Even Cataloging Sounds Better on Vinyl

Repurposing Metadata from Discogs.com to Catalog Sound Recordings at Michigan State University
1200 LPs donated to MSU

collection centered on Roma people and Gypsy stereotypes in popular culture

many musical styles:

• flamenco
• jazz
• opera (soooo much *Carmen*)
• folk

so many records, so little time
what about Discogs.com?

- expert community
- contributor guidelines
- entry review
- powerful metadata
- 40% of collection already there
Help, please!
You're serious about your music... and we are, too. All the info you need for finding, buying, selling, and cataloging music is right here.

I need help with...

Guidelines

**Basic Guidelines**

- Have the exact release in front of you when entering it to the database
- Only enter or change information that you can cite a trustworthy source for (stick to provable facts!)
- Check all the artist and labels you enter link to the correct page
- Use the search to find artist name variations, and use the ANV function to link them up
- Capitalize the first letter of every word in names and titles
- Follow the Quick Start Guide for your first few submissions

Related Help Documents

- Copy To Draft
- Database FAQ
- Database Glossary
- Database Guidelines 1. General Rules
- Database Guidelines 2. Artist
- Database Guidelines 3. Title

---

a parallel universe
original description

• on-call staff member Cathy
• Discogs metadata expertise
• original and enhanced entries
data elements
evaluated/enhanced/recorded

• artists
• composers
• genres/styles
• edition information
• album cover image

a parallel universe
Gipsy Kings – Bamboleo

Label: Elektra/Asylum Records – 0-66715
Format: Vinyl, 12", Single, 33 ⅓ RPM
Country: US
Released: 1988
Genre: Electronic, Rock
Style: House, Latin

Tracklist

A1 Bamboleo (Latin Extended Version) 7:15
   Remix – Arthur Baker
A2 Bamboleo (Latin Single) 3:45
B1 Bamboleo (Club Mix) 6:58
   Producer [Additional Post Production], Edited By – Francis Milano, Karl Siegel
   Remix – Arthur Baker

Companies, etc.

Copyright (c) – Elektra/Asylum Records
Phonographic Copyright (p) – Elektra/Asylum Records

Credits

Design – Carol Boboits
Photography By [Photo] – Jean-Jacques Damour
Gipsy Kings – Bamboleo

Genre: Latin
Style: Flamenco
Year: 1987

Tracklist

Bamboleo 3:25
Quiero Saber 4:10

Versions (34)

<table>
<thead>
<tr>
<th>Title (Format)</th>
<th>Label</th>
<th>Cat#</th>
<th>Country</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bamboleo (12&quot;)</td>
<td>P.E.M.</td>
<td>GKM 002</td>
<td>France</td>
<td>1987</td>
</tr>
<tr>
<td>Bamboleo (12&quot;)</td>
<td>Trans-Canada Disc</td>
<td>TCD-1298</td>
<td>Canada</td>
<td>1967</td>
</tr>
<tr>
<td>Bamboleo (7&quot;)</td>
<td>P.E.M.</td>
<td>GK 001</td>
<td>France</td>
<td>1987</td>
</tr>
<tr>
<td>Bamboleo (7&quot;)</td>
<td>CGD</td>
<td>INT 10813</td>
<td>Italy</td>
<td>1987</td>
</tr>
<tr>
<td>Bamboleo (7&quot;, Promo)</td>
<td>Elektra</td>
<td>7-59295</td>
<td>US</td>
<td>1987</td>
</tr>
</tbody>
</table>
A look under the hood...

- Let’s go

a parallel universe
you say tomato...

• Discogs “release”
  Versus: RDA “manifestation”
• Factors that trigger new release:
  o Promotional copies
  o Different colored vinyl
  o Different matrix numbers, mastering stamps, etc.
final collection

http://www.discogs.com/user/MSULibraries
The Discogs API v2.0 is a RESTful interface to Discogs data. You can access JSON-formatted information about Database objects such as Artists, Releases, and Labels. Your application can also manage User Collections and Wantlists, create Marketplace Listings, and more.

Database data is made available under the CC0 No Rights Reserved license, which means there are no restrictions on what you can do with it.

Your application can use “Discogs” in its name as long as it follows the Application Name and Description Policy.

Monthly data dumps are also available.

https://www.discogs.com/developers/
Python used to:

- query Discogs API
- convert JSON to generic XML
- retain JSON for a separate project (more later)
XSLT used to:

• convert generic XML to MARCXML
• search for LCNAF identifiers
• insert additional MARC elements
  o mapped LC terms
  o location codes
  o gift notes

digging out the metadata
Carlos Montoya

Profile: Spanish guitar player (1903 - 1993).


