# CATALOG ALL THE THINGS

LEVERAGING AUTOMATION TO CATALOG A MASSIVE AUDIO-VISUAL COLLECTION

Lucas Mak, Autumn Faulkner, Joshua Barton Michigan State University Libraries

slides: http://goo.gl/1rwkKC



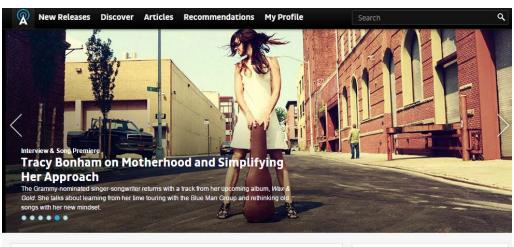
"Rovi is leading the way in the discovery and personalization of digital entertainment. Rovi helps power top brands around the world with marketleading guides, metadata, recommendations, audience analytics and advanced advertising solutions. With products deployed through an innovative cloud-based platform, Rovi is enabling customers worldwide to increase their reach, drive consumer satisfaction and create a better entertainment experience."

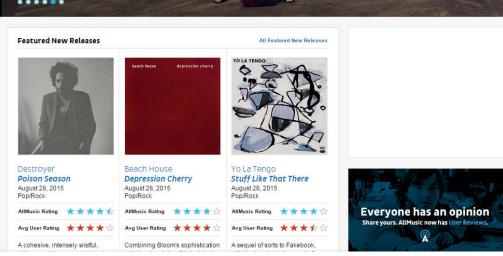
http://rovicorp.com/

### Data licensed as:

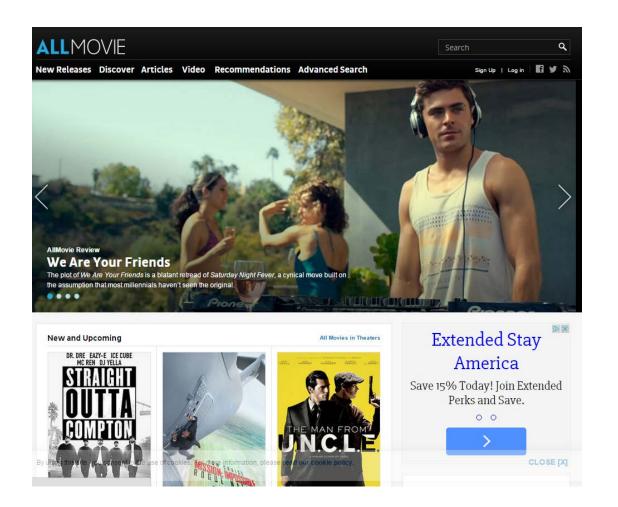
"guides, metadata, recommendations, audience analytics and advanced advertising solutions"

## **ALLMUSIC GUIDE**





## **ALLMOVIE GUIDE**



## **ALLGAME GUIDE (RIP)**

## Game over - AllGame is no more, but maybe one of our other sites will interest you:

Try AllMusic to find information about your favorite artists, albums and more.

On AllMovie you can find information on films, including what's new in theaters and on DVD.

Track episodes of your favorite TV shows and where to watch them online with Sidereel.



