

Die API des Deutschen Zeitungsportals / der Deutschen Digitalen Bibliothek

„6. Datenbezieher-Workshop“
Mittwoch, 11. März 2026

Michael Büchner
m.buechner@dnb.de



Historische Zeitungen
sind Lesestoff!



Geschichte aus erster Hand

Entdecken Sie historische Zeitungen aus den Jahren 1609 bis 1994



Zeitung über Titel
auswählen



Zeitung über Ort
auswählen



Zeitung über Jahr
auswählen



Das Zeitungsportal enthält **3.144 Zeitungen**,
bestehend aus **5.482.195 Ausgaben** mit
34.257.532 Seiten von **24 Datenpartnern.**

1609

1700

1800

1900

1994

Neuzugänge im Zeitungsportal

→ Allgemeines Noth- und Hülf-Blatt

→ Augsburgener allgemeine Zeitung, Wochenausgabe der Augsburgener allgemeinen Zeitung

General-Anzeiger. 1889-1945

Datengeber:

Universitäts- und Landesbibliothek der Rheinischen Friedrich-Wilhelms-Universität Bonn →

Suchergebnisse

Seite 12

Lockeneisen.15.20.30 usw.
 Welleneisen.30.40.75 usw.
Onduliereisen für den Bubenkopf .00.50 usw. Brennlampen zum Gebrauch von [...] aRrer Prriere... Alle Haararbeiten wie: Zöpfe, Unterlagen, Haar-Netze, Kopfwasser, Parfüms, Seifen, Bürstenwaren, Badehauben, Brenn- und **Onduliereisen**, Kämme, Haarschmuck, Hand- und Taschenspiegel, Kölnischwasser gegenüber und 4711 sind im Preise bedeutend herabgesetzt. Einige Beispiele: Stirnlocken aus echten Haaren, Paar.50 u.00

Verkaufe einzelne Artikel
unter Friedenspreis!

Alle Haararbeiten wie: Zöpfe, Unterlagen, Haar-Netze, Kopfwasser, Parfüms, Seifen, Bürstenwaren, Badehauben, Brenn- und **Onduliereisen**, Kämme, Haarschmuck, Hand- und Taschenspiegel, Kölnischwasser gegenüber und 4711 sind im Preise bedeutend herabgesetzt. Einige Beispiele:

Neuheit! Lockenkrause aus Syzig- und Kambodjischen...
Neuheit! Scheitel, Transformationen, Perücken sowie alle besseren Haararbeiten werden nach Maß und Probe angefertigt.

Haarnezze beste Qualität, aus echten Haaren.
 Brautkronen zur schone Wahl.
 Stütze-Brosche...
 Stütze-Flügel...
 Stütze-Wagen...
 Stütze-Flügel...
 Stütze-Wagen...
 Stütze-Flügel...
 Stütze-Wagen...

Kopfwasser besonders geeignet für Damen, Herren, Kinder...
 Birkwasser, Brennnesselbrühe, Eukalyptuswasser, Bayrum, Eau de Cologne...
 Fäulnis Tee u. Kamilleweisse, Löwe 2.00 Mk.

Keine grauen Haare mehr
 Keine grauen Haare mehr...
 Keine grauen Haare mehr...
 Keine grauen Haare mehr...

Damen- und Herren-Frisier- und Haarschneide-Salons.
 Spezialität: **Bubenkopfschneiden**
 in jeder geschmackvollen Form.
 Diese Arbeit ist bei sorgfältiger Behandlung zur Verfügung stellen, gegenüber...
 Können, am langen Haare, Friseur im Abwaschen sich besser bei: Bonn.

Haarhaus Schmitz



Historische Zeitungen sind
Forschungsquelle und
Massendaten

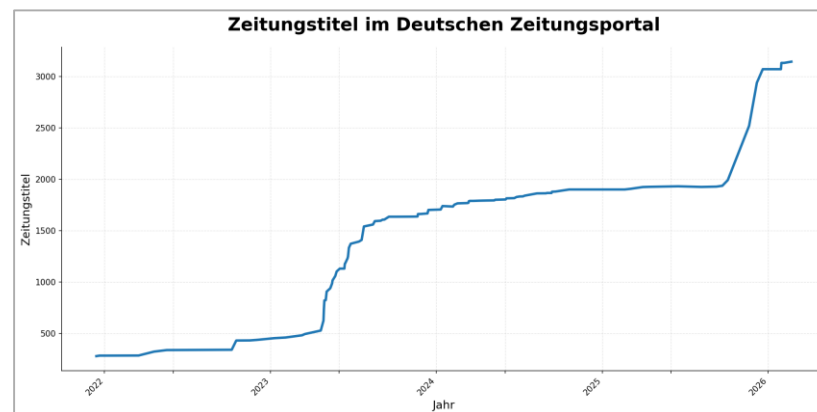


Warum eine API für historische Zeitungen?

