
MetaBrainz databases

Alastair Porter (MetaBrainz, MTG-UPF)
CIRMMT LinkedMusic Workshop
18 November 2022

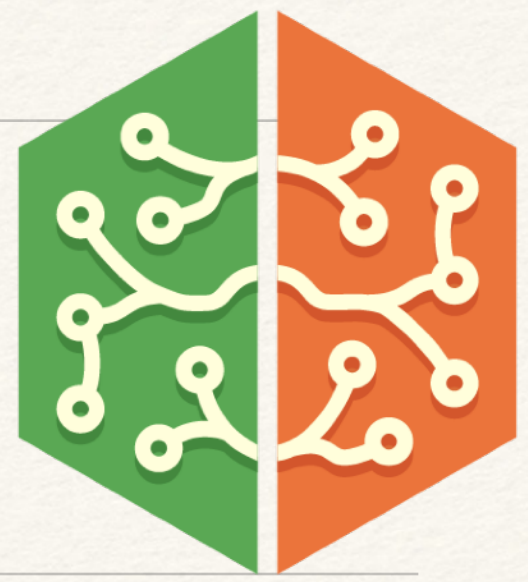
Overview

- ❖ The MetaBrainz Foundation
- ❖ MusicBrainz history and structure
- ❖ CritiqueBrainz
- ❖ AcousticBrainz
- ❖ ListenBrainz
- ❖ BookBrainz
- ❖ CoverArtArchive

A brief introduction

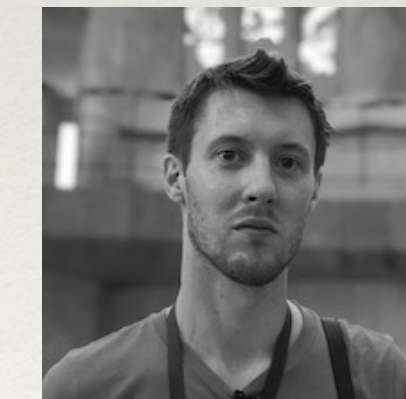
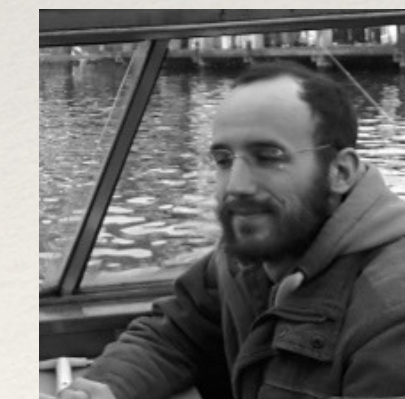
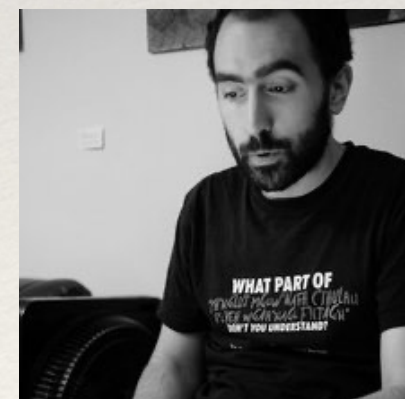
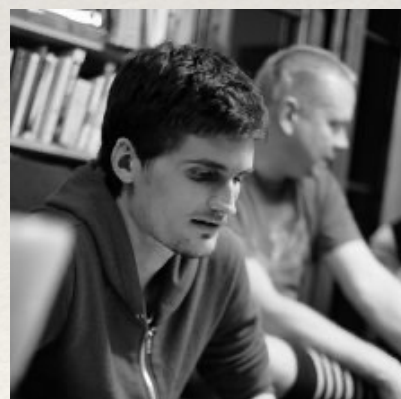
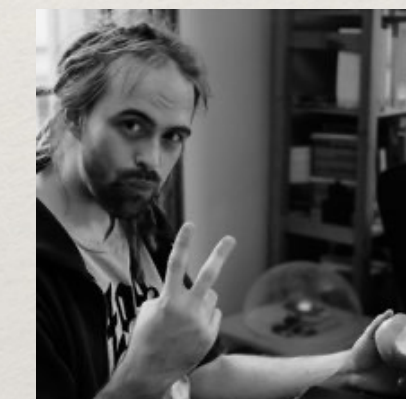
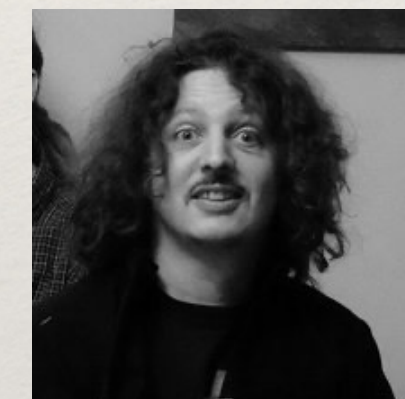
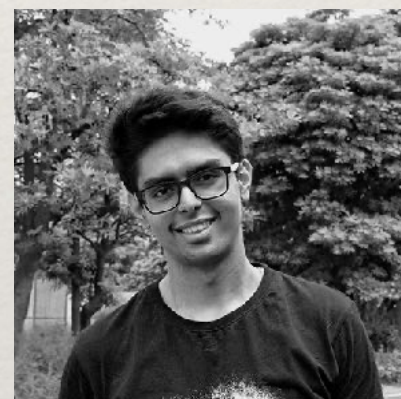
- ❖ Alastair Porter
- ❖ Researcher at Music Technology Group, Universitat Pompeu Fabra, Barcelona
- ❖ Freelance contractor with MetaBrainz foundation
- ❖ MA Music Technology (McGill) 2012
- ❖ Working with music databases for 15+ years

MetaBrainz foundation



<https://metabrainz.org>

- ❖ California non-profit, founded in 2004
- ❖ Umbrella organisation for all of the projects discussed here
- ❖ Paid team of 12, funded by license fees to the MusicBrainz database





