

Advances and Challenges in Large-Scale Music Information Retrieval

Colin Raffel and Dan Ellis

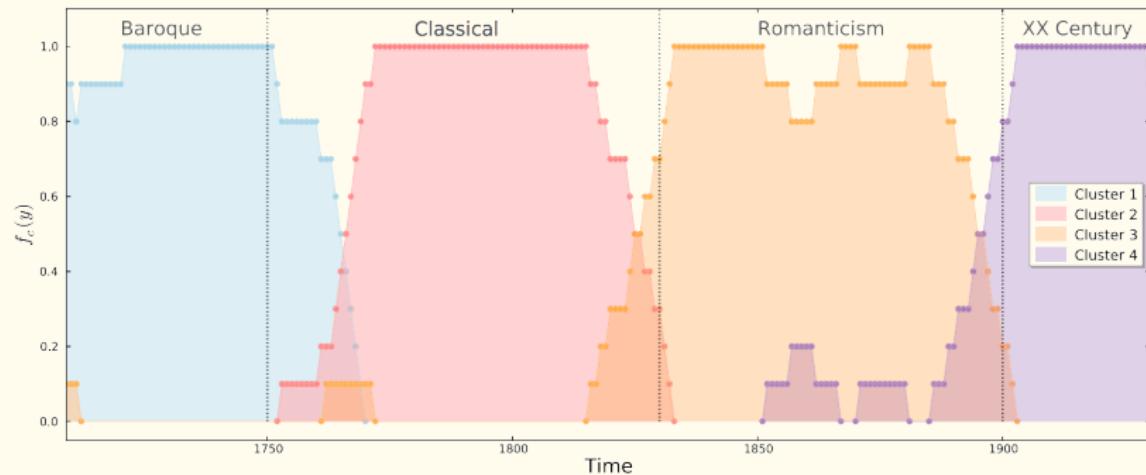
DMRN+8, December 17, 2013



IGERT Integrative Graduate
Education and Research Traineeship

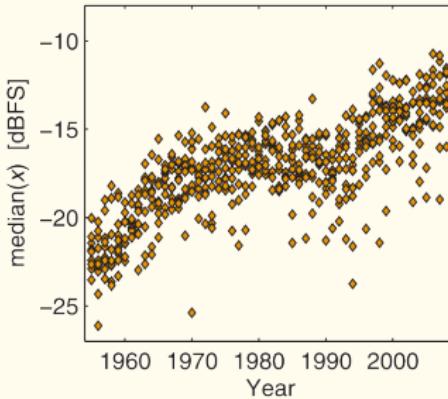
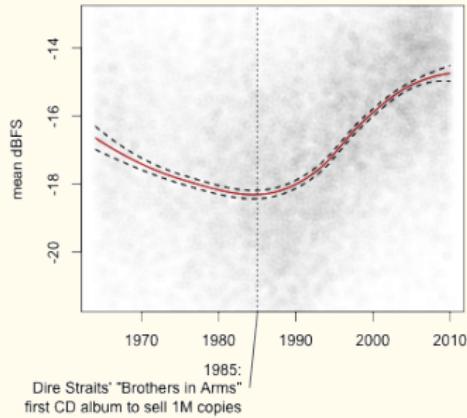


What can we do with lots of data?



Pablo H. Rodriguez Zivic et al. "Perceptual basis of evolving Western musical styles"

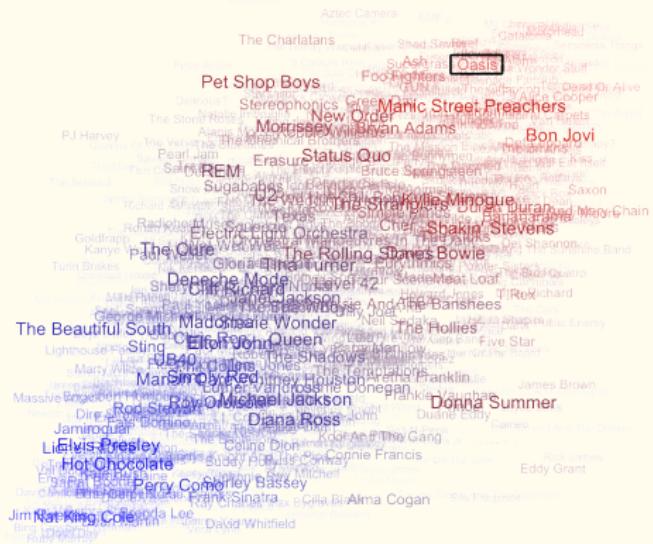
The Loudness War, four ways



Matthias Mauch, "Survival of the flattest"; Joan Serrà et al. "Measuring the Evolution of Contemporary Popular Music"; Paul Lamere, "The Loudness War Analyzed"; Shaun Ellis and Tom Engelhardt, "Visualizing a Hit"