- Das **Deutsche Zeitungsportal**

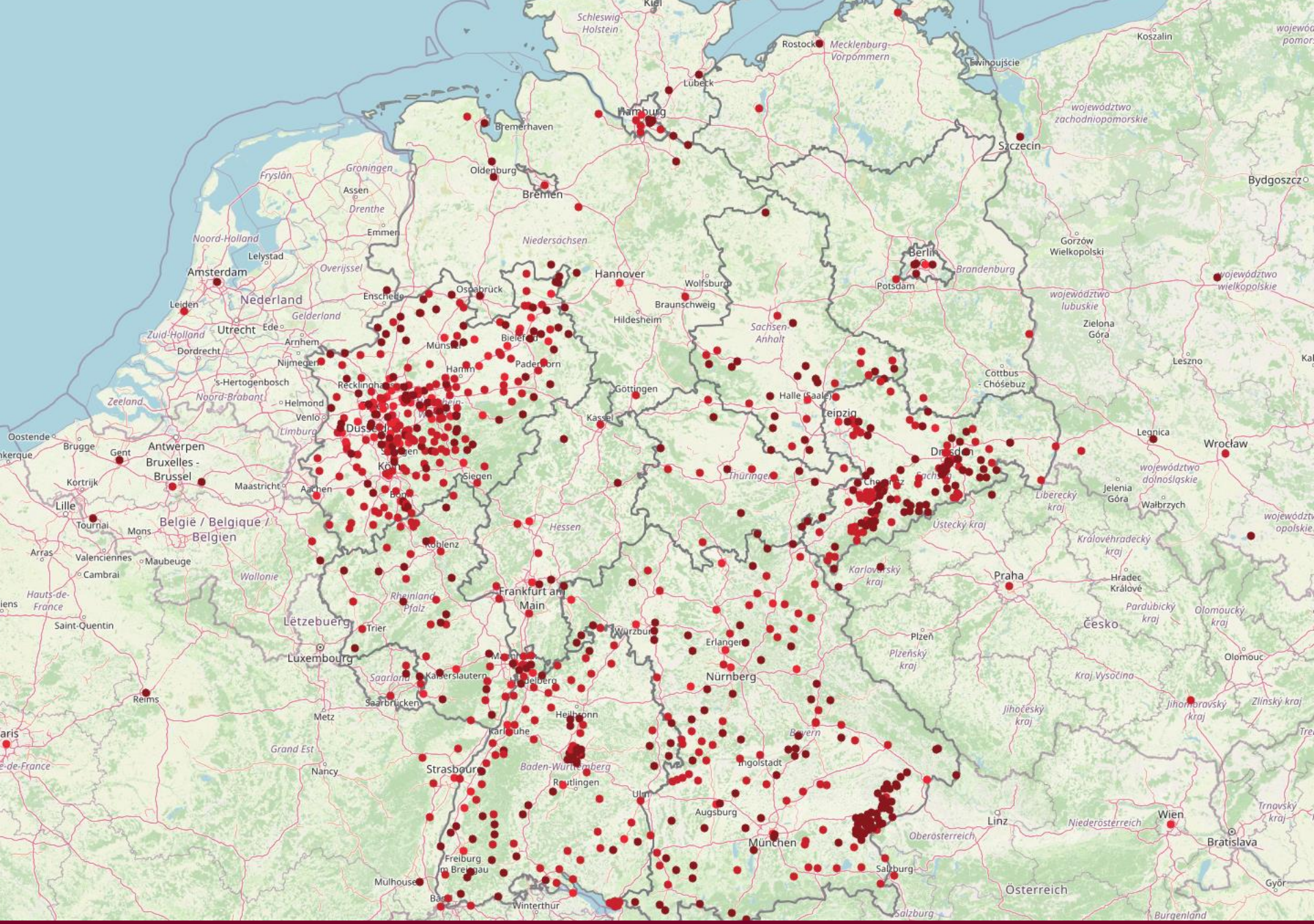
bündelt derzeit:

- 3.144 Zeitungen
- 5.482.195 Ausgaben
- 34.257.532 Seiten
- aus 24 Datenpartnern
- [Die Zeitschriftendatenbank (ZDB) weist 74.031 Zeitungen nach]



- Eine API macht diesen Bestand **systematisch, automatisiert und wiederholbar nutzbar**

- ...und das durch **jedermann**
- ...und jeder kann Anwendungen darauf aufbauen...

