

ORCASOUND

open source software for
catalyzing conservation of
**Southern Resident killer
whales (in real-time)**

Scott Veirs & Val Veirs on behalf of the
[Orcasound community](#)

sveirs@gmail.com | [Orcasound Slack](#)

orcasound.net/talks



Our orcas: the southern salmon seekers

Southern Resident Killer Whales (SRKWs) are:

- **Southern** = ranging from northern California to SE Alaska
- **Resident** = historically re-occurring within the Salish Sea (inland waters of WA and BC)
- **Killer** = apex predators, salmon specialists
- **Whales** = cultural icons, both historically & as modern “charismatic megafauna”



How can we collaborate across State and International boundaries to conserve these orcas?

Our orcas: recovery by a 1000 actions

Cumulative human impacts

NOAA recovery plan

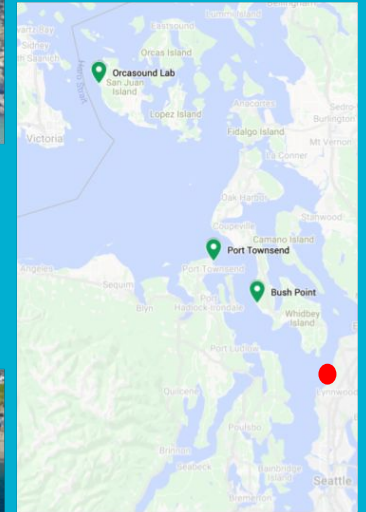
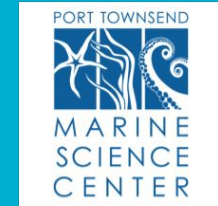
- 3 key risks:
 - Scarce salmon
 - Persistent pollutants in blubber
 - Vessel effects
- Vessel noise impacts:
 - Ship noise masks communication & echolocation
 - Boat interference lowers foraging efficiency



How can AI & human listeners work synergistically to conserve SRKWs (in real-time)?

How? cooperate as a hydrophone network

- 2002: Orcasound Lab (Val's "back yard")
- 2008-12 NOAA funding (expansion to 5 nodes)
- 2013-15 Philanthropy only (decline to 2 nodes)
- 2016-now Cooperative network (17 NGOs in 2021)



\$\$

- 2017 crowd-funding (\$20k Kickstarter)
- 2018 open-sourcing, crowd-sourcing, open data...
- 2020+ hackathons, philanthropy, Amazon+Microsoft cloud credits

How? collaborate openly in soft/hard-ware

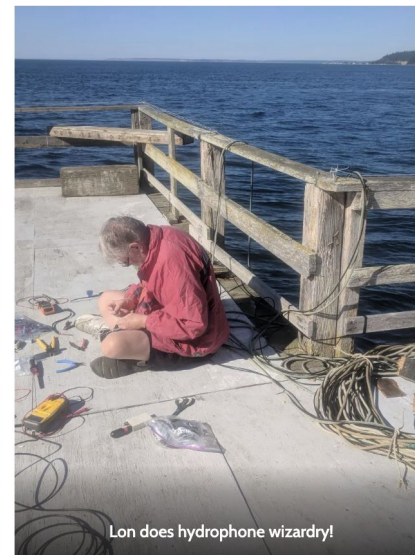
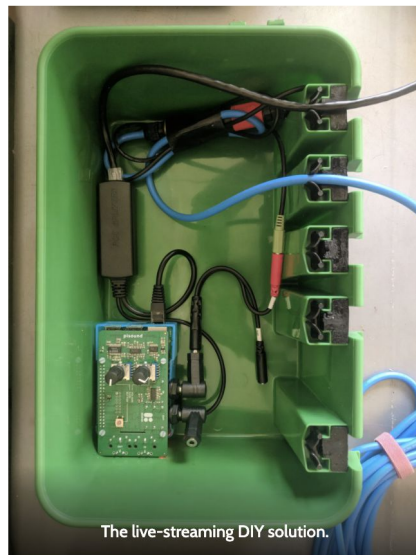
<2002-17

- humans listening via Shoutcast mp3 streams
- Val building custom software, alone
- Scott building web static web sites, Google sheets, manual Twitter/email notifications...

