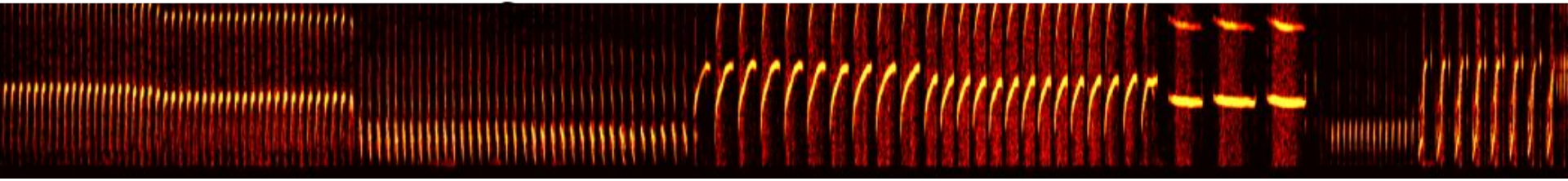


# Automated annotation of animal vocalizations (SciPy 2019 version)

**David Nicholson, Emory University, Biology Dept.**

**NickleDave** **@nicholdav**

# Acknowledgements


Gardner lab
- Yarden Cohen
- Alexa Sanchioni
- Emily Mallaber
- Vika Skidanova

Yarden


Emily


Alexa


Vika

Sober lab
- Jonah Queen


Jonah

hybrid-vocal-classifier + vak contributors: Varun Saravanan (Sober lab), Roman Ursu (Leblois), Bradley Colquitt + David Mets (Brainard), Ammon Perkes + Marc Badger (Schmidt)

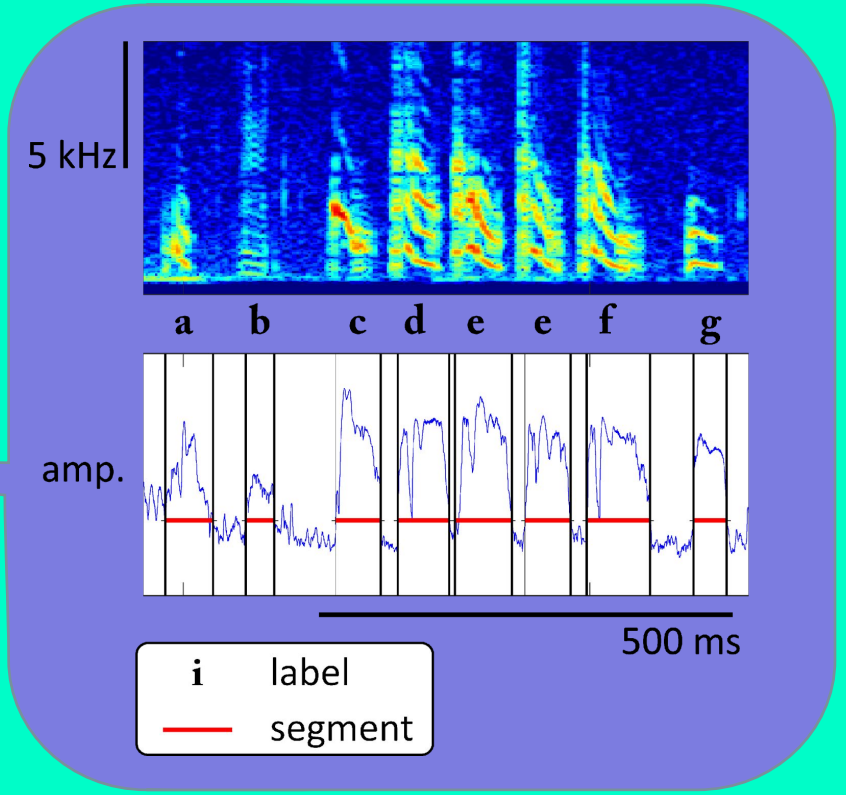
# Introduction

Why automate annotation of vocalizations?

1. save time
2. answer (new) research questions
    a. have all the statistical power
    b. measure things we couldn't measure before