Loudest Artists

loud



noisy

| Artist | dB |
|-------------------------|-------|
| Venetian Snares | -1.25 |
| Soulja Boy | -2.38 |
| Slipknot | -2.65 |
| Dimmu Borgir | -2.73 |
| Andrew W.K. | -3.15 |
| Queens of the Stone Age | -3.23 |
| Black Kids | -3.45 |
| Dropkick Murphys | -3.50 |
| All That Remains | -3.56 |
| Disturbed | -3.64 |
| Rise Against | -3.73 |
| Kid Rock | -3.86 |
| Amon Amarth | -3.88 |
| The Offspring | -3.89 |
| Avril Lavigne | -3.93 |
| MGMT | -3.94 |
| Fall Out Boy | -3.97 |
| Dragonforce | -4.02 |
| 30 Seconds To Mars | -4.08 |
| Billy Talent | -4.13 |
| Bad Religion | -4.13 |
| Metallica | -4.14 |
| Avenged Sevenfold | -4.23 |
| The Killers | -4.27 |
| Nightwish | -4.37 |
| Arctic Monkeys | -4.40 |
| Chromeo | -4.42 |
| Green Day | -4.43 |
| Oasis | -4.45 |
| The Strokes | -4.49 |
| System of a Down | -4.51 |
| Blink 182 | -4.52 |
| Bloc Party | -4.53 |

Matthias Mauch, "Survival of the flattest"; Paul Lamere, "The Loudness War Analyzed"

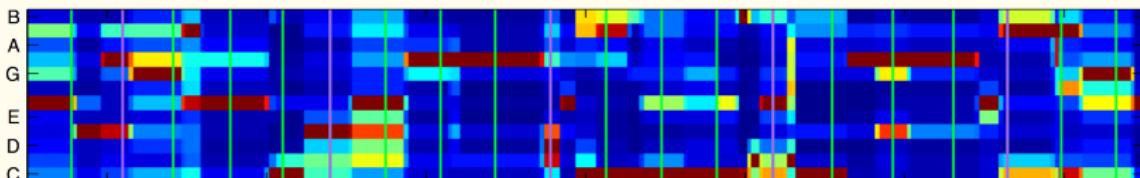
The Million Song Dataset

```
artist: 'Tori Amos'  
release: 'LIVE AT MONTREUX'  
title: 'Smells Like Teen Spirit'  
id: 'TRKUYPW128F92E1FC0'  
key: 5  
mode: 0  
loudness: -16.6780  
tempo: 87.2330  
time_signature: 4  
duration: 216.4502  
sample_rate: 22050  
audio_md5: '8'  
7digitalid: 5764727  
familiarity: 0.8500  
year: 1992
```

