

# *A Data Science Approach to Decoding Morse Code Audio*

*Matthew Younger*

*DSC105 @ Lyon College*

# Data is Everywhere

- **Audio is a bountiful source of data.**
- **R lacks native capabilities to study audio**
  - Complication:  
develop a DSP-specific pipeline and process for research  
Limitations of off-the-shelf tools (Jupyter Labs, GitHub, R Studio)
- **Solution:**
  - Careful discernment of best applications and methodology

## Choice of Tools:



GNU Emacs

Self-documenting Editor



Org Mode

Polyglot Notebook Platform



R-Lang

Quality Libraries, Plotting functionality



Gitea

On-premise, lacks restrictions



Debian Linux

Power of Linux + stability

# References:

## References

- [1] A. Eldridge, *Toward the extraction and evaluation of ecologically-meaningful sound-scape components using sparse coding methods*, <https://peerj.com/articles/2108/>, [Accessed 28-09-2025], 2016.
- [2] Y. Philips, *Visualization of environmental audio using ribbon plots and acoustic state sequences*, <https://ieeexplore.ieee.org/document/8114628>, [Accessed 28-09-2025], 2017.
- [3] M. Depraetere, *Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland*, <http://www.elsevier.com/locate/ecolind>, [Accessed 28-09-2025], 2011.
- [4] J. Sueur, *Rapid acoustic survey for biodiversity appraisal*, <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0004065>, [Accessed 28-09-2025], 2008.
- [5] T. Ishibashi, *Investigating audio data visualization for interactive sound recognition*, <https://dl.acm.org/doi/abs/10.1145/3377325.3377483>, [Accessed 28-09-2025], 2020.
- [6] Y. Zhao, *Diversity monitoring of coexisting birds in urban forests by integrating spectrograms and object-based image analysis*, [https://www.mdpi.com/1999-4907/13/2/264?type=check\\_update&version=2](https://www.mdpi.com/1999-4907/13/2/264?type=check_update&version=2), [Accessed 28-09-2025], 2022.
- [7] R. C. Team, *R: A language and environment for statistical computing*, <https://www.R-project.org/>, R Foundation for Statistical Computing, Vienna, Austria. Accessed September 15, 2025., 2021.
- [8] L. Villanueva, *R package with functions to calculate indices for soundscape ecology and other ecology research that uses audio recordings*, <https://ljvillanueva.github.io/soundecology/>, R package, Accessed September 26, 2025, 2016.
- [9] M. Niininen, *Morse learning machine - v1*, <https://kaggle.com/competitions/morse-challenge>, Kaggle competition. Accessed September 15, 2025., 2014.
- [10] E. Baker and Q. Geissman, *Sonicscrewdriver: Bioacoustic analysis and publication tools*, <https://cran.r-project.org/package=sonicscrewdriver>, R package. Accessed September 15, 2025., 2024.
- [11] U. Ligges and ptjers, *Tuner: Analysis of music and speech*, <https://cran.r-project.org/package=tuneR>, R package. Accessed September 15, 2025., 2024.
- [12] J. Sueur et al., *Seewave: Sound analysis and synthesis*, <https://cran.r-project.org/package=seewave>, R package. Accessed September 15, 2025., 2025.
- [13] M. Younger, *Audiocw: An exploration of morse code (cw) as a learning tool for computer science*, <https://github.com/allhailthetail/AudioCW>, GitHub repository. Accessed September 15, 2025., 2025.
- [14] M. Maynard, *K4icy's cw morse decoder for arduino*, [https://www.k4icy.com/cw\\_decoder.html](https://www.k4icy.com/cw_decoder.html), Accessed September 15, 2025., 2021.
- [15] M. Niininen, *Towards bayesian morse decoder*, <https://agile.blogspot.com/2013/01/towards-bayesian-morse-decoder.html>, Blog post. Accessed September 15, 2025., 2013.
- [16] D. Clink, *R for sound analysis tutorial*, <https://www.denaclink.com/post/20220317b-r-tutorial/>, Accessed September 15, 2025., 2022.
- [17] T. Weaver, *The path to morse code fluency*, <https://longislandcwclub.org/wp-content/uploads/2025/08/The-Path-to-Morse-Code-Fluency.pdf>, Long Island CW Club. Accessed September 15, 2025., 2025.
- [18] T. Weaver, *Executive summary koch 1936*, <https://longislandcwclub.org/wp-content/uploads/2024/01/EXECUTIVE-SUMMARY-KOCH-1936-1.pdf>, Long Island CW Club. Accessed September 15, 2025., 2024.
- [19] Tenacity Community, *Tenacity: An easy-to-use audio editor and recorder*, <https://tenacityaudio.org/>, Accessed September 15, 2025., 2025.
- [20] R. Ihaka, R. Gentleman, et al., *R: Programming language for statistical analysis*, <https://www.r-project.org/>, Accessed September 15, 2025., 2024.
- [21] Free Software Foundation, *Gnu emacs*, Version 29.1. Accessed September 16, 2025., 2023.
- [22] Org mode community, *Gnu org mode: Organize your life in plain text*, Version 9.7.6. Accessed September 16, 2025., 2024.



Project & Abstract



Recorded Presentation