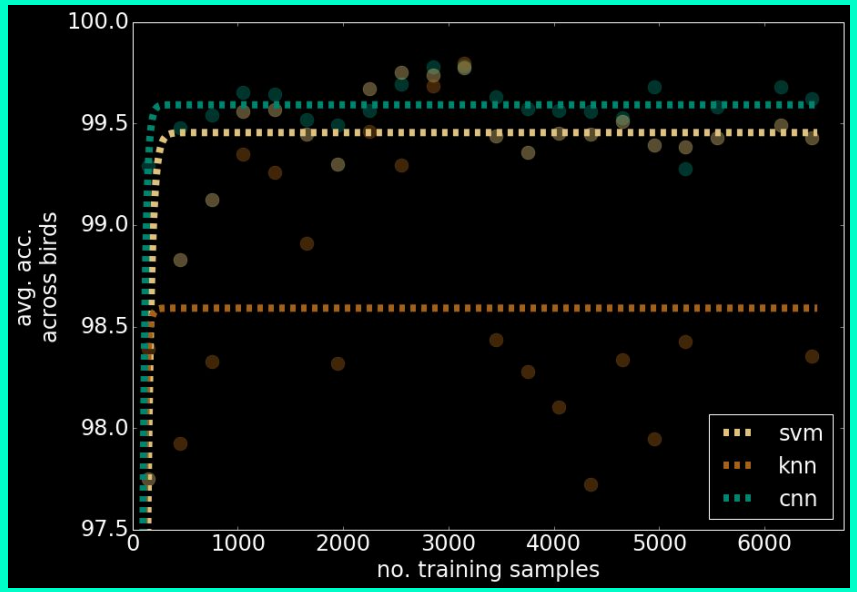
# Introduction

- What do I mean by **annotate vocalizations**?
- Birdsong
  - consists of elements called syllables
  - segment sound file into syllables by threshold crossings of amplitude



# Introduction

**A simple convolutional neural network outperforms other algorithms (SVM, k-NN) that rely on engineered features extracted from audio segments**




**hybrid-vocal-classifier**

github.com/NickleDave/hybrid-vocal-classifier

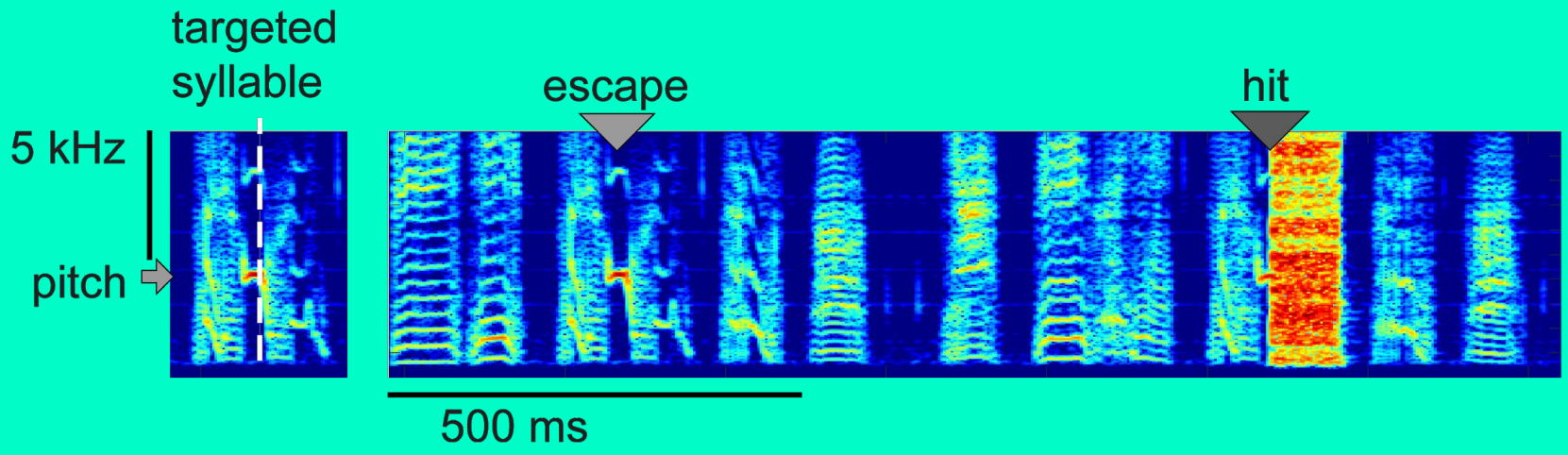
https://hybrid-vocal-classifier.readthedocs.io

https://youtu.be/BwNeVNou9-s

https://github.com/NickleDave/ML-comparison-birdsong

# Introduction

**There are many cases in which segmenting is noisy, e.g. because of actual noises, and so methods that extract features from segments will fail**



# Introduction

What would a good auto-annotater do for us?

| Criterion | Software we developed to meet this criterion |
| --- | --- |
| • **segment** audio into vocalizations (birdsong syllables, speech syllables, whatever)<br>• **predict labels** for segments | **TweetyNet (neural network)** |
| | |
| | |

# Introduction

What would a good auto-labeler do for us?

| Criterion | Software we developed to meet this criterion |
| --- | --- |
| • **segment** audio into vocalizations (birdsong syllables, speech syllables, whatever)<br>• **predict** labels for segments | **TweetyNet (neural network)** |
| • make it easy for anyone to use | **vak (library)** |
| | |

