



# Discographic Metadata as a Research Resource for Studying the Development of Popular Music Genres and Cultures

**JADA WATSON, UNIVERSITY OF OTTAWA**

SIMSSA WORKSHOP XVII: INFRASTRUCTURE FOR MUSIC DISCOVERY

# THEORY & METHODOLOGY

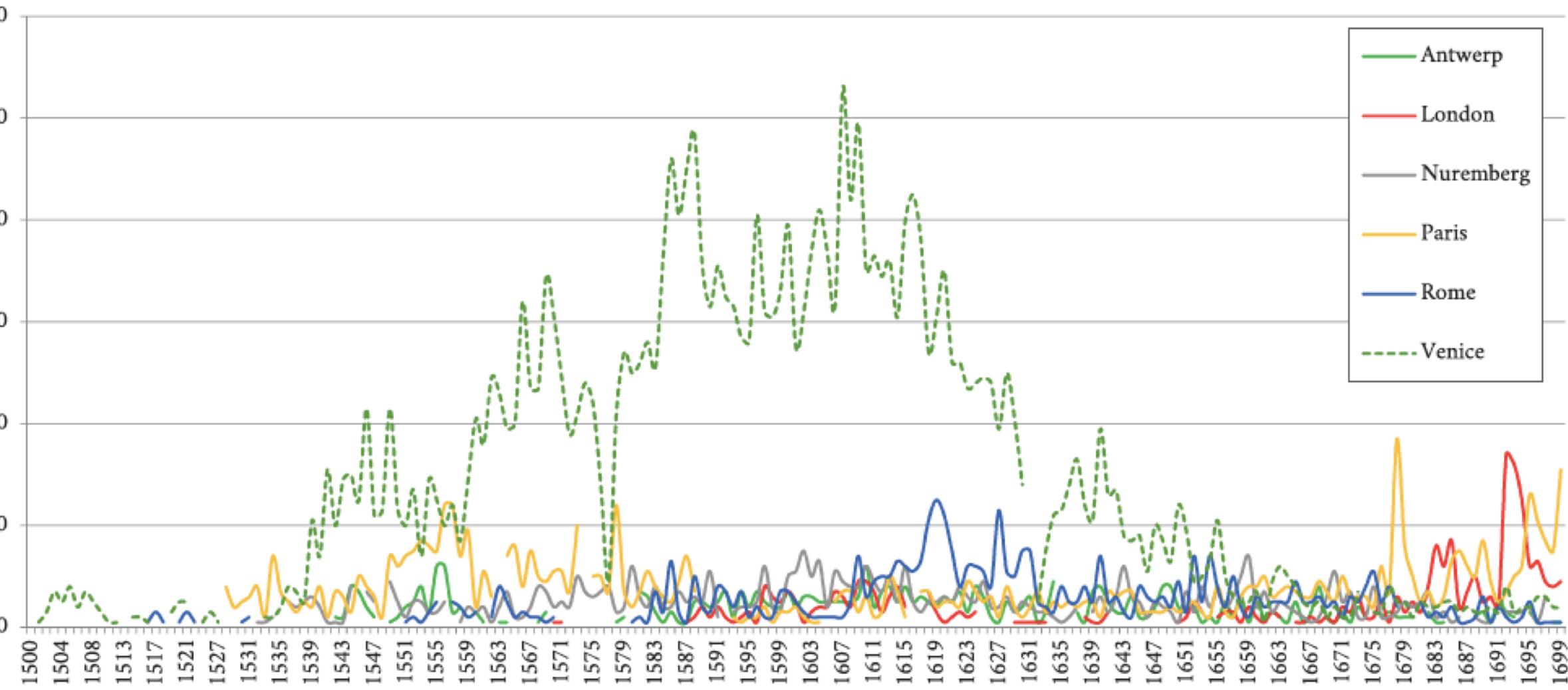
PROSOPOGRAPHY – DISTANT READING

**Prosopography** is “about what the analysis of the sum of data about many individuals can tell us about the different types of connection between them, and hence about how they operated within and upon the institutions – social, political, legal, economic, intellectual – of their time” (Keats-Rohan 2000).

As a methodology, **prosopography** “gathers and digest information about the individual persons who are attested in a particular historical period; as well as uncovering specific careers and relationships, it may also provide a tool for the broader detection of historical trends” (Cameron 2003).