2017-18: software + hardware (Kickstarter for v1 web app), live.orcasound.net launched Nov '18

2019-20: v2 UI beta-tested in Nov '19, launched in May, 2020

2021: v3 UI + a proliferation of related projects (24 public [Orcasound Github repositories](#))

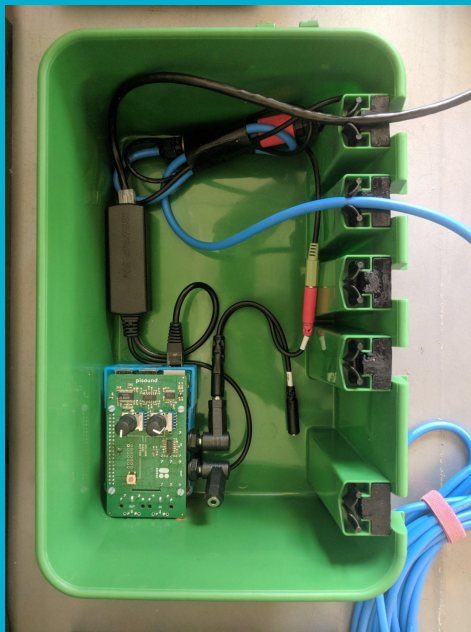


ORCASOUND

Easy to listen live. Scales inexpensively. Lets user tag data.

- 3 cabled nearshore sites streaming 24/7 (2018-21)
- Streaming hardware cost: ~U\$300 per node
- Hydrophone \$300-1500
- [PiSound HAT](#) = 2 channel, 24 bit, 48/96/192 kHz
- Free open source code!

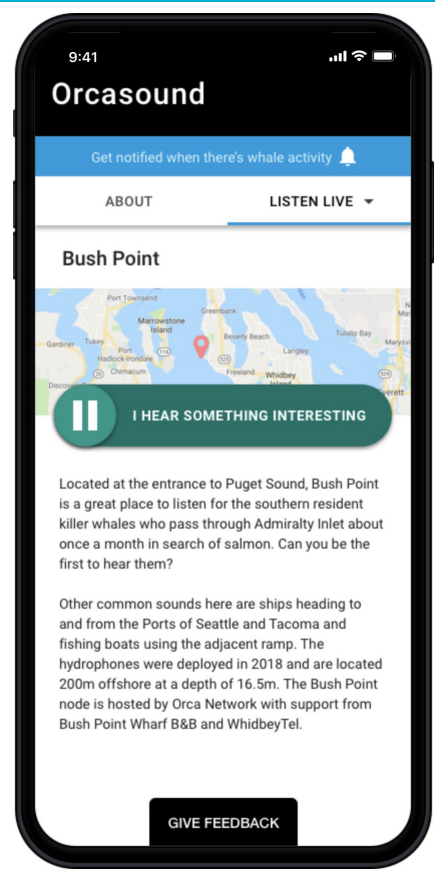
Community scientists detect **orca** & **novel sounds** in real-time via a web app -- live.orcasound.net



Orca sounds

[orcanode code](#)

[orcasite code](#)



How? open access raw data

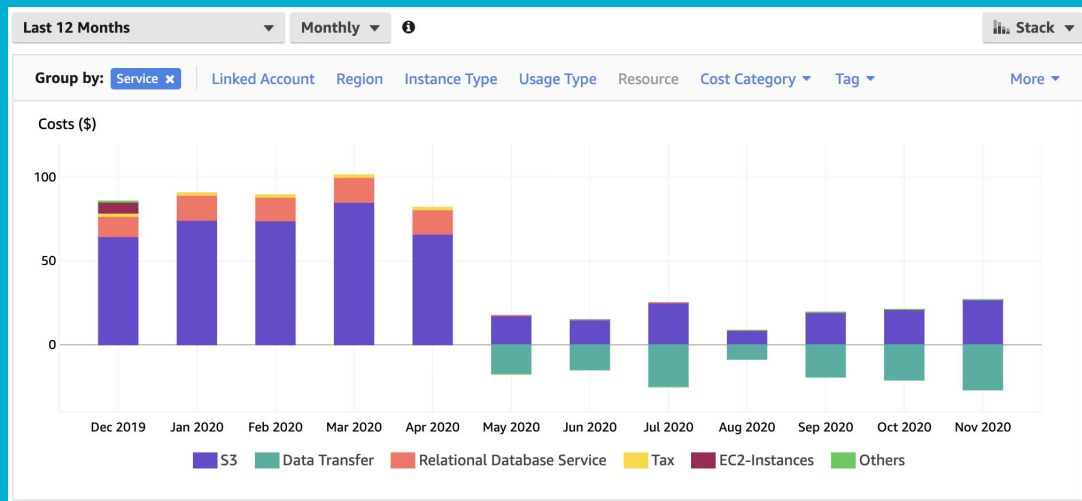
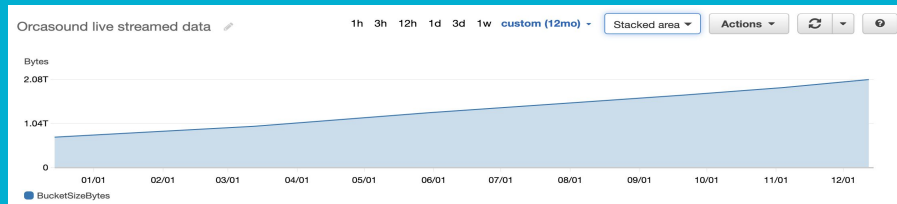
```
aws --no-sign-request s3 sync streaming-orcasound-net
```