# Introduction

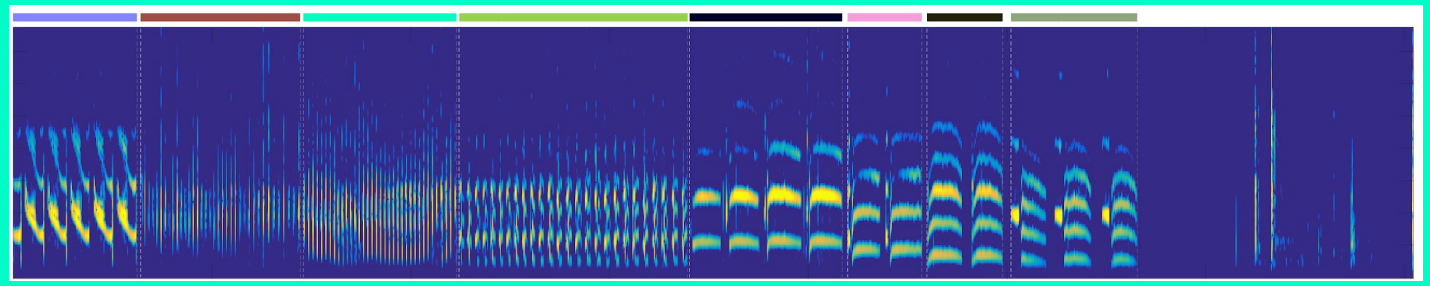
What would a good auto-labeler do for us?

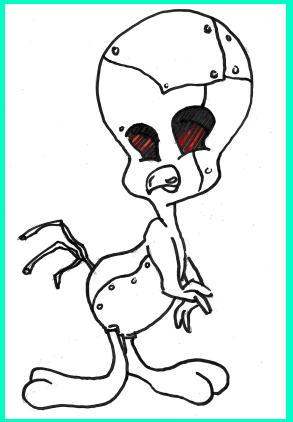
| Criterion | Software we developed to meet this criterion |
| --- | --- |
| • **segment** audio into vocalizations (birdsong syllables, speech syllables, whatever)<br>• **predict** labels for segments | **TweetyNet (neural network)** |
| • make it easy for anyone to use | **vak (library)** |
| • work with many different data formats | **vak, crowsetta (libraries)** |

# Methods

**`TweetyNet:`** a hybrid convolutional-recurrent neural network that segments and labels birdsong and other vocalizations

