

Corpus-based Musicological Research

Xavier Serra, Barış Bozkurt

Agenda

- Introductory presentations (Xavier, Martin, Baris) [30 min.]
- Musicological perspective (Martin) [30 min.]
- **Corpus-based research (Xavier, Baris) [30 min.]**

----- break -----

- Rhythm analysis (Martin, Baris) [30 min.]
- Tuning analysis (Baris) [30 min.]
- Closing remarks (Xavier) [10 min.]
- Open discussion [20 min.]

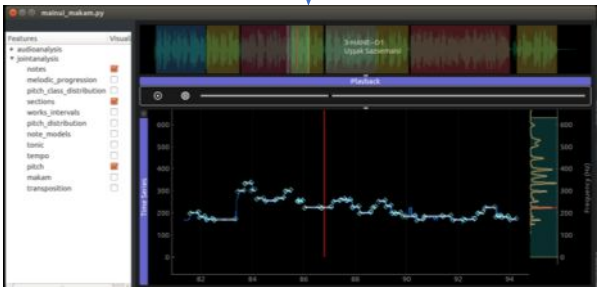
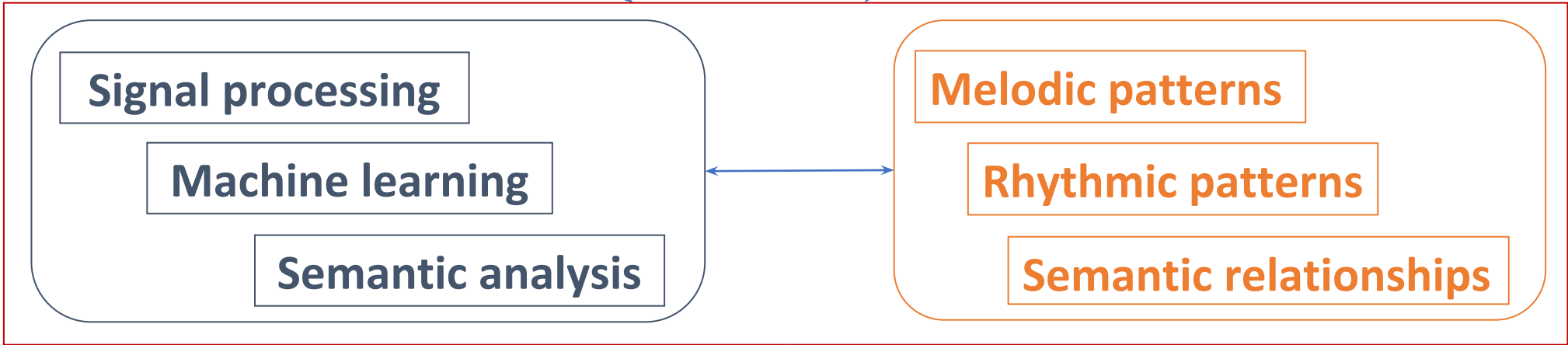
Corpus-based musicological research

- *Object of study:* music data
- *Methodologies:* statistics, signal processing, natural language processing, machine learning, ...
- *Goal:* understand music