Data volume

- 3.2 TB (HLS audio only)
- 0.4 TB/yr/node
- Increasing soon...
...FLAC, 48/96 kHz

Data costs

- <\$100/mo w/3 nodes
- ~30\$/mo/node, but scalable
- Free thus far with cloud credits!
- [AWS open data registry](#) + [Quilt](#)



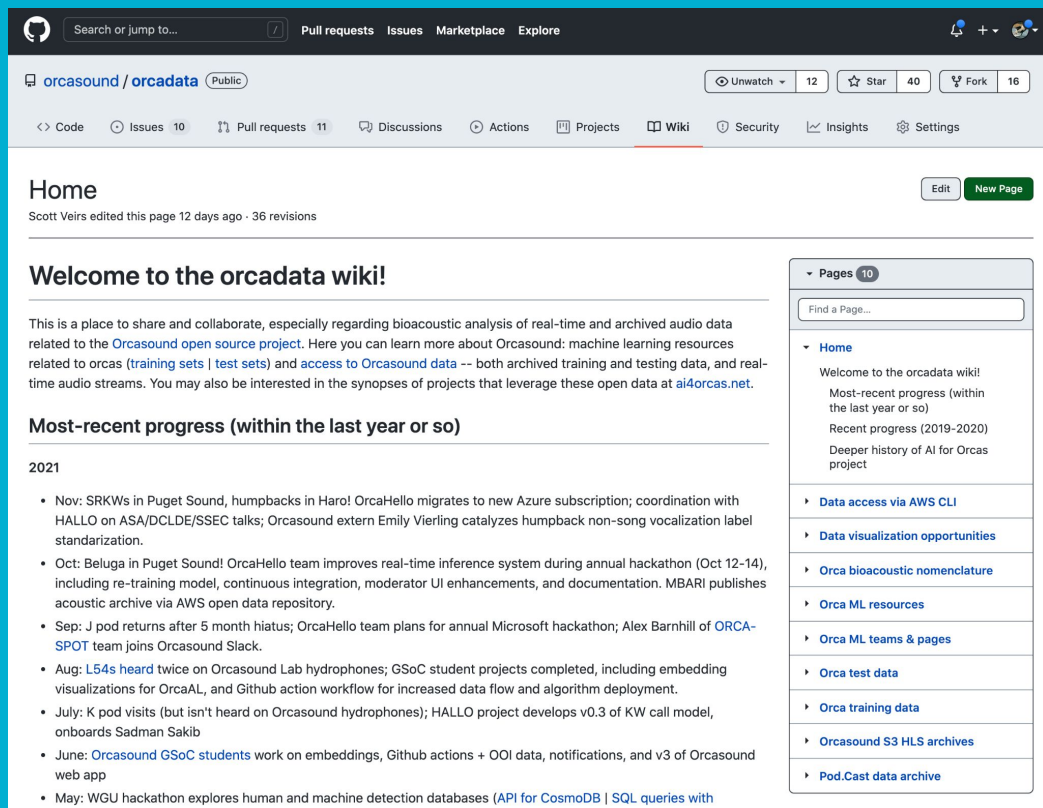
How? open labeled data & labeling tools

Labeled data ([orcadata wiki](#))

- SRKW calls
- 2022: signals of other SRKW sounds, Bigg's KW, and humpback

Labeling tools

- 2000 Orcasound human labels (~100/month, free text)
- Pod.Cast (2019)
- OrcaHello tags (2020+) w/D. Bain
- Orca Active Learning (2020+)
- Audacity (expert labeled test sets)
- 2022: HALLO annotation tool



The screenshot shows the GitHub repository page for 'orcadata/wiki'. The repository is public and has 12 unwatched items, 40 stars, and 16 forks. The 'Wiki' tab is selected, showing the 'Home' page. The page content includes a welcome message, a description of the project's purpose (sharing and collaborating on bioacoustic analysis), and a list of 'Most-recent progress' items for the year 2021. A sidebar on the right contains a search bar and a list of navigation links.

