



MoisesDB - A Dataset for Source Separation Beyond 4-stems



Igor Pereira, Felipe Araújo, Filip Korzeniowski, Richard Vogl

igor@moises.ai

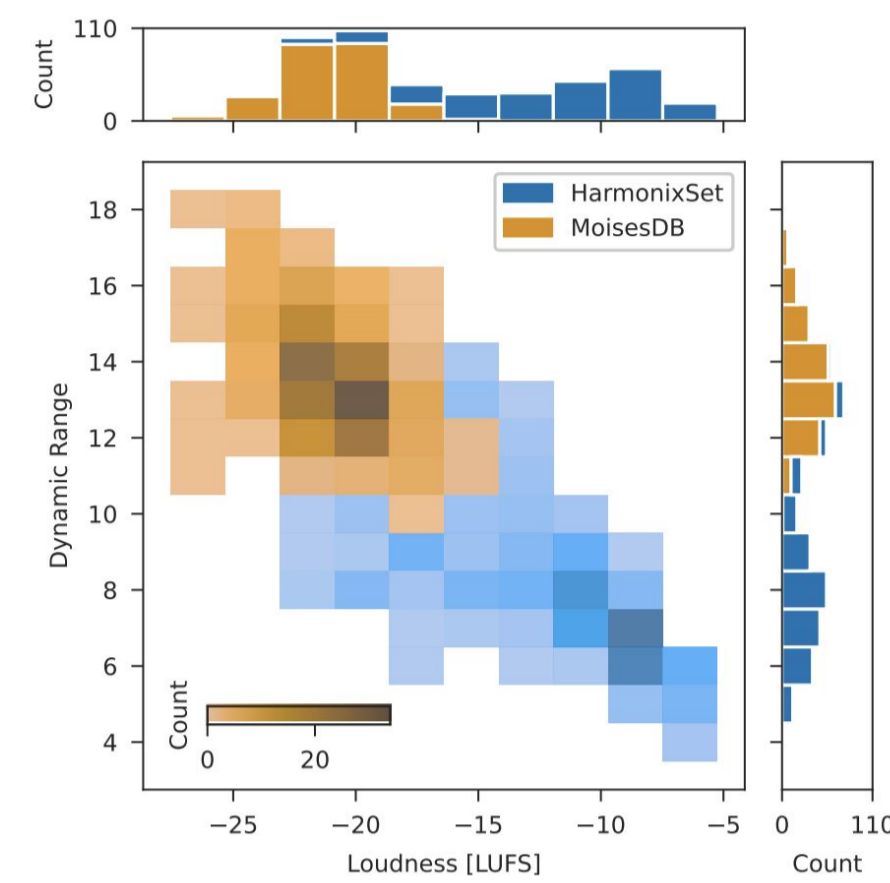
1. Objectives & Highlights

Stem	Track
Bass	Bass Guitar, Bass Synthesizer, Contrabass
Bowed Strings	Cello, Cello Section, Other Strings, String Section, Viola Section, Viola Solo
Drums	Cymbals, Drum Machine, Full Acoustic Drumkit, Hi-Hat, Kick Drum, Overheads, Snare Drum, Toms
Guitar	Acoustic Guitar, Clean Electric Guitar, Distorted Electric Guitar
Other	Fx
Other Keys	Organ, Electric Organ, Other Sounds, Synth Lead, Synth Pad
Other Plucked	Banjo/Mandolin/Ukulele/Harp
Percussion	A-Tonal Percussion, Pitched Percussion
Piano	Electric Piano, Grand Piano
Vocals	Background Vocals, Lead Female Singer, Lead Male Singer, Other
Wind	Brass, Flutes, Other Wind, Reeds

Build **more-than-4-stems** source separation systems and **overcomes limitation** of using only four stems!

- **240 tracks** from 45 artists across twelve musical genres.
- Total duration: **14+ hours**.
- Stem-base, **two-level hierarchical taxonomy** of stems.