Variations: Viewing At Carlos Montoya
Carlos Montoya Y Su Grupo Flamenco

External links [edit]

- Some photos of LP covers (Oviatt Library Digital Collections)

Authority control

WorldCat • VIAF: 39566363 • LCCN: n82052134

d241d75d-921f-4dfb-b2e0-d977f5f679d

Mapping to Authorized Forms
subject/genre access

Discogs genre/style: Hip Hop/Instrumental

LCGFT:  

Hip-hop (Music)
USE Rap (Music)

mapping to authorized forms
equivalency issues

Discogs: Blues/Country Blues

LC:
- 655 _7 Blues (Music) $2 lcgft
- 655 _7 Country music $2 lcgft

mapping to authorized forms
equivalency issues

Discogs: Hip Hop/Favela Funk

Funk carioca, favela funk and, elsewhere in the world, baile funk, is a type of dance music from Rio de Janeiro, derived from Miami bass and African style music.[1][2]

LC:
- 655 _7 Dance music $2 lcgft
- 655 _7 Funk (Music) $2 lcgft
- 650 _0 Rap (Music) $z Brazil
equivalency issues

Discogs: Folk, World, & Country/Gamelan

LC:
- 655 _7 Folk music $2 lcgft
- 650 _0 Folk music $z Indonesia
- 650 _0 Gamelan music
- 382 __ gamelan

mapping to authorized forms
noticeable gaps

Discogs → country music genre/subgenres

LCGFT → electronica, alternative rock, and punk subgenres

mapping to authorized forms
<table>
<thead>
<tr>
<th>Styles</th>
<th>Genres</th>
<th>MARC form of comp</th>
<th>LCGFT 1 (655 _7 $2 lcgraf)</th>
<th>LCGFT 2 (655 _7 $2 lcgraf)</th>
<th>LCSH 1 (650 _0)</th>
<th>LCSH 2 (650 _0)</th>
<th>LCMP (382 la)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock</td>
<td>Hardcore</td>
<td>rc</td>
<td>Hardcore (Music)</td>
<td></td>
<td></td>
<td></td>
<td>electronics</td>
</tr>
<tr>
<td>Rock</td>
<td>Heavy Metal</td>
<td>rc</td>
<td>Heavy metal (Music)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Folk, World, &amp; Country</td>
<td>Highlife</td>
<td>pp</td>
<td>Highlife (Music)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Folk, World, &amp; Country</td>
<td>Hindustani</td>
<td>fm</td>
<td>Hindustani music</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic</td>
<td>Hip Hop</td>
<td>pp</td>
<td>Electrornica (Music)</td>
<td>Underground dance music</td>
<td></td>
<td></td>
<td>electronics</td>
</tr>
<tr>
<td>Electronic</td>
<td>Hip-House</td>
<td>pp</td>
<td>House music</td>
<td>Rap (Music)</td>
<td></td>
<td></td>
<td>electronics</td>
</tr>
<tr>
<td>Hip Hop</td>
<td>Hiplife</td>
<td>pp</td>
<td>Rap (Music)</td>
<td>Popular music</td>
<td>Rap (Music) iz Ghana</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Folk, World, &amp; Country</td>
<td>Honky Tonk</td>
<td>pp</td>
<td>Honky-tonk music</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip Hop</td>
<td>Horrorcore</td>
<td>pp</td>
<td>Rap (Music)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic</td>
<td>House</td>
<td>pp</td>
<td>House music</td>
<td></td>
<td></td>
<td></td>
<td>electronics</td>
</tr>
<tr>
<td>Hip Hop</td>
<td>Hyphy</td>
<td>pp</td>
<td>Rap (Music)</td>
<td>Underground dance music</td>
<td>Rap (Music) iz San Francisco Bay Area (Calif,)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic</td>
<td>IDM</td>
<td>pp</td>
<td>Intelligent dance music</td>
<td></td>
<td></td>
<td></td>
<td>electronics</td>
</tr>
<tr>
<td>Electronic</td>
<td>Illbient</td>
<td>pp</td>
<td>Electrornica (Music)</td>
<td>Ambient music (Electronica)</td>
<td></td>
<td></td>
<td>electronics</td>
</tr>
<tr>
<td>Classical</td>
<td>Impressionist</td>
<td>mu</td>
<td>Art music</td>
<td>Impressionism (Music)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Folk, World, &amp; Country</td>
<td>Indian Classical</td>
<td>mu</td>
<td>Art music</td>
<td>Popular music</td>
<td>Music iz India</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pop</td>
<td>Indie Pop</td>
<td>pp</td>
<td>Alternative rock music</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rock</td>
<td>Indie Rock</td>
<td>rc</td>
<td>Alternative rock music</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic</td>
<td>Industrial</td>
<td>pp</td>
<td>Industrial music</td>
<td>Electrornica (Music)</td>
<td></td>
<td></td>
<td>electronics</td>
</tr>
<tr>
<td>Rock</td>
<td>Industrial</td>
<td>rc</td>
<td>Industrial music</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip Hop</td>
<td>Instrumental</td>
<td>pp</td>
<td>Rap (Music)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Music</td>
<td>Interview</td>
<td>nn</td>
<td>Sound recordings</td>
<td>Interviews</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic</td>
<td>Italo House</td>
<td>pp</td>
<td>House music</td>
<td>Disco (Music)</td>
<td>Underground dance music iz Italy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic</td>
<td>ItaloDance</td>
<td>pp</td>
<td>Electrornica (Music)</td>
<td>Underground dance music</td>
<td>Electronic (Music) iz Italy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic</td>
<td>Italo-Disco</td>
<td>pp</td>
<td>Electrornica (Music)</td>
<td>Disco (Music)</td>
<td>Electronic (Music) iz Italy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic</td>
<td>Jazzdance</td>
<td>pp</td>
<td>Electrornica (Music)</td>
<td>Jazz</td>
<td></td>
<td></td>
<td>electronics</td>
</tr>
<tr>
<td>Jazz</td>
<td>Jazz-Funk</td>
<td>iz</td>
<td>Jazz</td>
<td>Funk (Music)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
unique access points for works not possible