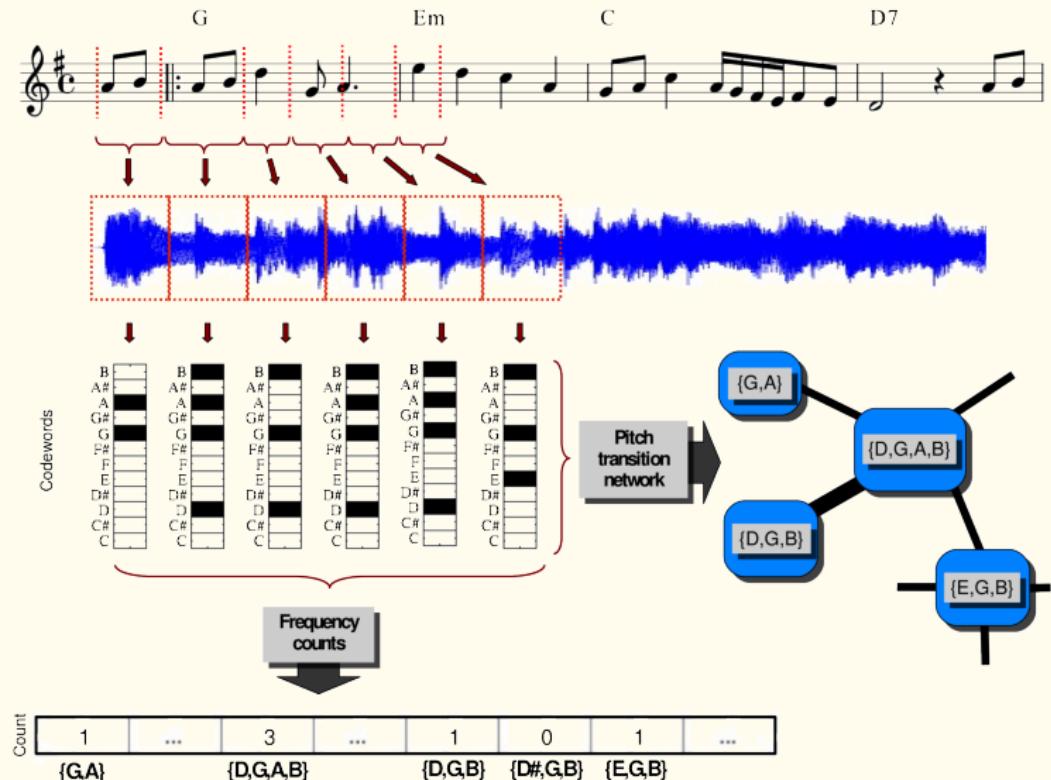
| | |
|---------------------------|------------------------|
| 100.0 - cover | 5.0 - cover songs |
| 57.0 - covers | 4.0 - soft rock |
| 43.0 - female vocalists | 4.0 - nirvana cover |
| 42.0 - piano | 4.0 - Mellow |
| 34.0 - alternative | 4.0 - alternative rock |
| 14.0 - singer-songwriter | 3.0 - chick rock |
| 11.0 - acoustic | 3.0 - Ballad |
| 8.0 - tori amos | 3.0 - Awesome Covers |
| 7.0 - beautiful | 2.0 - melancholic |
| 6.0 - rock | 2.0 - k001 chix |
| 6.0 - pop | 2.0 - indie |
| 6.0 - Nirvana | 2.0 - female vocalist |
| 6.0 - female vocalist | 2.0 - female |
| 6.0 - 90s | 2.0 - cover song |
| 5.0 - out of genre covers | 2.0 - american |

```
#5489,4468, Smells Like Teen Spirit  
TRTUOVJ128E078EE10 Nirvana  
TRFZJ0Z128F4263BE3 Weird Al Yankovic  
TRJHCKN12903CDD274 Pleasure Beach  
TRELTOJ128F42748B7 The Flying Pickets  
TRJKBXL128F92F994D Rhythms Del Mundo feat. Shanade  
TRIHLAW128F429BBF8 The Bad Plus  
TRKUYPW128F92E1FC0 Tori Amos
```

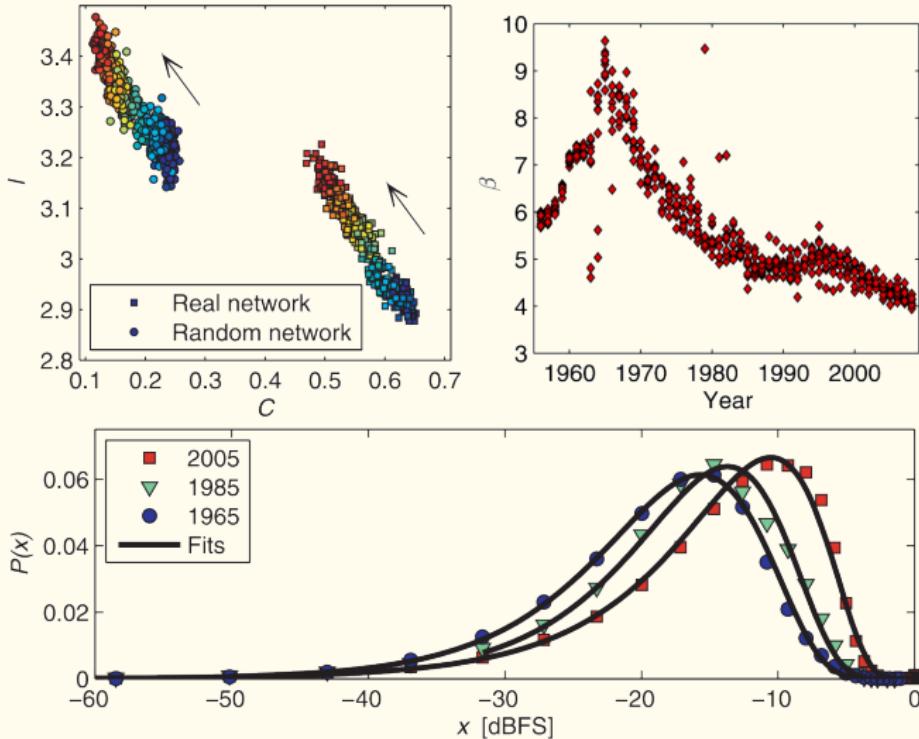
| | | |
|----------|-------------|----------|
| 12 hello | 6 here | 3 is |
| 11 i | 6 us | 3 with |
| 10 a | 6 entertain | 3 oh |
| 9 and | 4 the | 3 out |
| 7 it | 4 feel | 3 an |
| 6 are | 4 yeah | 3 light |
| 6 we | 3 to | 3 less |
| 6 now | 3 my | 3 danger |



