

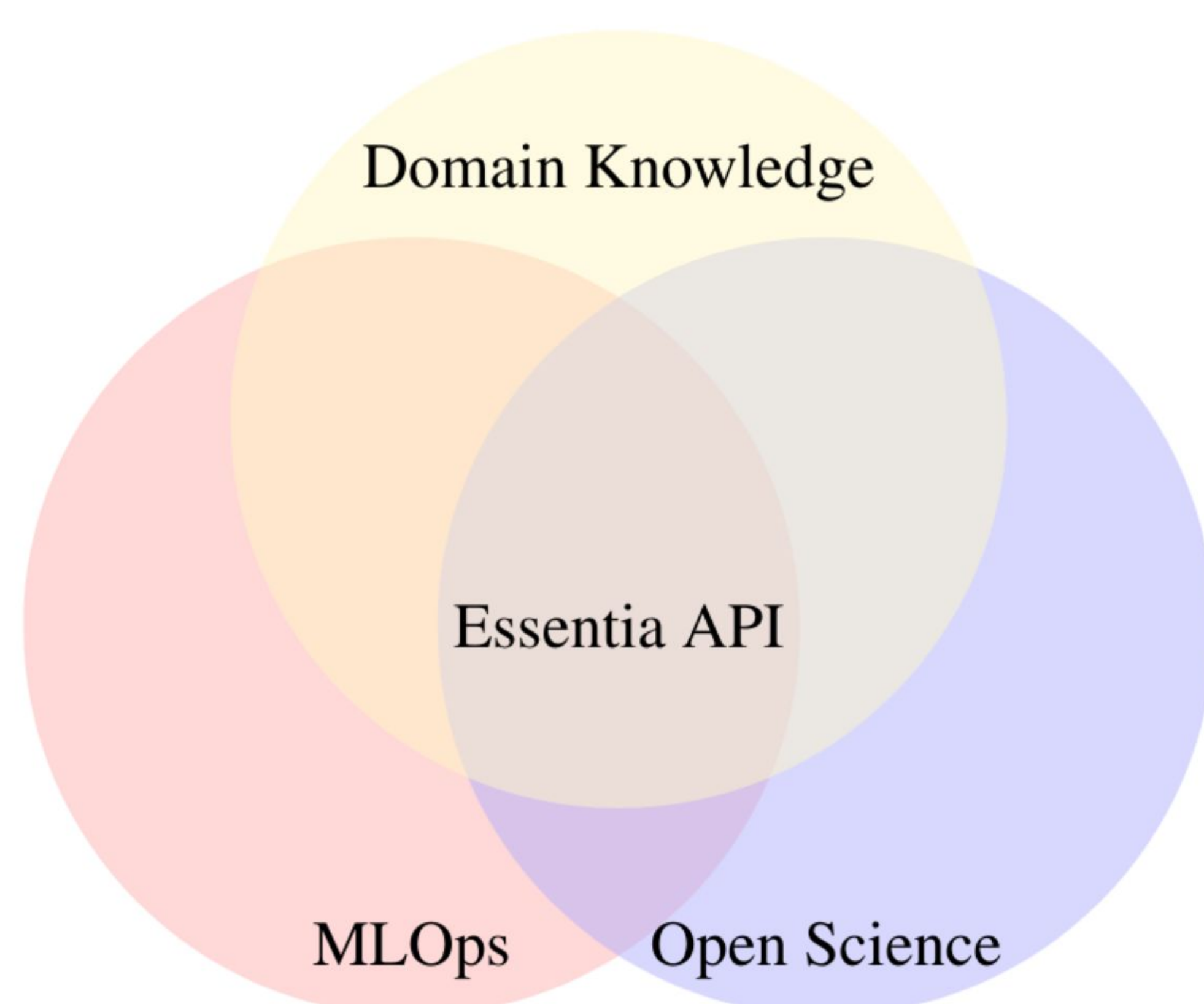
Essentia API: a Web API for Music Audio Analysis

Albin Correy Dmitry Bogdanov Pablo Alonso-Jiménez Xavier Serra

Music Technology Group, Universitat Pompeu Fabra

Motivation

- ❖ Easier access to music audio analysis tools for a broader community of developers.
- ❖ Less technical debt for the users.
- ❖ Transparency. Access to recent open-source models and algorithms developed by researchers at MTG.