4. Audio Properties

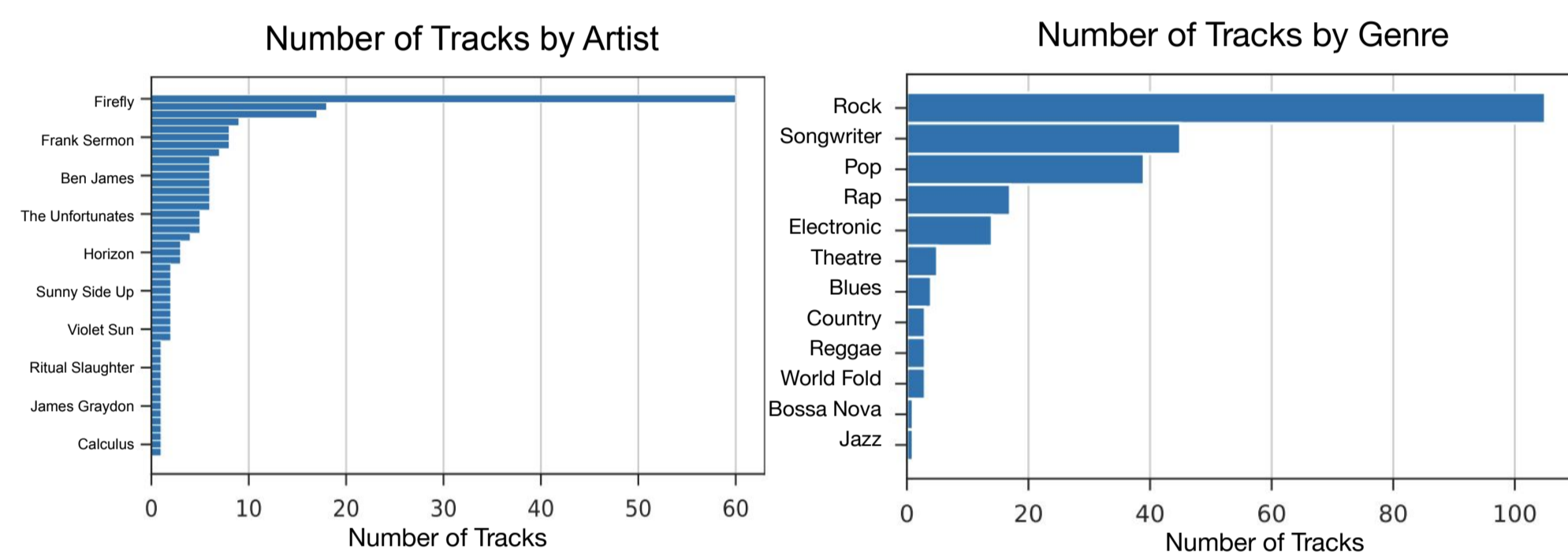


Mastering of the audio tracks:

- Different distributions of both LUFS and Dynamic Range when compared to pop songs.
- Loudness normalization mainly fixes these issues.

▶ If your application is *loudness sensitive*:
Normalize the audio!

2. Artists and Genre Distributions



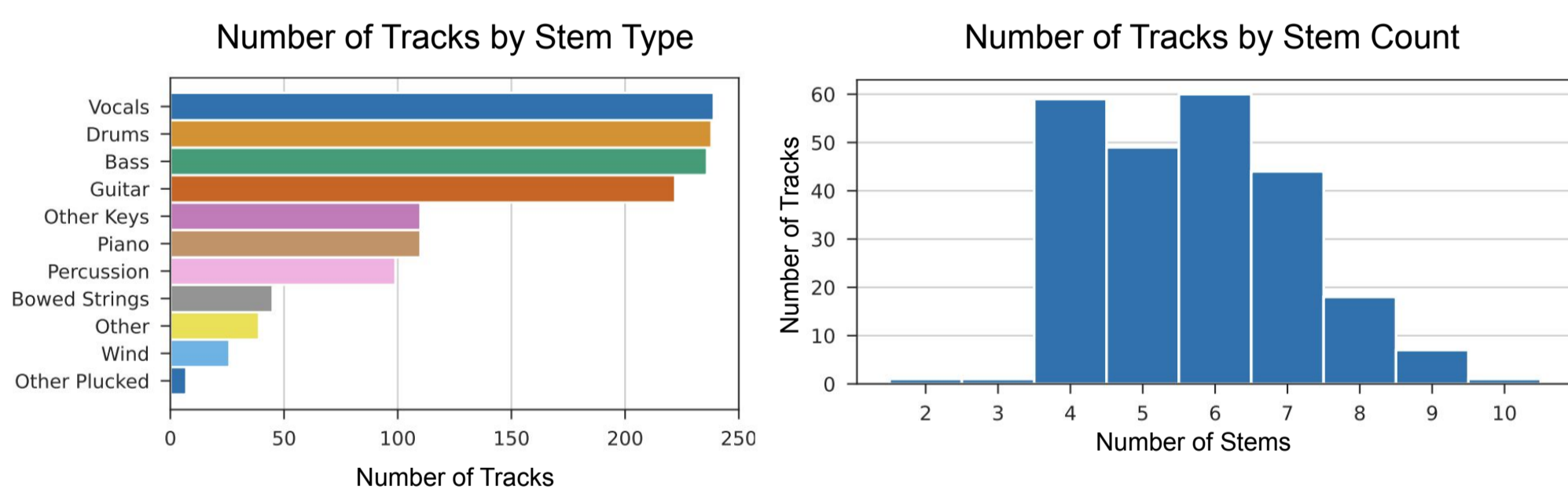
- Contains mostly Rock and Pop.
- Includes Reggae, Bossa Nova and Jazz.