Using Features: Music Evolution



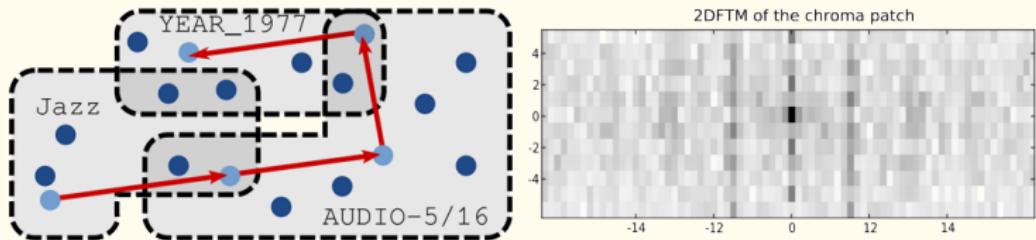
Measuring The Evolution... Results



(Partial) Python implementation available:

<http://github.com/craffel/music-evolution>

Linking with existing data



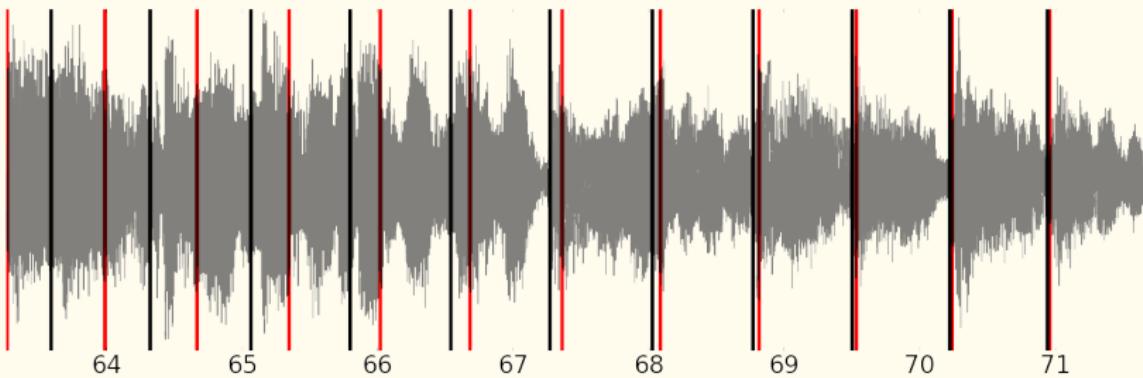
| Lowest Word | Highest Word | Lowest Tag Feature | Highest Tag Feature |
|-------------|--------------|--------------------|---------------------|
| Word Weight | Weight | Weight | Weight |
| Death | -0.0346 | Love | 0.0572 |
| Dead | -0.0272 | Baby | 0.0382 |
| Hate | -0.0220 | Blue | 0.0179 |
| Burn | -0.0216 | Girl | 0.0171 |
| Pain | -0.0191 | People | 0.0147 |
| Dream | -0.0161 | Hit | 0.0372 |
| Love | -0.0131 | Man | 0.0343 |
| Heart | -0.0123 | Rock | 0.0320 |
| Home | -0.0114 | Girl | 0.0291 |
| Sad | -0.0111 | Baby | 0.0283 |
| Metal | | -0.0581 | FVs |
| DM | | -0.0517 | Pop |
| Brutal DM | | -0.0510 | Classic Country |
| Melodic BM | | -0.0348 | Oldies |
| Metalcore | | -0.0267 | Soul |
| Beautiful | | -0.0183 | Hip Hop |
| FVs | | -0.0130 | Hip-Hop |
| Ambient | | -0.0113 | Rap |
| Christian | | -0.0112 | Gangsta Rap |
| Mellow | | -0.0106 | UHH |