Search or jump to... Pull requests Issues Marketplace Explore

orcasound / **orcadata** (Public) Unwatch 12 Star 40 Fork 16

<> Code Issues 10 Pull requests 11 Discussions Actions Projects Wiki Security Insights Settings

Home

Scott Veirs edited this page 12 days ago · 36 revisions

Welcome to the orcadata wiki!

This is a place to share and collaborate, especially regarding bioacoustic analysis of real-time and archived audio data related to the [Orcasound open source project](#). Here you can learn more about Orcasound: machine learning resources related to orcas ([training sets](#) | [test sets](#)) and [access to Orcasound data](#) -- both archived training and testing data, and real-time audio streams. You may also be interested in the synopses of projects that leverage these open data at [ai4orcas.net](#).

Most-recent progress (within the last year or so)

2021

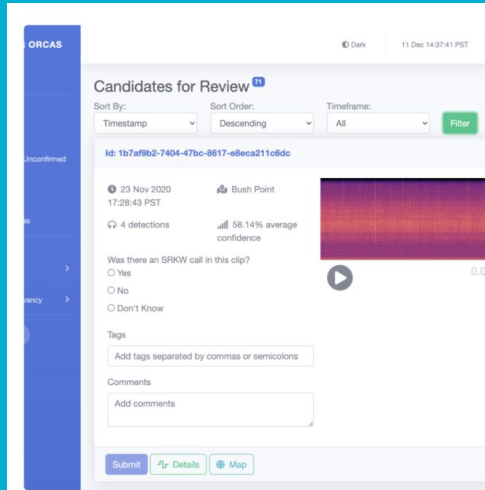
- Nov: SRKWs in Puget Sound, humpbacks in Haro! OrcaHello migrates to new Azure subscription; coordination with HALLO on ASA/DCLDE/SSEC talks; Orcasound extern Emily Vierling catalyzes humpback non-song vocalization label standardization.
- Oct: Beluga in Puget Sound! OrcaHello team improves real-time inference system during annual hackathon (Oct 12-14), including re-training model, continuous integration, moderator UI enhancements, and documentation. MBARI publishes acoustic archive via AWS open data repository.
- Sep: J pod returns after 5 month hiatus; OrcaHello team plans for annual Microsoft hackathon; Alex Barnhill of ORCA-SPOT team joins Orcasound Slack.
- Aug: L54s heard twice on Orcasound Lab hydrophones; GSoC student projects completed, including embedding visualizations for OrcaAL, and Github action workflow for increased data flow and algorithm deployment.
- July: K pod visits (but isn't heard on Orcasound hydrophones); HALLO project develops v0.3 of KW call model, onboards Sadman Sakib
- June: [Orcasound GSoC students](#) work on embeddings, Github actions + OOI data, notifications, and v3 of Orcasound web app
- May: WGU hackathon explores human and machine detection databases ([API for CosmoDB](#) | [SQL queries with](#)

Pages 10

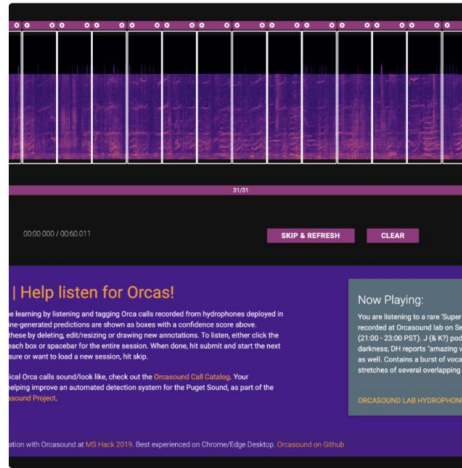
Find a Page...

- ▼ Home
 - Welcome to the orcadata wiki!
 - Most-recent progress (within the last year or so)
 - Recent progress (2019-2020)
 - Deeper history of AI for Orcas project
- ▶ [Data access via AWS CLI](#)
- ▶ [Data visualization opportunities](#)
- ▶ [Orca bioacoustic nomenclature](#)
- ▶ [Orca ML resources](#)
- ▶ [Orca ML teams & pages](#)
- ▶ [Orca test data](#)
- ▶ [Orca training data](#)
- ▶ [Orcasound S3 HLS archives](#)
- ▶ [Pod.Cast data archive](#)

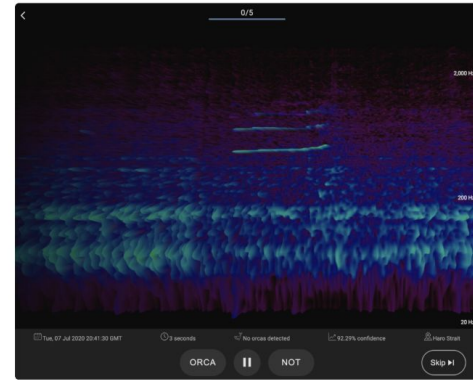
Artificial Intelligence for orcas



OrcaHello live inference system



Pod.Cast annotation system



Orca Active Learning

AI for orcas (#ai4orcas) -- ai4orcas.net -- [OrcaHello](#) | [Pod.Cast](#) | [OrcaAL](#)

towards (more) open (marine) bioacoustic data science...