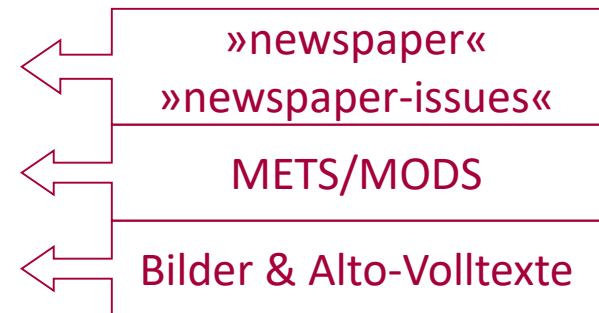


Verbreitungsorte der Zeitungen im Deutschen Zeitungsportal

Die API im Überblick



- Öffentliche **REST-API** der Deutschen Digitalen Bibliothek
 - **OpenAPI**-dokumentiert
<https://api.deutsche-digitale-bibliothek.de/2/q/openapi>
- Zugriff auf:
 - **Suche** in Apache-Solr-Indizes
 - **Metadaten**
 - **Digitalisate**
- Für Zeitungskontexte relevant:
 - **Zeitungstitel (title)** → »newspaper«
 - **Ausgaben (issue) / Seiten (page)** → »newspaper-issue«



Suchen über die API



- DDB-API erlaubt vollen Zugriff auf die **Solr-Suchindizes**
 - **Solr-Schemata** der Indizes
<https://dev.fiz-karlsruhe.de/stash/projects/DDB/repos/ddb-backend-solr/browse/conf>
 - typische **Solr-Parameter**:
 - q = Suchanfrage
 - fq = Filter (Facetten)
 - fl = zurückgelieferte Felder
 - rows / start = Paging
 - sort = Sortierung
- *Achtung: dezentrale Datenhaltung*, d.h. digitale Objekte (z.B. Zeitungsseiten) liegen i. d. R. auf den Servern unserer Datenpartner

<https://api.deutsche-digitale-bibliothek.de>

GET <https://api.deutsche-digitale-bibliothek.de/2/search/index/newspaper/select>

Senden

Speichern

Parameter 7 Anfragekörper Header Autorisierung Pre-Request-Skripte Post-request Script

Variablen

Abfrageparameter

🔍 🗑️ ✎️ +

Parameter	Wert	Beschreibung		
q	*.*	Beschreibung	🟢	🗑️
fq	hasLoadedIssues:true	Beschreibung	🟢	🗑️
fq	title:*Frankfurt*	Beschreibung	🟢	🗑️
fl	id,title,location	Beschreibung	🟢	🗑️
rows	100	Beschreibung	🟢	🗑️
wt	csv	Beschreibung	🟢	🗑️
sort	sort.title ASC	Beschreibung	🟢	🗑️
Schlüssel	Wert	Beschreibung	🟢	🗑️