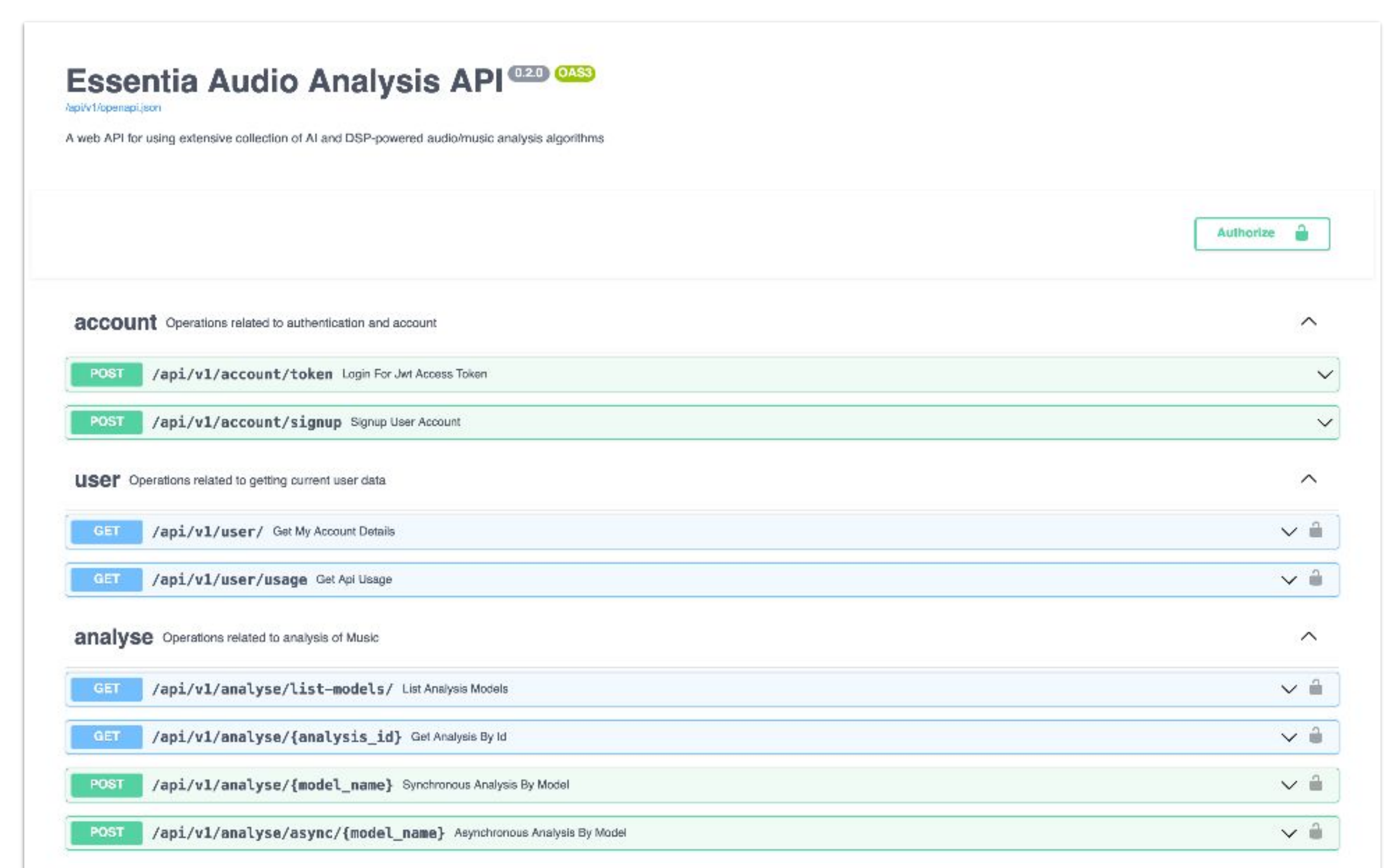


Algorithms

- ❖ Music **style** classification (400 music styles)
- ❖ Music auto-tagging (~230 tags covering **genre, mood, epoch, instrumentation**, etc.)
- ❖ Single-label music classification (genre, mood, **danceability, voice, instrumentation**, engagement, approachability, etc.)
- ❖ Tempo estimation (**BPM**).
- ❖ Musical **key** estimation (multiple profiles).
- ❖ Other semantic features (**arousal/valence, approachability, engagement, danceability**).
- ❖ Deep audio **embeddings**
- ❖ **Loudness** (EBU R 128 loudness standard metrics).

API design

- ❖ Based on music audio analysis algorithms and ML models in **Essentia** (<https://essentia.upf.edu/>).
- ❖ Synchronous and asynchronous end-points.
- ❖ Queries with a custom list of algorithms.
- ❖ The algorithms are optimized in terms of latency and computational efficiency.
- ❖ The analysis is done on a track level. Each analysis request is related to a single track.
- ❖ All common audio formats are accepted. There is a limit on the maximum upload file size per query.
- ❖ Analysis results include an overall description summarized over time as well as a more detailed time-aligned description for some of the algorithms.
- ❖ Currently, analysis results are available with a short retention period, while the audio is never stored on the server.
- ❖ User-level authorization using JSON Web Tokens.
- ❖ Backward compatibility is ensured by versioning of models and algorithms.



Async Query
example

```
curl -X 'POST' \
  'http://localhost:8080/api/v1/analyse/async/voice_instrumental?model_version=1' \
  -H 'accept: application/json' \
  -H 'audio-url: https://github.com/MTG/essentia-audio/blob/master/recorded/flamenco.mp3?raw=true' \
