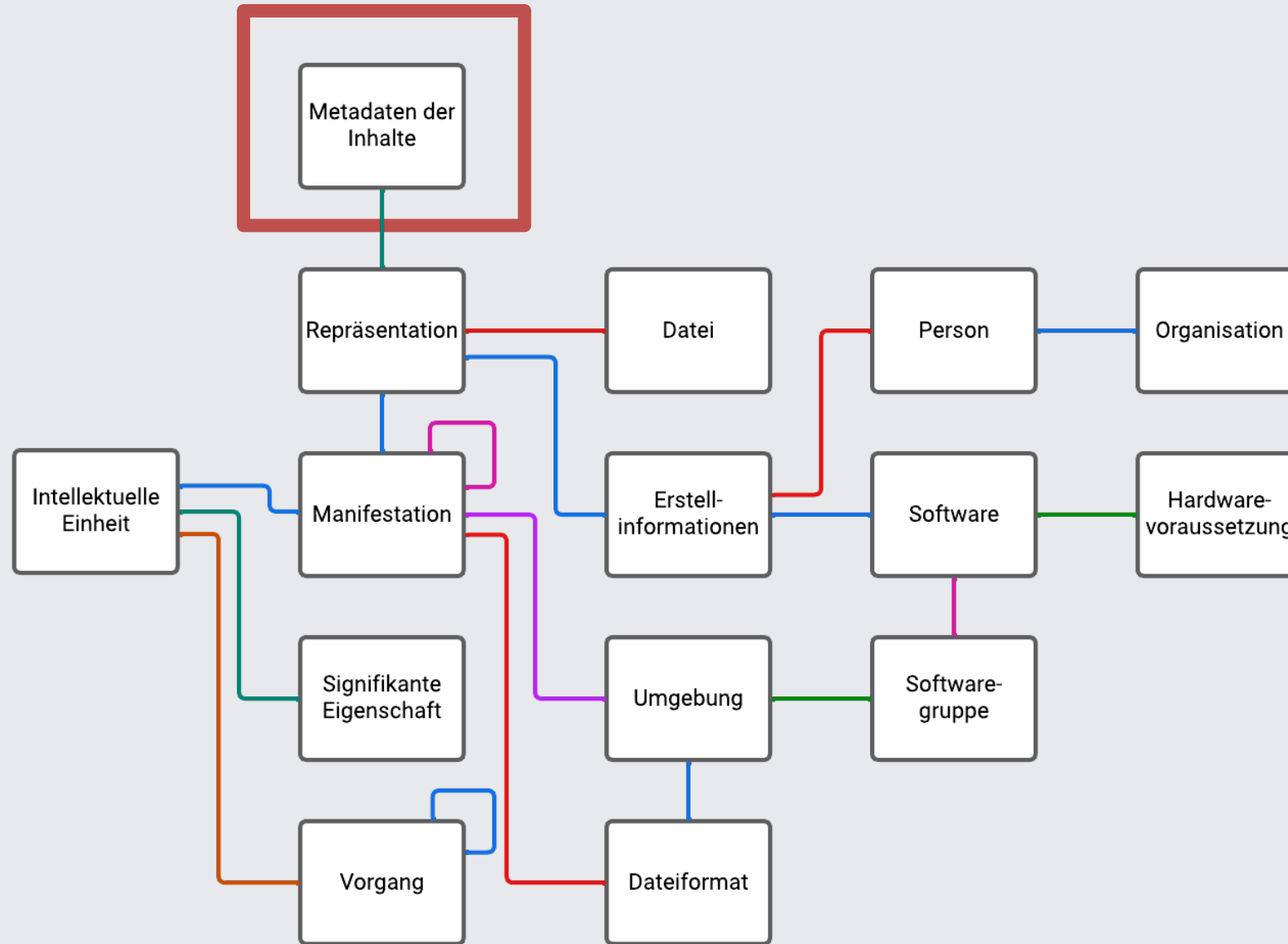
Vorschau der Einreichung Vereinfachte Ansicht

```
550e8400-e29b-11d4-a716-446655440000
├── bag-info.txt
├── bagit.txt
├── data
│   ├── files
│   │   ├── Analysen
│   │   │   ├── T550__Spanischer_Mantel_R_4__SP_1__Abb_1.png
│   │   │   ├── T550__Spanischer_Mantel_R_4__SP_1__Abb_1.tif
│   │   │   ├── T550__Spanischer_Mantel_R_4__SP_1.docx
│   │   │   └── T550__Spanischer_Mantel_R_4__SP_1.pdf
│   │   ├── Berichte
│   │   │   ├── T550__Spanischer_Mantel_R_4__RB_1__Abb_1.jpg
│   │   │   ├── T550__Spanischer_Mantel_R_4__RB_1__Abb_1.tif
│   │   │   ├── T550__Spanischer_Mantel_R_4__RB_1__Abb_2.jpg
│   │   │   ├── T550__Spanischer_Mantel_R_4__RB_1__Abb_2.tif
│   │   │   ├── T550__Spanischer_Mantel_R_4__RB_1__Abb_3.jpg
│   │   │   ├── T550__Spanischer_Mantel_R_4__RB_1__Abb_3.tif
│   │   │   ├── T550__Spanischer_Mantel_R_4__RB_1.docx
│   │   │   └── T550__Spanischer_Mantel_R_4__RB_1.pdf
│   └── meta
│       ├── cidoc-crm.rdf
│       ├── dublin-core.xml
│       ├── lido.xml
│       ├── mets-mods.xml
│       ├── oai-pmh.xml
│       ├── premis.xml
│       ├── tei.xml
│       ├── xdomea.xml
│       └── xmp.xml
├── manifest-sha256.txt
├── manifest-sha512.txt
├── tagmanifest-sha256.txt
└── tagmanifest-sha512.txt
```

```
Bag-Software-Agent: bagit.py v0.0.dev0 <https://github.com/LibraryOfCongress/ba
Source-Organization: GNM (IKK)
Bagging-Date: 2023-06-28
Contact-Name: Lora Ipsum
Contact-Email: l.ipsum@gnm.de
Payload-Oxum: 2219660.21
Bag-Size: 2268 KB
```

