

Constituer une Base de Données du Rap Français

Arnault Chatelain

CREST, Ecole Polytechnique

12/12/2023

Table des Matières

Introduction

Une Base de Données du Rap Français

Statistiques Descriptives

Table des Matières

Introduction

Une Base de Données du Rap Français

Statistiques Descriptives

Contexte

Cette présentation est une version éditée d'une communication lors de la journée d'étude "Les territoires du rap: Hiérarchies, autorités, rapport de pouvoir" organisée 12/12/2023 à Rennes, par B. Lechaux, A. Henneguella et A. Allemand

Avertissement

Si constituer cette base est un exercice de collecte de traces numériques intéressant, son utilisation dans le cadre d'un travail de recherche scientifique nécessite quelques précautions:

- ▶ Cette base n'est pas exhaustive (surtout dans les premières années)
- ▶ Cette base n'est pas une sous-population tirée au hasard dans la population des rappers/rappeuses. Tout travail d'inférence statistiques se doit de le reconnaître
- ▶ Il existe déjà une littérature scientifique sur le rap français à la fois en sciences sociales et en humanités qu'il peut être judicieux de consulter en amont de tout travail

Table des Matières

Introduction

Une Base de Données du Rap Français

Statistiques Descriptives

Créer la Base

Point de départ: Genius

Créer la Base

Point de départ: Genius → Genius France: Discographie du Rap français



The screenshot shows a yellow-themed page on the Genius website. On the left, there is a large 'G' logo and the text 'GENIUS Discographie du rap'. On the right, the title 'Discographie rap - 1999' is displayed, followed by a breadcrumb trail: 'Genius France • Track 18 on Discographie du Rap français'. Below the title, there is a paragraph of text: 'Discographie – Année 1999 Cette page a pour but de recenser au maximum les albums, EPs, mixtapes et maxis de rap francophone sortis et prévus pour l'année 1999. Pour bien... [Read More](#)'. At the bottom of the text block, it shows 'Jan. 1, 1999', '1 Viewer', and '8.8K Views'.

Discographie rap - 1999 Lyrics  [19 Contributors](#)

Janvier :

- 08/01 : Les X-Men - *Jeunes, coupables et libres*
- * 12/01 : Beedjy - *Pile ou femme*
- * 18/01 : Malédiction du Nord - *Les raisons de la colère*
- * 26/01 : Kent-zo - *L'Olympe*

Février :

- * 02/02 : Zoxea - *A mon tour d'briller*
- * 08/02 : 3ème Oeil - *Hier, Aujourd'hui, Demain*
- * 17/02 : Diam's - *Premier Mandat*
- * 17/02 : Faouzi Tarkhani - *Guerrier pour la paix*
- * 23/02 : Bisso Na Bisso - *Racines*
- * XX/02 : D Abuz System - *Le syndikat*

Créer la Base

Point de départ: Genius → Genius France: Discographie du Rap français

Janvier :

- 08/01 : Les X-Men - *Jeunes, coupables et libres*
- * 12/01 : Beedjy - *Pile ou femme*
- * 18/01 : Malédiction du Nord - *Les raisons de la colère*
- * 26/01 : Kent-zo - *L'Olympe*

Février :

- * 02/02 : Zoxea - *A mon tour d'briller*
- * 08/02 : 3ème Oeil - *Hier, Aujourd'hui, Demain*
- * 17/02 : Diam's - *Premier Mandat*
- * 17/02 : Faouzi Tarkhani - *Guerrier pour la paix*
- * 23/02 : Bisso Na Bisso - *Racines*
- * XX/02 : D Abuz System - *Le syndikat*

Mars :

- * 05/03 : Prodiges Namor - *L'heure de vérité*
- * 16/03 : Artistes multiples - *Compilation Delabel Hostile Rap*
- * 18/03 : Artistes multiples - *Compilation "Première classe vol.1"*
- 22/03 : Freeman - *Combien j'ai ramé (Maxi)*
- 23/03 : Freeman - *L'Palais de Justice*
- 29/03 : Fonky Family - *Hors Série Volume 1*
- * 30/03 : Artistes multiples - *Face cachée de Mars*

Avril :

- * 09/04 : 3ème type - *Ovniprésent*
- * 10/04 : Kamnouze - *La technique du globule noir*
- * 17/04 : Artistes Multiples - *15 Balles Perdues*

Genius Annotation

1 contributeur



Jeunes, coupables et libres

Morceaux :

- 01 : [Intro](#)
- 02 : [Dieu a béni mon clan](#)
- 03 : [Score](#)
- 04 : [One One One](#)
- 05 : [Bla Bla Bla](#)
- 06 : [Jeunes, coupables et libres](#)
- 07 : [Même les meufs escroquent](#)
- 08 : [Casse la baraque](#)
- 09 : [Laisse les bons briller \(feat. Celsius, Watchos & Kamal\)](#)