Discogs does not attribute principal artist or creator in same way as RDA

Can’t reliably create 1XX

RDA and access point issues
7XX creator access points

- Can create these
- Discerning between personal name and corporate body is problematic
- Must eliminate Discogs disambiguation qualifiers in final MARC

Los Chicos (7) – Santa Maria

bad news first...
7xx creator access points

- Discogs “roles” versus RDA “relationship designators”
  - Requires a new mapping
- Not all creators in a Discogs release have a “role” assigned
  - None in “Artist” element - the main one!
Django Reinhardt – Djangology 49

Label: RCA – 9988-1-RB
Format: Vinyl, LP
Country: US
Released: 1990
Genre: Jazz
Style: Gypsy Jazz

Credits

Bass – Carlo Pecori
Compilation Producer – Orrin Keepnews
Drums – Aurelio De Carolis
Executive Producer – Steve Backer
Guitar – Django Reinhardt
Piano – Gianni Safred
Violin – Stéphane Grappelli
non-Roman scripts

- Inconsistent practice in Discogs
- What to do when no transliteration?
- Experimenting with additional step:
  - Bumping Cyrillic script records out for manual review, transliteration in Connexion
AAPs for musical works

- Cannot automate the creation of unique access points
- Requires manual review
  - ...start with the Classical stuff?

<table>
<thead>
<tr>
<th>t</th>
<th>245</th>
<th>0</th>
<th>I Love You</th>
<th>/cElek Bacsik</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>264</td>
<td>1</td>
<td>US</td>
<td>bBob Thiele Music, c1974</td>
</tr>
<tr>
<td>p</td>
<td>264</td>
<td>4</td>
<td></td>
<td>c©1974</td>
</tr>
</tbody>
</table>

bad news first...
hybrid records

- defaulted to no 1XX
- inability to automate unique access points for works
- thus, records are formulated under AACR2/RDA hybrid guidelines
harmonizing with WorldCat

- Or just getting OCLC numbers
- Fingers crossed for Batch Services help

bad news first...
making the tradeoff

• We’re feeling okay about it

• Leveraging Discogs community for:
  Language expertise
  Obscure format expertise
  Collection PR
making the tradeoff

- The chance to experiment
- Developing relationship w/ Discogs
- Sharpening metadata expertise
- Creating visualizations
- Creating alternative displays
D3.js process:

- Discogs API ->
- Python code to parse data ->
- D3.js library to present output
visualization:
  • http://msu-libraries.github.io/discoqs/

code:
  • https://github.com/MSU-Libraries/discogs
larger context

• Will this help us with Rovi?
  ○ Not sure yet
  ○ Good intellectual work done:
    • Discogs-to-LC mapping
    • Coding to retrieve, transform data
  ○ Need to find ways to identify Rovi holdings in Discogs to do data extracts