allmovie



## THE ROVI COLLECTION

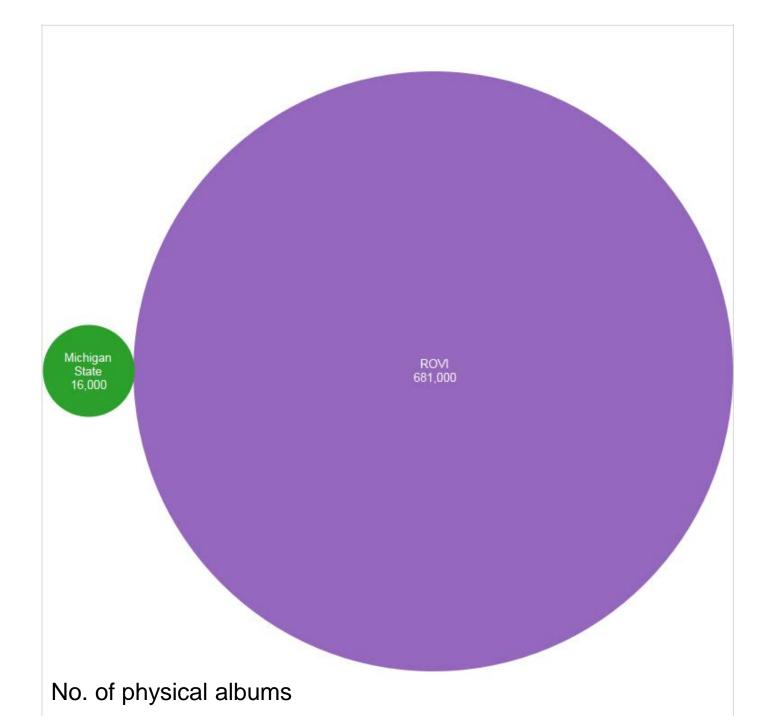
 Physical archive of nearly one million CDs, DVDs, Blu-Rays and Video Games



## THE ROVI COLLECTION: MUSIC

- Spans mid-1980s to 2014
- American and some international markets
- 681,000 CDs

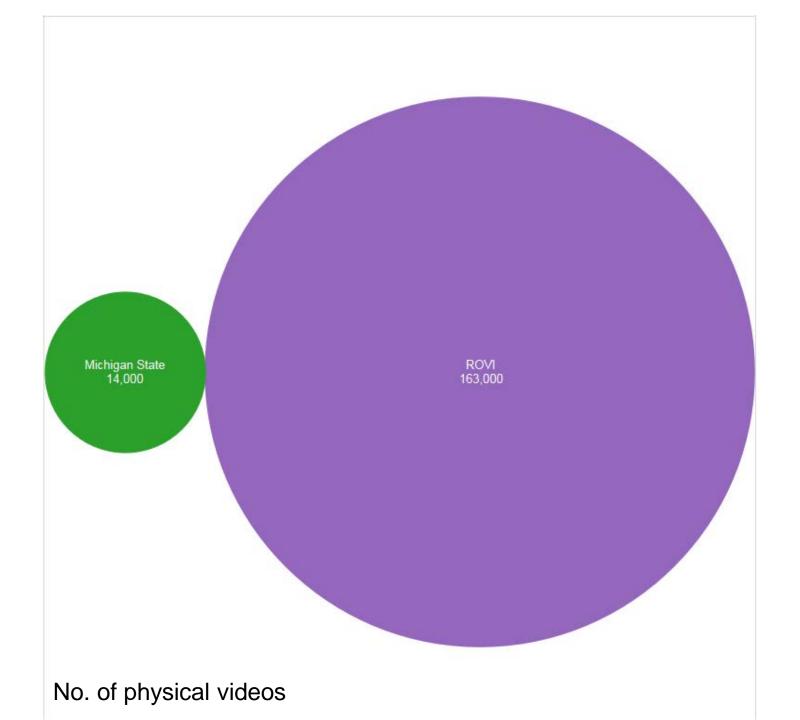




## THE ROVI COLLECTION: MOVIES

- Spans late 1990s to 2014
- 163,000 titles, DVD and Blu-Ray





## THE ROVI COLLECTION: VIDEO GAMES

- Spans 1983-2014
  - Bulk of titles mid-1990s onward
- 17,000 titles





## **ROVI METADATA**

#### Provided ROVI with metadata "wishlists"

Desired elements for music, movies, games

#### Received brief metadata records from donor

Selections from our wishlists

### Estimated cost of manual cataloging

\$20-25 million over 20 years



# ROVI VIDEO METADATA

As their metadata is proprietary, this archived copy of our presentation has been redacted per our agreement with Rovi.