# PROSOPOGRAPHY

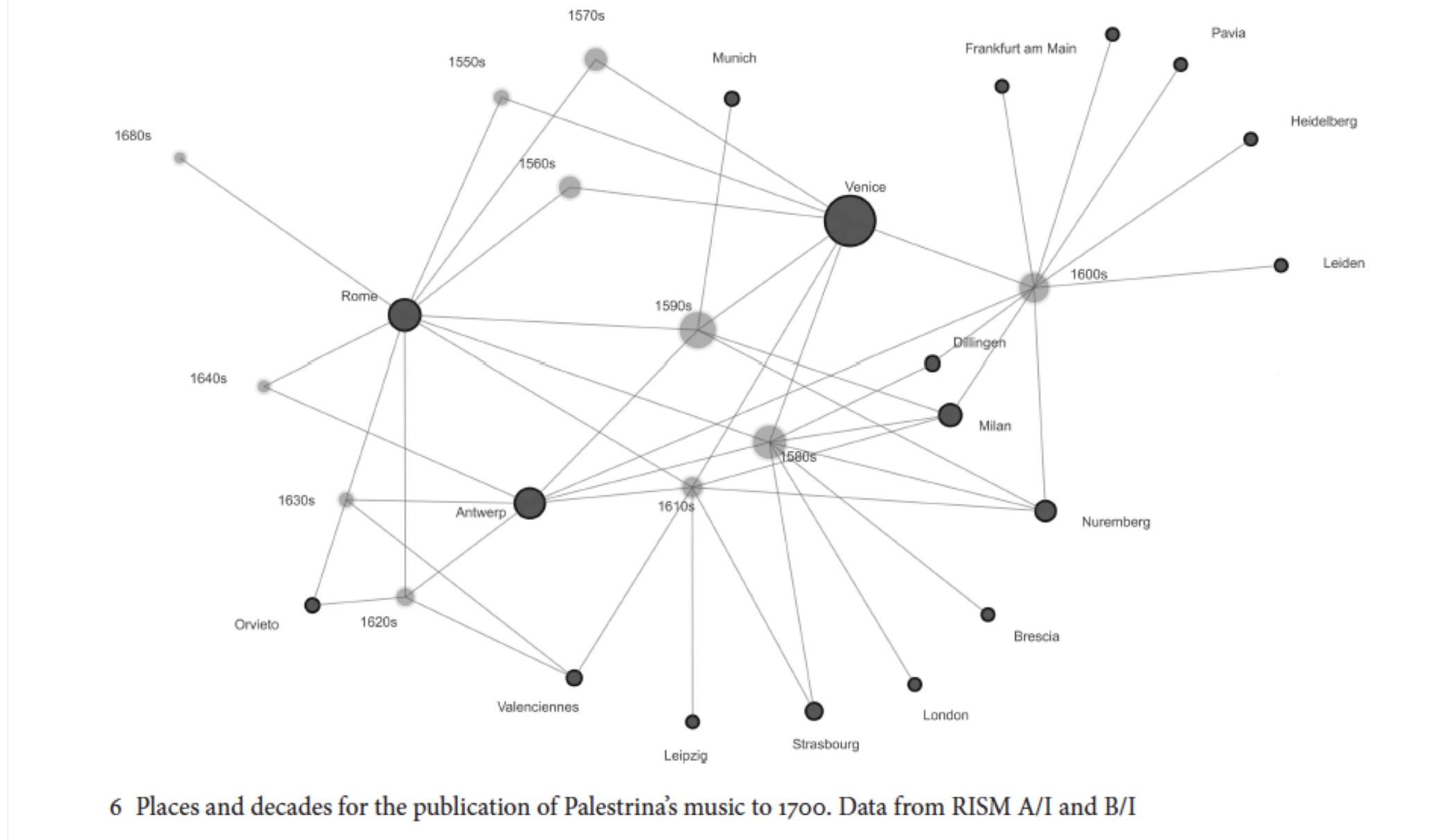
*influenced by Katherine Keats-Rohan (2000) and Averil Cameron (2003)*



3 Annual output of printed music for six major cities, 1500–1699. Data from RISM A/I and B/I

## DISTANT READING OF MUSIC DATA

“Writing a Big Data History of Music” in *Early Music* XLIII, no. 4 (2015) by Stephen Rose, Sandra Tuppen, Loukia Drosopoulou.



## DISTANT READING OF MUSIC DATA

“Writing a Big Data History of Music” in *Early Music* XLIII, no. 4 (2015) by Stephen Rose, Sandra Tuppen, Loukia Drosopoulou.

# DATA CURATION & MODELING

BILLBOARD HOT COUNTRY SONGS, 1944-2016

- **Discography:**
  - 19,844 singles (including covers, re-releases, etc.)
- **Biography:**
  - 4,127 Performers (solo artists and members of ensembles)
  - 8,193 songwriters
  - ~1,500 producers
- **Industry:**
  - ~1,187 labels and partnerships



There is overlap here: some artists in this dataset perform more than one function – and often are all three (performers, songwriters and producers).

## **DATASET: HOT COUNTRY SONGS, 1944-2016**

## 1 ▪ Discography → three main phases

1. Augmented dataset to include names of songwriters and label catalogue numbers for each record → DONE
2. Added all records charting between 1944 and 1958 (not included in core Billboard dataset) → DONE
3. Adding all corresponding album details (title, release date, producer) → *in progress*

## 2 ▪ Biography → three main phases

1. Created a dataset with all of the biographic details for every charting performer (solo and ensemble) → DONE
2. Create a subset with the biographic details for all songwriters → *in progress*
3. Create a subset with the biographic details for all producers → *in progress*

## 3 ▪ Industry → one phase

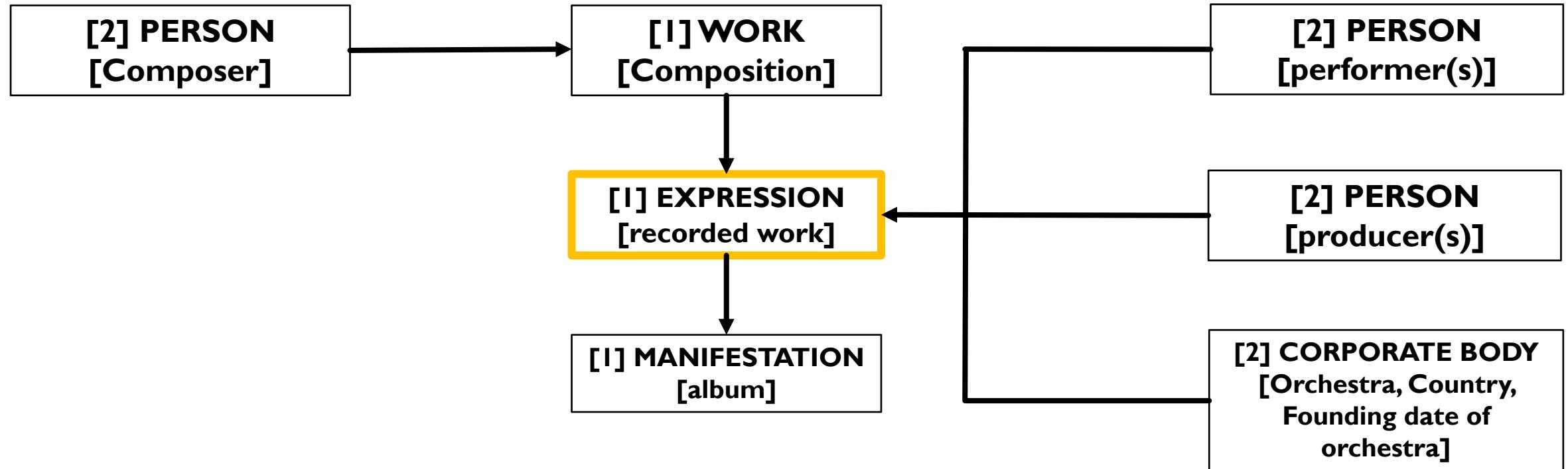
1. Dataset with the labels producing these singles → *in progress*