A partir des listes de discographie (1982-2025):

1. Récupérer tous les liens d'artistes (et groupes), albums et chansons accessibles sur les pages

A partir des listes de discographie (1982-2025):

1. Récupérer tous les liens d'artistes (et groupes), albums et chansons accessibles sur les pages
2. Pour les albums et chansons: retrouver les artistes qui les chantent

A partir des listes de discographie (1982-2025):

1. Récupérer tous les liens d'artistes (et groupes), albums et chansons accessibles sur les pages
2. Pour les albums et chansons: retrouver les artistes qui les chantent

5 733 pages uniques d'artistes retrouvées sur Genius,

A partir des listes de discographie (1982-2025):

1. Récupérer tous les liens d'artistes (et groupes), albums et chansons accessibles sur les pages
2. Pour les albums et chansons: retrouver les artistes qui les chantent

5 733 pages uniques d'artistes retrouvées sur Genius,
dont 3 106 avec une description

A partir des listes de discographie (1982-2025):

1. Récupérer tous les liens d'artistes (et groupes), albums et chansons accessibles sur les pages
2. Pour les albums et chansons: retrouver les artistes qui les chantent

5 733 pages uniques d'artistes retrouvées sur Genius,

dont 3 106 avec une description

dont 1 507 contenant les suites de caractères "rap fr, french rap, rapeu, rappeu"

A partir des listes de discographie (1982-2025):

1. Récupérer tous les liens d'artistes (et groupes), albums et chansons accessibles sur les pages
2. Pour les albums et chansons: retrouver les artistes qui les chantent

5 733 pages uniques d'artistes retrouvées sur Genius,
dont 3 106 avec une description
dont 1 507 contenant les suites de caractères "rap fr, french rap,
rapeu, rappeu"

La base de données se concentrera sur ces 1507 artistes

Pour ces artistes:

3. Télécharger toutes les paroles de chansons disponibles sur Genius pour chaque artiste (au 08/12/2023)

Pour ces artistes:

3. Télécharger toutes les paroles de chansons disponibles sur Genius pour chaque artiste (au 08/12/2023)
→ Attribuer les featuring à l'artiste principal(e)

Pour ces artistes:

3. Télécharger toutes les paroles de chansons disponibles sur Genius pour chaque artiste (au 08/12/2023)
→ Attribuer les featuring à l'artiste principal(e)

62 059 chansons (inclus freestyle) retrouvées

Pour ces artistes:

3. Télécharger toutes les paroles de chansons disponibles sur Genius pour chaque artiste (au 08/12/2023)
→ Attribuer les featuring à l'artiste principal(e)

62 059 chansons (inclus freestyle) retrouvées

4. Utiliser les liens des pages Genius, Facebook, Twitter et Instagram (disponibles sur Genius) pour les retrouver sur Musicbrainz [Musicbrainz ?](#)

Pour ces artistes:

3. Télécharger toutes les paroles de chansons disponibles sur Genius pour chaque artiste (au 08/12/2023)
→ Attribuer les featuring à l'artiste principal(e)

62 059 chansons (inclus freestyle) retrouvées

4. Utiliser les liens des pages Genius, Facebook, Twitter et Instagram (disponibles sur Genius) pour les retrouver sur Musicbrainz [Musicbrainz ?](#)

423 artistes retrouvés sur Musicbrainz

Tout le processus de constitution de la base est “automatisé”:

- ▶ De nouveaux liens sur les pages discographies du rap français

Tout le processus de constitution de la base est “automatisé”:

- ▶ De nouveaux liens sur les pages discographies du rap français
→ 3h pour ré-extraire tous les liens

Tout le processus de constitution de la base est “automatisé”:

- ▶ De nouveaux liens sur les pages discographies du rap français
→ 3h pour ré-extraire tous les liens
- ▶ De nouveaux artistes ou de nouvelles chansons pour certains artistes sur Genius

Tout le processus de constitution de la base est “automatisé”:

- ▶ De nouveaux liens sur les pages discographies du rap français
→ 3h pour ré-extraire tous les liens
- ▶ De nouveaux artistes ou de nouvelles chansons pour certains artistes sur Genius → Variable de 1h à 10-11h

Tout le processus de constitution de la base est “automatisé”:

- ▶ De nouveaux liens sur les pages discographies du rap français
→ 3h pour ré-extraire tous les liens
- ▶ De nouveaux artistes ou de nouvelles chansons pour certains artistes sur Genius → Variable de 1h à 10-11h
- ▶ Mise à jour de la base MusicBrainz (plus d'informations, nouveaux artistes, nouveaux liens)