CompMusic

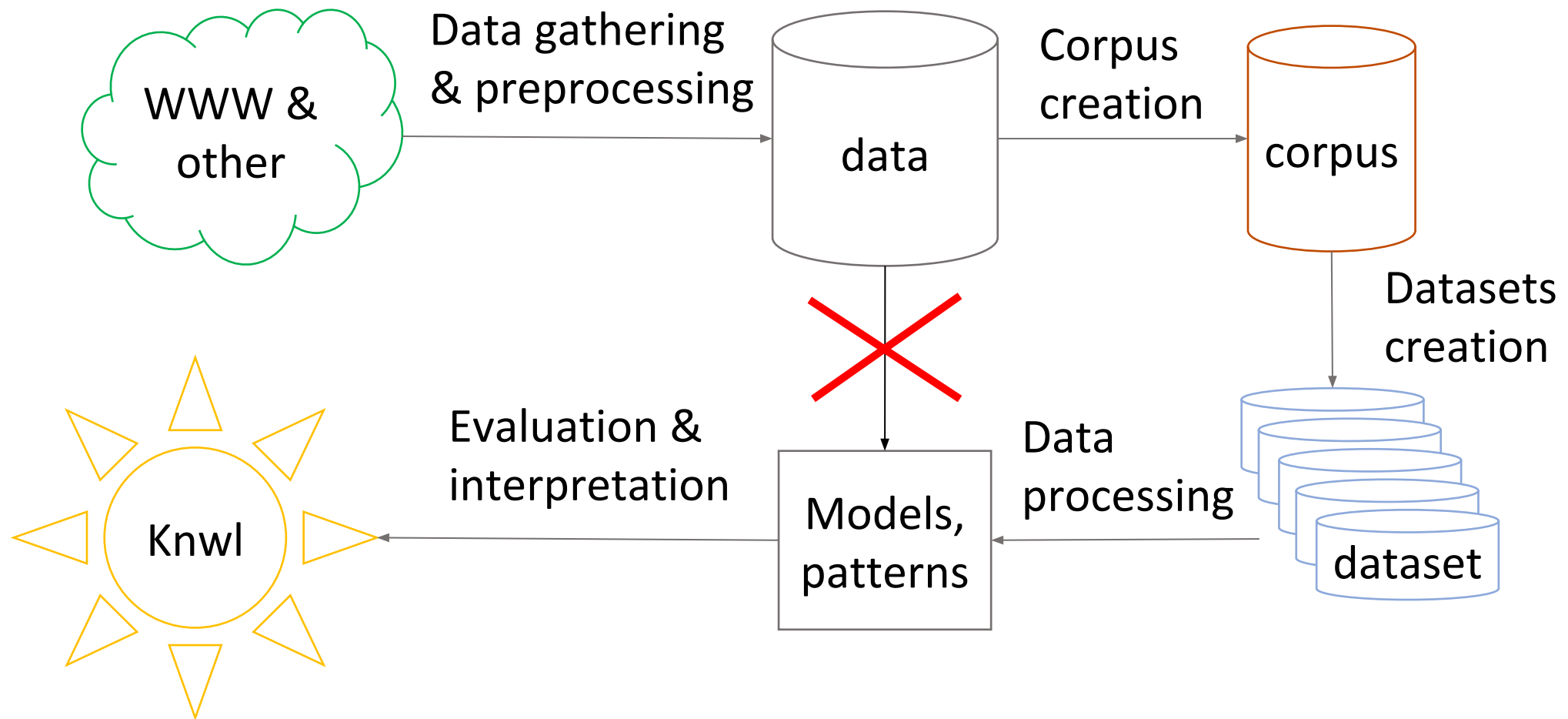


Release	Artist	Format	Tracks	Date	Country	Label	Catalog#
All India Radio Archival Release Vol. 3	Vijayal Ramesh Khan	CD	2	1960	IN	T-Series	SVCCD 126
Echoes of a Golden Voice	Jasraj	CD	4	1982	IN	Saraswathi Enterprises	P801
A Life in Music	Anur Khan	CD	2	1984	IN	Saregama	CINFL50641A02
An Undying Parampara Vol. 1	Omkarnath Thakur	CD	3	1987	IN	Rhythm House	CD 240361
The Sandoz of Parampara	Harna Bhande	CD	2	1987	IN	Rhythm House	CD 240367
Tribute To Rishi Tansen Vol-1	Vasna Sahasrabudhe	CD	2	1987	IN	Rhythm House	CD 240 369
Tribute To Rishi Tansen Vol-2	Vasna Sahasrabudhe	CD	2	1987	IN	Rhythm House	CD 240 370
An Undying Parampara Vol. 2	Omkarnath Thakur	CD	3	1988	IN	Rhythm House	CD 240374
An Undying Parampara Vol. 3	Omkarnath Thakur	CD	3	1989	IN	Rhythm House	CD 240381