canary song segmented into phrases

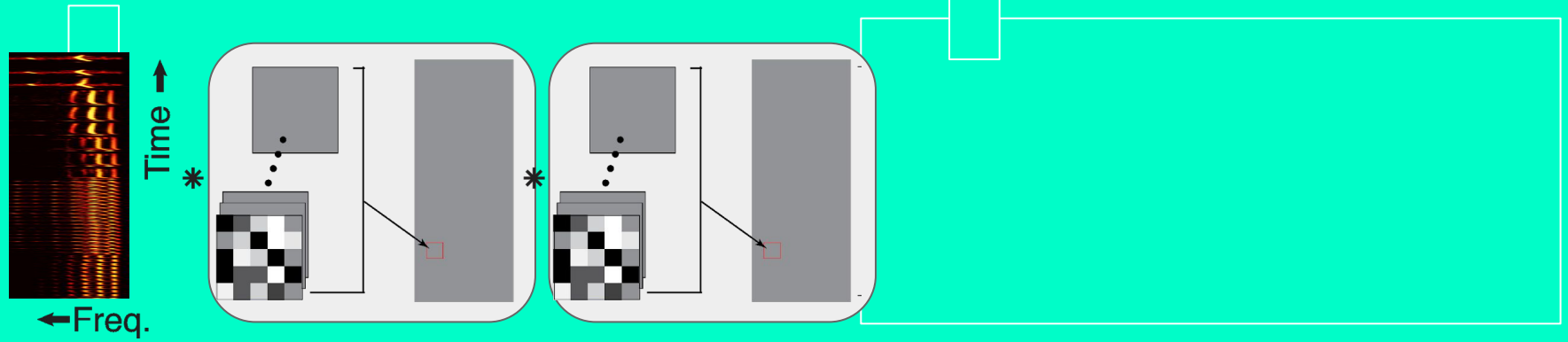




`https://github.com/yardencsGitHub/tweetynet`

# METHODS

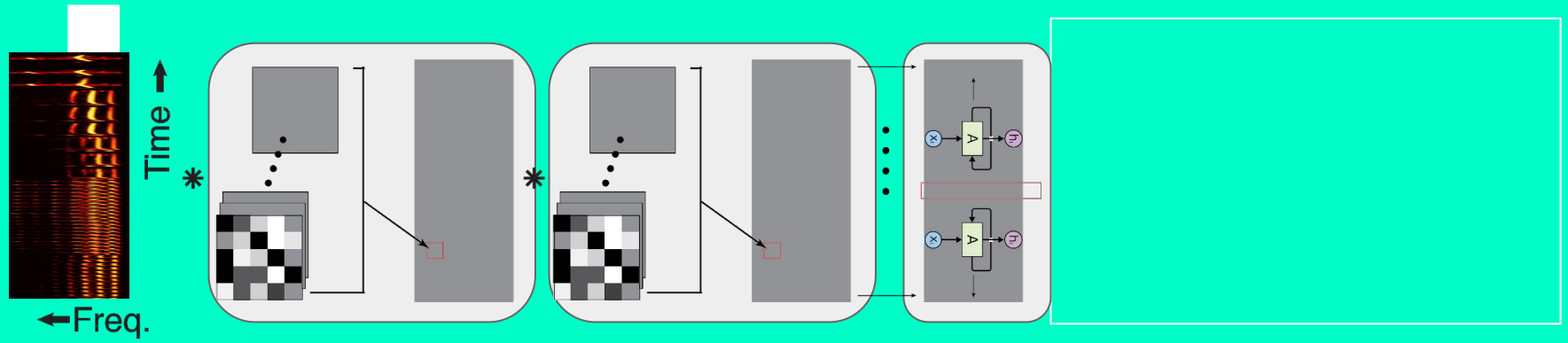
**TweetyNet:** a hybrid convolutional-recurrent neural network that segments and labels birdsong and other vocalizations




https://github.com/yardencsGitHub/tweetynet

# METHODS

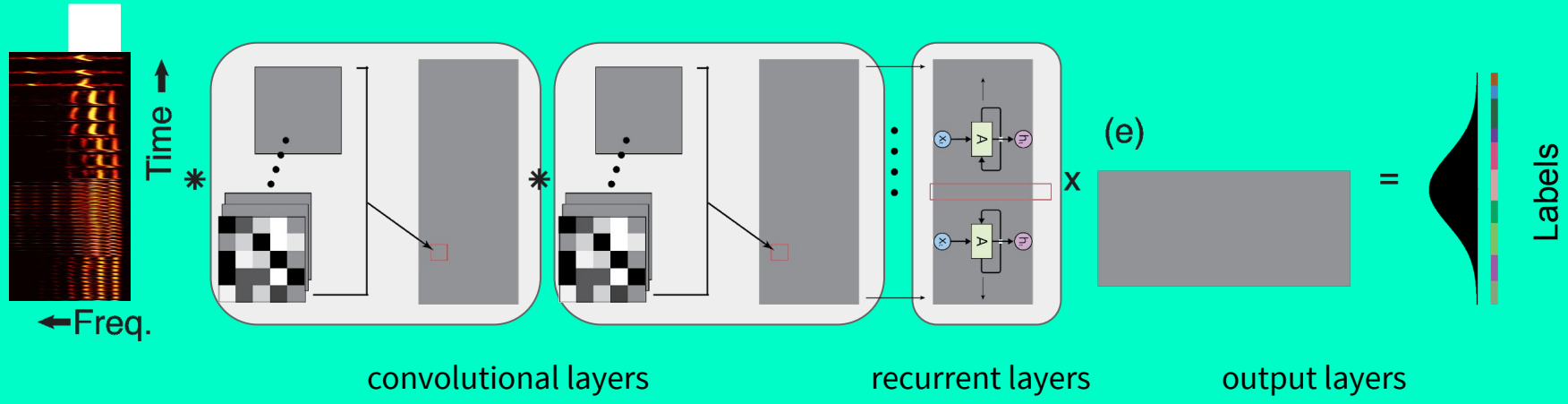
**TweetyNet:** a hybrid convolutional-recurrent neural network that segments and labels birdsong and other vocalizations