# CURATION PHASES

# 2

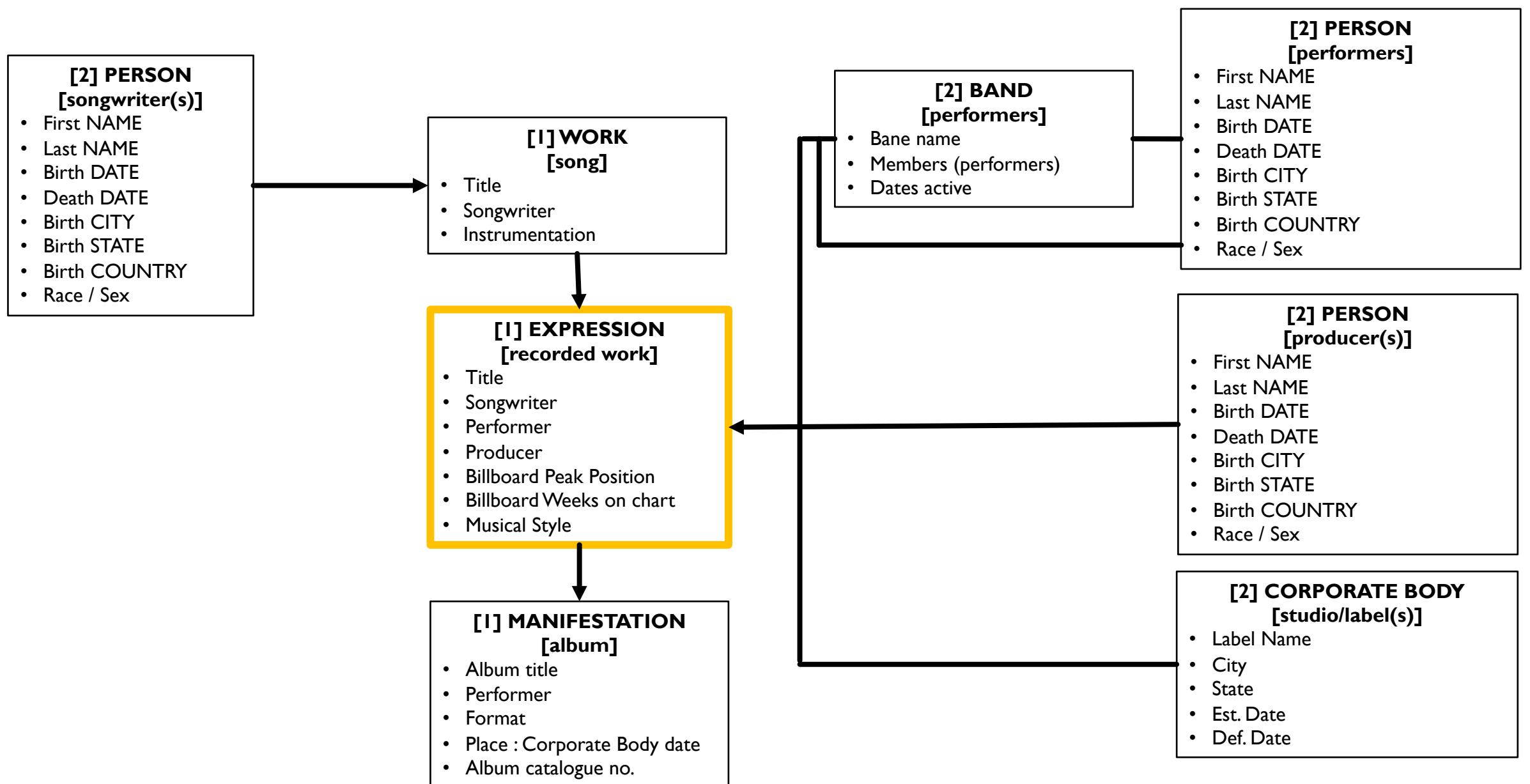
- **Discography → BGSU MARC Records, WorldCat and Wikipedia**
  - **BGSU Bill Schurk Sound Archives** → ~15,000 country album MARC records from collection
  - **Cascading Search in WorldCat** → using ISBN and/or album titles
  - **Search API** → Wikipedia to pull discographic data from album pages
  
- **Biography → Library of Congress, WorldCat and Wikipedia**
  - **MARC Edit** → to pull URI records for Name Authority files from the Library of Congress
  - **Virtual International Authority File (VIAF)** → to pull URI records for Name Authority files from WorldCat
  - **Search API** → Wikipedia to pull biographic details for songwriters / producers with Wikipedia pages