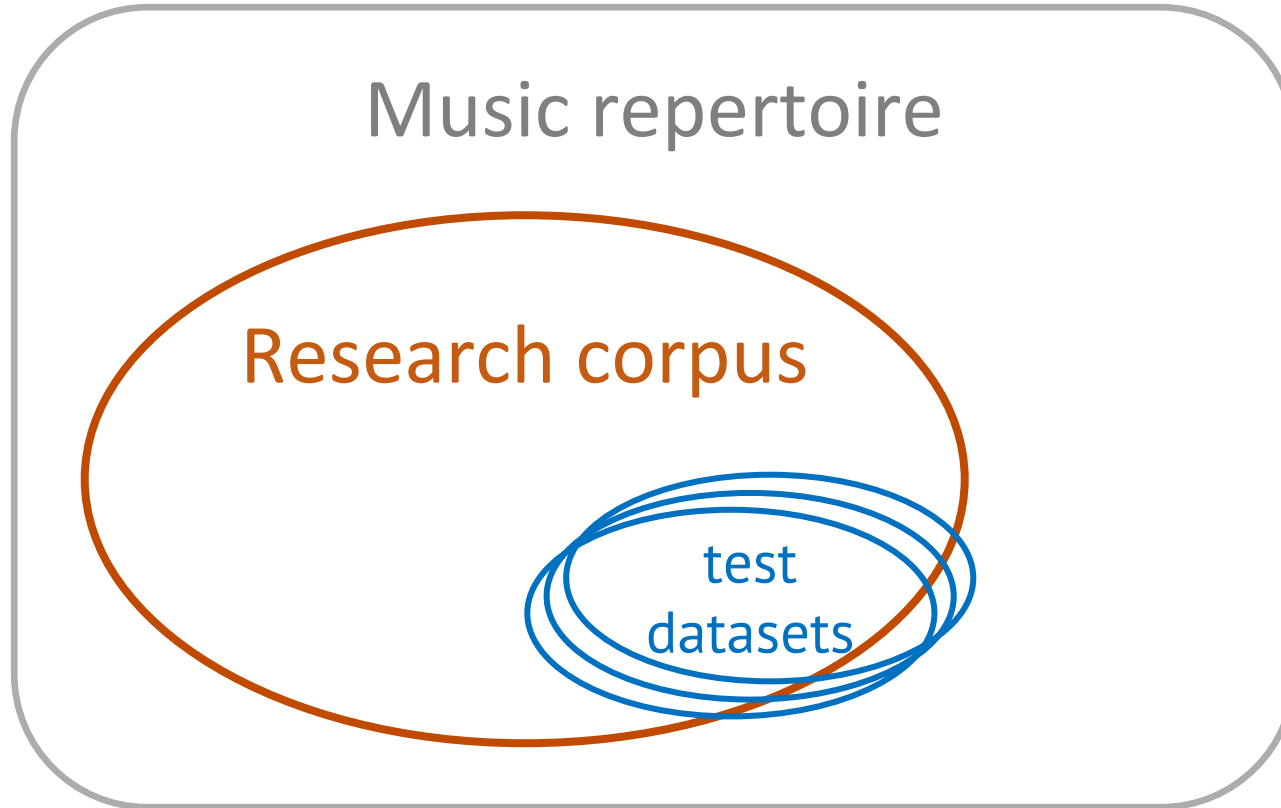
ISMIR 2018 Tutorial

X. Serra, "A Multicultural Approach in Music Information Research." ISMIR 2011.



X. Serra. "The computational study of a musical culture through its digital traces." Acta Musicologica 2017.


Corpora and datasets




- Purpose
- Coverage
- Completeness
- Quality
- Reusability

X. Serra. "Creating Research Corpora for the Computational Study of Music: the case of the CompMusic Project." AES 2014.

Data: audio + content features

[About](#) [Downloads](#) [API/Data](#) [Datasets](#) [Sign in](#)



Welcome to AcousticBrainz!