https://github.com/yardencsGitHub/tweetynet

# METHODS

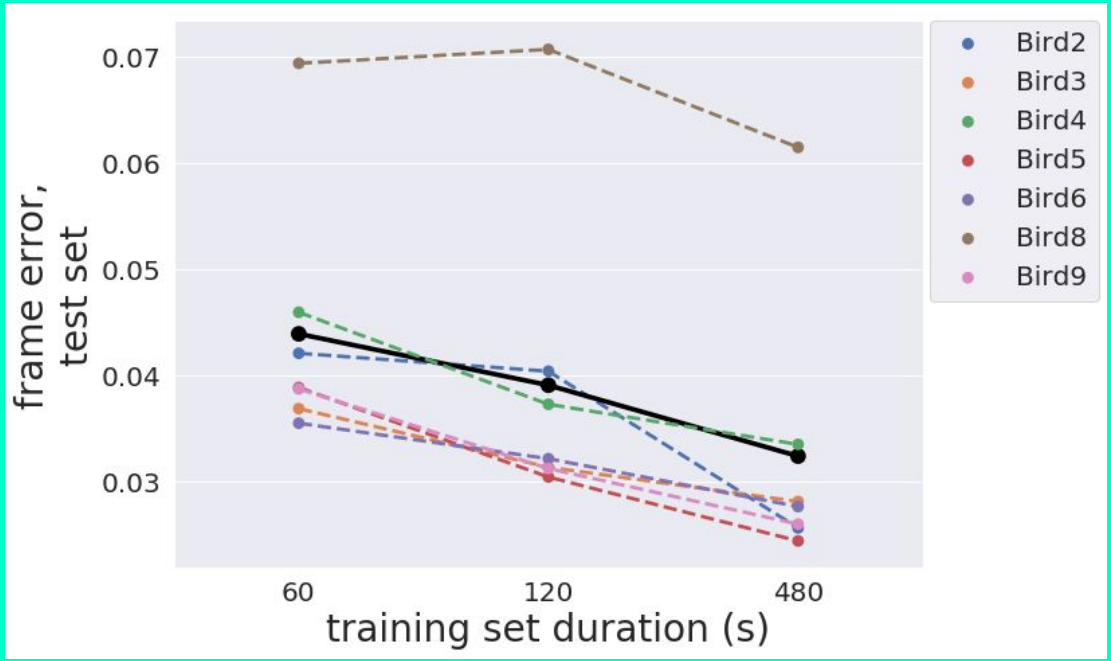
**TweetyNet:** a hybrid convolutional-recurrent neural network that segments and labels birdsong and other vocalizations




https://github.com/yardencsGitHub/tweetynet

# Results

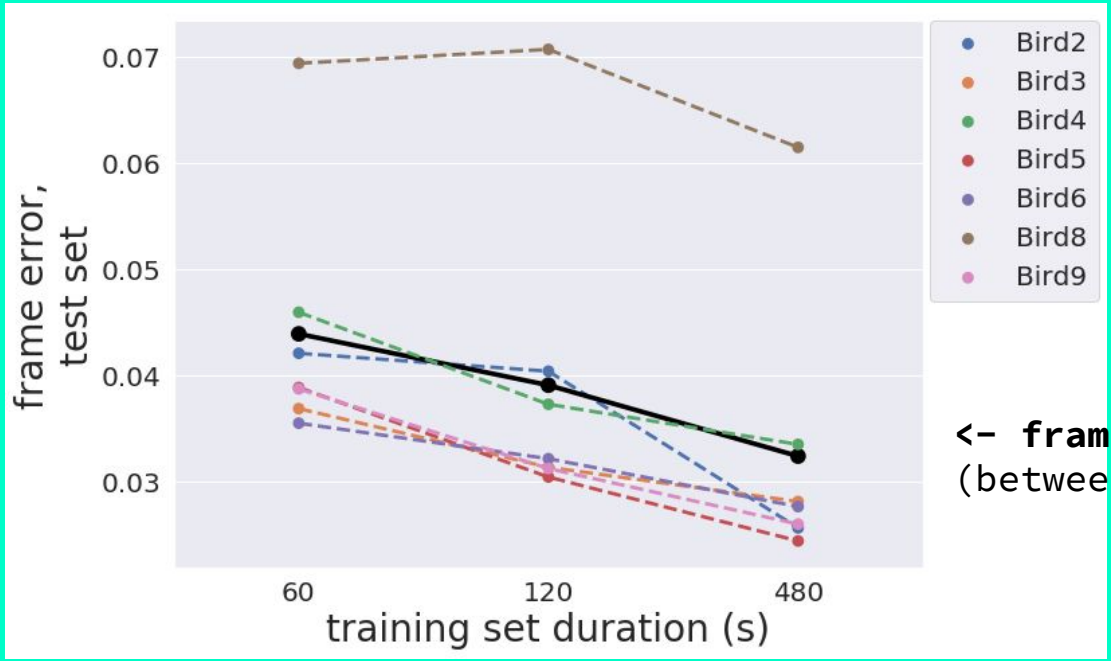
TweetyNet achieves low frame error **across individuals**