## OVERCOMING CHALLENGES



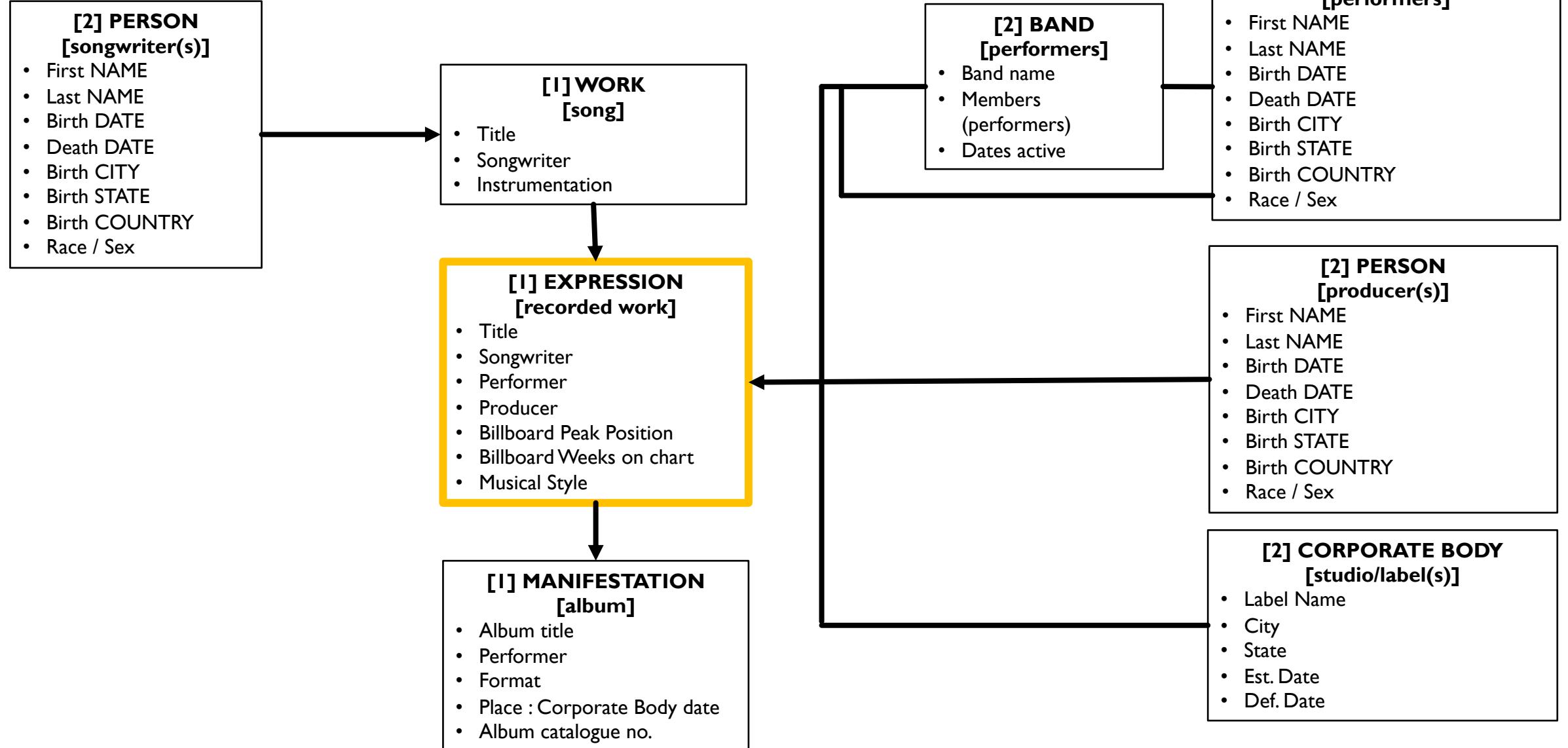
# ENTITY-RELATION MODEL FOR VARIATIONS3

influenced by FRBR (Riley, Hunter, Colvard, Berry, Indiana University; 2007)

