"Empowering Music Creators via Open-Source ML"

Scott H. Hawley

**Professor of Physics
Senior Fellow, BDC
Technical Fellow, Stability AI / Harmonai**

@drscotthawley

- Research Fellow, Mathematics Group,
  Kirby Laing Centre for Public Theology (Cambridge UK)
- Founding Member, AI and Faith
- Board Member, AI Theology
- Board / Faculty Advisor, Waveshaper.ai (Montreal)

Before Belmont (< '06): Supercomputing, Astrophysics, Nonlinear Dynamics
At Belmont: Acoustics, Audio, Machine Learning – esp. for Musical Audio Production.
Teaching: Physics for Audio Engineering, Deep Learning & AI Ethics

What I like....

Interdisciplinary talk (esp. on classification), learning via interactive data visualization (esp. 3D), & Jupyter notebooks!
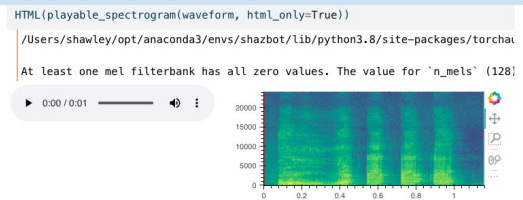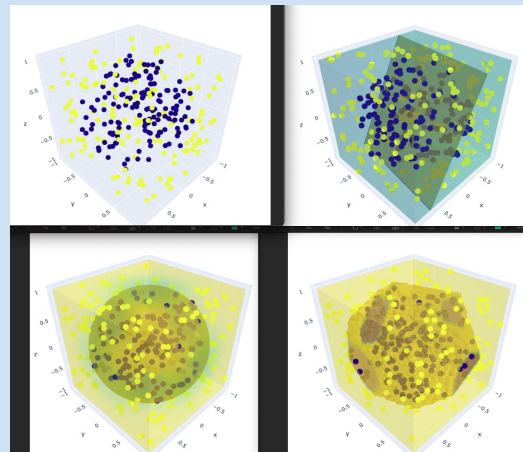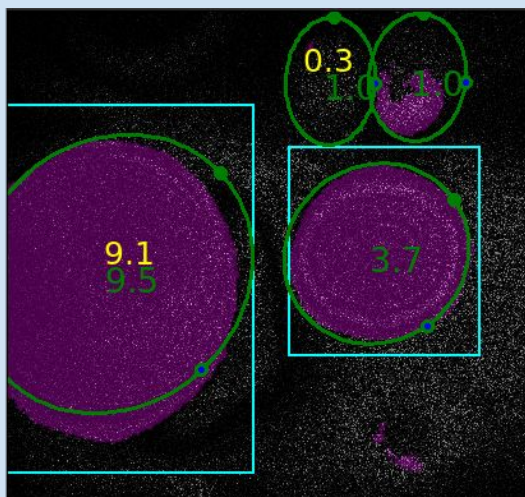

AI ART
@xaltint
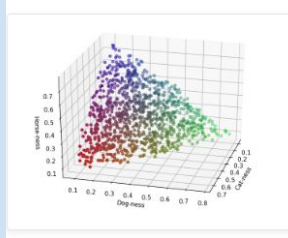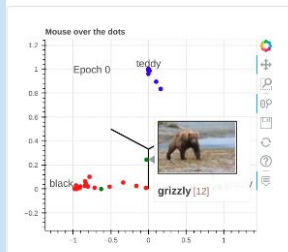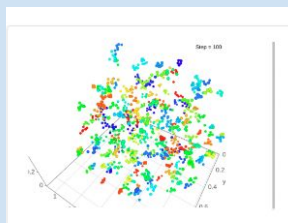
botograder

An autograder. There are many like it but this one is mine.

Downloads and runs parts of Google Colab notebooks when they change. (Ideally) Emails students their scores.
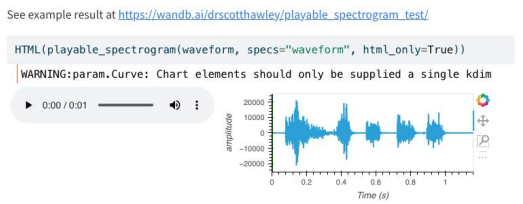
Requirements:

```
pip install pydrive2 jupytext yagmail
```

HTML(playable_spectrogram(waveform, html_only=True))

```
/Users/shawley/opt/anaconda3/envs/shazbot/lib/python3.8/site-packages/torchau
```

At least one mel filterbank has all zero values. The value for `n_mels` (128)

Sample usage with WandB:

```
wandb.init(project='audio_test')
wandb.log({"playable_spectrograms": playable_spectrogram(waveform)})
wandb.finish()
```