  -H 'Authorization: Bearer eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJzdWIiOiJhbmVpbGlzImV4cCI6IjI0MTY2OTkyMjI0IiwiaWF0IjoiMTY2OTkyMjI0Iiwidm91dGEiOiJhbmVpbGlzIn0.eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJzdWIiOiJhbmVpbGlzImV4cCI6IjI0MTY2OTkyMjI0IiwiaWF0IjoiMTY2OTkyMjI0Iiwidm91dGEiOiJhbmVpbGlzIn0' \
  -d ''

Request URL
http://localhost:8080/api/v1/analyse/async/voice_instrumental?model_version=1

Server response
Code    Details
200

Response body
{
  "id": "06ce6fe9-4a66-4682-86c2-decb6308288c",
  "status": "PENDING",
  "filename": "https://github.com/MTG/essentia-audio/blob/master/recorded/flamenco.mp3?raw=true",
  "model_name": "voice_instrumental"
}
```

```
curl -X 'GET' \
  'http://localhost:8080/api/v1/analyse/06ce6fe9-4a66-4682-86c2-decb6308288c' \
  -H 'accept: application/json' \
  -H 'Authorization: Bearer eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJzdWIiOiJhbmVpbGlzImV4cCI6IjI0MTY2OTkyMjI0IiwiaWF0IjoiMTY2OTkyMjI0Iiwidm91dGEiOiJhbmVpbGlzIn0.eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJzdWIiOiJhbmVpbGlzImV4cCI6IjI0MTY2OTkyMjI0IiwiaWF0IjoiMTY2OTkyMjI0Iiwidm91dGEiOiJhbmVpbGlzIn0'

Request URL
http://localhost:8080/api/v1/analyse/06ce6fe9-4a66-4682-86c2-decb6308288c

Server response
Code    Details
200

Response body
{
  "id": "06ce6fe9-4a66-4682-86c2-decb6308288c",
  "status": "SUCCESS",
  "results": {
    "model_version": "1",
    "model_name": "voice_instrumental",
    "correlation_id": "d6f4a191-b08c-44af-8da9-1d55ee1d4a8",
    "results": {
      "instrumental": {
        "instrumental": 0.9908175875531006,
        "voice": 0.009182412446899394
      }
    },
    "error": null
  }
}
```