Status: 200 • OK Zeit: 391 ms Größe: 1.16 KB

Rohdaten Header 2 Testergebnisse

Antworttext

☰ ⬇️ 📄 📄 ...

```
1 id,title,location
2 2796284-2, Frankfurter Journal, Frankfurt am Main
3 2791135-4, "Frankfurter Journal, Frankfurter Journal / Handels-Beilage zum Frankfurter Journal", Frankfurt am Main
4 3192457-8, "Frankfurter Journal, Frankfurter Journal / Verloosungs-Beilage zum Frankfurter Journal", Frankfurt am Main
5 2759882-2, Frankfurter Kayserl. Reichs-Ober-Post-Amts-Zeitung, Frankfurt am Main
6 2795022-0, Frankfurter Konversationsblatt, Frankfurt am Main
7 2751567-9, Frankfurter Ober-Post-Amts-Zeitung, Frankfurt am Main
8 2796279-9, Frankfurter Postzeitung, Frankfurt am Main
9 2938421-7, "Frankfurter Staats-Ristretto oder Kurzgefaßte Erzählung der neuesten und merkwürdigsten Nachrichten und Begebenheiten der Europaischen Staaten\, wie auch der
Wissenschaften\, der Kuenste und nuetzlichen Erfindungen", Frankfurt am Main
10 2959566-6, Frankfurter Volksbote, Frankfurt am Main
11 2256037-3, "Monatsblätter des jüdischen Kulturbundes Bezirk Rhein-Main\, Frankfurt am Main",
12 2796233-7, Neue Frankfurter Zeitung, Frankfurt am Main
13 2748922-X, Neues Frankfurter Museum,
14 2793403-2, Zeitung des Großherzogthums Frankfurt, Frankfurt am Main
15
```

<https://hopp.sh/r/nFhi6WHSSog5>

GET https://api.deutsche-digitale-bibliothek.de/2/search/index/newspaper-issues/select

Senden

Speichern

Parameter 5 Anfragekörper Header Autorisierung Pre-Request-Skripte Post-request Script

Variablen

Abfrageparameter			🔍	🗑️	✍️	+
q	*.*	Beschreibung	✓	🗑️		
fq	type:page	Beschreibung	✓	🗑️		
fq	plainpagefulltext:Ondolierreisen	Beschreibung	✓	🗑️		
fl	id,paper_title,publication_date,pagenummer	Beschreibung	✓	🗑️		
rows	100	Beschreibung	✓	🗑️		
Schlüssel	Wert	Beschreibung	✓	🗑️		

Status: 200 • OK Zeit: 282 ms Größe: 4.17 KB

JSON Rohdaten Header 2 Testergebnisse

Antworttext

```
1 {
2   "responseHeader": {
3     "zkConnected": true,
4     "status": 0,
5     "QTime": 111,
6     "params": {
7       "q": "*.*",
8       "fl": "id,paper_title,publication_date,pagenummer",
9       "fq": [
10        "type:page",
11        "plainpagefulltext:Ondolierreisen"
12      ],
13       "rows": "100"
14     }
15  },
16  "response": {
17    "numFound": 14,
18    "start": 0,
19    "maxScore": 1.0,
20    "numFoundExact": true,
21    "docs": [
22      {
```

<https://hopp.sh/r/Nm1rjG0oJ9g7>

Und die Metadaten?

```
<mets:fileSec>
  <mets:fileGrp
    USE="PRESENTATION">
    <mets:file ID="FILE_0001_PRESENTATION" MIMETYPE="image/tiff">
      <mets:FLocat LOCTYPE="URL" xmlns:xlink="http://www.w3.org/1999/xlink"
        xlink:href="https://img.sub.uni-hamburg.de/kitodo/PPN1012330699_18080905/00000001.tif" />
    </mets:file>
    <mets:file ID="FILE_0002_PRESENTATION" MIMETYPE="image/tiff">
      <mets:FLocat
        LOCTYPE="URL" xmlns:xlink="http://www.w3.org/1999/xlink"
        xlink:href="https://img.sub.uni-hamburg.de/kitodo/PPN1012330699_18080905/00000002.tif" />
      </mets:file>
      <!-- ... -->
    </mets:fileGrp>
    <mets:fileGrp USE="FULLTEXT">
      <mets:file ID="FILE_0001_FULLTEXT" MIMETYPE="text/xml">
        <mets:FLocat LOCTYPE="URL" xmlns:xlink="http://www.w3.org/1999/xlink"
          xlink:href="https://img.sub.uni-hamburg.de/kitodo/PPN1012330699_18080905/00000001.xml" />
      </mets:file>
      <mets:file ID="FILE_0002_FULLTEXT" MIMETYPE="text/xml">
        <mets:FLocat LOCTYPE="URL" xmlns:xlink="http://www.w3.org/1999/xlink"
          xlink:href="https://img.sub.uni-hamburg.de/kitodo/PPN1012330699_18080905/00000002.xml" />
      </mets:file>
      <!-- ... -->
    </mets:fileGrp>
    <mets:fileGrp USE="THUMBS">
    <mets:fileGrp USE="DOWNLOAD">
    <mets:fileGrp USE="IIIF">
    <mets:fileGrp USE="ZOOM">
    <mets:fileGrp USE="DEFAULT">
    <mets:fileGrp USE="DDB_FULLTEXT">
  </mets:fileSec>
```

Vielen Dank!