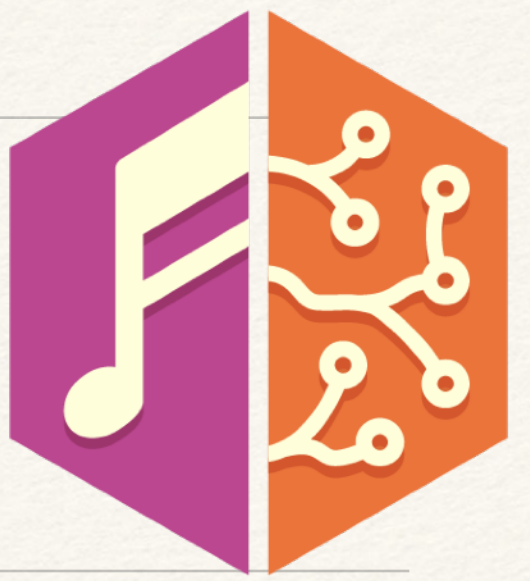
CDDDB

- ❖ Do you remember CDs?
- ❖ Put a CD into your computer and get the metadata!
 - ❖ No CDText
 - ❖ MP3s were new!
 - ❖ Artist, Album, Song
- ❖ Open database, shared to the public

The life and death of CDDDB

- ❖ The originally open CDDDB Database gets sold
 - ❖ Company gets named Gracenote
 - ❖ Eventually bought by Sony and then bounced around a few other companies
- ❖ Database gets closed and made private, locked behind a license agreement

MusicBrainz starts



<https://musicbrainz.org>

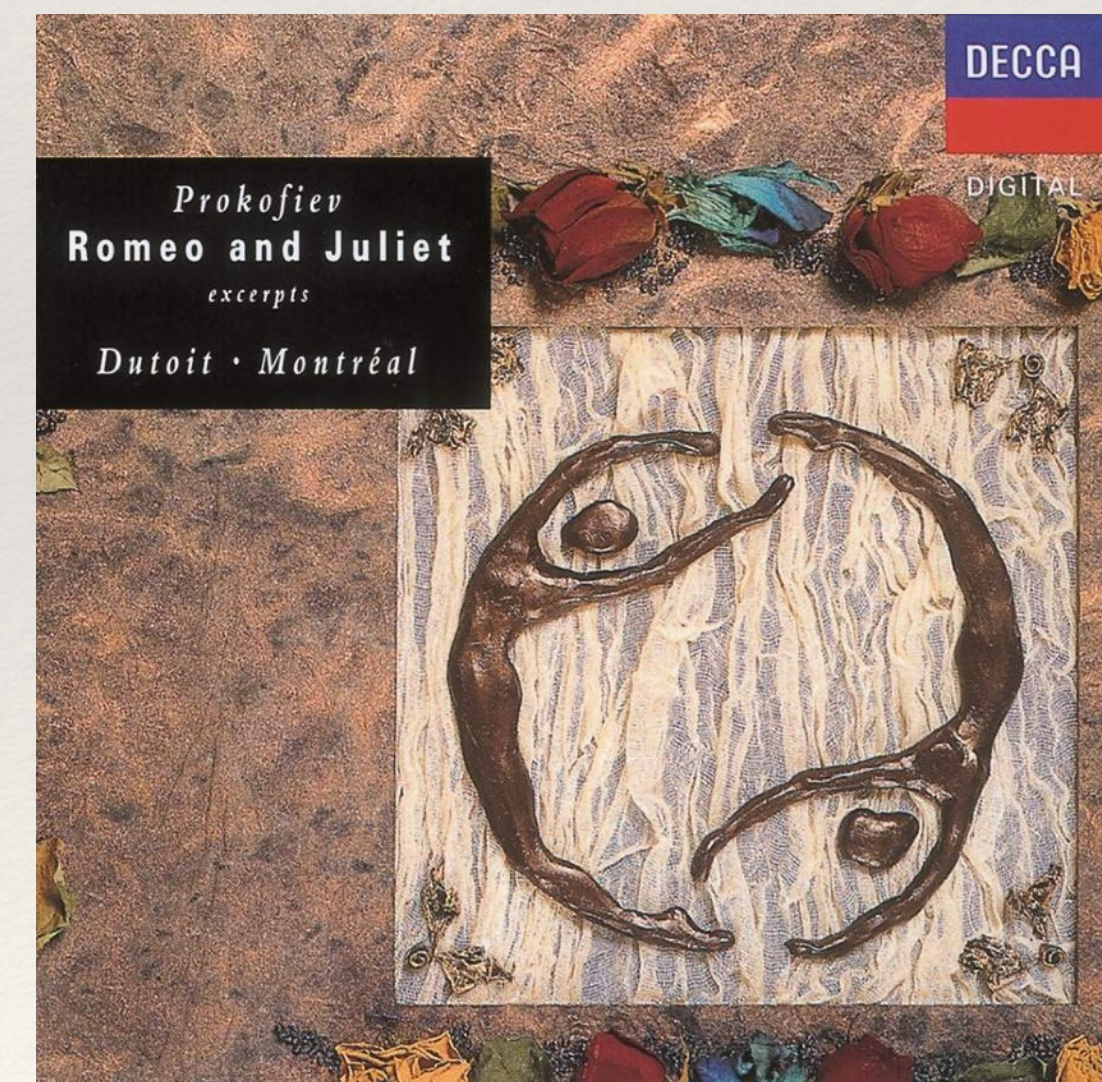
- ❖ Open source hackers bristle at the prospect of their open database being made closed
- ❖ So they make their own
- ❖ freedb, gnudb, MusicBrainz (2000)...
- ❖ <https://www.eff.org/deeplinks/2021/06/organizing-public-interest-musicbrainz>