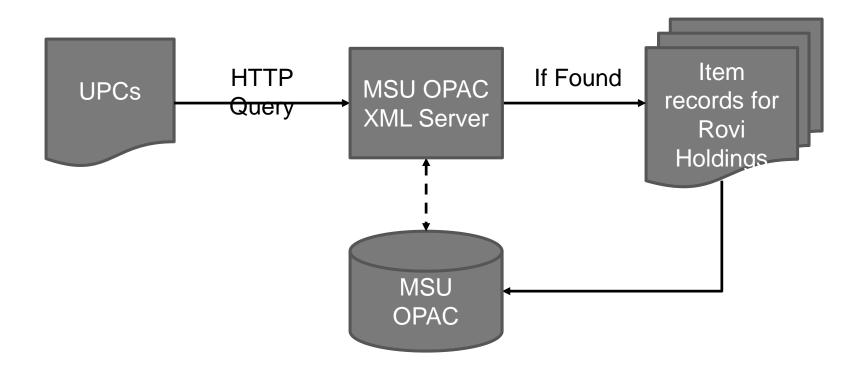


# ROVI MUSIC METADATA

As their metadata is proprietary, this archived copy of our presentation has been redacted per our agreement with Rovi.

# PHASED CATALOGING PROCESS

Phase 1 – Local Holdings Lookup



## SAMPLE PHASE 1 RECORD

Title Philadelphia [videorecording] / TriStar Pictures presents a Clinica Estetico Production.

Publisher Burbank, CA: Columbia TriStar Home Video, c2000.

Citation Cite this item (from WorldCat) (click on "Cite" in WorldCat)

LOCATION	CALL # (CLICK TO BROWSE SIMILAR ITEMS)	STATUS
MSU G.M.KLINE DIGITAL/MEDIA DVD, 4 WEST  Map It	PN1995.9.A434 P47 2000 VideoDVD	AVAILABLE
MSU ROVI MOVIE COLLECTION	AF2 D0010644 VideoDVD	AVAILABLE Request ROVI

Edition Widescreen version.

Description 1 videodisc (ca. 125 min.): sd., col.; 4 3/4 in.

Series Rovi Media Collection. Movies.

Note Originally released as a motion picture in 1993.

Contents "1993 Academy Award Winner (Best Actor: Tom Hanks ; Best Original Song: Bruce Springsteen ("Streets of Philadelphia")"

Credits Music, Howard Shore; producers, Edward Saxon, Jonathan Demme; director, Jonathan Demme.

Cast Tom Hanks, Denzel Washington, Jason Robards, Mary Steenburgen, Antonio Banderas.

Summary Powerful story of two lawyers who join together to sue a prestigious Philadelphia law firm when the firm fires one of them b

Audience Rated PG-13.

System Req DVD, 5.1 Dolby digital and Dolby surround.

Note Closed-captioned.

Includes English, French and Spanish language soundtracks, with optional subtitles in English or Korean.

Awards Academy Awards, 1993. Best actor, Tom Hanks ; Best original song, Bruce Springsteen.

Subject AIDS (Disease) -- Drama.

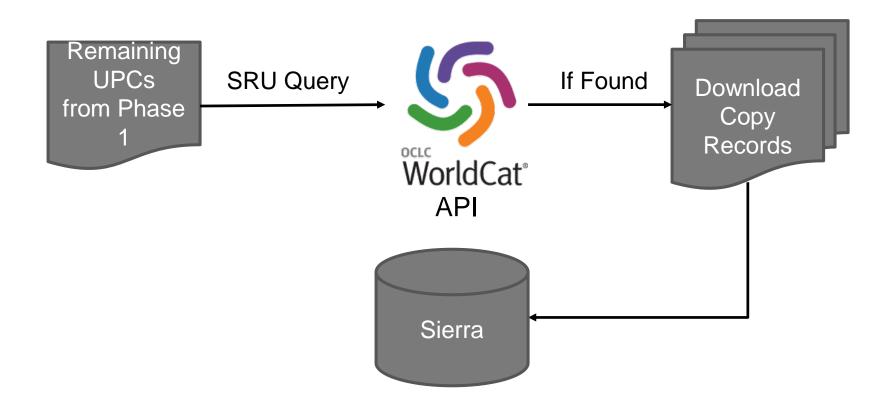
Genre Feature films.

Fiction films.

Video recordings for the hearing impaired.