See example result at https://wandb.ai/drscotthawley/playable_spectrogram_test/

HTML(playable_spectrogram(waveform, specs="waveform", html_only=True))

```
WARNING:param.Curve: Chart elements should only be supplied a single kdim
```

🎅 Naughty by Numbers:
Classifications at Christmas

▶ First Time Here? (click to expand)

▶ Optional: Math Discussion (click to expand)

Two Analytics Li'l Helpers, Molasses Slow Elf 🧝 and Brilliant Celerity Elf 🧝, are tasked with creating predictive models to map a child's SANTA score into a verdict of Naughty (lower half of the graph) or Nice (upper half) by fitting an S-shaped curve to scores for children judged Naughty 😈 or Nice 😇 in the past. Let's see how their models compare...

Train the Models   Get New Data

▶ Controls (click to expand)

Data & predictions:

My Role as a BDC Senior Fellow…

…what I was already doing? ML-Audio / AI-Ethics / Interdisc.

- **Foster Campus-Wide/Interdisciplinary Discussion/Training/Events**
  - re. the impact of data on our fields, our world, esp. our undergraduate students' lives & futures
  - **November 28!** Stable Diffusion "AI Art" convo w/ WCoA
  - Spring '23: Teaching Honors Symposium on "Classification" (book in progress)

- **Pursue research project**

- **Project Belmont's data-scholarly presence in region, nation & world, e.g. giving talks, serving on boards & networking…**

CODE	SECT	TITLE	HRS	INSTR
HON	1895.01	Honors Colloquium Full Term Classification: Aristotle to Algorithims	1.00	Scott Hawley

ML-Audio Collaboration: Harmonai.org 🔵

- Harmonai leader Zach Evans (formerly Microsoft Senior Programmer) & I introduced in May…

- "Open Source Collective" of DJs, developers, music producers, researchers,… from stadium-entertainers to Google employees to indie artist/coders

- Goal is to produce tools to *empower* musicians, not replace them!

- We've had huge buy-in from music community: artists *giving us* their catalogs to train models on!

- Focussed on building tools/apps rather than writing papers, but grad students very involved, being mentored & mentoring others – e.g. Belmont's Caleb Koch!

ML-Audio Collaboration: Harmonai
.org

"Stability AI Family":

- Open-Source AI: models released to world for free
  - GPT-J (NLP)
  - Stable Diffusion* (text-to-image)
- Huge AWS GPU Cluster (450+ nodes x 8 A100s)
  - Research compute grants available!

*All background images made via Stable Diffusion!

eleutherai

A decentralised collective focusing on AI alignment.

LAION

A community-driven non-profit organisation devoted to making large-scale AI openly accessible.

DreamStudio

Open-sourced image generation model that cultivates autonomous freedom to produce incredible imagery.

CarperAI

Minimize the burden of understanding human preferences through automation.

CAIAC

Collective & Augmented Intelligence Against COVID-19.

Harmonai
COMING SOON

A community-driven organization releasing open-source generative audio tools to make music production more accessible and fun for everyone.

OpenBioML
COMING SOON

A brand new community focused on fair and open research for artificial intelligence in biology.

DeepFloyd
COMING SOON

An independent Modular AI Research Lab. Discovering new facets of Multimodality.

If you don't know Stability: raised > $100 million

Launch Event, Oct 17

Forbes

September 9

VENTURE CAPITAL • EDITORS' PICK

Startup Behind AI Image Generator Stable Diffusion Is In Talks To Raise At A Valuation Up To $1 Billion

Stable Diffusion "a landmark event in the history of AI"
- Andrew Ng

Project: "SHAZBOT" (working title)

with **Joe Baldridge**

Project: "SHAZBOT" (working title)

- **"SHAZAM for Stems"**: provide a way to **track provenance** of parts of songs (e.g., drum or bass lines) that get remixed into new music works

- To provide an **intervention to benefit the local Nashville community** of musicians as well as **worldwide**

- By offering an additional **revenue stream** (for "neighboring rights")

- => Addresses **structural inequality** in the music industry ("slush fund")

- …Using Deep Learning, without BlockChain, without audio steganography

- *Similar to detections in* **_bioacoustics_** *&* ***gravitational wave astro***.

- Baby steps: "**Audio Algebra**" for mixing & effects using "embeddings"

Campus-Wide Initiative: DATA CONSENT!

Joe & I are building a dataset: "BDCT"

….We're ALL going to be making datasets.

Which means we ALL need to get data-ethics-savvy fast.

e.g.: Student data-harvesting should be subject to IRB.

University-wide language / guidance re. collection, storage, sharing, licensing,...etc.

("Creative Works" page is vague)

cf. "Rights to Use of Student Work" of SCAD.edu – thanks Trudy @ WCoA!