Original MusicBrainz Schema

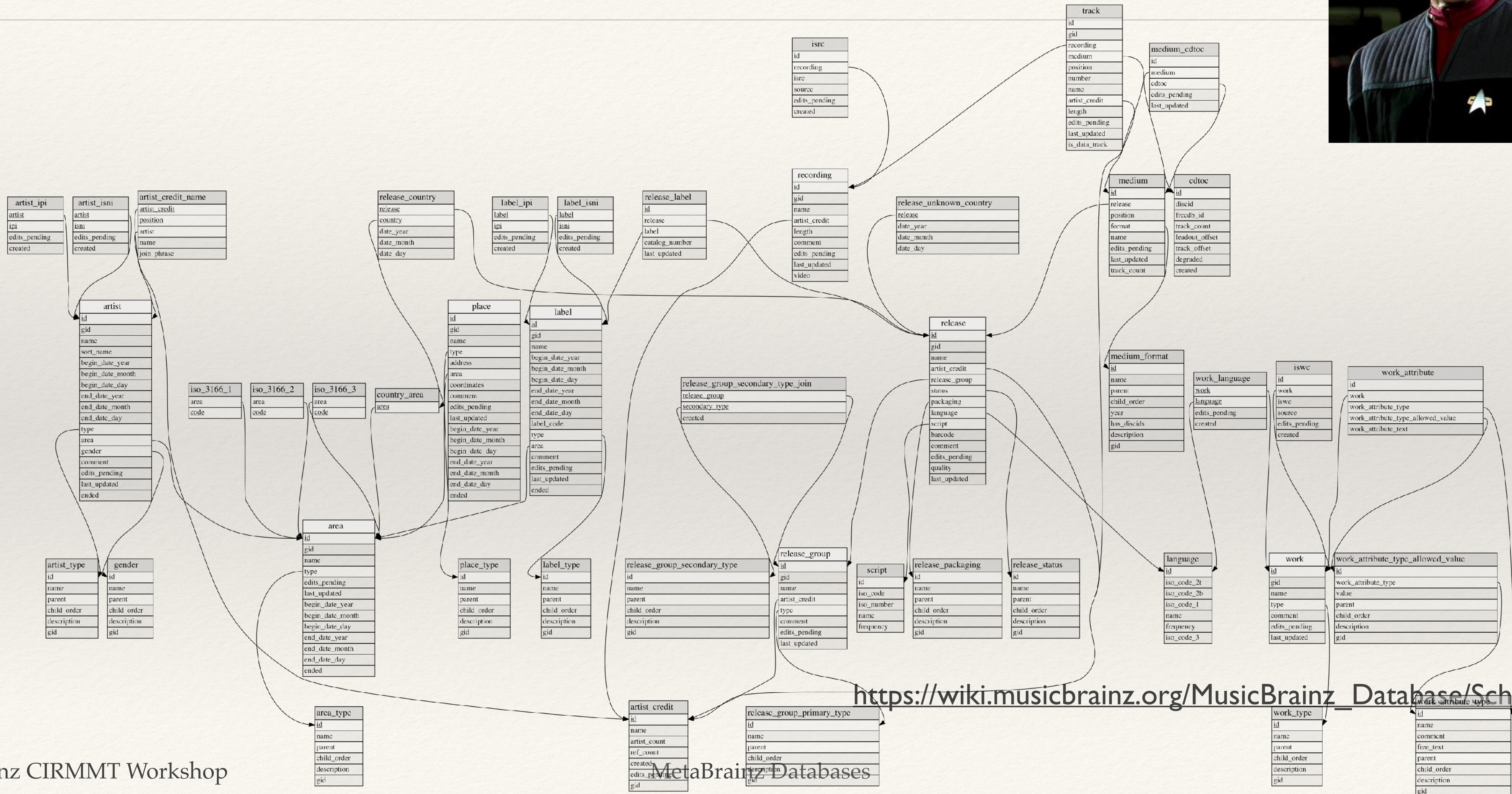
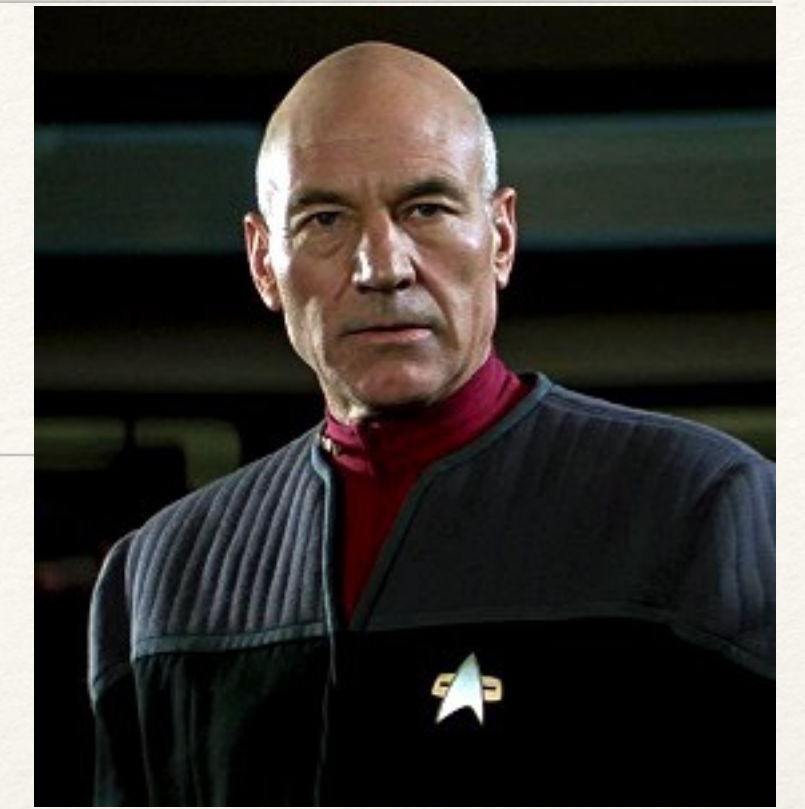
- ❖ Originally a copy of the CDDDB structure: Artist, Album, Track
- ❖ Give each entity a MBID (MusicBrainz ID), a random UUID
- ❖ Web interface to add data
- ❖ API to retrieve data

