



Großes Bild – Lewis R: <https://www.pexels.com/de-de/foto/kunst-vereinigte-staaten-von-amerika-usa-strassenfotografie-12341819>

Kleines Bild – Sami Abdullah: <https://www.pexels.com/de-de/foto/wasser-pool-urlaubsort-ferienort-7313084>



# Anreicherung von Musiktonträgerkatalogisaten durch nutzergenerierte Inhalte

Projekt MIAMI

(Metadaten IntegrAtion MusIk)

Deutsche Nationalbibliothek

<https://www.dnb.de>

1. Projektziel

2. Vorgehen

3. Trackdatensätze

4. Fazit / Ausblick

- Dr. Angela Vorndran

Team Datenmanagement, IT

- Simon Zetzsche

Automatische Erschließungs-  
verfahren, Netzpublikationen

# 1. Projektziel

- Projektantrag: 2015, Aktualisation: 2018
- regelmäßige retrospektive Anreicherung von Musik-Metadaten, insbesondere Trackdaten
- Metadaten freier Musikdatenbanken (MusicBrainz, ...)
- jeder Track inkl. zugehöriger Metadaten wird in einen eigenen Datensatz gespielt
- Ziel v.a. verbesserte Recherche, Benutzung im Audioplayer, Datenaustausch mit Datensätzen digitaler Audiofiles

## 1.1 Musiktonträger im DMA (DNB) (Stand: Mai 2024)

- 1,38 Mio. Musiktonträger
- Trackinformationen bis 2016 - an verschiedenen Stellen zu finden:
  - mit Tonträgern verknüpfte untergeordnete Trackdatensätze
    - ca. 180.000 Fälle, die mit 56.000 Tonträger-Sätzen verknüpft sind
  - Eintragen der Tracks in ein Feld im Datensatz des Tonträgers
    - 311.850 Sätze haben eingetragene Tracktitel

## 1.2 MusicBrainz (Stand: Mai 2024)

- Offenes (CC0-Lizenz), nutzergeneriertes Webangebot, das Metadaten für Tonträger sammelt und zur Verfügung stellt
- ca. 2 Mio. Künstler\*innen / Musikgruppen
- 27 Mio. Recordings
- 37 Mio. Tracks
- 1,2 Mio. Barcodes (GTIN/EAN/UPC)
- eigene Identifier (MBID) für Künstler\*innen, Aufnahmen und Tracks

## 2. Vorgehen

- Tracktitel aus MusicBrainz sollen für DNB-Titel übernommen werden
- Abgleich von DNB-Tonträgern mit MusicBrainz
  - Erkennen gleicher Tonträger via GTIN/EAN/UPC
  - Bedingung: Information erfasst
- Tracktitel übernehmen
  - Relevante Elemente in MusicBrainz identifizieren
  - Neues einheitliches Trackdatensatzformat definieren
- Verlinkung zum Titeldatensatz (Zusammenstellung)

## 2.1 API-Abfrage über Identnummer

- Anhand der Global Trade Item Number (GTIN) / European Article Number (EAN) o. dem Universal Product Code (UPC) werden Musik-Tonträger aus dem DNB-Bestand in MusicBrainz gesucht
- DMA-Bestand:
  - ca. 320.700 GTINs / EANs / UPCs in 1,38 Mio. G\*-Sätzen
  - davon 213.500 ohne Trackinformationen
- 99.000 Barcodes in MusicBrainz gefunden u. abgerufen

## 2.1 API-Abfrage über Identnummer (Abgleich der Identnummer)

 **Greatest Hits I, II & III: The Platinum Collection** (2011 remaster)  
~ Veröffentlichung von Queen (alle Versionen dieser Veröffentlichung ansehen, 16 vorhanden)

**Übersicht** Disc IDs (3) Cover-Art (119) Aliasse Tags Details Bearbeiten Beziehungen bearbeiten

### Anmerkung

discs come in individual Super Jewel Boxes:  
CD1: 602527583648 / 2758364  
CD2: 602527583655 / 2758365  
CD3: 602527735368 / 277 353 6

Anmerkung zuletzt geändert am 2018-11-09.