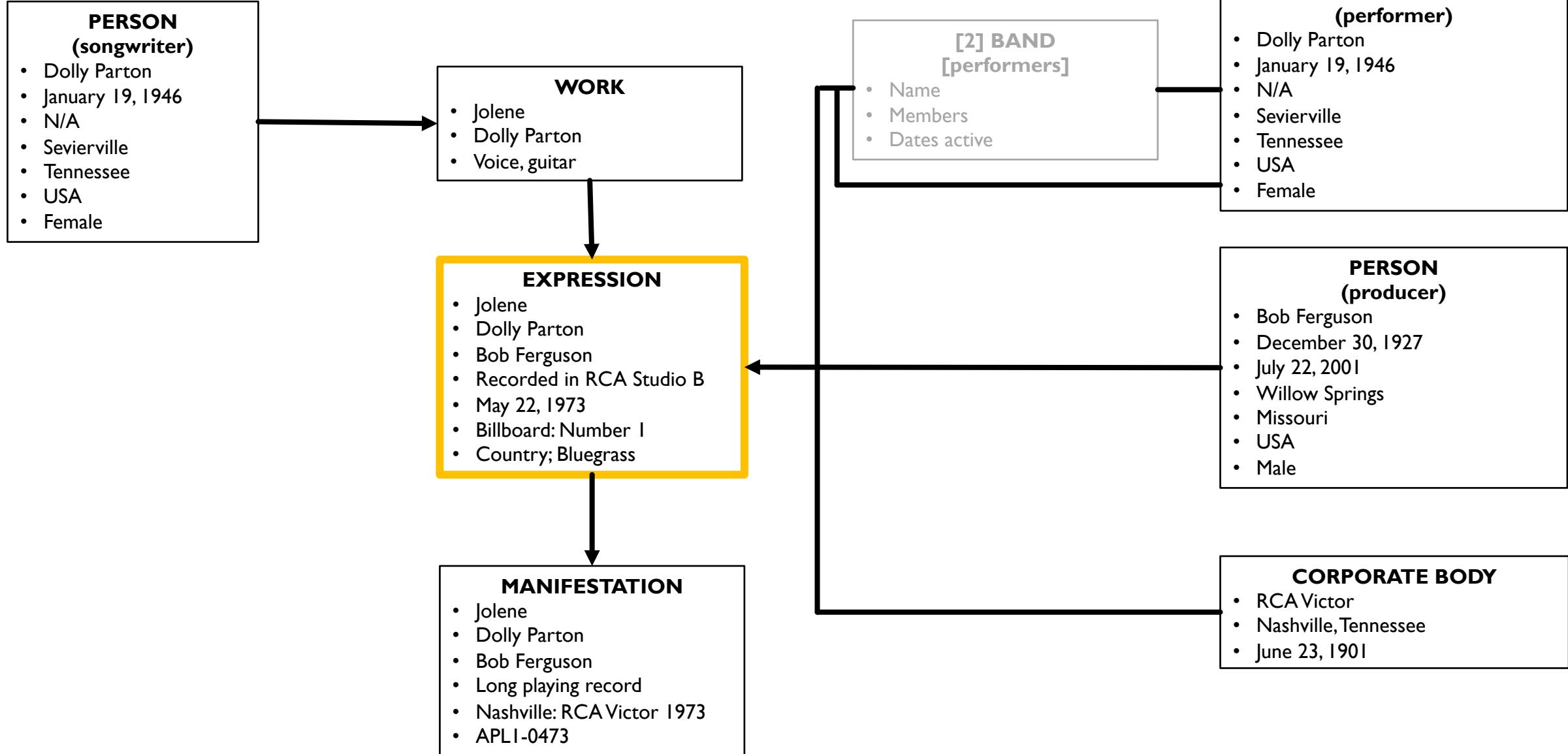


# CONCEPTUAL MODEL – ADAPTED TO SONGDATA

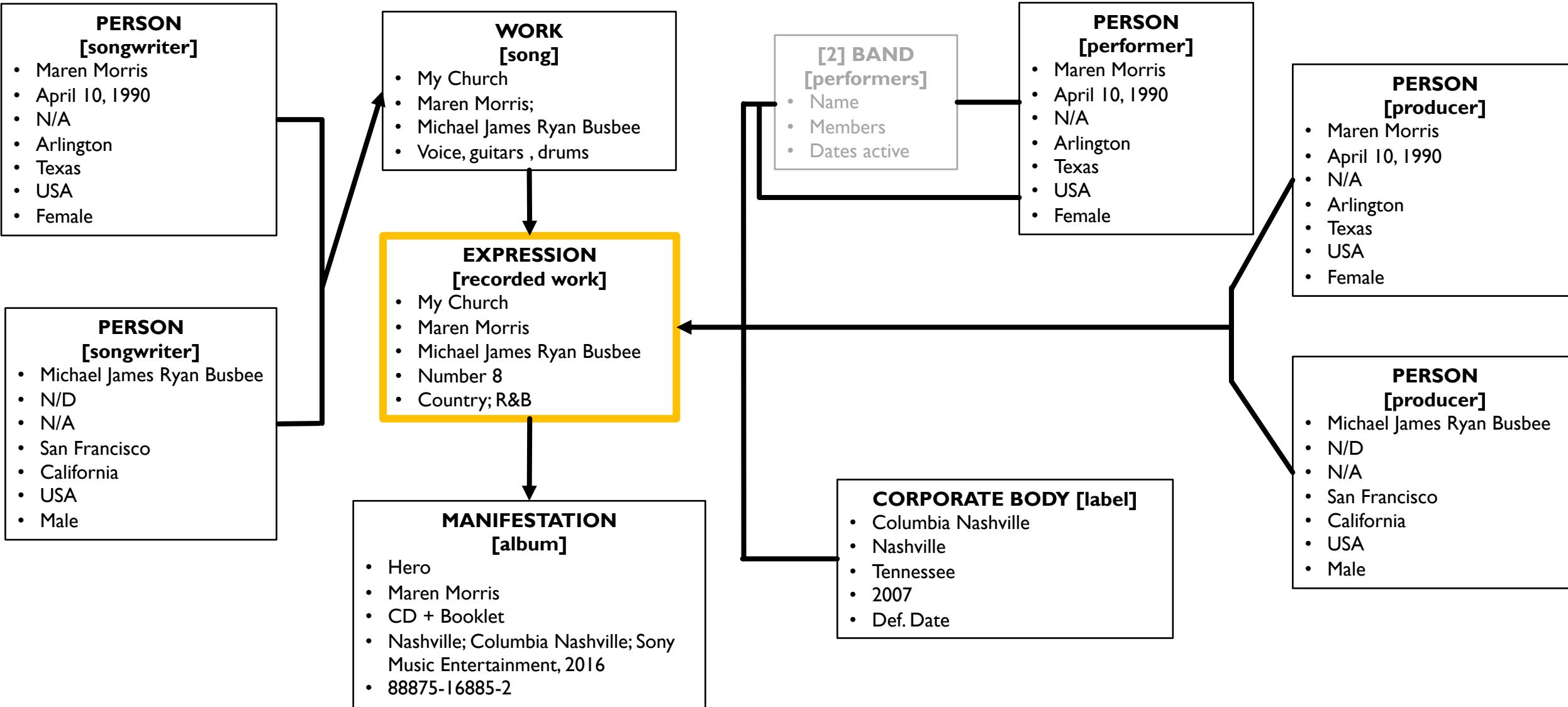
influenced by FRBR and Variations3 catalogue at Indiana University (Riley, Hunter, Colvard, Berry 2007)