Shortcomings of the schema

- ❖ The same recording on different albums is in the database twice
- ❖ Multi-CD albums get entered as different albums
- ❖ The "Classical Music Problem"



The Next Generation Schema



https://wiki.musicbrainz.org/MusicBrainz_Database/Schema

Improvements

- ❖ Recordings can be shared between multiple albums
- ❖ Artist Credits: More than one person on an album, or with a different name as a special case
- ❖ Relationships: more detailed information between recordings and artists (and other types of entities)

Relationships

orchestra: Orchestre symphonique de Montréal (from 1989-10-04 until 1989-10-05)

conductor: Charles Dutoit (conductor) (from 1989-10-04 until 1989-10-05)

recording of: Romeo and Juliet, op. 64: Act I. Scene I. No. 1. Introduction

Related works

Romeo and Juliet, op. 64: Act I. Scene I. No. 1. Introduction

composer: Сергей Сергеевич Прокофьев (Russian composer)

part of: Romeo and Juliet, op. 64: Act I (order: 2)

<https://musicbrainz.org/recording/f07a6189-6ae1-4440-8ca9-10cb7f69f6d8>

Under Pressure
~ Recording by Queen & David Bowie

Relationships

producer: David Bowie
 Brian May (Queen guitarist)
 Queen (UK rock group)
 Roger Taylor (Queen drummer)

bass guitar and guitar: John Deacon (from 1981-06 until 1982-03)

drums (drum set): Roger Taylor (Queen drummer) (from 1981-06 until 1982-03)

electric guitar: Brian May (Queen guitarist) (from 1981-06 until 1982-03)

finger snaps and handclaps: John Deacon (from 1981-06 until 1982-03)
 Brian May (Queen guitarist) (from 1981-06 until 1982-03)
 Freddie Mercury (from 1981-06 until 1982-03)
 Roger Taylor (Queen drummer) (from 1981-06 until 1982-03)

finger snaps and handclaps: David Bowie (from 1981 until 1982-03)

keyboard: John Deacon (from 1981-06 until 1982-03)
 Freddie Mercury (from 1981-06 until 1982-03)

piano: Freddie Mercury (from 1981-06 until 1982-03)

synthesizer: David Bowie (from 1981-06 until 1982-03)

background vocals: Freddie Mercury (from 1981-06 until 1982-03)
 Roger Taylor (Queen drummer) (from 1981-06 until 1982-03)

lead vocals: David Bowie (from 1981-06 until 1982-03)
 Freddie Mercury (from 1981-06 until 1982-03)

performer: David Bowie
 Queen (UK rock group)

phonographic copyright (©) by: David R. Jones (Jones Music subsidiary) (in 1981)
 Mainman S.A. (not for release label use! © & © rights use only) (in 1981)
 Queen Productions Ltd. (company and copyright holder, do not use as an imprint or release label) (in 1981, in 2011)
 Raincloud Productions Ltd. (in 1981)

recorded at: Mountain Studio in Montreux, Vaud, Switzerland (from 1981-06 until 1982-03)
 Musicland Studios (Munich) in Munich, Bayern, Germany (from 1981-06 until 1982-03)

has remixes: Under Pressure (Knebworth mix)
 Under Pressure (Mike Spencer mix)
 Under Pressure (Rah mix)

sampled by: Come Again by The Kleptones
 Ice Ice Baby by Vanilla Ice

mash-ups: Funk Funk Baby (Daft Punk vs. Vanilla Ice vs. David Bowie & Queen) by DJ Tripp
 Immaterial and Under Pressure by Triple-Q
 Jacko Under Pressure (Michael Jackson vs. Queen & David Bowie) by Go Home Productions
 Mouth Pressure by Neil Cicierega
 Under Ice by Stark Effect
 Under Pressure, Building [French remix] (Muse vs Vanilla Ice & Queen featuring David Bowie) by Oki
 Jingle Pressure (Queen & Bowie vs. Smokey Robinson) by dj BC (in 2012)

part of: The Rolling Stone Magazine's 500 Greatest Songs of All Time: 2021 edition (number: 429) (order: 429)

stream for free: <https://open.spotify.com/track/2fuCquhmrzHpu5xcA1ci9x> [info]

recording of: Under Pressure (from 1981-06 until 1982-03)

<https://musicbrainz.org/recording/32c7e292-14f1-4080-bddf-ef852e0a4c59>

Web interface to add/edit data

The image displays three overlapping screenshots of the MusicBrainz web interface, illustrating the process of editing data for the release "The Game" by Queen.

Top Screenshot: Overview
The browser address bar shows the URL: `musicbrainz.org/release/61b54635-a8b0-45df-84f4-286a3056c960/edit`. The page title is "The Game" and it is identified as a release by Queen. The "Edit" button is highlighted in orange.