The AcousticBrainz project aims to crowd source acoustic information for all music in the world and to make it available to the public. This acoustic information describes the acoustic characteristics of music and includes low-level spectral information and information for genres, moods, keys, scales and much more. The goal of AcousticBrainz is to provide music technology researchers and open source hackers with a massive database of information about music. We hope that this database will spur the development of new music technology research and allow music hackers to create new and interesting recommendation engines.

AcousticBrainz is a joint effort between [Music Technology Group](#) at [Universitat Pompeu Fabra](#) in Barcelona and the [MusicBrainz](#) project. AcousticBrainz was originally envisioned by Xavier Serra, the founder and head of the MTG. At the heart of this project lies the [Essentia toolkit](#) from the MTG – this open source toolkit enables the automatic analysis of music. The output from Essentia is collected by the AcousticBrainz project and made available to the public.




AcousticBrainz organizes the data on a recording basis, indexed by the [MusicBrainz ID](#) for [recordings](#). If you know the MBID for a recording, you can easily fetch from AcousticBrainz. For details on how to do this, visit our [API documentation](#).


All of the data contained in AcousticBrainz is licensed under the [CC0 license](#) (public domain).

Examples

If you're wondering what this collected data actually looks like, have a look at the last 5 recordings that have been submitted:

1. [Magic Slim and the Teardrops / 1823 South Michigan Avenue](#)
2. [Magic Slim and the Teardrops / Wonder Why](#)
3. [Bobby Caldwell / Open Your Eyes](#)
4. [Magic Slim and the Teardrops / Make My Dreams Come True](#)
5. [BOBBY CALDWELL / Angel Eyes](#)

**AcousticBrainz**
@AcousticBrainz
We have put in pl
users will need to
log in. For more d
compliance, pleas
rk
 

**AcousticBrainz**
[Embed](#)

ESSENTIA

Essentia

Open-source library and tools for audio and music analysis, description and synthesis

GET STARTED


DOWNLOAD

Data Statistics	
	Lossy
Unique	2,783,181
All	6,695,605
updated e last updated: more on the d	

Tweets by @AcousticB

 **AcousticBrainz**
@AcousticBrainz
We have put in pl
users will need to
log in. For more d
compliance, pleas
rk

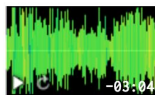
 **AcousticBrainz**
[Embed](#)

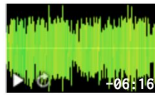


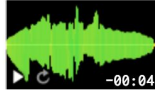
[Home](#) [Sounds](#) [Forums](#) [People](#) [Help](#)


search sounds

previous next 1 2 | 7290 sounds

**carnatic-violin.wav** ★★★★★
carnatic violin solo improvisation, Alap section, played by Lavanya Kothandaraman on the raga Kharaharapriya as part of a concert of ...
India Lavanya-Kothandaraman music CompMusic violin Chennai Alap Carnatic
23 more results in the same pack "carnatic music"

**Carnatic varnam by Prasa...** ★★★★★
Varnam is a compositional form of **carnatic** music, rich in melodic nuances. This is a recording of a varnam, titled ...
carnatic music varnam compmusic iit-madras carnatic-varnam-dataset
27 more results in the same pack "Carnatic Varnams"

**violin carnatic phrase A...** ★★★★★
Short **carnatic** phrase in A-minor, played on the violin
carnatic phrase violin A-minor hpr-model legato bow string