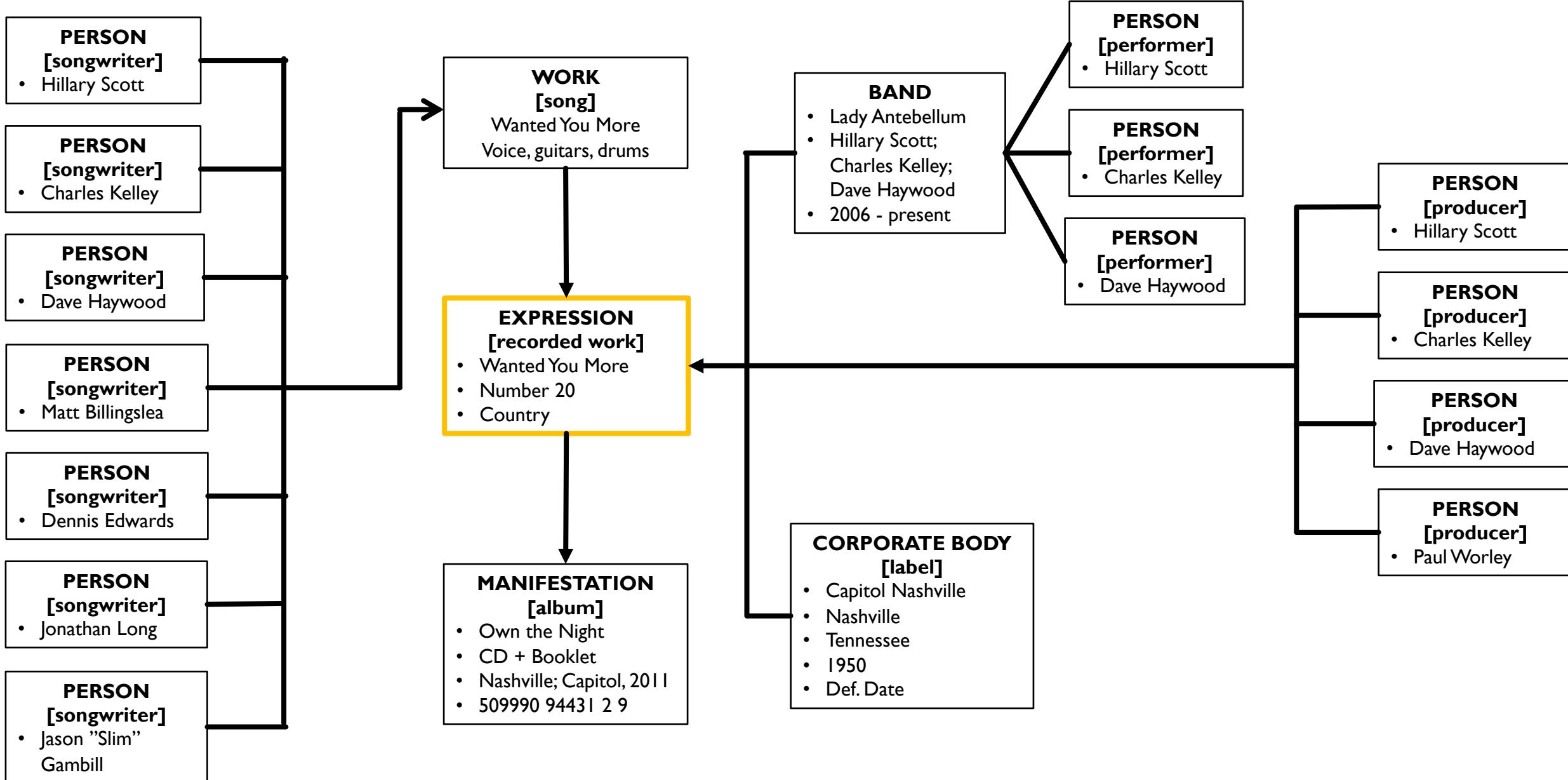
# CONCEPTUAL MODEL – ADAPTED TO SONGDATA



**CONCEPTUAL MODEL – DOLLY PARTON’S “JOLENE”**



# CONCEPTUAL MODEL – MAREN MORRIS’S “MY CHURCH”



# GENDER REPRESENTATION IN THE COUNTRY MUSIC INDUSTRY, 1996-2016

**PILOT STUDY**

*RECENTLY PUBLISHED IN JOURNAL OF POPULAR MUSIC AND SOCIETY (43, NO. 1).*

Total number of debuting songs: **4,276**

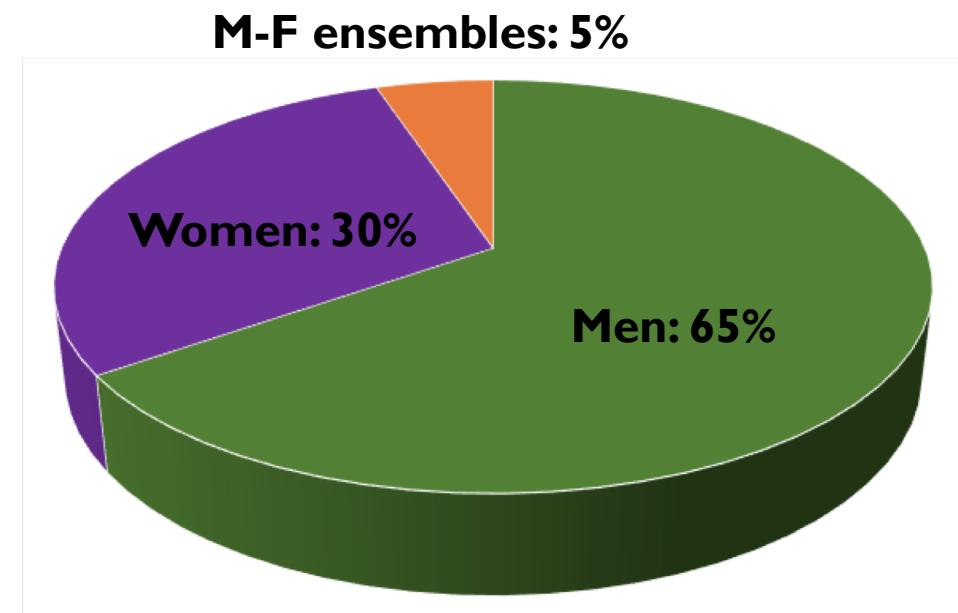
Total number of songs peaking at #1: **432**