dataset: https://figshare.com/articles/BirdsongRecognition/3470165

# Results

TweetyNet achieves low frame error **across individuals**

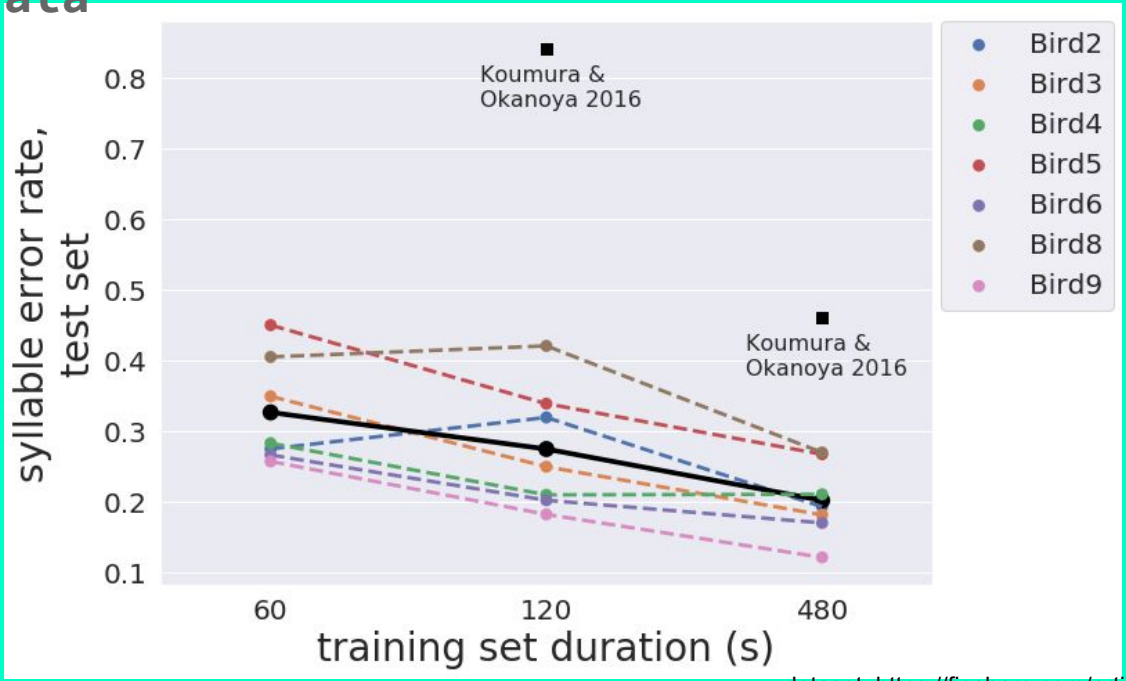



**<- frame error!**
(between 0 and 1)

dataset: https://figshare.com/articles/BirdsongRecognition/3470165

# Results

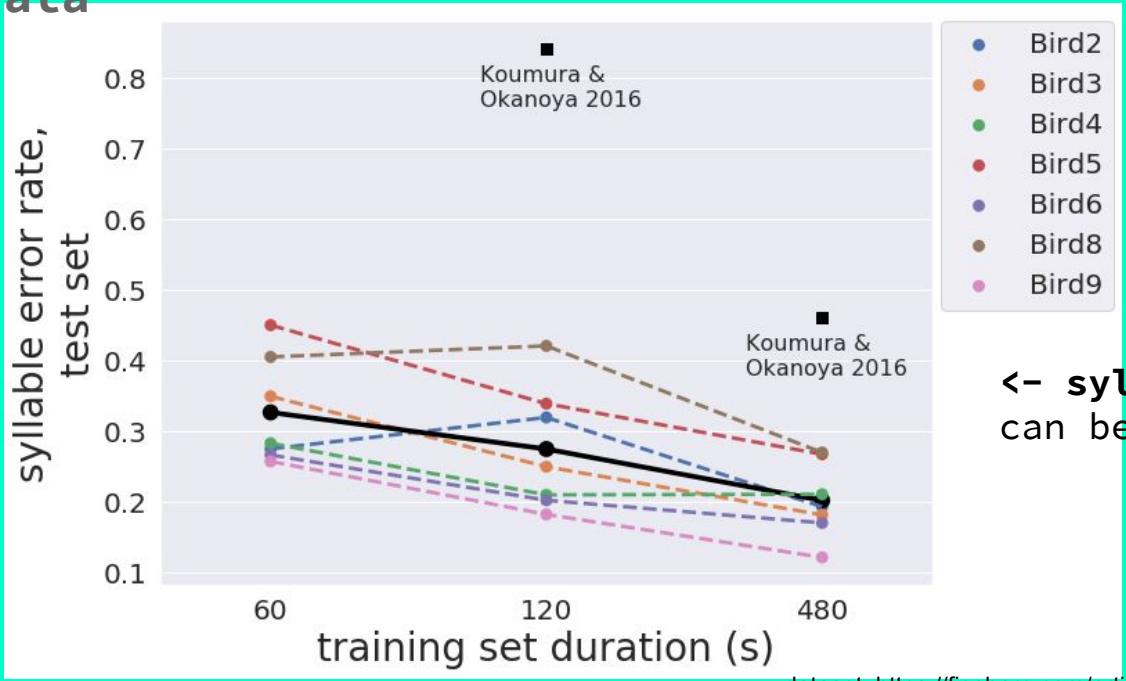
TweetyNet achieves lower syllable error rate with **less training data**