McFee and Lanckriet, "Hypergraph Models of Playlist Dialects"

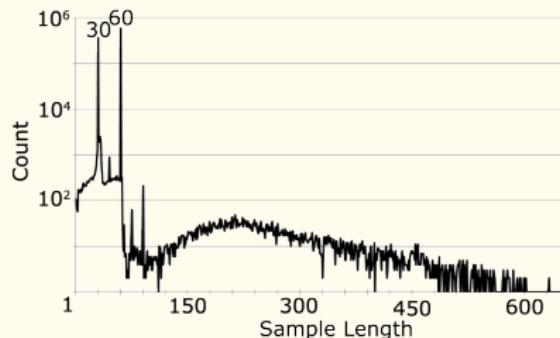
Mahieux and Ellis, "Large-Scale Cover Song Recognition Using the 2D Fourier Transform Magnitude"

McVicar and De Bie, "CCA and a Multi-way Extension for Investigating Common Components between Audio, Lyrics and Tags"

Ground Truth?



Audio? One solution:

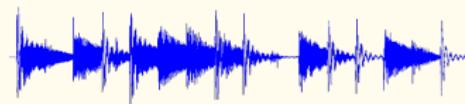
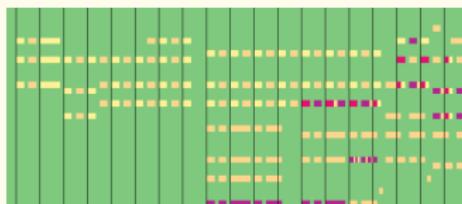
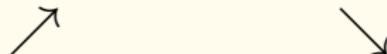


| Samplerate | | |
|-----------------------------|---------|--------|
| 22 | 768,710 | 77,26% |
| 44 | 226,169 | 22,73% |
| other | 81 | 0,01% |
| Bitrate | | |
| 128 | 646,120 | 64,94% |
| 64 | 343,344 | 34,51% |
| other (VBR) | 5,494 | 0,55% |
| Channels | | |
| Mono | 6,342 | 0,64% |
| Stereo | 150,779 | 15,15% |
| Joint stereo / dual channel | 837,839 | 84,21% |

Schindler et al. "Facilitating Comprehensive Benchmarking Experiments on the Million Song Dataset"

MIDI + Audio + MSD

```
artist: 'Tori Amos'  
release: 'LIVE AT MONTREUX'  
title: 'Smells Like Teen Spirit'  
id: 'TRKUXPW128F92E1FC0'  
duration: 216.4502  
sample_rate: 22050  
audio_md5: '8'  
7digitalid: 5764727  
year: 1992
```



Matching by Text



```
artist: 'Tori Amos'  
release: 'LIVE AT MONTREUX'  
title: 'Smells Like Teen Spirit'  
duration: 216.4502
```

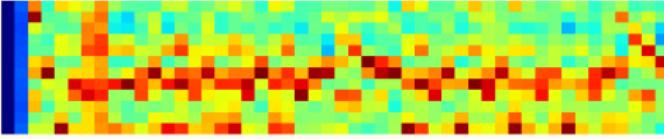
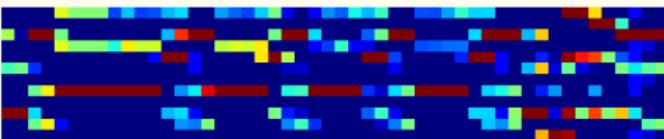
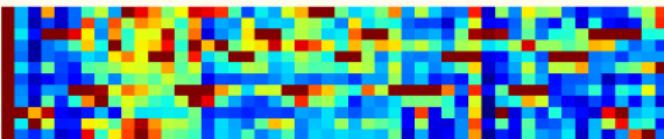


```
/nirvana/smells_like_teen_spirit.mid
```