< ZurückWeiter >

Inhaltsbeschreibung



Inhaltsbeschreibung

CIDOC-CRM

PREMIS

CRMdig

LRMoo

```
<owl:Ontology rdf:about="http://dlza.gnm.de/ontology/">
<owl:versionIRI rdf:resource="http://dlza.gnm.de/ontology/1.0"/>
<rdfs:comment xml:lang="de">Diese Ontology erweitert das CIDOC CRM 7.1.2. (
https://www.cidoc-crm.org/cidoc-crm/7.1.3) und das kompatible Modell des CRMdig 4.0 (
https://cidoc-crm.org/extensions/crmdig/4.0) um Konzepte der digitalen
Langzeitarchivierung am Germanischen Nationalmuseum. Dabei werden Konzepte und
Beziehungen aus der PREMIS Ontologie Version 3 (http://www.loc.gov/premis/rdf/v3/) und
der Formatspezifikation des BagIt-Formats Version 1.0 (https://doi.org/10.17487/RFC8493)
implementiert. Autor: Robert Nasarek (r.nasarek@gnm.de). Repository: https://github.com/
rnsrk/dlza-ontology.</rdfs:comment>
<rdfs:comment xml:lang="en">Ontology for longterm-preservation at the Germanic National
Museum Nuremberg.</rdfs:comment>
<rdfs:comment xml:lang="en">This ontology extends the CIDOC CRM 7.1.2. (
https://www.cidoc-crm.org/cidoc-crm/7.1.2) and the compatible model of CRMdig 4.0 (
https://cidoc-crm.org/extensions/crmdig/4.0) by concepts of digital long-term archiving
at the Germanisches Nationalmuseum. Concepts and relationships from the PREMIS ontology
version 3 (http://www.loc.gov/premis/rdf/v3/) and the format specification of the BagIt
format version 1.0 (https://doi.org/10.17487/RFC8493) are implemented. Author: Robert
Nasarek (r.nasarek@gnm.de). Repository: https://github.com/rnsrk/dlza-ontology.</rdfs:
comment>
<rdfs:label xml:lang="de">Ontology für die digitale Langzeitarchivierung am Germanischen
Nationalmuseum Nürnberg</rdfs:label>
<owl:versionInfo rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">1.0</owl:
versionInfo>
</owl:Ontology>
```

Digitale Langzeitarchivierung

Mikro-Services (DROID, TrID, VeraPDF...)

CIDOC-CRM + PREMIS

Drupal-Modul auf Basis BagItPHP

Digitales Langzeitarchiv

Identifikation der Dateiformate

Anreicherung mit Metadaten

Erstellen von Archivpaketen

Validierung der Dateien und Formate

Einreichen von Daten

Migration in LZA-Formate

Open Archival Information System

Ausliefern von Archivpaketen

Mikro-Services (ImageMagick, PDF24, ffmpeg)

Drupal/ WissKI + React Extension

Mikro-Services (VeraPDF, checksum, autocopy...)