AI+human detection is optimal

For many end-users, expert validation of acoustic &/or visual detections is still important.

2021 results:

- 71 hr SRKW bouts
- 30 hr Bigg's bouts
- 18 hr humpback bouts
- Of 41 known SRKW transits:
AI detected 54%; humans 81%; and 100% combined
- AI outage in Oct missed 12 (31%)

The screenshot shows the 'AI FOR ORCAS' web application interface. The top navigation bar includes a 'Dark' mode toggle, the date '11 Dec 14:37:41 PST', and a user profile for 'sveirs'. The main content area is titled 'Candidates for Review' with a notification badge for 71 items. Below the title are filters for 'Sort By' (Timestamp), 'Sort Order' (Descending), and 'Timeframe' (All), along with a 'Filter' button. The candidate entry has the ID '1b7af9b2-7404-47bc-8617-e8eca211c6dc'. It shows a timestamp of '23 Nov 2020 17:28:43 PST', a location 'Bush Point', and '4 detections' with a '58.14% average confidence'. A spectrogram is displayed with a play button and a duration of '0.00 / 00.00'. Below the spectrogram is a question: 'Was there an SRKW call in this clip?' with radio button options for 'Yes', 'No', and 'Don't Know'. There is a 'Tags' section with a text input field 'Add tags separated by commas or semicolons' and a 'Comments' section with a text input field 'Add comments'. At the bottom, there are buttons for 'Submit', 'Details', and 'Map'.

Technical and scientific challenges

Where are they now?

Applied conservation problems

Can we automate reliable real-time notification of SRKW presence, or will humans remain in the loop?

1. Notify as soon as possible, or only when “sure” of SRKW?
2. What is end-user’s false positive tolerance?
3. The conservation challenge: how to inspire users to act for SRKWs when empathy is maximized during a live acoustic event?

What are they saying?

Basic biology problems

Bioacoustics topics our data can inform:

- marine mammal communication systems
- biosonar
- marine acoustic ecology
- soundscape analysis

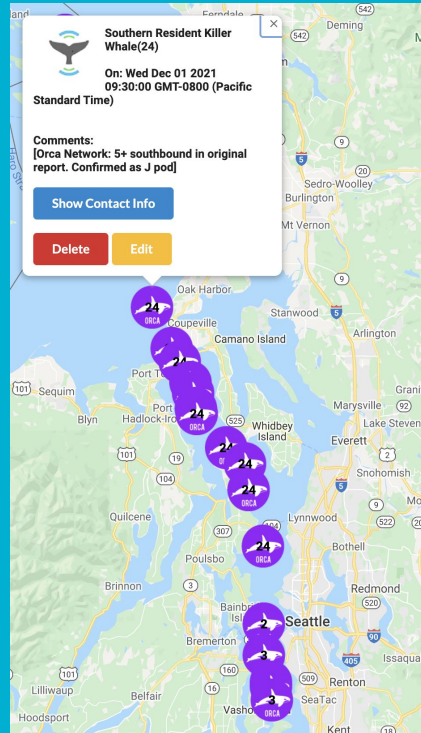
Lots of visualization opportunities!

Sighting & listening together

Great synergies lie in integrating acoustic and visual real-time observations!

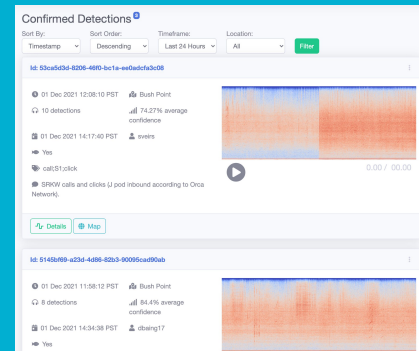
2022 goals:

- Data cooperative for sharing:
 - vetted sightings
 - verified human detections
 - moderated OrcaHello detections
- Public API & Creative Commons license
- Notifications API
- [Orca Network Facebook group](#)