### Titelliste

#### ▼ CD 1: Greatest Hits I

Nr.	Titel
1	<b>Bohemian Rhapsody</b> zusätzliches Sound-Engineering:  leitende(r) Techniker(in):  Mike S zusätzliche(r) Techniker(in):  Kris leitende(r) Produzent(in):  Brian I Produzent(in):  Roy Thomas Bake Bassgitarre:  John Deacon (von 1 Gitarre:  Brian May (Queen guitar gong, Pauke und Schlagzeug:  Rc Klavier und Sologesang und sonstig sonstiger Gesang [operatic vocals]:

	
<b>Link zu diesem Datensatz</b>	<a href="https://d-nb.info/1027716822">https://d-nb.info/1027716822</a>
<b>Titel</b>	Greatest hits : the platinum collection
<b>Organisation(en)</b>	Queen (Ausführender)
<b>Verlag</b>	Berlin : Universal Music Entertainment Berlin : Universal Music Entertainment
<b>Zeitliche Einordnung</b>	Erscheinungsdatum: P 2011
<b>Umfang/Format</b>	3 CDs in Schuber ; 12 cm + Beih.
<b>Persistent Identifier</b>	URN: urn:nbn:de:101:1-2016033113770
<b>Bestellnummer(n)</b>	Universal (LC 01846) 277 241 7
<b>UPC</b>	602527724171
<b>Beziehungen</b>	Greatest hits : 1 - 3
<b>Sachgruppe(n)</b>	780 Musik



 Spiele auf ListenBrainz

**Veröffentlichungsinformationen** [Anzeigen](#)

**Strichcode:** 602527724171

**Format:** 3×CD

**Länge:** 3:28:10

#### Zusätzliche Details

**Typ:** Album + Kompilation

**Verpackung:** Super Jewelcase

**Status:** Offiziell

**Sprache:** Englisch

**Schrift:** Lateinisch

**Datenqualität:** Normal



## 2.1 API-Abfrage (Übernahmeoptionen)

- Antwort der API-Abfrage als XML-Dokument
- Relevante Angaben, z.B.:
  - Titel
  - Position des Tracks
  - MusicBrainz-ID
  - Dauer in Millisekunden
  - Künstler\*in, Gruppe (Name, MusicBrainz-ID)
  - Erstveröffentlichungsdatum

```

<metadata>
  <release id="c07df500-5eef-47cb-a449-807e8fa64473">
    <title>From Paris With Love</title>
    <status id="4e304316-386d-3409-af2e-78857eec5cfe">Official</status>
    <quality>normal</quality>
  </text-representation>
  <language>eng</language>
  <script>Latin</script>
  </text-representation>
  <artist-credit>
    <name-credit>
      <artist id="0aad8921-dfaf-42b0-bb1a-1a2d6cf97b9b" type="Person" type-id="b6e035f4-3ce9-331c-97df-83397230b0df">
        <name>David Buckley</name>
        <sort-name>Buckley, David</sort-name>
        <disambiguation>English score composer</disambiguation>
        <alias-list count="7">
          <alias sort-name="Buckley" type="Search hint" type-id="1937e404-b981-3cb7-8151-4c86ebfc8d8e">Buckley</alias>
          <alias sort-name="Dave Buckley" type="Search hint" type-id="1937e404-b981-3cb7-8151-4c86ebfc8d8e">Dave Buckley</alias>
          <alias sort-name="David Buckeley" type="Search hint" type-id="1937e404-b981-3cb7-8151-4c86ebfc8d8e">David Buckeley</alias>
          <alias sort-name="Buckley, David" type="Legal name" type-id="d4dcd0c0-b341-3612-a332-c0ce797b25cf">David Buckley</alias>
          <alias sort-name="David Buckley" type="Search hint" type-id="1937e404-b981-3cb7-8151-4c86ebfc8d8e">David Buckley</alias>
          <alias sort-name="David Buckley" type="Search hint" type-id="1937e404-b981-3cb7-8151-4c86ebfc8d8e">David Buckley</alias>
          <alias sort-name="David Burckley" type="Search hint" type-id="1937e404-b981-3cb7-8151-4c86ebfc8d8e">David Burckley</alias>
        </alias-list>
      </artist>
    </name-credit>
  </artist-credit>
  <date>2010-02</date>
  <release-event-list count="1">
    <release-event>
      <date>2010-02</date>
    </release-event>
  </release-event-list>
  <barcode>886976489120</barcode>
  <asin>B0034BW9DS</asin>
  <cover-art-archive>
    <artwork>true</artwork>
    <count>2</count>
    <front>true</front>
    <back>true</back>
  </cover-art-archive>
  <label-info-list count="1">
    <label-info>
      <label id="011d1192-6865-45bd-85c4-0400dd45693e" type="Imprint" type-id="b6285b2a-3514-3443-80df-fcfc528824ded">
        <name>Columbia</name>
        <sort-name>Columbia</sort-name>
        <disambiguation>
          Sony Music, worldwide except JP, formerly owned by CBS between 1938-1990 within US/CA/MX
        </disambiguation>
        <label-code>162</label-code>
        <alias-list count="5">
          <alias sort-name="Columbia">Columbia</alias>
          <alias sort-name="Columbia Records">Columbia Records</alias>
          <alias sort-name="Columbia">Columbia</alias>
          <alias sort-name="Columbia US">Columbia US</alias>
          <alias sort-name="columbia">columbia</alias>
        </alias-list>
      </label>
    </label-info>
  </label-info-list>
  <medium-list count="1">