Tout le processus de constitution de la base est “automatisé”:

- ▶ De nouveaux liens sur les pages discographies du rap français → 3h pour ré-extraire tous les liens
- ▶ De nouveaux artistes ou de nouvelles chansons pour certains artistes sur Genius → Variable de 1h à 10-11h
- ▶ Mise à jour de la base MusicBrainz (plus d'informations, nouveaux artistes, nouveaux liens) → Immédiat mais il me faudrait 2h pour retélécharger la base et remplacer les bons fichiers

Table des Matières

Introduction

Une Base de Données du Rap Français

Statistiques Descriptives

Sur Genius

Décomposition de l'attribution de chansons aux artistes

| Artiste principal(e) | | |
|----------------------|--------|-------|
| ≥ 1 fois | Jamais | Total |
| 1 428 | 79 | 1 507 |
| % 94.7 | 5.3 | 100 |

Décomposition de l'attribution de chansons aux artistes

| Artiste principal(e) | | |
|----------------------|--------|-------|
| ≥ 1 fois | Jamais | Total |
| 1 428 | 79 | 1 507 |
| % 94.7 | 5.3 | 100 |

Qui sont ces artistes sans morceaux?

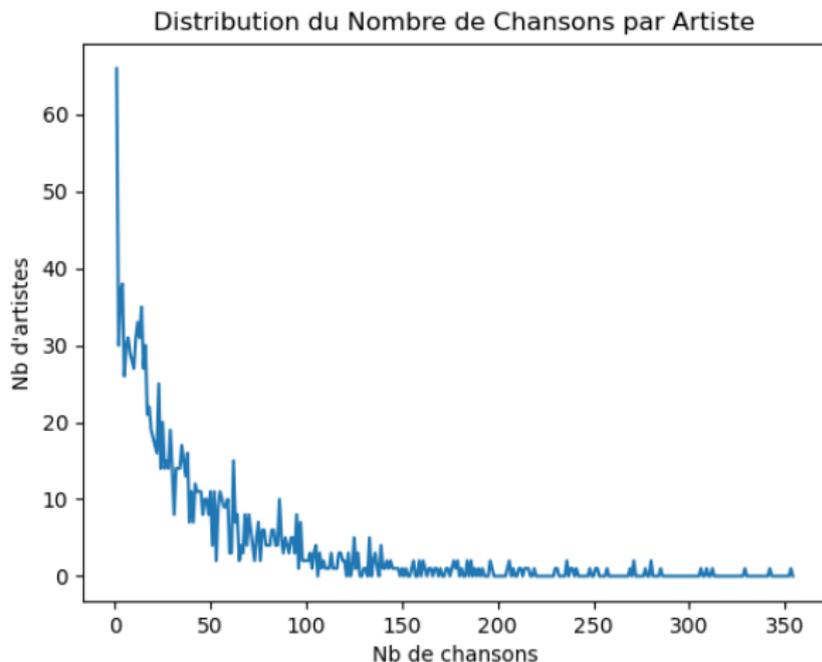
- ▶ Chansons documentés mais pas les paroles (e.g. BRK (FRA))
- ▶ Artiste membre de groupe et/ou participant à des featuring mais sans morceau solo (e.g. Fonky Flav (1995))

Une Distribution Chansons/Artistes Asymétrique

| N | Moy. | Ecart-type | Min | 25% | Médiane | 75% | Max |
|-------|------|------------|-----|-----|---------|-----|-----|
| 1 428 | 44,8 | 56,2 | 1 | 11 | 26 | 59 | 920 |

Une Distribution Chansons/Artistes Asymétrique

| N | Moy. | Ecart-type | Min | 25% | Médiane | 75% | Max |
|-------|------|------------|-----|-----|---------|-----|-----|
| 1 428 | 44,8 | 56,2 | 1 | 11 | 26 | 59 | 920 |



Informations Biographiques

423/1 507= **28%** des artistes retrouvé(e)s sur MusicBrainz.

Type d'Artiste et Genre

| Type / Genre | F | H | N/A | Total |
|--------------|----|-----|-----|-------|
| Solo | 15 | 337 | 9 | 361 |
| % | 4 | 93 | 2 | 100 |
| Groupe | | | | 46 |
| N/A | | | | 16 |
| Total | | | | 423 |

Informations Biographiques

423/1 507 = 28% des artistes retrouvé(e)s sur MusicBrainz.

Type d'Artiste et Genre

| Type / Genre | F | H | N/A | Total |
|--------------|----|-----|-----|-------|
| Solo | 15 | 337 | 9 | 361 |
| % | 4 | 93 | 2 | 100 |
| Groupe | | | | 46 |
| N/A | | | | 16 |
| Total | | | | 423 |

MusicBrainz propose aussi:

Informations Biographiques

423/1 507 = 28% des artistes retrouvé(e)s sur MusicBrainz.