Detections

ID	Node	Listeners	Description	Timestamp
4086	bush-point	94	Dolphin or whale clicks and whistles	12/1/2021, 12:03:17 PM
4085	bush-point	80	Orcas!!!	12/1/2021, 12:01:35 PM
4084	bush-point	80	Orca	12/1/2021, 11:59:18 AM
4083	bush-point	76	Orca	12/1/2021, 11:57:21 AM
4082	bush-point	63	Orca calls and echolocating at 11:58.	12/1/2021, 11:56:51 AM
4081	bush-point	56	Squeaking & clicking	12/1/2021, 11:55:58 AM



2022 challenge: Saving SRKW's & salmon

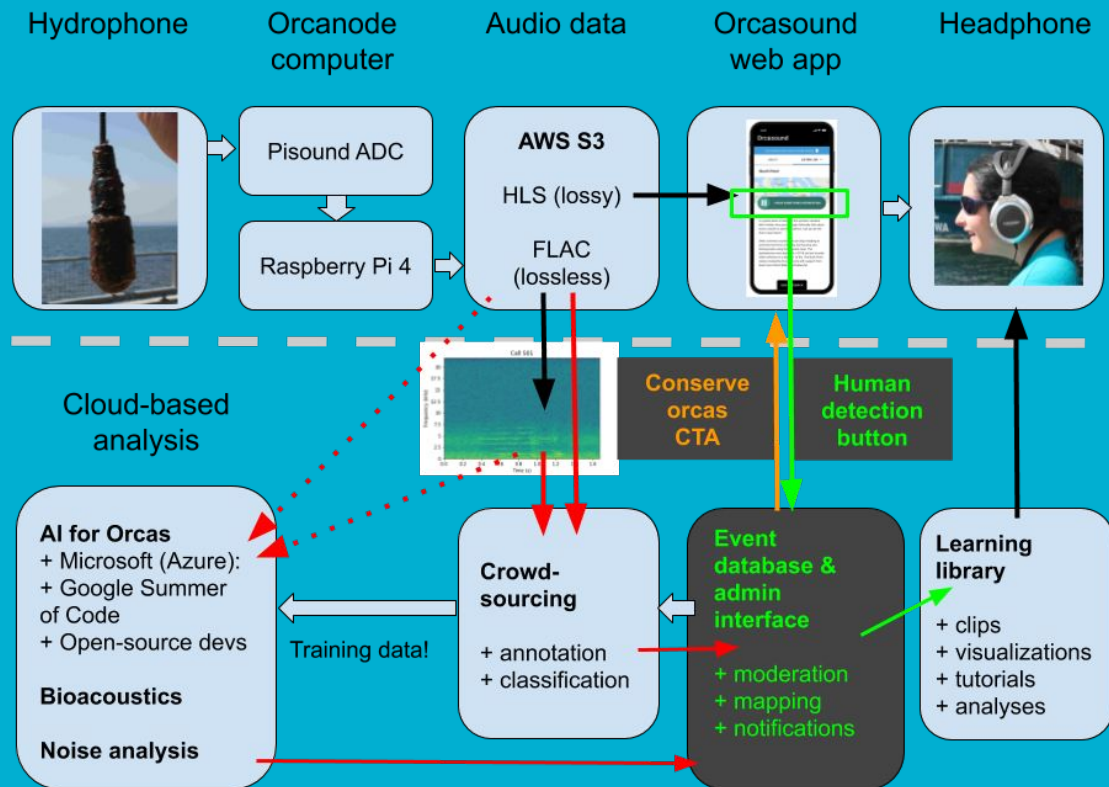
Will users label, heed a

Conservation Call To Action, or both?

- 1800 subscribers
- Baseline survey via subscription form
- Conservation effect survey

Guiding principles:

- User-centered design
- Open roadmap (Trello)



Grand challenge: sharing solutions globally

There are lots of soniferous species that need a voice!

Expansion possibilities for our open source solutions:

- More SRKW nodes (U.S.)
- SRKW nodes in Canada
- Separate networks in [other regions of the world \(map\)](#) for other species? (belugas?!)



Acknowledgements & links

Give orcas a voice!
live.orcasound.net

Thanks to all our collaborators!

- The [Orcasound open source community's volunteer hackers](#)
- The many NGOs & volunteers who maintain the hydrophones at each node
- [Google Summer of Code](#)
- [Microsoft AI for Earth](#)

More info:

- orcasound.net
- ai4orcas.net
- github.com/orcasound



Extra slides...

From other Orcasound talks...

Extra discussion topics:

Practical detection vs miss rates

From candidate table

“False alarm”

Talk about false positive rate (10 per day)

- About right winter 2021
- Uncomfortable in spring (pigeon guillemot)
- Too low in summer (ACI bugs)
- Still too low/slow in fall?
- Just right this winter?