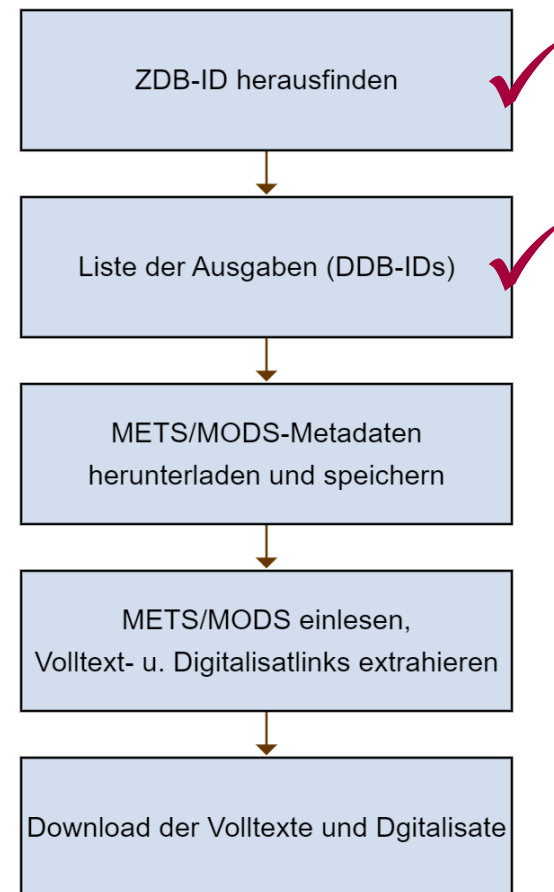
Haben Sie Fragen?



Download eines Zeitungstitels: Ablauf

Wie lädt man denn nun einen gesamten Zeitungstitel herunter?

1. **ZDB-ID** des gewünschten Zeitungstitels sollte bekannt sein
 - Suchindex „newspaper“ → „2149754-0“ für „La otra Alemania“
2. Erstellung einer **Liste der Ausgaben** (mit DDB-IDs)
 - Suchindex „newspaper-issues“
→ Suche nach „zdb_id:2149754-0 AND type:issue“
3. Download der **METS/MODS-Metadaten** aller Ausgaben
4. Extraktion der Links zu **Volltexten und Digitalisaten**
5. **Download** und Abspeichern der Volltexte und Digitalisate



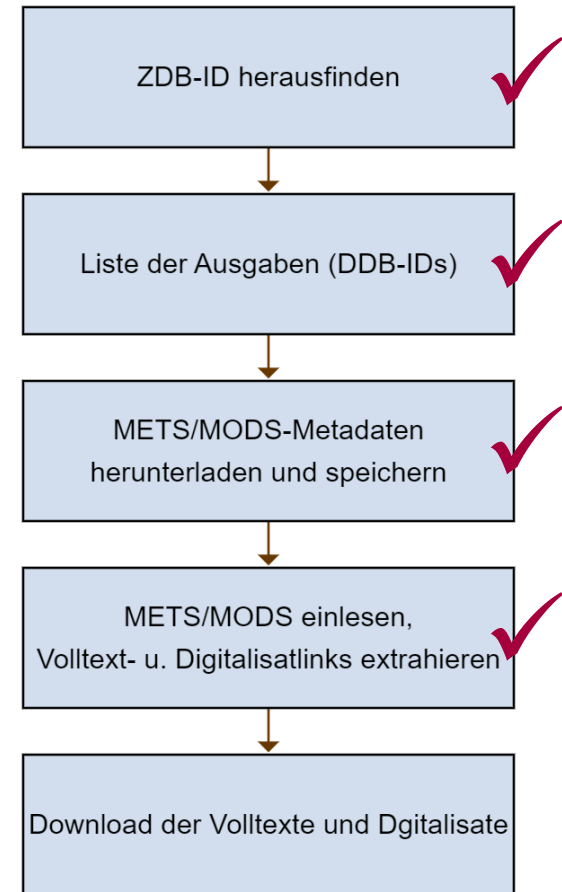
Extraktion der Links

- **METS/MODS**

- **METS:** Ein XML-Standard zur Strukturierung und Übermittlung digitaler Objekte in Archiven und Bibliotheken
- **MODS:** Ein XML-Format zur Beschreibung bibliografischer Daten, oft als Ergänzung zu METS genutzt