Type d'Artiste et Genre

| Type / Genre | F | H | N/A | Total |
|--------------|----|-----|-----|-------|
| Solo | 15 | 337 | 9 | 361 |
| % | 4 | 93 | 2 | 100 |
| Groupe | | | | 46 |
| N/A | | | | 16 |
| Total | | | | 423 |

MusicBrainz propose aussi:

- ▶ Les dates de naissance (ou de formation) des groupes

Informations Biographiques

423/1 507 = 28% des artistes retrouvé(e)s sur MusicBrainz.

Type d'Artiste et Genre

| Type / Genre | F | H | N/A | Total |
|--------------|----|-----|-----|-------|
| Solo | 15 | 337 | 9 | 361 |
| % | 4 | 93 | 2 | 100 |
| Groupe | | | | 46 |
| N/A | | | | 16 |
| Total | | | | 423 |

MusicBrainz propose aussi:

- ▶ Les dates de naissance (ou de formation) des groupes
- ▶ Leur lieu de naissance et d'activité

Représentativité?

Difficile d'évaluer à quel point la base est complète

Représentativité?

Difficile d'évaluer à quel point la base est complète

- ▶ Des données parcellaires jusqu'en 2010

Représentativité?

Difficile d'évaluer à quel point la base est complète

- ▶ Des données parcellaires jusqu'en 2010
- ▶ Qui sont les contributeurs de Genius? Quel biais dans la représentativité cela introduit-il?

Représentativité?

Difficile d'évaluer à quel point la base est complète

- ▶ Des données parcellaires jusqu'en 2010
- ▶ Qui sont les contributeurs de Genius? Quel biais dans la représentativité cela introduit-il?

Quant à compléter la base

- ▶ Ce qui est automatisable: transcription des paroles (Whisper, OpenAI), extraction des métadonnées à partir d'un album

Représentativité?

Difficile d'évaluer à quel point la base est complète

- ▶ Des données parcellaires jusqu'en 2010
- ▶ Qui sont les contributeurs de Genius? Quel biais dans la représentativité cela introduit-il?

Quant à compléter la base

- ▶ Ce qui est automatisable: transcription des paroles (Whisper, OpenAI), extraction des métadonnées à partir d'un album
- ▶ Ce qui ne l'est pas: trouver les albums (notamment les albums des années 1990)

Conclusion

En conclusion je souhaiterais partager quelques questions que je recommande à chacun de se poser en travaillant cette base:

- ▶ Une analyse de contenu qui ignore aussi bien l'histoire que les enjeux spécifiques de domination et d'altérisation du rap est-elle pertinente?
- ▶ Pourquoi vouloir étudier le rap en particulier?

“MusicBrainz est une encyclopédie musicale libre qui collecte des métadonnées musicales et les propose au public.”



Laylow (French hip-hop)
Personne

[Vue d'ensemble](#) [Parutions](#) [Enregistrements](#) [Œuvres](#) [Événements](#) [Relations](#) [Allas](#) [Tags](#) [Critiques](#) [Détails](#) [Modifier](#)

Nom d'interprète de : [Jey Anderson](#)

Wikipédia

Laylow, né Jérémy Larroux le 12 mars 1993 à Toulouse, est un rappeur et chanteur français.

Il a sorti cinq EP et deux albums studio. Sa musique se caractérise par sa dimension futuriste.

[Lire la suite sur Wikipédia...](#) Le contenu Wikipédia est fourni selon les modalités de la licence [Creative Commons BY-SA](#)

Discographie

[Filtrer](#)

Album

| Année | Titre | Artiste | Évaluation | Parutions |
|-------|--|------------------------|------------|-----------|
| 2020 | TRINITY | Laylow | ***** | 2 |
| 2021 | L'Étrange Histoire de Mr. Anderson | Laylow | ***** | 2 |

Single

| Année | Titre | Artiste | Évaluation | Parutions |
|-------|------------------------|-------------------------------------|------------|-----------|
| 2020 | Démons | Joanna feat. Laylow | ***** | 2 |
| 2022 | Stars | Doums feat. Laylow | ***** | 1 |

EP

Informations sur l'artiste

Type : Personne
Genre : Homme
Date de naissance : 1993-03-12 (Il y a 30 ans)
Région : France

Évaluation

Tags

Genres
(aucun)

Autres tags
(aucune)

[Voir tous les tags](#)

Liens externes

- [@Jey.laylow](#)
- [@JeyLaylow](#)
- Diffusion en flux [Deezer](#)
- Diffusion en flux [Spotify](#)
- [Discogs](#)
- Wikidata: [Q96386530](#)
- [YouTube](#)

[Voir toutes les relations](#)