Middle Screenshot: Tracklist
The "Tracklist" tab is active, showing a table with columns for track number, track name, recording name, and duration. Each row includes an "Edit" button for the recording.

track	recording
1 Play the Game (3:31)	Play the Game (3:32)
2 Dragon Attack (4:19)	Dragon Attack (4:19)
3 Another One Bites the Dust (3:34)	Another One Bites the Dust (3:35)
4 Need Your Loving Tonight (2:48)	Need Your Loving Tonight (2:49)
5 Crazy Little Thing Called Love (2:44)	Crazy Little Thing Called Love (2:44)
6 Rock It (Prime Jive) (4:33)	Rock It (Prime Jive) (4:33)
7 Don't Try Suicide (3:53)	Don't Try Suicide (3:53)
8 Sail Away Sweet Sister (3:32)	Sail Away Sweet Sister (3:33)
9 Coming Soon (2:51)	Coming Soon (2:51)
10 Save Me (3:48)	Save Me (3:49)

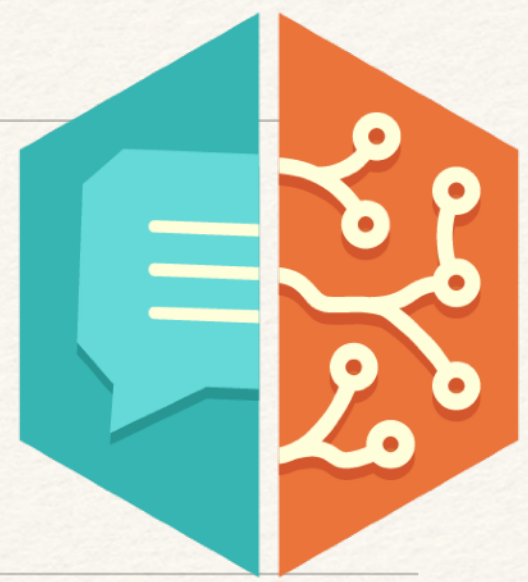
Bottom Screenshot: Relationships
The "Edit Relationships" page is shown. It includes instructions on how to edit relationships (yellow), remove them (red), or add them (green). It features a "Track Relationships" section with checkboxes for "Recording" and "Related Works".

- Recording (0 recordings selected):** 12" Vinyl 1
- Related Works (0 works selected):** Play the Game (recording of) (1980-02 - 1980-05) ×
 - composer: Freddie Mercury
 - later translated versions: Juégale, apuéstale
 - lyricist: Freddie Mercury
 - part of: We Will Rock You (Queen musical)
 - publisher: Beechwood Music Corp. (1980)
 - publisher: EMI Music Publishing Ltd.
 - publisher: Queen Music Ltd. (1980)

MusicBrainz learning points

- ❖ Use UUIDs for everything (cool URLs don't change)
- ❖ Be able to retrieve all data via APIs (XML, JSON, JSON-LD)
- ❖ Always use MBIDs when referring to items in the API response
 - ❖ This allows us to make changes/restructure the database while keeping the public IDs stable (*Andrew: We reserve the right to change the structure of our DB)
- ❖ Always keep the data open and freely available
- ❖ The Community is key to the project (we have a team of 12 going on 20,000)

CritiqueBrainz



<https://critiquebrainz.org>

- ❖ Wikipedia rules say that content has to be from a Neutral point of View
- ❖ So, how do you write about a new album without being a music critic?
- ❖ CC licensed, write whatever you want
- ❖ Initially only albums, but now all "core" MB entities
- ❖ Refer to items in MusicBrainz, use MBIDs
- ❖ API to read (and write)
- ❖ Interface to write reviews directly from MusicBrainz, ListenBrainz, BookBrainz
- ❖ 10,500 reviews (~400-500/year)

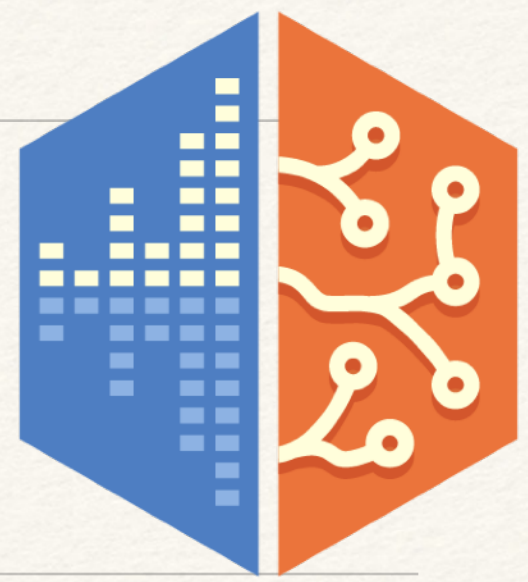
A screenshot of the CritiqueBrainz website. The browser address bar shows 'critiquebrainz.org'. The website has a navigation bar with 'Sign in', 'Reviews', and 'Statistics' links, along with search and artist filters. The main content area features a 'Welcome to CritiqueBrainz!' message, a brief description of the site, and a 'Statistics' table. Below this is a 'Popular reviews' section with several album reviews, each with a cover image and a short description.

Statistics	
Description	Number
Number of Reviews	10,399
Number of Users	2,722

Popular reviews

- RAYFORCE by ZUNTATA**
Review by UCHUBomb
One of the most unique shmup OST's for one of the most unique arcade shmups ever made, and arguably Tamayo Kawamoto's...
- Spaceball: The American Radio Sessions by ...**
Review by Jon Lusk
Subtitled The American Radio Sessions, this double-disc set collects largely acoustic versions of classic T. Rex songs performed in...
- Kamaal the Abstract by Q-Tip**
Review by Stevie Chick
That Q-Tip spent his own money to purchase the rights to this previously-shelved album suggests how close it is to his heart. That it...
- Odd and Awkward by Nils Wogram**
Review by Peter Marsh
Trombonist Nils Wogram is pretty much unknown over here but has made a name for himself on the Eurojazz circuit, and this...
- Summer Days (and Summer Nights!!) by The ...**
Review by Chris Roberts
Released just four months after Today! in July 1965, The Beach Boys' ninth album was at first deemed by some to be a regression. Its...
- Bongo Bong**
Review by rob
This track has long been the stand-out track that exemplifies Manu Chao for me -- I've been loving this song ever since it came out ...

AcousticBrainz



<https://acousticbrainz.org>

- ❖ The Echo Nest (2005) builds a huge database of music metadata, semantic graph, and automatically extracted features (MFCCs, beats, danceability, popularity)
- ❖ 2014: Spotify buys the Echo Nest
- ❖ Open source hackers bristle at the prospect of a database being closed down
 - ❖ Rob comes to my office and says "hey, so..."
 - ❖ Prototype built in an afternoon

Recording "You Look Like I Need a Drink" by Against Me!