5. Python Library

```
from moisesdb.dataset import MoisesDB

db = MoisesDB(data_path='./moises-db-data')
n_songs = len(db)
track = db[0]
# mix multitracks to stems
stems = track.stems
# stems = {
#   'vocals': np.ndarray (stem audio data),
#   'bass': np.ndarray (stem audio data),
#   ...
# }
# mixture: np.ndarray
mixture = track.audio
# save mixed stems
track.save_stems('./stems/track_0')
```

3. Stem Distributions



- **Vocals, Bass** and **Drums** dominates the stems distribution followed closely by **Guitars**.
- Also includes other stems and **individual sources** in a reasonable amount.
- Many tracks contains **more than 4-stems**.

6. SDR for Baseline Systems

	4 stems (N = 235)											
	HT-Demucs		Spleeter		IBM		IRM		MWF			
	Mean ± Std	Mdn	Mean ± Std	Mdn	Mean ± Std	Mdn	Mean ± Std	Mdn	Mean ± Std	Mdn	Mean ± Std	Mdn
vocals	10.05 ± 2.48	9.62	7.61 ± 2.45	7.27	9.02 ± 2.13	8.67	10.72 ± 2.03	10.37	10.72 ± 2.11	10.27		
bass	11.64 ± 3.35	11.99	6.46 ± 2.26	6.57	6.46 ± 2.08	6.31	8.43 ± 2.03	8.20	8.68 ± 2.07	8.38		
drums	10.94 ± 2.30	10.91	6.65 ± 1.72	6.64	7.33 ± 1.77	7.30	8.98 ± 1.68	8.92	9.01 ± 1.67	8.83		
other	7.00 ± 2.76	7.30	4.45 ± 2.26	4.69	5.77 ± 1.72	5.61	7.74 ± 1.65	7.57	7.90 ± 1.65	7.79		
overall	9.91 ± 3.27	9.69	6.29 ± 2.47	6.24	7.14 ± 2.28	6.99	8.97 ± 2.16	8.81	9.08 ± 2.15	8.87		
5 stems (N = 104)												
vocals			6.99 ± 1.97	6.74	8.29 ± 1.66	8.08	9.94 ± 1.59	9.75	10.01 ± 1.71	9.68		
bass			6.26 ± 2.27	6.28	6.13 ± 2.15	5.86	8.02 ± 2.07	7.82	8.32 ± 2.08	8.03		
drums			6.89 ± 1.88	6.97	7.67 ± 1.94	7.87	9.29 ± 1.84	9.34	9.32 ± 1.84	9.36		
other			1.97 ± 1.76	2.09	4.04 ± 1.47	4.13	6.00 ± 1.44	6.01	6.10 ± 1.48	6.19		
piano			1.17 ± 1.86	0.75	3.04 ± 2.37	2.55	4.99 ± 2.32	4.60	5.30 ± 2.46	4.79		
overall			4.66 ± 3.20	5.02	5.12 ± 2.81	4.87	7.65 ± 2.66	7.60	7.81 ± 2.66	7.83		
6 stems (N = 88)												
vocals			9.55 ± 1.87	9.39	8.09 ± 1.51	7.98	9.73 ± 1.46	9.61	9.81 ± 1.49	9.61		
bass			11.93 ± 2.87	12.13	6.04 ± 1.98	5.83	7.92 ± 1.93	7.73	8.24 ± 1.96	8.03		
drums			11.02 ± 2.44	11.28	7.58 ± 1.96	7.79	9.19 ± 1.86	9.21	9.23 ± 1.85	9.25		
other			0.28 ± 1.84	0.39	2.85 ± 1.76	2.74	4.67 ± 1.76	4.57	4.72 ± 1.82	4.55		
piano			1.60 ± 1.68	1.64	2.78 ± 1.61	2.49	4.71 ± 1.61	4.47	4.97 ± 1.74	4.70		
guitar			3.07 ± 1.81	3.16	3.35 ± 1.54	3.44	5.28 ± 1.54	5.36	5.41 ± 1.65	5.46		
overall			6.24 ± 5.17	6.05	5.12 ± 2.81	4.87	6.91 ± 2.70	6.69	7.06 ± 2.73	6.89		

7. Conclusion

- MoisesDB fills a gap in publicly available data for source separation research.
- Provides opportunity for improvements on other tasks, such as:
 - Instrument recognition, automatic mixing etc.
- We've included all details as well as benchmarking with ideal masks and two source separation models.
- A python package was released to help parsing the tracks metadata and create stems versions of the dataset.



<https://github.com/moises-ai/moises-db>

- Sharing of recordings for **individual sources** enables application beyond initial scope!
- Includes individual common sources for stems as well as unconventional sources
 - Such as Cello, Viola, Brass, Reeds, etc.