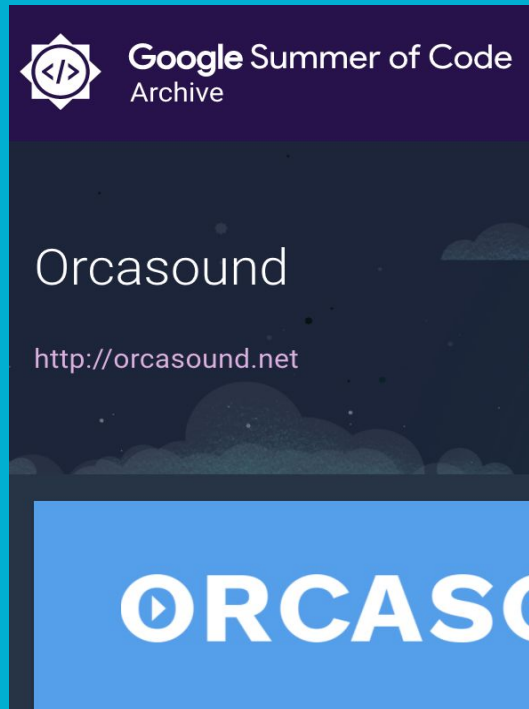
User info

Adrian plots (e.g. subscriber growth?)

Geographic partitioning

Orcasound & Google Summer of Code

- GSoC: supports students & open-source software projects, like Orcasound
- 2019: OrcaCNN for Alaskan killer whale calls
- 2020: Build a tool to speed up the labeling of SRKW calls

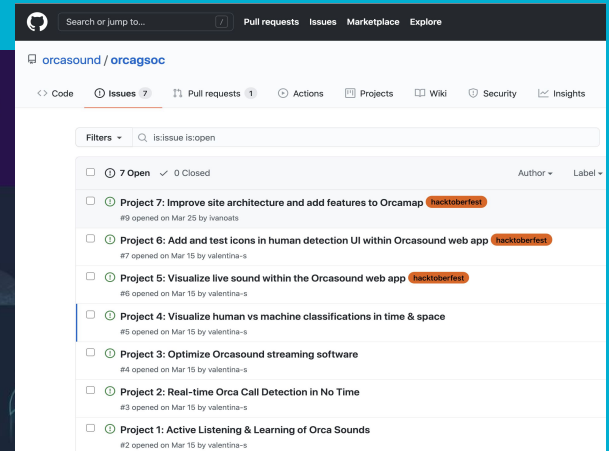


Google Summer of Code Archive

Orcasound

<http://orcasound.net>

ORCASOUND



Search or jump to...

Issues Marketplace Explore

orcasound / orcagsoc

Code Issues Pull requests Actions Projects Wiki Security Insights

Filters is:issue is:open

7 Open 0 Closed Author Label

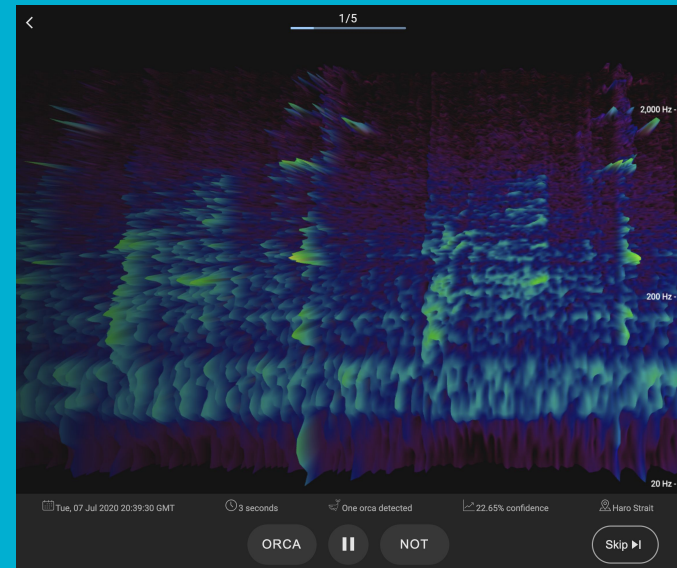
- Project 7: Improve site architecture and add features to Orcamap **hacktoberfest**
#9 opened on Mar 25 by ivanoits
- Project 6: Add and test icons in human detection UI within Orcasound web app **hacktoberfest**
#7 opened on Mar 15 by valentina-s
- Project 5: Visualize live sound within the Orcasound web app **hacktoberfest**
#6 opened on Mar 15 by valentina-s
- Project 4: Visualize human vs machine classifications in time & space
#5 opened on Mar 15 by valentina-s
- Project 3: Optimize Orcasound streaming software
#4 opened on Mar 15 by valentina-s
- Project 2: Real-time Orca Call Detection in No Time
#3 opened on Mar 15 by valentina-s
- Project 1: Active Listening & Learning of Orca Sounds
#2 opened on Mar 15 by valentina-s

