

literature analysis with spatial help

Introduction

Make use of the combination of NLP (NER) and spatial analysis in the humanities.

- Big-Ish Data
- Data Scraping
- Data Cleaning
- Tools Integration
- NLP
- Cartography
- Visualisation
- Sharing

Web Scraping

discography - HTML parser {Beautiful Soup / Rvest}

Secure | https://en.wikipedia.org/wiki/Bob_Dylan_discography

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page
Print/export

Bob Dylan discography

From Wikipedia, the free encyclopedia

American singer-songwriter **Bob Dylan** has released 36 studio albums, 91 singles, 26 notable extended plays, 40 music videos, 12 live albums, 13 volumes comprising *The Bootleg Series*, 19 compilation albums, 13 box sets, 7 soundtracks as main contributor, 5 music home videos and 2 non-music home videos. Dylan has starred in 6 documentaries, 3 theatrical films, appeared in an additional 8 films and 10 home videos and is the subject of the biographical tribute film *I'm Not There*. He has written and published lyrics, artwork and memoirs in 11 books and 2 of his songs have been made into children's books. He has done numerous collaborations, appearances and tribute albums. The albums *Planet Waves* and *Before the Flood* were initially released as double albums.

Dylan has won many awards for his entire body of work.^[1] For a list of the bootlegged; for an examination of this

Contents [hide]

- Albums
 - 1.1 Studio albums
 - 1.2 Live albums
 - 1.3 Compilation albums
 - 1.4 Box sets
 - 1.5 The Bootleg Series
- Singles
- Extended plays
- Music videos
- Collaborations and other appearances

Year	Album details	Peak chart positions			Certifications (sales thresholds)
		US	AUS	UK	
1962	<i>Bob Dylan</i> - Release: March 19, 1962 - Label/Catalog No.: Columbia CL-1779 (mono)/Columbia CS-8579 (stereo) - Format: vinyl (mono/stereo)	—	—	13	UK: Silver ^[2]
1963	<i>The Freewheelin' Bob Dylan</i> - Release: May 27, 1963 - Label/Catalog No.: Columbia CL-1986 (mono)/Columbia CS-8786 (stereo) - Format: vinyl (mono/stereo)	22	—	1	US: Platinum ^[3] UK: Gold ^[2]
1964	<i>The Times They Are a-Changin'</i> - Release: January 13, 1964 - Label/Catalog No.: Columbia CL-2105 (mono)/Columbia CS-8905 (stereo) - Format: vinyl (mono/stereo)	20	—	4	US: Gold ^[3] UK: Gold ^[2]
	<i>Another Side of Bob Dylan</i> - Release: August 8, 1964 - Label/Catalog No.: Columbia CL-2193 (mono)/Columbia CS-8993 (stereo) - Format: vinyl (mono/stereo)	43	—	8	US: Gold ^[3] UK: Silver ^[2]
	<i>Bringing It All Back Home</i> - Release: March 22, 1965 - Label/Catalog No.: Columbia CL-2328 (mono)/Columbia CS-9128 (stereo) - Format: vinyl (mono/stereo)	6	—	1	US: Platinum ^[3] UK: Gold ^[2]

Bob Dylan discography



Studio albums	36
Live albums	12
Compilation albums	19
Box sets	11
Music videos	40
Extended plays	26
Singles	91
Soundtrack albums	7
Tribute albums	29

Web Scraping - cont

- Genius api



`'https://api.genius.com/search?q='`

Data Cleaning

```
all_songs <- apply(album_data_frame, 1 ,FUN = function(x)
  tryCatch(
    genius_album(artist = "Bob Dylan",album = x[1]),
    error = function(e) print(e)))

album_names <- gsub(" ", "_", album_name[2:37])
# sort all lyrics
for (i in 1:36){
  assign(x = paste0(trimws(album_names[i])),
    tryCatch(
      aggregate(data.frame(all_songs[[i]]), list(data.frame(all_songs[[i]]),1)),
      FUN = function(x) paste0(unique(x))),
      error = function(e) print(e)))
  print(album_names[i])
}
```

Integration

NER - Name Entity Recognition

- Polyglot
- integrating python in to R



Interactive Table

Search:

	album	track_title	place
1	Bob_Dylan	Highway 51	wisconsin
2	Bob_Dylan	House of the Risin' Sun	orleans
3	Bob_Dylan	Man of Constant Sorrow	colorado
4	Bob_Dylan	Pretty Peggy-O	texas
5	Bob_Dylan	Pretty Peggy-O	louisiana
6	Bob_Dylan	Talkin' New York	greenwich
7	The_Freewheelin'_Bob_Dylan	Down the Highway	italy
8	The_Freewheelin'_Bob_Dylan	Gamblin' Willie's Dead Man's Hand (Limited Release Only)	orleans
9	The_Freewheelin'_Bob_Dylan	Let Me Die In My Footsteps (Limited Release Only)	nevada
10	The_Freewheelin'_Bob_Dylan	Let Me Die In My Footsteps (Limited Release Only)	mexico