See Similar

Metadata	value
MBID	4ad2ca2d-5a12-4142-8371-e402b3401be8
title	You Look Like I Need a Drink
artist	Against Me!
release	As the Eternal Cowboy
track number	9 / 11
track length	2:14

Submission #1 out of 18 →

Low-level information [Summary](#)

Tonal & Rhythm	value
key	A# minor (46.0%)
chords key	D minor
danceability	1.04
bpm	143
beat count	302

High-level information

Voice, timbre, gender, etc.	value	probability
Voice	Instrumental	100%

AcousticBrainz

- ❖ Collaboration with UPF
- ❖ Uses Essentia feature extractor (<https://essentia.upf.edu/>)
- ❖ Audio descriptors (MFCCs), Tonal features (key, chords), Rhythmic Features (bpm)
- ❖ Support Vector Machines to compute features using common MIR tasks
 - ❖ e.g. Using MFCC, rhythmic summaries or other computed features
 - ❖ Genre, Mood, Instrumentation

AcousticBrainz

- ❖ Some problems with the data
 - ❖ Uncertainty about the accuracy of some features (30m submissions, which ones are correct?)
 - ❖ Lower quality ML models (Deep Learning got big about the time we released, summary MFCCs became less useful)
 - ❖ We hoped that we would encourage development and collaboration, but it didn't eventuate
- ❖ Shut down at the beginning of 2022
- ❖ Data still available for download

ListenBrainz



<https://listenbrainz.org>

- ❖ last.fm (2002) pioneered a way to make a record of the songs you are listening to, and a community of users passionate about discussing their music
- ❖ In 2007 it was bought by CBS
- ❖ Open source hackers bristle at the prospect of their database disappearing
- ❖ Over a weekend in September 2015 a bunch of us get together and built a prototype

Core aspects of ListenBrainz

- ❖ A "listen", your record of listening to a song
- ❖ Submit to the API
- ❖ Give us a Recording MusicBrainz ID if you have one
- ❖ If you don't, we generate a "MessyBrainz" id, based on the metadata you send
 - ❖ And then match it to a MusicBrainz Recording
- ❖ Import from Spotify, or send data from a music player on your computer

User "alastairp" - ListenBrainz x +

listenbrainz.org/user/alastairp/

ListenBrainz Home Feed Recent Data Explore About

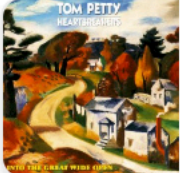


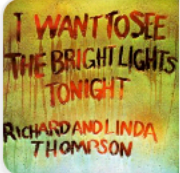



Search Listener alastairp

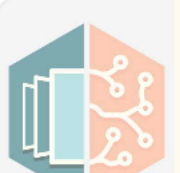
See profile on MusicBrainz

alastairp

Listens Reports Charts Love/Hate Pins Playlists Recommendations Collaborations

Recent listens

	Learning To Fly 4:02 Tom Petty and the Heartbreakers	5 minutes ago	♡ ♡ ⋮
	Roll Um Easy 2:31 Little Feat	9 minutes ago	♡ ♡ ⋮
	Get It On 4:22 T. Rex	12 minutes ago	♡ ♡ ⋮
	I Want To See The Bright Lights Tonight 3:07 Richard & Linda Thompson	16 minutes ago	♡ ♡ ⋮
	Blue Red And Grey 2:46 The Who	19 minutes ago	♡ ♡ ⋮
	Hazy Jane II 3:46 Nick Drake	22 minutes ago	♡ ♡ ⋮
	Heaven on the Beach with You 2:02 Robert Lester Folsom	26 minutes ago	♡ ♡ ⋮