**Monotonized Carnatic mus...** ★★★★★
This sound is a synthesized sound where the loudness and the timbre (weights) of the

xserra
July 26th, 2012
234 downloads
5 comments

gopalkoduri
March 19th, 2014
7 downloads
0 comments

violinsimma
May 21st, 2018
9 downloads
0 comments

sankalp
August 25th, 2016

licenses

- Attribution (115)
- Attribution Noncommercial (7162)
- Creative Commons 0 (13)

tags

carnatic compmusic drum h4n-zoom icassp2013-dataset iitm indian-percussion mridangam mridangam-dhin-dsh mridangam-num-dsh mridangam-stroke-dataset mridangam-ta-b mridangam-ta-c mridangam-ta-csh mridangam-ta-d mridangam-ta-b mridangam-tha-c mridangam-tha-csh mridangam-tha-d mridangam-tha-dsh mridangam-tha-e mridangam-thi-b mridangam-thi-c mridangam-thi-csh mridangam-thi-d mridangam-thi-dsh mridangam-thi-e mridangam-thom-b percussion sm-58

Data: text + metadata

MusicBrainz

Log In

Create Account

Search

Artist

About Us

Products

Search

Documentation

English

Ravi Shankar

~ Person

Overview

Releases

Recordings

Works

Events

Relationships

Aliases

Tags

Details

Edit

Wikipedia

Ravi Shankar (Bengali: রবী শঙ্কর) (IPA: [ˈrobi ˈʃɔŋkor]; 7 April 1920 – 11 December 2012), born **Rabindra Shankar Chowdhury**, his name often preceded by the title Pandit ('Master'), was a Indian musician and a composer of Hindustani classical music. He was one of the best-known proponents of the sitar in the second half of the 20th century and influenced many other musicians throughout the world. In 1999, Shankar was awarded India's highest civilian honour, the Bharat Ratna.

Show more...

Continue reading at Wikipedia...Wikipedia content provided under the terms of the [Creative Commons BY-SA license](#)

Discography

Previous

1

2

Next

Album

Year	Title	Artist	Rating	Releases
1956	Music Of India (Three Classical Ragas)	Ravi Shankar	★★★★★	3
1958	India's Master Musician	Ravi Shankar	★★★★★	2
1958	The Sounds of India	Ravi Shankar	★★★★★	5
1959	Music Of India - Rāgas And Tālas	Ravi Shankar	★★★★★	1
1962	Music of India	Ravi Shankar	★★★★★	1
1962	Ravi Shankar In Concert	Ravi Shankar	★★★★★	3
1963	In London	Ravi Shankar	★★★★★	3

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

Create a book

Download as PDF

Artist info

Sort name:

Type: Person


Gender: Male

Born: 1920

Born in: Utt

Died: 2012

Died in: San

Area:  India

Rating

★★★★★

Tags

Indian classical

sitar player,

Article

Talk

Read

Edit

View history

Search Wikipedia

Ravi Shankar

From Wikipedia, the free encyclopedia

For other people named Ravi Shankar, see [Ravi Shankar \(disambiguation\)](#).


Ravi Shankar (Bengali: রবী শঙ্কর) (IPA: [ˈrobi ˈʃɔŋkor]; 7 April 1920 – 11 December 2012), born **Rabindra Shankar Chowdhury**,^[2] his name often preceded by the title **Pandit** ('Master'), was a **Indian musician** and a composer of **Hindustani classical music**. He was one of the best-known proponents of the **sitar** in the second half of the 20th century and influenced many other musicians throughout the world. In 1999, Shankar was awarded India's highest civilian honour, the **Bharat Ratna**.

Shankar was born to a **Bengali** family in India,^[3] and spent his youth touring India and **Europe** with the dance group of his brother **Uday Shankar**. He gave up dancing in 1938 to study sitar playing under court musician **Allauddin Khan**. After finishing his studies in 1944, Shankar worked as a composer, creating the music for the ***Apu Trilogy*** by **Satyajit Ray**, and was music director of **All India Radio**, New Delhi, from 1949 to 1956.