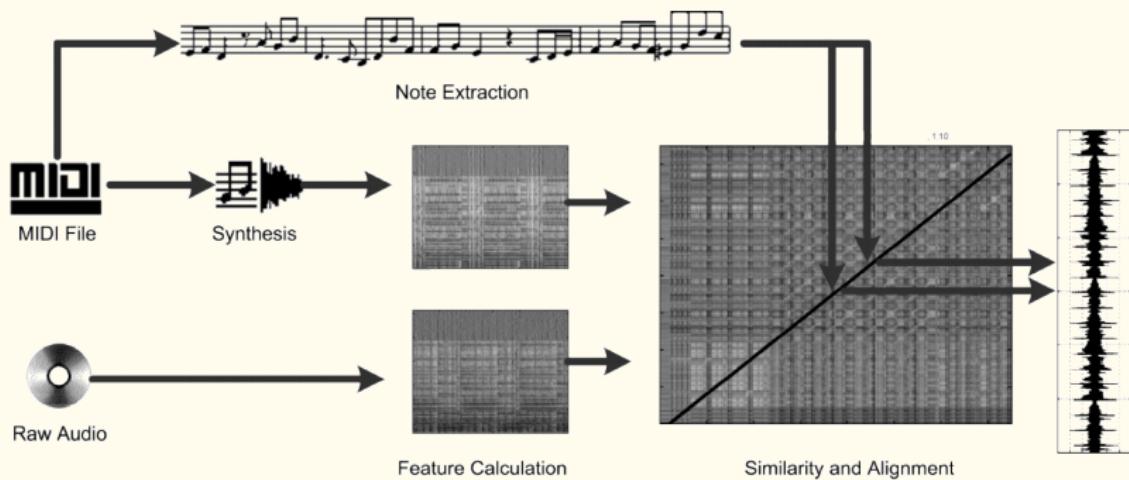


```
artist : Amos, Tori  
release : Live 91/92  
title : Smells Like Teen Spirit (Nirvana cover)  
duration : 216.3
```

Matching by Content



Alignment



Turetsky and Ellis, "Ground-Truth Transcriptions of Real Music from Force-Aligned MIDI Syntheses"

Scale and Access



300 GB

Open access



3 GB

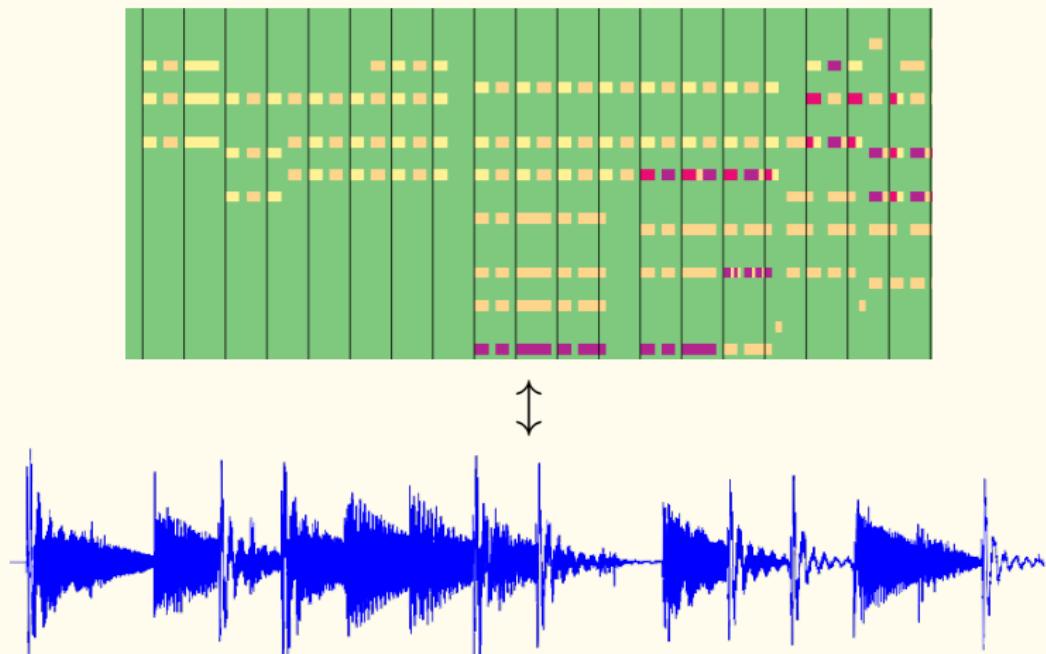
Open access?



200 TB

Restricted access

What does aligned MIDI get us?



Thanks!



The
MIDI
ORGANIZER

YOUR CHOICE. YOUR FAMILY. YOUR MIDI.

CALL NOW TOLL NOT INCLUDED

1-800-555-0199

CINCO

ALLOW 10-12 WEEKS FOR DELIVERY
NO MONEY BACK GUARANTEE

craffel@gmail.com