# PHASED CATALOGING PROCESS

**Phase 2 – Locating Copy Records** 



### Adding Rovi-data to OCLC copy

- Disc count
  - 866 holding info for multi-disc set
  - "discs 1-n" as call number suffix
- Format
  - As Call number suffix
     e.g. Blu-ray/DVD → Blu-ray Video & Video DVD

## **SAMPLE PHASE 2** RECORD

Title A certain scientific railgun. Part one, episodes 1-12 [videorecording] / executive producer, Gen Fukunaga; director, Tatsuyuki Nagai.

Publisher [Tokyo, Japan]: Geneon Universal Entertainment; [Flower Mound, Tex.]: FUNimation Productions, Ltd.; [2013?]

Citation Cite this item (from WorldCat)

LOCATION	CALL # (CLICK TO BROWSE SIMILAR ITEMS)	STATUS
MSU ROVI MOVIE COLLECTION	CY3 D0148969 VideoDVD discs 1-2	AVAILABLE Request ROVI
MSU ROVI MOVIE COLLECTION	AR3 D0153121 VideoDVD discs 1-2	AVAILABLE Request ROVI

Edition [Ltd. ed.].

Description 2 videodiscs (300 min.): sd., col.; 4 3/4 in.

Series Rovi Media Collection. Movies.

Library Has disc 1-disc 2

Note Originally broadcast on television in 2009 as individual episodes of the television series.

Extras include selected episode comentaries and textless songs.

Contents Disc one. 01, Electromaster -- 02, When working under a hot sun, rehydration is essential -- 03, Tokiwadai is targeted -- 04, Urban legends -- 05, A certain pair of

new trainees -- 06, Everyone is proactive when it comes to this -- Disc two. 07, Abilities and powers -- 08, Level upper -- 09, Majority report -- 10, Silent majority

-- 11, Dr. Kiyama -- 12, AIM burst.

Credits Producers, Carly Hunter, Justin Cook; original story, Kazuma Kamachi, Motoi Fuyukawa; anime produced by J.C. Staff.

Audience Rated: TV14.

"Misaka's electro-manipulation abilities and delightfully destructive railgun projectile move make her a rock star in Academy City. The techno metropolis is packed with supernaturally powered students known as espers, including Misaka's flirty friend and roommate, Kuroko. She uses her teleportation skills as a member of the iudament law enforcement team, fighting crime alongside her fellow agent Uiharu. Joined by their friend Saten, a spunky level 0 esper, Misaka, Kuroko and

Uiharu have a blast taking on danger whenever and wherever it arises"--Container.

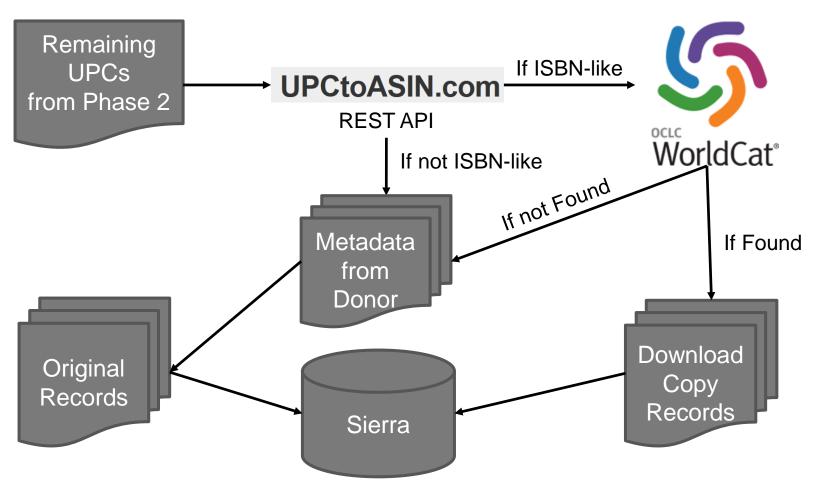
System Req DVD, region 1, NTSC; 16:9; 5.1 Dolby surround sound (English), stereo. (Japanese).

Note Soundtrack in English and Japanese with optional English subtitles.