Orca active learning (OrcaAL)

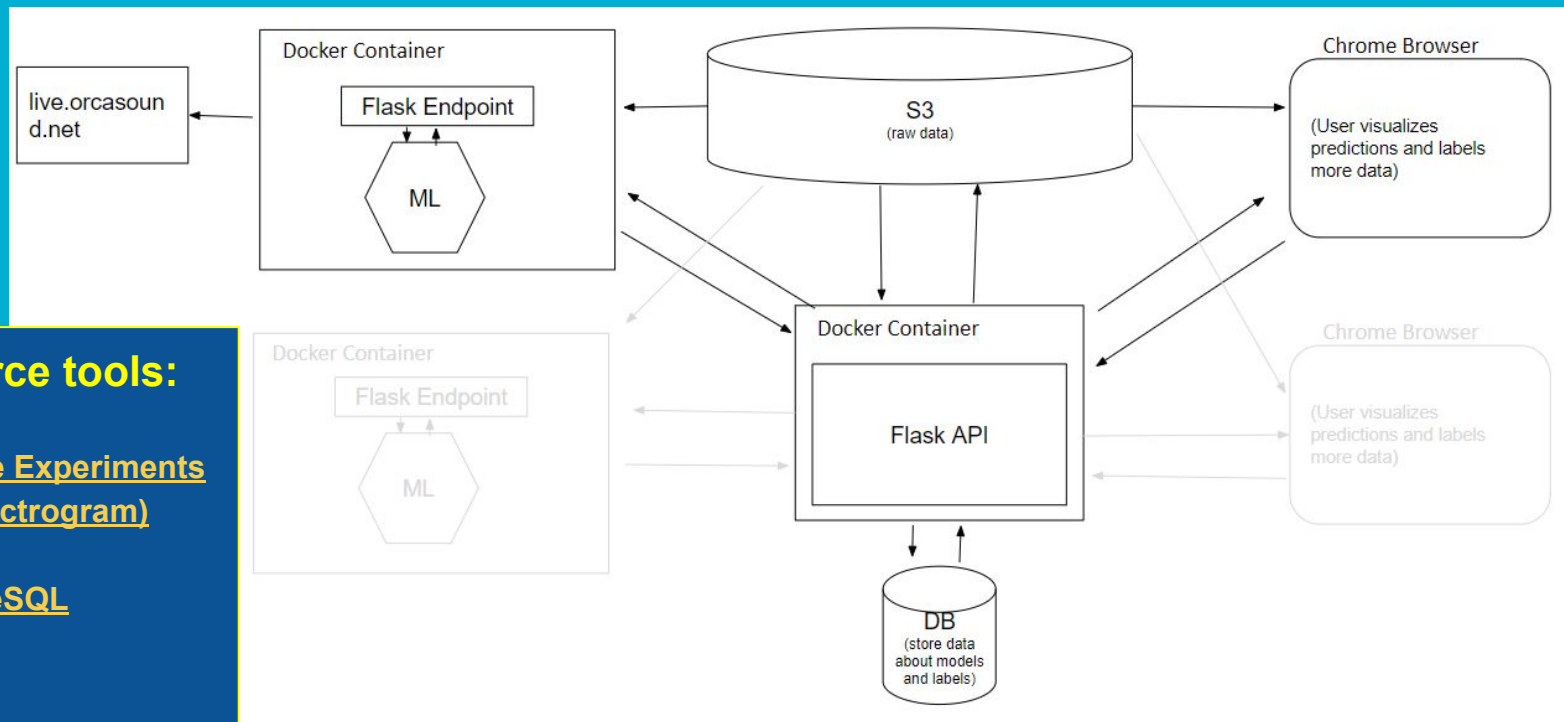
- Train a deep learning algorithm on an initial small labeled dataset
- Use “uncertainty sampling” strategy to label only those samples for which the model is most uncertain (i.e. samples with confidence near 0.5)
- Integrate model training, sample selection, and annotation in an AL tool.

Motivating question: How much does labelling a subset of the samples increase accuracy?

Demo of the new citizen-science labeling tool!



OrcaAL's architecture



Open-source tools:

- [Chrome Experiments \(3D spectrogram\)](#)
- [Docker](#)
- [PostgreSQL](#)
- [Flask](#)