Let Me Roll It - 2010 Remaster 4:49
Wings 🎵 Listening now —

You have listened to

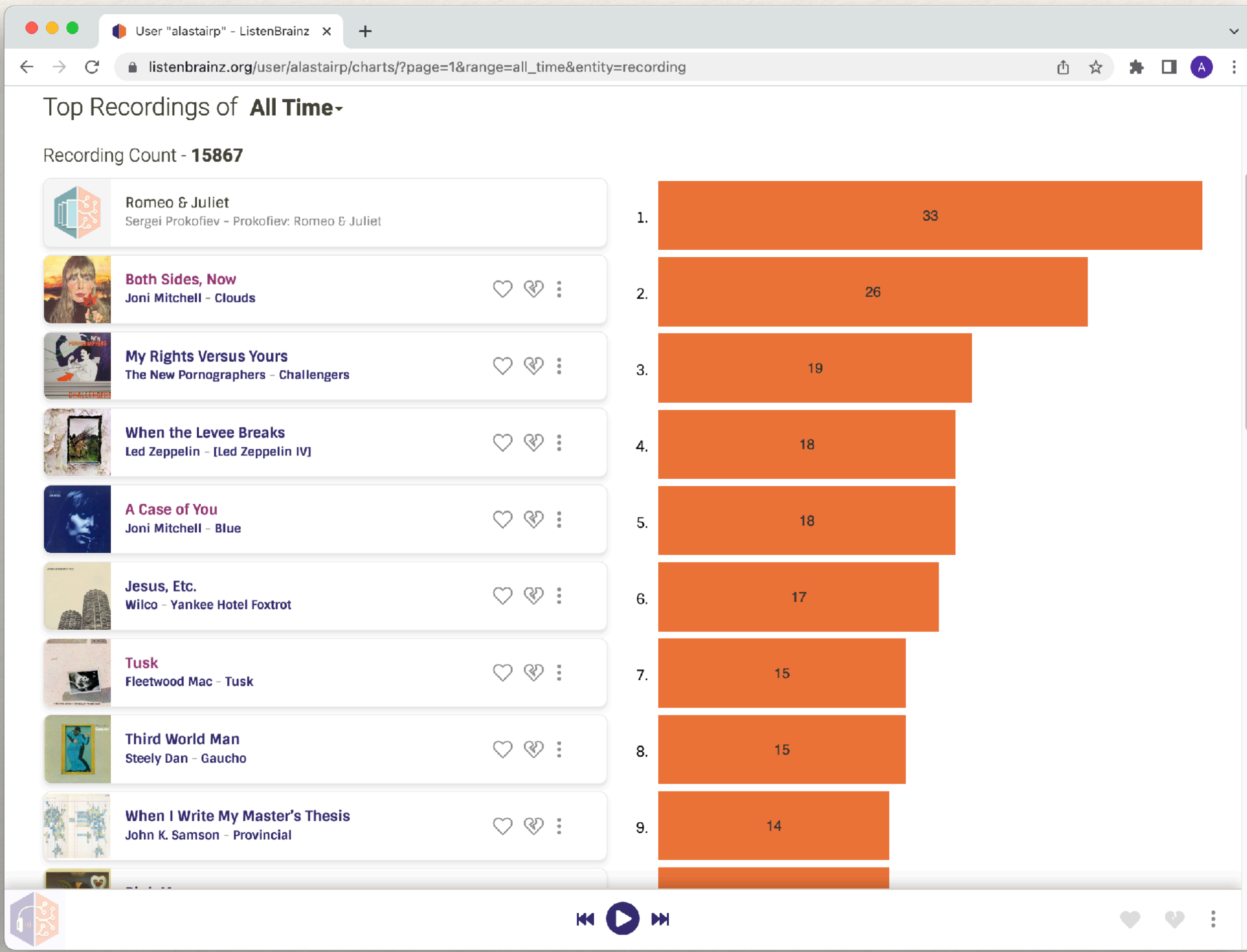
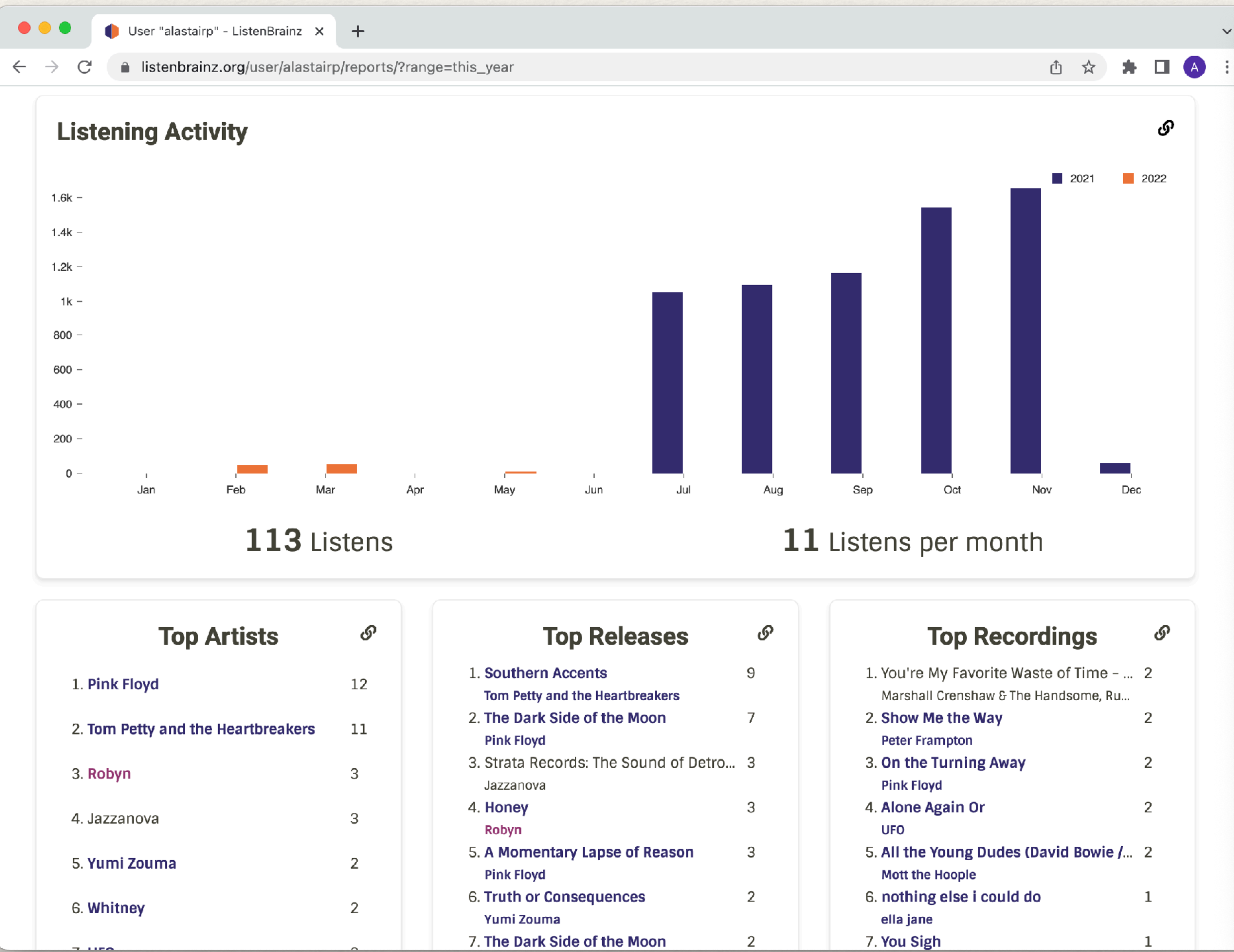
25,297
songs so far