- **Extraktion** von Daten aus XML

- **XPath:** Eine Abfragesprache, die zum Navigieren und Auswählen von Elementen in XML-Dokumenten verwendet wird
- **Pfad-Ausdrücke:** XPath ermöglicht es, Teile eines XML-Dokuments mithilfe von Pfaden präzise zu lokalisieren und zu extrahieren
- siehe <https://en.wikipedia.org/wiki/XPath>



```
<?xml version="1.0" encoding="UTF-8" ?>
<?xml-stylesheet href="http://www.loc.gov/METS/" type="text/xsl" ?>
<?xml-stylesheet href="http://www.loc.gov/mods/v3" type="text/xsl" ?>
<!-- Header -->
```

```
<!-- Agent Information -->
<!-- Descriptive Metadata -->
```

```
<!-- File Section -->
```

```
<!-- MODS Metadata -->
```



```
//mets:mets/mets:fileSec/mets:fileGrp[@USE="DEFAULT"]/mets:file/mets:FLocat/@xlink:href
```

```
<!-- MODS Metadata -->
```

```
</mets:mets>
```



Sinkkasten Vergangenen Samstag Abends mit Wasserverschluss für Höfe, Brauereien, Waschküchen etc. dem Unterzeichneten 2 Schafe zu Grosses Lager in Frankfurt a. M. und Köln. Ä Ä Ä Wachene JRoten abgeho werden Frankfurter Baubanke - ger - (7683a) Abtheilung Canalbau & Wasserleitung. # Müller, * - - Hoher Feiertage wegen komme ich Mitt-- F. ### woch den 8. April nicht nach Freifing. Da- | |E ZI gegen bin ich B - EFSF J. - - 2 -ES ## # Mittwoch den 15. April - # # # - - 5 S- F > und von diefer Zeit an an jedem Markttag S° S 55 wieder regelmäßig hier. F FE - 2 sR S ## E. Landauer - =5 sank- & Wechselgefchäft. F##### (742 3b) f - Fé - # # ### # FFFFÄ# I "S ÄFFFSF F S I- 9 = EZ T | -- -, s° FS Perkaufsangebot. E ##### - Die den Erben des verlebten Herrn Stadtpfarrers Plank gehörigen S9 de° ### Häufer und Grundftücke, nämlich das ehemalige Hafnerhaus in der Lucken- K ### gaffe daher und ein kleines Haus mit Garten in Marzling, dann mehrere - A-, S = = » Torf- und Mooswiefen werden hiemit zum Verkaufe aus freier Hand x- Äg angeboten; eben # eine kleine Kirchenorgel zu verkaufen. A- = == Die Grben. S ### FTTWRITSTÄT - - - - „ . . B. - tritte # Mittelpreis Ä - *** GefliegTGefäll | xw°..... * °fCl. - TETTET TETF (IE M kulatur-Papie TTTTTTTTTTTTTTTTTTTTTTTF T 4 -- per Pfund 6kr Korn - 1010 10-24- 21 402112-19— Franz Paul Datterer Gerfte 8 84 92 6 19 ### TÄT = -- T° Haber 2 41 43 4 - 14 41 D 1 - T- reife d. k. Wal ü Sc Freifinger Schranne v°n STÄFTS7IT Ä J. Dav. Ä nFre A. Verkauf nach P - = = j Vor Zu-Schr-Ver- ñhr Mt.-Niedr Telpreis. . r Getreide- Reft fuhr. ftand kauft j Kr° Aus Weizen. T Gattung - Ä- -- ...15 _ _2Hect2Hect2Hect2Hect. - T.Tr. I. Tr-T- Dr. T.1 IT- . 0 ... Waizen . . 7259 266 266 + " + | | Korn. 348 5149 L) # Gerfte 239 41 39 2 | 19221816446 j" Haber -236 236 236 - 12201137957 mehl „44. . Reps - ----| - - - - - - Ä - - "Ä. - 10 10 10 - -1458— ele - - - ttungen - - - - - gen. B. Verkauf Ä - - - - - g Etr. - Etr. | Ctr. „Etr. | Etr. f.kr. f.kr. f.fr. Nr. 1. Waizen ##### Korn . - - - - - - - - - - h 3 Wicken2c.. - - - - - = - - - - - 5 *