# PHASED CATALOGING PROCESS

Phase 3 – Original Record Generation (Video)



# GENRE MAPPING (BY AUTUMN FAULKNER)

[Column of Rovi genre terms]

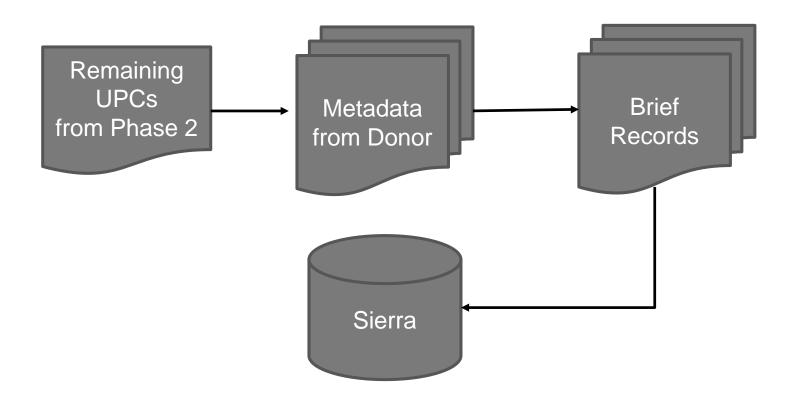
As their metadata is proprietary, this archived copy of our presentation has been redacted per our agreement with Rovi.

В	С	D	Е	F
LCGFT 1 (655 _7 \$2 lcgft)	LCGFT 2 (655 _7 \$2 lcgft)	LCSH 1 (650 _0)	LCSH 2 (650 _0)	Unauthorized (655 _4)
Action and adventure films				
Pornographic films				
Action and adventure films				
Experimental films				
Motion pictures		Motion picture industry  z India  z Mumbai	Feature films  z India	
Industrial films				
Children's films				
Comedy films				
Comedy films				
Crime films				
Motion pictures				
Dance films	Filmed dance			
Motion pictures				
Educational films				
Epic films				
Motion pictures		Interpersonal relations		
Fantasy films				
Motion pictures				Entertainment industry
Motion pictures		Physical fitness	Health	
Historical films				
Nonfiction films				
Horror films				
Motion pictures		Language and languages	Literature	
Motion pictures	Recreational works			
Concert films				
Musical films				
Detective and mystery films				
Nature films				



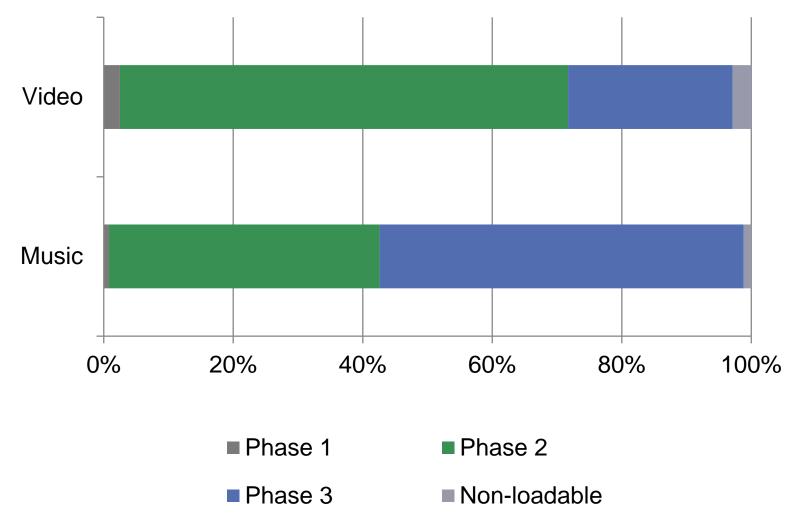
# PHASED CATALOGING PROCESS

Phase 3 – Brief Record Generation (Music)





## RECORD COUNT



## **LIMITATIONS**

### False Negative in Matching against local catalog & OCLC

Not all records have UPC

#### **Error in ROVI data**

- Wrong disc count
- Wrong format info
  - Conflicting data within a bib record
     e.g. "1 disc" in 300 but "2 discs" in 866

## **LIMITATIONS**

#### Circulation

- All-or-nothing for multi-disc sets
  - 1 accession number assigned per title
  - Accession number used as barcode for circulation