**Total number of artists active: 718**

Men: **469**

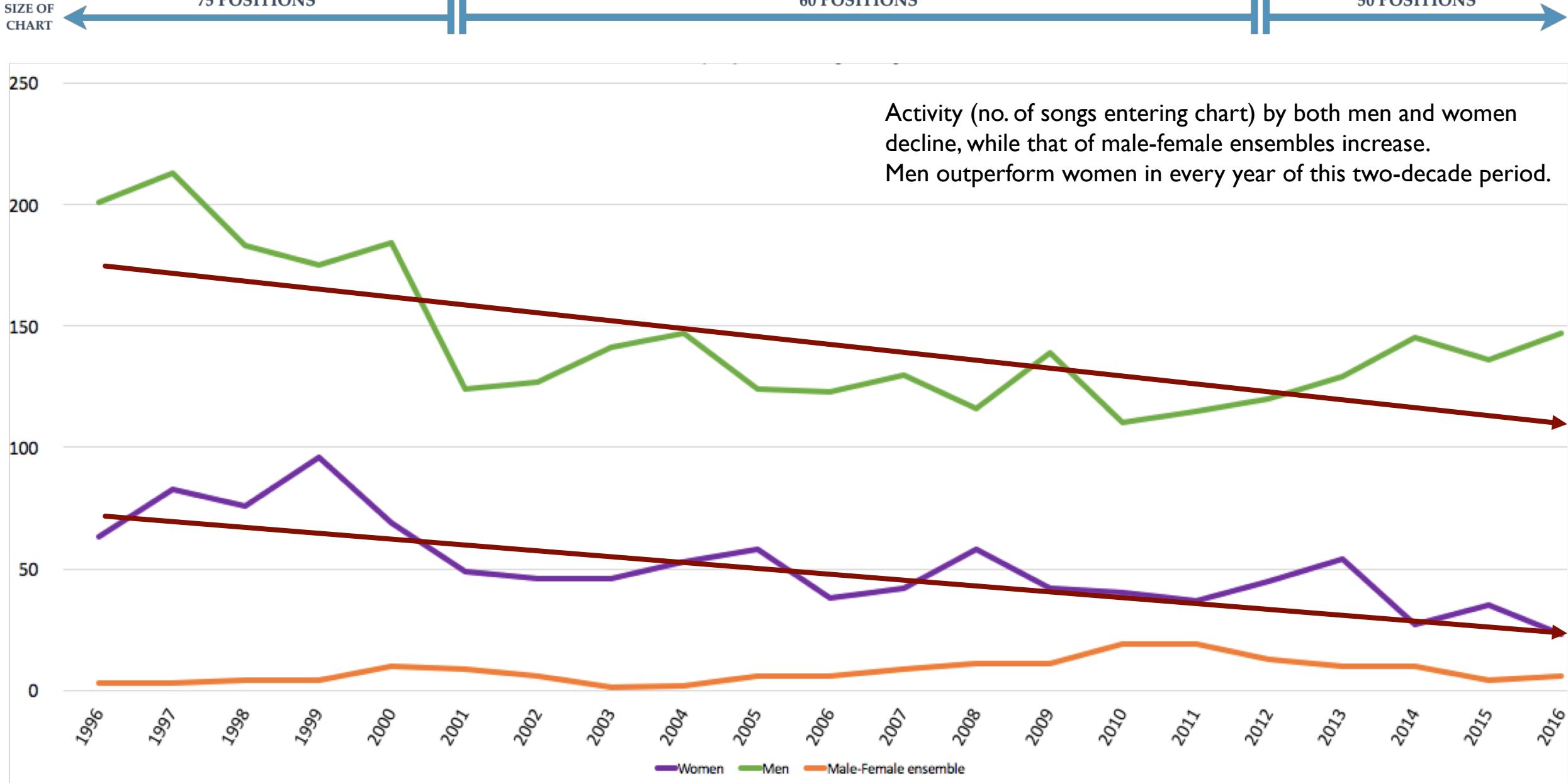
Women: **212**

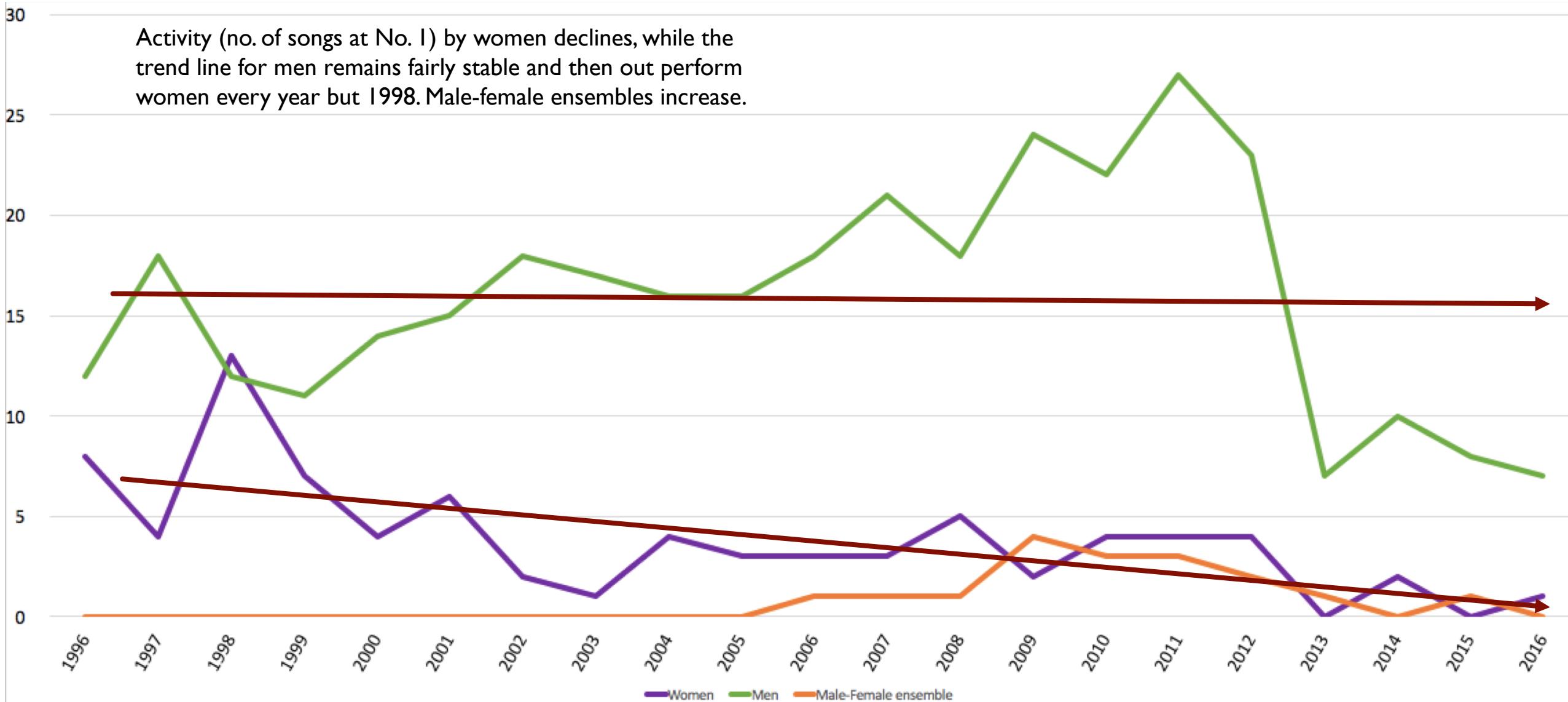
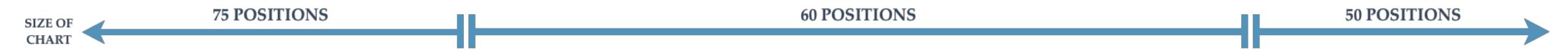
Male-female ensembles: **37**