Showing 1 to 10 of 61 entries

Previous

1

2

3

4

5

6

7

Next

Cartography

Geocode

Nominatim - a search engine for OpenStreetMap data

google maps api 2,500 a day vs nominatim 1 a second (if needed possible to install locally)



Visualisation

Slide with Leaflet map



Shiny Visualization

opening
 Geographical Digital Humanities
eliavs@gmail.com

album
album
 Bob_Dylan
 The_Times_They_Are_a-
 Changin'

	Data	map		
Show	10 ▼	entries	Search:	
	album	track_title	place	
1	Bob_Dylan	Highway 51	wisconsin	
2	Bob_Dylan	House of the Risin' Sun	orleans	
3	Bob_Dylan	Man of Constant Sorrow	colorado	
4	Bob_Dylan	Pretty Peggy-O	texas	
5	Bob_Dylan	Pretty Peggy-O	louisiana	
6	Bob_Dylan	Talkin' New York	greenwich	
12	The_Times_They_Are_a- Changin'	The Lonesome Death of Hattie Carroll	baltimore	
13	The_Times_They_Are_a- Changin'	The Lonesome Death of Hattie Carroll	maryland	

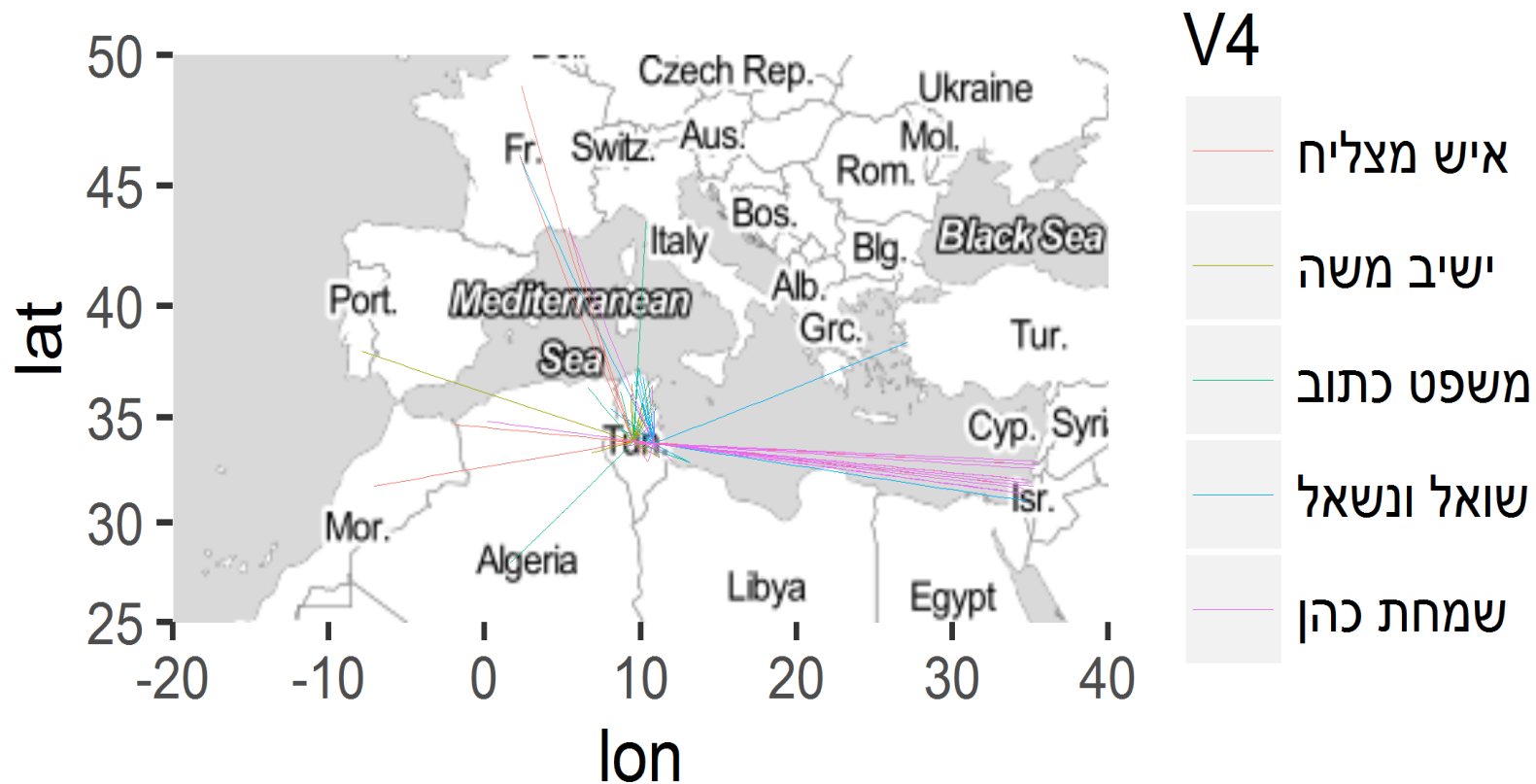
Showing 1 to 8 of 8 entries

Previous

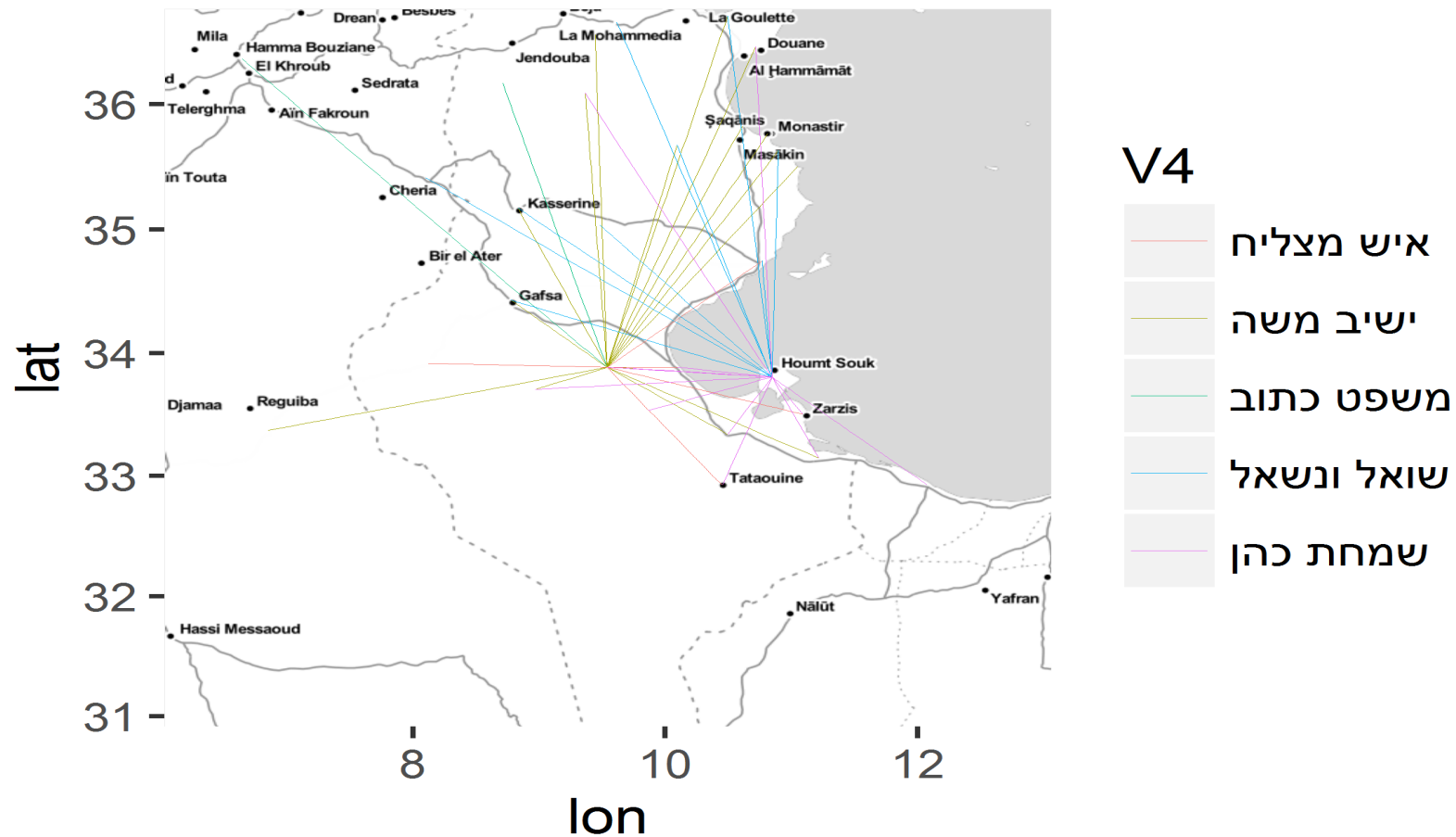
1

 Next

More examples - Responsa



More examples - Responsa



Sharing

- shiny
- Rmarkdown
- Digital Ocean
- Github

More Examples General solution

General Solution on a private server (<http://104.236.8.216:3838/>)