#### Dark Shadows: The Complete Original Series (Deluxe Edition) Jonathan Frid (Actor), Lara Parker (Actor), Dan Curtis (Director) Rated: NR (Not Rated) Format: DVD ★★★★ ▼ 545 customer reviews DVD \$405.79 Additional DVD options Edition Discs Price New from Deluxe Edition \$405.79 \$405.79 DVD 131 (Jul 10, 2012) DVD Limited Edition 131 (Apr 10, 2012)



Start Your 30-day Free Trial

Try Prime FREE to stream thousands of movies & TV shows. Learn more

## LIMITATIONS

#### Phase 3 music records

- Corporate name treated as personal names
  - No differentiation in original metadata
- No composer names for classical titles
  - "Artist" in data means "performer"

## UNINTENDED MESS-UP

#### **Mismatches**

- UPC was the only match point
  - Multiple hits pick the longest record
- Cataloging practices
  - UPCs of individual vol. recorded in set record
  - Set UPC in separately cataloged individual vol. record
- Consequences
  - Single-disc title → multi-disc set record
  - Multi-disc set title → Single-disc record
  - Totally different title (shared UPC code)

## UNINTENDED MESS-UP

### "Duplicate" OCLC records

- Records for same title merged on OCLC but not in local catalog
  - Existing local catalog record has an obsolete OCLC number (but different UPC code)
  - Phase 1 did not find the existing holding because of the unmatched UPC
  - Downloaded a OCLC copy (Phase 2) with an obsolete OCLC number (019) that matching the one already in local catalog
  - ILL request comes in requesting the ROVI copy → creating confusion to ILL staff since ROVI holding is not yet on OCLC

## **STAFF IMPACT**

#### **Patrons love Rovi!**

- receiving hundreds of MSU patron and interlibrary loan requests each week
- had to eventually cap fulfillments at 100 per day for outside requests

### Requests driving additional work in other units

- Interlibrary services staff pulling/refiling Rovi items
- Catalog maintenance staff processing/packaging items for borrowing
- Lucas and AV catalogers correcting records, working on special clean up projects

## REMEDIES

## Continue correcting bib records both on-the-fly and in special projects

workflows being streamlined

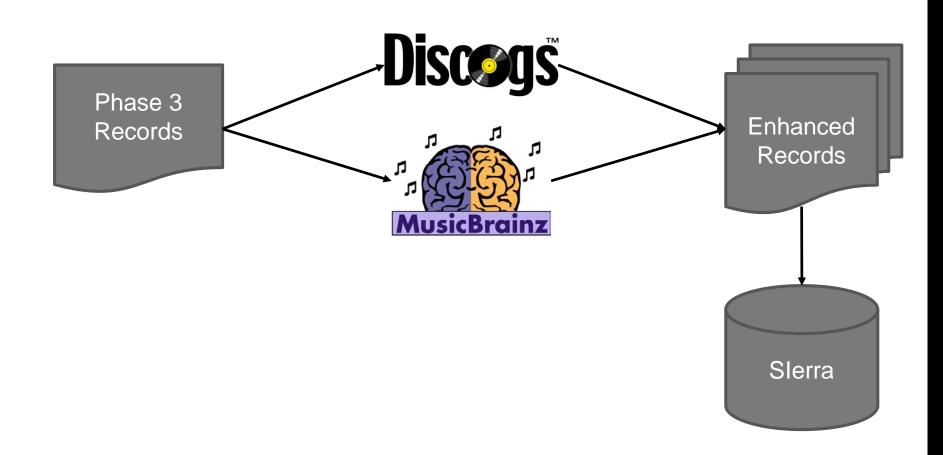
### Adjusting barcode number for circulation

 Adding disc number as suffix to allow circulating individual discs from a set separately



## **REMEDIES**

### **Record Enhancement (music)**





### **Authorized Access Point Lookup**

LIBRARY OF CONGRESS LINKED DATA SERVICE





DIGITA

#### The Library of Congress > Linked Data Service > LC Name At

Montoya, Carlos, 1903-1993

From Library of Congress Name Authority File



Visualization

**Montoya**, Carlos, 1903-1993

#### URI(s)

http://id.loc.gov/authorities/names/n82025134

#### Instance Of

- MADS/RDF PersonalName
- MADS/RDF Authority
- SKOS Concept 2

#### Scheme Membership(s)

> Library of Congress Name Authority File

#### Collection Membership(s)

- > Names Collection Authorized Headings
- > LC Names Collection General Collection





### **Carlos Montoya**

Profile: Spanish guitar player (1903 - 1993).