Student Audio Data Release Consent Form

This form concerns the release of anonymized data from student assignments in the AET 3190 "Audio 2" course, in the form of Pro Tools session files including raw audio, automation controls, and mixes.

The purpose of this is an academic study on the possibility of using Deep Learning models to provide attribution to remixed audio parts ("stems") which appear in later recordings, i.e., to provide an additional royalty-tracking and therefore revenue stream for independent musicians who supply parts such as drums and bass lines.

In order to train the Deep Learning model, a dataset is required. In keeping with academic standards of reproducibility, the dataset should be released publicly. Thus, student consent is necessary for the release of their (anonymized) recording projects. The dataset would be "available upon request." See the NYU MedleyDB Dataset (raw audio only) for precedent.

Three options are provided for student consent. *In all 3 cases, students retain the rights to their own recordings*; the differences take the form of *what license is granted to others*. Further information on the license options is provided in the links. Select the option below as per your choice.

1) Creative Commons 4.0: Full release of rights commercial

[_] I, _____ BUID _____ hereby grant Belmont University all rights to use my creative work from Audio 2 recoding project class studio assignment to distribute, remix, adapt, and build upon the material in any medium or format, so long as

Outputs So Far…

- Progress w/ **diffusion models** for musical audio gen: **Dance Diffusion!** (DD, next slide)

- Support code: Two libraries **"aeiou"** & **"prefigure"** on GitHub & PyPI. **aeiou.viz** supported dev. of **DD "Harmonai Hangouts"** Seminar Series

- **Destructo!** Colab notebook for **"audio algebra"**

- Draft of Consent Form re. use of student work - lawyers. Need more help to push for university-wide policy

Dance Diffusion

Generative musical audio model: Shapes noise into music, e.g., dubstep, metal, old-timey stuff. Best for short samples, not whole songs. 100% consensual/PD training data!

Weights & Biases Interview!

Weights & Biases used DD for Podcast Theme Song!
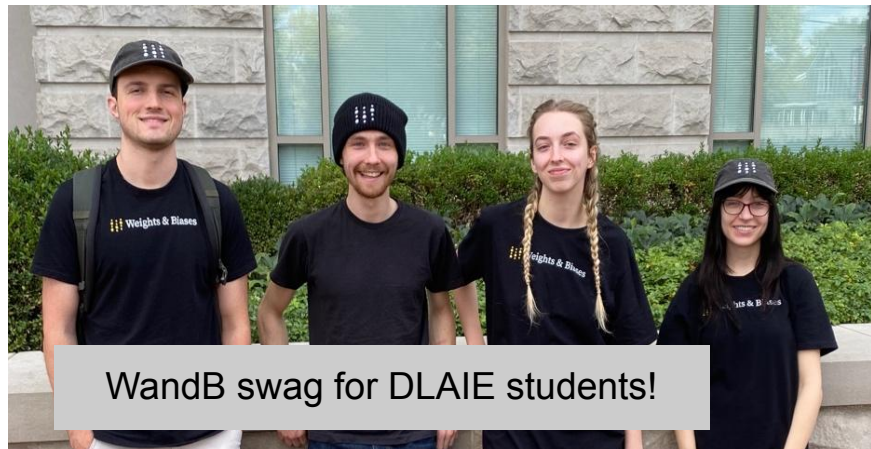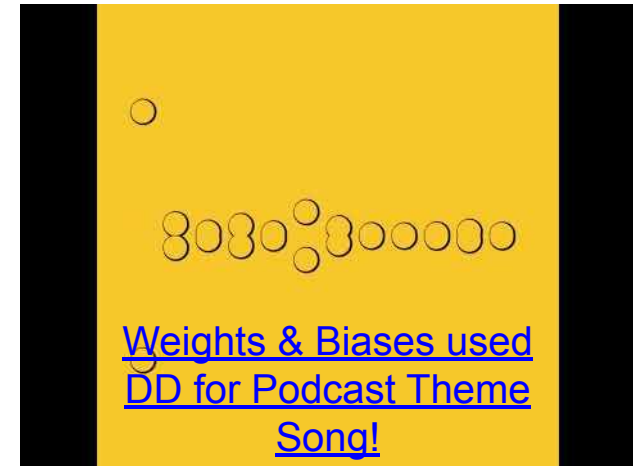
WandB swag for DLAIE students!

Demo Page

Any remaining time:

- Talk about **CAMPUS-WIDE CLOUD COMPUTE needs?** (e.g., Eleuther donated CoreWeave GPUs for DLAIE class)
- Talk about the **NEED FOR MORE PHYSICS FACULTY** ...or…
- **Blast the Stable Diffusion launch video? (2 min)**

…Thanks! -@drscotthawley