Followers (6) **Following (5)**

mr_monkey	Following
iliekcomputers	Following
troi-bot	Following

⏮ ⏪ ⏩ ⏭ ♡ ♡ ⋮

❖ Incoming stream of songs



❖ Statistics

- ❖ What you've been listening to
- ❖ When
- ❖ How much
- ❖ Where it's from

- ❖ 600,000,000 total submissions
- ❖ 400,000 per day (mostly from Spotify)

❖ Recommendations

- ❖ Collaborative filtering based on listening history in LB
- ❖ Make a Playlist
- ❖ Listen on LB or export to Spotify
- ❖ API to create playlists, others can contribute

The screenshot shows a web browser window with the URL `listenbrainz.org/user/alastairp/recommendations/`. The page header includes the ListenBrainz logo and navigation links: Home, Feed, Recent, Data, Explore, and About. A search bar and the user name 'alastairp' are visible. Below the header, the user's name 'alastairp' is displayed, along with a link to 'See profile on MusicBrainz'. A horizontal menu contains tabs for Listens, Reports, Charts, Love/Hate, Pins, Playlists, Recommendations (which is active), and Collaborations. The main content area is titled 'Recommendation playlists created for alastairp' and includes a note: 'These playlists are ephemeral and will only be available for a month. Be sure to save the ones you like to your own playlists !'. Below this, there is a grid of 15 recommendation playlist cards. Each card displays the playlist title, the date and day of creation, and a 'Save' button. The cards are arranged in three rows: the first row has 5 cards, the second row has 5 cards, and the third row has 5 cards. The titles of the cards include 'Top 100 similar artists recommendations from model...', 'Daily Jams for alastairp, 2022-11-18 Fri', and 'Top 100 top artists recommendations from model...'. The 'Created' and 'Last Modified' dates are consistently 17/11/2022, 16/11/2022, and 15/11/2022 respectively for the different days.

Playlist Title	Created	Last Modified
Top 100 similar artists recommendations from model listenbrainz-recommendation-model-0cc49569-145d-47a3-9b70-	17/11/2022	17/11/2022
Top 100 top artists recommendations from model listenbrainz-recommendation-model-0cc49569-145d-47a3-9b70-	17/11/2022	17/11/2022
Daily Jams for alastairp, 2022-11-18 Fri	17/11/2022	17/11/2022
Top 100 similar artists recommendations from model listenbrainz-recommendation-model-135d5e45-ad90-40b6-94e0-	17/11/2022	17/11/2022
Top 100 top artists recommendations from model listenbrainz-recommendation-model-135d5e45-ad90-40b6-94e0-	17/11/2022	17/11/2022
Daily Jams for alastairp, 2022-11-17 Thu	16/11/2022	16/11/2022
Top 100 similar artists recommendations from model listenbrainz-recommendation-model-60cbb211-f0dd-4f23-9a33-	16/11/2022	16/11/2022
Top 100 top artists recommendations from model listenbrainz-recommendation-model-60cbb211-f0dd-4f23-9a33-	16/11/2022	16/11/2022
Daily Jams for alastairp, 2022-11-16 Wed	15/11/2022	15/11/2022
Top 100 similar artists recommendations from model listenbrainz-recommendation-model-6e387505-23c2-4e00-b54f-	15/11/2022	15/11/2022
Top 100 top artists recommendations from model		
Daily Jams for alastairp, 2022-11-15 Tue		
Daily Jams for alastairp, 2022-11-14 Mon		
Top 100 similar artists recommendations from model		
Top 100 top artists recommendations from model		

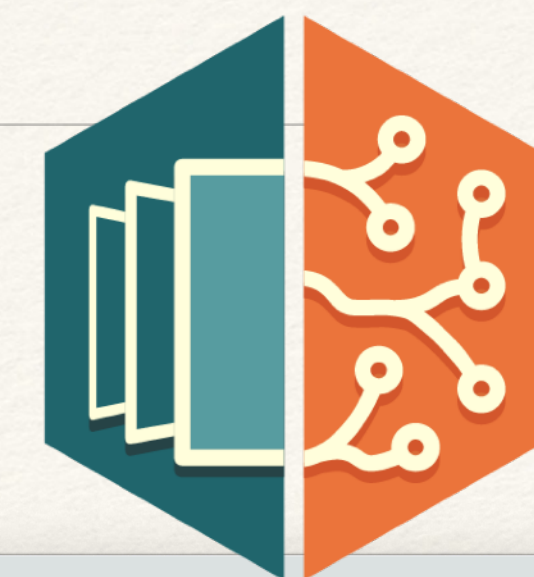
BookBrainz



<https://bookbrainz.org>

- ❖ What if you had MusicBrainz, but for books instead?
- ❖ Originally a community-driven project, now an official part of MetaBrainz
- ❖ Deals with entities such as authors, works, editions, series, publisher
- ❖ Like MusicBrainz, all entities have a BBID
- ❖ Discogs used to have a partner project *Bookogs*, but it was closed at the end of 2020
 - ❖ Much of the community moved to BookBrainz
- ❖ Write reviews about these things on CritiqueBrainz!
- ❖ Work: 37220, Author: 14635, Edition: 7286

CAA



- ❖ <https://coverartarchive.org/>
- ❖ Store images at the Internet Archive
- ❖ Direct integration to MusicBrainz to add images
- ❖ API gives direct URLs to download

The screenshot shows the ListenBrainz website interface. At the top, there's a navigation bar with 'ListenBrainz' logo and links for 'Home', 'Feed', 'Recent', 'Data', 'Explore', and 'About'. A search bar is on the right. The main content area is titled 'Huesound^{beta}' and features a grid of 20 album covers. To the left of the grid is a color wheel with various color segments. At the bottom, a video player is visible, showing a video titled 'Curmudgeon - Flaccid' with a progress bar at 0:04 / 2:05.

Thanks

<https://metabrainz.org>

<https://musicbrainz.org>

<https://twitter.com/musicbrainz>