Sites: Wikipedia

Variations:

Viewing All

Carlos Montoya

Carlos Montoya Y Su Grupo Flament



The Free Encyclopedia

### External links [edit]

Authority control

WorldCat @ • VIAF: 39566363 @ • LCCN: n82025134 @ • I d241d75d-921f-4cfe-b2e0-da977f5f679d @



- BBC Music
- **Discogs**
- ₫ IMDb
- os Eagles
- 🗗 LyricWiki
- \*\*\* eaglesmusic
- ☑ MusicMoz
- Rateyourmusic
- Musik-Sammler
- 5 WorldCat
- iTunes
- SecondHandSongs
- **Y** @TheEaglesBand
- EaglesBand
- Stream at Spotify
- VIAF: 139561748
- IIII Q189635
- W en: Eagles (band)
  View all relationships



# MORE MAPPING (AUTUMN)

Discogs - general	Discogs - specific 1	LCGFT 1 (655 _7 \$2 lcgft) 💌	LCGFT 2 (655 _7 \$2 lcgft	LCSH 1 (650 _0)	LCSH 2 (650 _0)
Latin	Beguine	Beguines (Music)	]		
Electronic	Berlin-School	Electronica (Music)			
Folk, World, & Country	Bhangra	Bhangra (Music)			
Jazz	Big Band	Jazz		Big bands	
Electronic	Big Beat	Techno (Music)	House music		
Rock	Black Metal	Black metal (Music)			
Folk, World, & Country	Bluegrass	Bluegrass music			
Rock	Blues Rock	Blues-rock music			
Latin	Bolero	Boleros (Music)			
Pop	Bollywood	Motion picture music	Popular music	Motion picture industry Iz India	Popular music Iz India
Hip Hop	Bongo Flava	Rap (Music)		Rap (Music) Iz Tanzania	
Latin	Boogaloo	Boogaloos (Music)			
Funk / Soul	Boogie	Boogie woogie (Music)			
Blues	Boogie Woogie	Boogie woogie (Music)			
Hip Hop	Boom Bap	Rap (Music)		Rap (Music) Iz United States	
Jazz	Вор	Bop (Music)			
Jazz	Bossa Nova	Bossa nova (Music)			
Latin	Bossanova	Bossa nova (Music)			
Hip Hop	Bounce	Bounce (Music)			
Brass & Military	Brass Band			Brass band music	
Electronic	Breakbeat	Electronica (Music)			
Electronic	Breakcore	Electronica (Music)	Underground dance music		
Electronic	Breaks	Electronica (Music)			
Rock	Brit Pop	Popular music		Popular music Iz Great Britain	
Hip Hop	Britcore	Rap (Music)		Rap (Music) Iz Great Britain	
Electronic	Broken Beat	Electronica (Music)			
Folk, World, & Country	Cajun	Cajun music			
Reggae	Calypso	Calypso (Music)			
Folk, World, & Country	Cambodian Classical	Art music		Music Iz Cambodia	
Folk, World, & Country	Canzone Napoletana	Villanelle (Songs)			
Jazz	Cape Jazz	Jazz		Jazz Iz South Africa	

# IF WE COULD DO IT AGAIN...

### Use regular barcode for circulation

- Allow circulation of individual discs separately from the get-go
- Have to print a new barcode label anyway

#### **Additional match points**

- Label number (music) and name of publisher?
- Maybe including "disc count" in addition to UPC as match point to avoid mismatches??



## **QUESTIONS?**

<u>bartonjp@msu.edu</u> | <u>makw@msu.edu</u> | <u>autumn@msu.edu</u>