Active artists between 1996-2016

**HOT COUNTRY SONGS, 1996-2016**





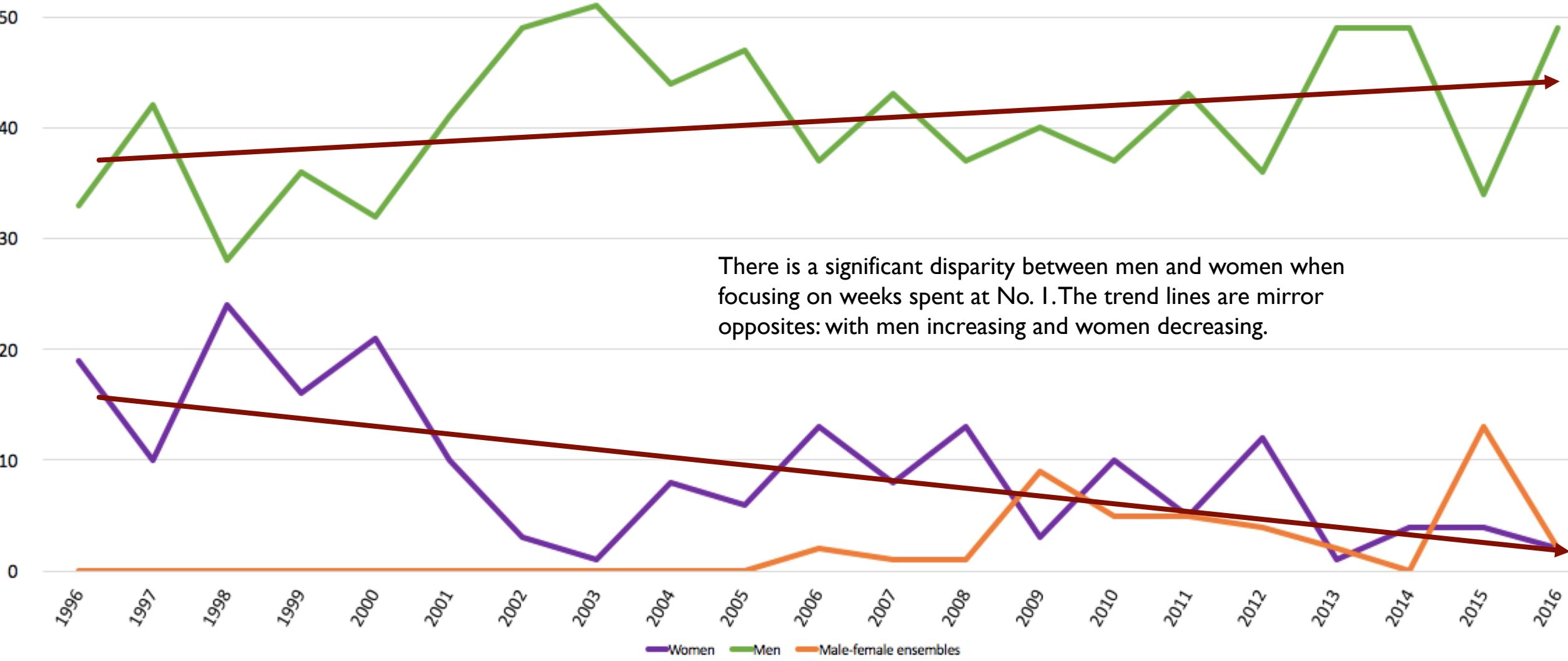
AUDIENCE-IMPRESSIONS

SIZE OF  
CHART

75 POSITIONS

60 POSITIONS

50 POSITIONS



There is a significant disparity between men and women when focusing on weeks spent at No. 1. The trend lines are mirror opposites: with men increasing and women decreasing.

METHOD  
CHANGE

TOTAL DETECTIONS

AUDIENCE-IMPRESSIONS

HOT 100-HYBRID



billboard



class  
networked cultures  
categories

prosopography  
radio metadata spaces commercialization

subcultures race sexuality

historical processes

networks communities scenes

discography identity biography

geography genres records industry

category classification gender data

culture





# Discographic Metadata as a Research Resource for Studying the Development of Popular Music Genres and Cultures

**JADA WATSON, UNIVERSITY OF OTTAWA**

Website: [www.SongData.ca](http://www.SongData.ca)

Contact: [jada.watson@uottawa.ca](mailto:jada.watson@uottawa.ca)



Social Sciences and Humanities  
Research Council of Canada

Conseil de recherches en  
sciences humaines du Canada

Canada