dataset: https://figshare.com/articles/BirdsongRecognition/3470165

# Results

TweetyNet achieves lower syllable error rate with **less training data**

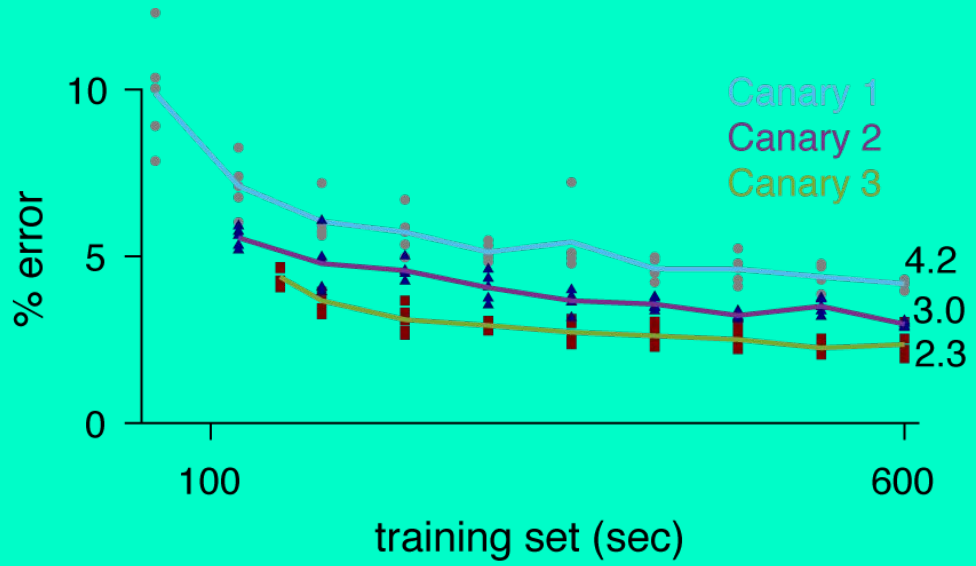



**<- syllable error!**
can be greater than 1

dataset: https://figshare.com/articles/BirdsongRecognition/3470165

# Results

TweetyNet is accurate across large datasets of **canary song with many syllables and lengthy bouts**