```

## 2.2 Übernahmeoptionen

- Im 1. Gesamtlauf (2023) könnten für 50.444 Tonträger 947.363 Trackdatensätze erstellt werden
  - ca. 21.000 DNB-Datensätze davon ohne Trackinformationen
- Zur Qualitätssicherung der Ergebnisse eingesetzte Prüfungen:
  - Barcode, Label und Erscheinungsjahr (+/- 1) stimmen überein oder
  - Barcode, Label und Titel stimmen überein und es liegt kein Veröffentlichungsdatum vor
  - Digitale Medien und mehrfaches Vorkommen von Barcodes werden ausgeschlossen
  - Reduzierung der Filter? (Ergebnismenge vs. Fehleranfälligkeit)

## 3. Trackdatensätze

- Neue Struktur zur Erfassung von Trackdatensätzen in CBS konzipiert
- Für jeden Track wird ein eigener Trackdatensatz erstellt, der mit dem Datensatz des Tonträgers verknüpft wird

```

Eingabe: 1147:24-06-22 Änderung: 1147:28-06-22 08:21:18 Status: 1147:24-06-22
0500 Gifm
0501 aufgeführte Musik$bpmm
0502 audio$bs
0503 Audiodisk$bsd
0550 $cMusicBrainz
0551 X$bmm
0600 tm
1100 2011
1500 /1eng
1700 /1XA-DE
2054 GBCEE0100128
2054 USRH10652606
2240 DNB:1283310104
3000 Mercury, Freddie$BKomponist$4cmp$BTextdichter$4lyr$BSänger$4sng$BInstrumentalmusiker$4itr
3010 Deacon, John$BInstrumentalmusiker$4itr
3010 May, Brian$BInstrumentalmusiker$4itr
3010 Taylor, Roger$BInstrumentalmusiker$4itr
3110 Queen$BAusführender$4prf
4000 We Are the Champions / Queen
4030 Berlin : Universal Music Entertainment
4060 Track$d3 min 07 sec
4070 $h17$17
4201 John Deacon, Bassgitarre. Brian May (Queen guitarist), E-Gitarre und Gitarre. Freddie Mercury, Klavier und Sologesang.
4241 In:1027716822!-Gaxm--: Greatest hits
4243 Erscheint auch als $nOnline-Ausgabe1283313081Mercury, Freddie$BKomponist: We Are the Champions / Queen
4700 |AEN|*MusicBrainz-Test – Track als unselbstständiger Teil
4700 |AEN|*https://musicbrainz.org/recording/3604eb06-4bc2-4416-9b31-ceadae51bc70
5050 780
101011|einzig dnh |101a|

```

## 4. Fazit / Ausblick

- Mit den Daten wird weiter gearbeitet
  - Umgang mit Dubletten
  - Parallelabgleich via ISRC
- Erfahrungsaufbau für weitere Datenbanken (Discogs)
- Verzeichnung weiterer Identifier, Vernetzung
- evt. Übernahmen in übergeordnete Titeldatensätze
- Artist-Informationen für GND

# Vielen Dank für Ihre Aufmerksamkeit

Simon Zetzsche  
<mailto:s.zetzsche@dnb.de>



Angela Vorndran  
<mailto:a.vorndran@dnb.de>