In 1956, Shankar began to tour Europe and the Americas playing **Indian classical music** and increased its popularity there in the 1960s through teaching, performance, and his association with violinist **Yehudi Menuhin** and **Beatles** guitarist **George Harrison**. His influence on the latter helped popularize the **use of Indian instruments in pop music** throughout the 1960s. Shankar engaged Western music by writing compositions for sitar and orchestra, and toured the world in the 1970s and 1980s. From 1986 to 1992, he served as a nominated member of Rajya Sabha, the upper chamber of the **Parliament of India**. He continued to perform until the end of his life.

Pandit

Ravi Shankar



Shankar in 1988

Background information

Birth name

Rabindra Shankar Chowdhury

Born

7 April 1920

Benares, Benares State, British Raj

(now Varanasi, Uttar Pradesh, Republic of India)

Died

11 December 2012 (aged 92)

San Diego, California, United States

Genres

Indian classical music

Occupation(s)


Musician · composer

Instruments

Sitar · veena

ISMIR 2018 Tutorial

Data: scores

 Search for sheet music

Strings ▾ 🔍

Go Pro

Upload

Log In

Find Music For: **Strings** Piano Film, TV & Video Games Saxophone Percussion Trumpet Other...


⏮ ⏭ 00:00 🎹 🎵 🔔 ⋮ 🗨


Quddam Maya

Tetouan Orchestra Traditional

♩ = 120

Violin





MTG - Music Technology Group

329 0 0

Download

Share

Favorite

UPLOADED ON NOV 26, 2014

This is a transcription of an Arabic recording of the Tetouan Orchestra. [Show more](#)

THIS SCORE APPEARS IN:

Arab-Andalusian (set)

YOUR COMMENT

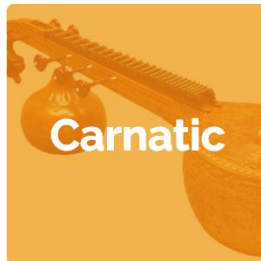
```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE score-partwise PUBLIC "-//Recordare//DTD MusicXML 2.0 Partwise//EN"
    "http://www.musicxml.org/dtds/partwise.dtd">
<score-partwise version="2.0">
  <movement-title>Mshalia Isbahan_Tetuán</movement-title>
  <identification>
    <encoding>
      <software>Finale 2009 for Windows</software>
      <software>Dolet Light for Finale 2009</software>
      <encoding-date>2017-02-01</encoding-date>
      <supports attribute="new-system" element="print" type="yes" value="yes"/>
      <supports attribute="new-page" element="print" type="yes" value="yes"/>
    </encoding>
  </identification>
  <defaults>
    <scaling>
      <millimeters>7.2319</millimeters>
      <tenths>40</tenths>
    </scaling>
    <page-layout>
      <page-height>1545</page-height>
      <page-width>1194</page-width>
      <page-margins type="both">
        <left-margin>140</left-margin>
        <right-margin>70</right-margin>
        <top-margin>88</top-margin>
        <bottom-margin>88</bottom-margin>
      </page-margins>
    </page-layout>
    <system-layout>
      <system-margins>
        <left-margin>0</left-margin>
        <right-margin>0</right-margin>
      </system-margins>
      <system-distance>121</system-distance>
      <top-system-distance>70</top-system-distance>
    </system-layout>
  </defaults>
</score-partwise>
```



[Contact](#) [About](#) [Register](#) [Developers](#)

Dunya comprises the music corpora and related software tools that have been developed as part of the CompMusic project. These corpora have been created with the aim of studying particular music traditions and they include audio recordings plus complementary information that describes the recordings. Each corpus has specific characteristics and the developed software tools allow to process the available information in order to study and explore the characteristics of each musical repertoire.