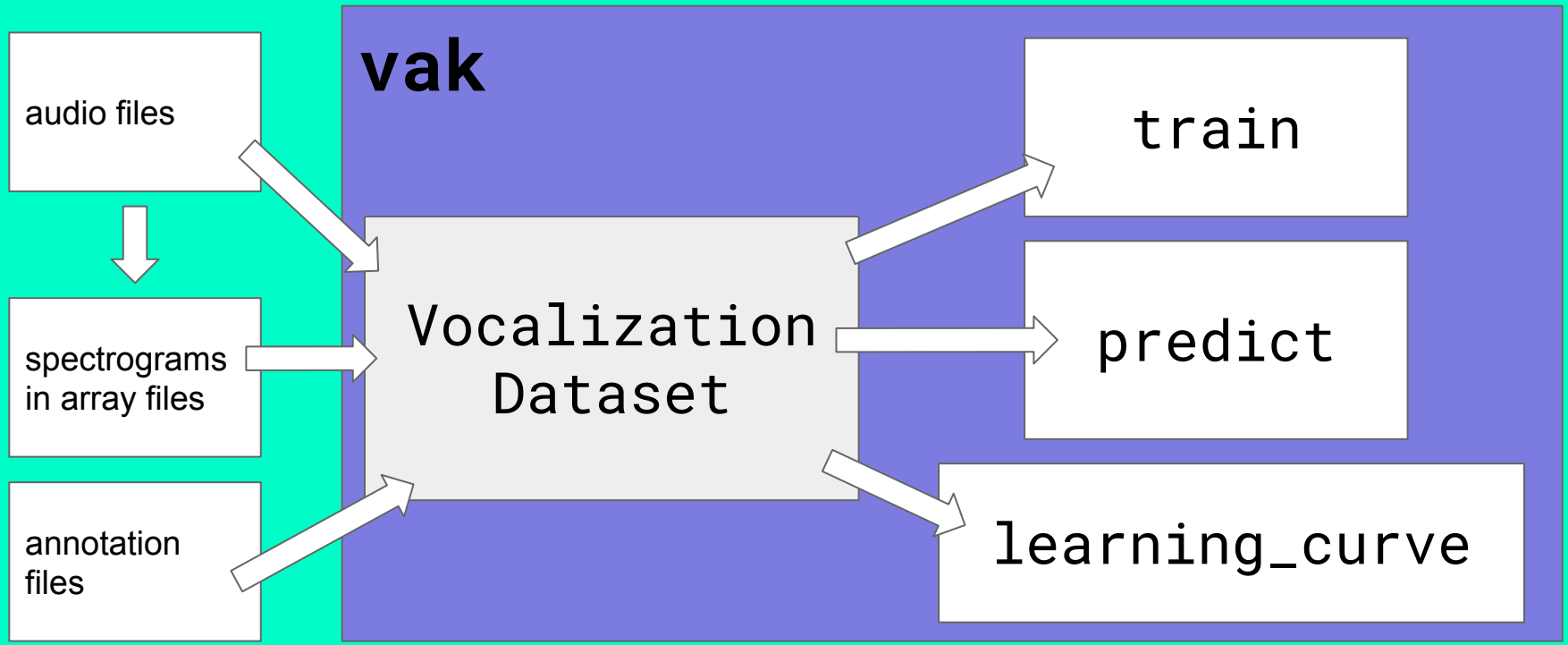
# DISCUSSION

Question:

how do I use TweetyNet? Doing science is already hard enough, I don't want to have to learn how to program neural networks on top of that

# DISCUSSION

**vak:** automated annotation of vocalizations for everybody
https://github.com/NickleDave/vak

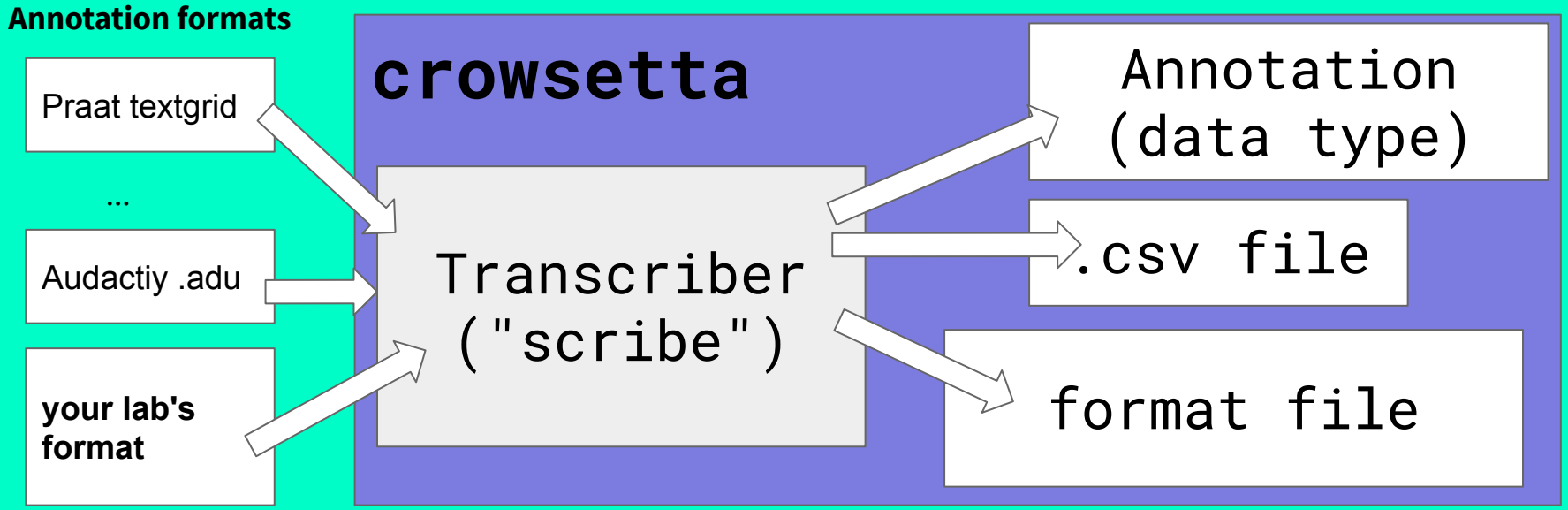


# DISCUSSION

**crowsetta**

A tool to work with any format for annotating vocalizations

**https://crowsetta.readthedocs.io/en/latest/**



# Conclusion

Open-source community-developed tools provide opportunity for high throughput automated annotation of vocalizations

Development on Github

- https://github.com/yardencsGitHub/tweetynet
- https://github.com/NickleDave/vak
- https://crowsetta.readthedocs.io/en/latest/

Next version out in time for Neuroscience 2019 in Chicago!