Explore our collections



Carnatic



Hindustani



Makam

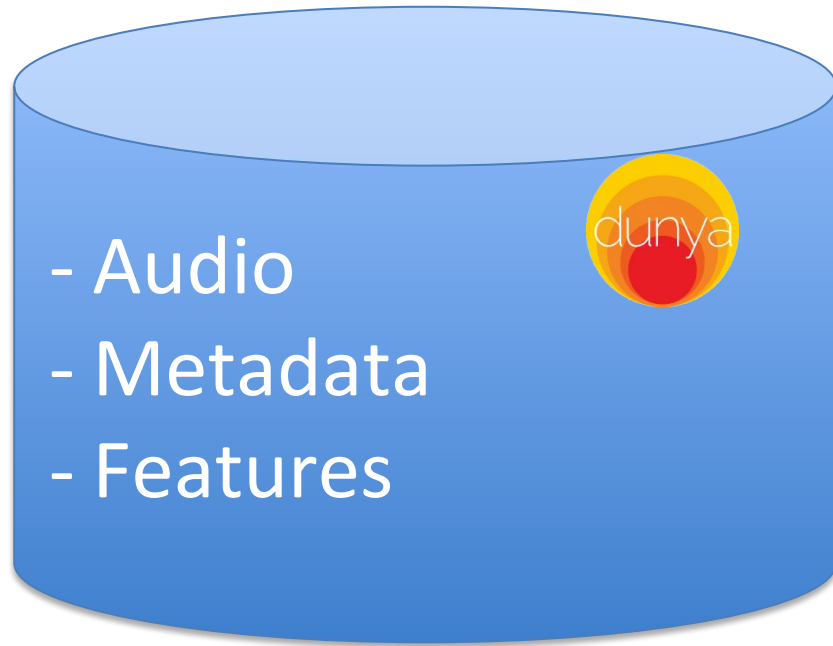


Jingju



Andalusian

Dunya Server



pycompmusic

Python tools for accessing audio, annotations and features.

<https://github.com/MTG/pycompmusic>

Features extracted using culture-specific tools designed by MTG researchers

HANDS ON - Set up



(! Titles in red refer to slides of hands on parts where we go through actual code)

```
$ git clone https://github.com/MTG/ismir2018TutorialNotebooks.git
```

```
$ cd ismir2018TutorialNotebooks/
```

! Mac users should launch Docker before this command

```
$ docker-compose up
```

Access <http://localhost:8888> using a browser, password: *mir*

Docker image: <https://github.com/MTG/MIR-toolbox-docker>

Contains:

Common data-science packages: numpy, scikit-learn, pandas, etc.

MIR tools: mir-eval, music21

MTG tools: Essentia, freesound-python, pycompmusic

Jupyter server

Accessing Dunya corpora using the API

Hands on: DownloadDataFromDunya_noToken.ipynb

This notebook demonstrates downloading data using pycompmusic API.

- downloading a single file using recording's MusicBrainz ID
 - downloading files of a CompMusic dataset
- (https://github.com/MTG/otmm_makam_recognition_dataset)

Music Brainz ID for a recording is simply the code you can retrieve from a MusicBrainz link to a recording.

Example: the recording

<https://musicbrainz.org/recording/e666ec52-b752-492d-9423-24e1c7bffb7c>

has a Music Brainz ID : 'e666ec52-b752-492d-9423-24e1c7bffb7c'

Accessing features, metadata, annotations

Hands on: `downloadAllSARAGAContent.ipynb`

This notebook demonstrates the use of Dunya api for downloading Saraga dataset files which includes manual annotation files accompanying recordings with Creative Commons licence(in mp3 format).

Saraga dataset is also available for direct download on
: <https://doi.org/10.5281/zenodo.1256126>

Saraga dataset is composed of two collections:

- [Hindustani collection](#)
- [Carnatic collection](#)

The notebook creates two subfolders and saves all data in these folders. Each annotation is saved in a separate text file.

Corpus-based Musicological Research

Xavier Serra, Barış Bozkurt