

Danelle Cline Sorted By Project

300057 - Mooring Maintenance

- Attended meeting preparing for the June OASIS3 mooring turn. It looks like the instrument is pulled together now for testing. Started on automated release to version OASIS changes. (Week 19)
- Weekly meetings with Rich. OASIS turn is coming up quickly. (Week 22)

500055 - Video Annotation DM

- On call while Brian was on vacation. All was quiet with no issues. (Week 22)

901009 - CANON

901200 - AMP

901502 - Passive Acoustic Sensing

- Project kick-off meeting (Week 1)
- Decimated 512 kHz data from the 2020 experiment. This will fill in holes in the time series for October. (Week 1)
- Setup replacement HTCondor VM that reached its end of life. The latest version of HTCondor is incompatible with older HTCondor versions in the pool so more work to be done to bring the compute pool back to life. (Week 1)
- Reviewing COVID Marine Frontiers paper (Week 1)
- Testing new decimation with optimized filters (the third iteration) that the team has converged on with the help of Yanwu, Paul McGill, and John Ryan. (Week 2)
- Standing on the pillar of the two giants Google and AWS, working with John to test data access in the AWS Data Sharing program with a Colab notebook. Jupyter notebooks will come next. (Week 2)
- Updating wav metadata generator in preparation for adding it to soundscape workflow. Computed wav metadata, including time corrections for all 5 years of recordings. (Week 2)
- Added compute VM gizo.shore.mbari.org to the HTCondor compute pool. (Week 2)
- Verified Matlab overlap-add decimation algorithm correctness and converted to python production code. Added metadata for AWS open data sharing. (Week 5)
- Downloaded Blue whale D call results and aggregating (Week 5)
- Team meeting (Week 5)
- Decimation optimization work continues with John and Yanwu for the AWS Open Data Sharing program. It looks like the Matlab code is not handling timekeeping correctly. (Week 8)
- Restarted the soundscape live-stream (Week 8)
- AWS open data sharing scripts to upload 2018-2020 data for COVID paper requirements (Week 9)
- Debugging 1-second offset decimation scripts for 2020-08 data (Week 9)
- Uploaded 2018-2020 2kHz data set and working on 16kHz data sets (Week 11)
- Started on python notebooks for upcoming ESMO time-series workshop. (Week 11)
- Intern selections (Week 11)
- All 100 TB or so of data has been uploaded to AWS (Week 19)
- Upgraded recording computer to Windows 10 (thank you IS) and syncing resumed (Week 19)
- Working on tutorials notebooks for AWS data sharing (Week 19)
- Presentation planned for Soquel High School next week (Week 19)

- COVID paper is available through open access. Reduction of Low-Frequency Vessel Noise in Monterey Bay National Marine Sanctuary During the COVID-19 Pandemic. (Week 22)
- Summer intern planning (Week 22)
- Porting classifier code to train in SageMaker (Week 22)
- John and Danelle gave an outdoor COVID safe presentation to Soquel High School in Santa Cruz. Special thanks to Dave French for rebuilding the mobile sound exhibit which we gave a test drive at this event. (Week 22)
- (Week 22)
- 2022 proposal (Week 27)
- Worked on AWS batch process testing of both Cloudformation-based ECS + SimpleQueueService (SQS) and AWS Batch + Fargate. (Week 27)
- Moved the live stream over to Icecast per changes in the live streaming service we use for improved performance (Week 44)
- Responding to licensing and general data inquiries on AWS open data access; we are getting a number of external groups interested in discussing how to facilitate sharing models, training data, etc. (Week 44)
- Testing streamable PCEN detection on blue D calls (Week 44)
- Looked at AWS S3 storage lens results for monitoring open data use. Free tier has limitations, however, so may need to upgrade for more detailed reporting. .022 cents per GB. (Week 48)
- Running clusterGAN model sweeps on blue D calls to separate training data by pose (Week 48)
- Waiting for blue A call detections to finish to run batch classification; AWS Batch queues are setup and ready to go (Week 48)
- EMSO participant follow-up discussion on potential CNN model/training data sharing (Week 48)
- Working with J.Ryan and W. Oestrich to analyze potential fin 40 Hz labels generated from clustering. Continued work on cluster sweeps to find the best parameters. (Week 51)
- Blue A calls are almost completed and ready for AWS batch processing. (Week 51)

901602 - Data TAG

901603 - Video TAG: Visualizing Ocean Change

901915 - Improving MBARI Data Impact

902005 - VARS Annotation Assistance

- Installing GraphQL API to serve Visual Event Tracks and other machine learning data to the new CentOS7 VM deepsea-ai.shore.mbari.org. This has uncovered some unfortunate truths about the limitations of CentOS as a replacement OS for MBARI. (Week 1)
- Updating deepsea-ai GraphQL API to correctly reference VARs video references and adding query and mutation capability as needed for Cthulu testing. (Week 2)
- Deepsea-ai.shore.mbari.org setup is pending request to move to Ubuntu and docker (instead of CentOS7 and podman which has been fraught with issues). (Week 2)
- After countless interruptions to fix an issue with wandb logging which interfered with SPOT training, I think I've finally solved it and I'm ready to do cloud training. (Week 2)
- Project planning (Week 5)
- Onboarding discussion with Carlos Rueda (Week 5)
- Working on week one of a two-week sprint with the team to stage training data. Data registry tool is taking shape and data storage and cost tagging has started. Coordination on tasks through ClickUp is good so far. (Week 8)
- Weekly meeting with CSUMB capstone team. Students are getting to know the data and running through tutorials to familiarize themselves with the tasks. (Week 8)

- Week 2 of 2 week sprint going well. Training data staged on Atlas and scripts for generation and upload in progress with AWS tagging proposal. (Week 9)
- Fixed issues with deesea-tfrecord versioning and correct DVC registry build/delete in vaaregistry. Learned more than I wanted to know about python packages, docker tests and DVC. (Week 11)
- Started on Sprint 2 to train in AWS using the TensorFlow implementation of Google's Efficient Detection model. (Week 11)
- Cost reporting for jobs is getting reported here, and met with Carlos and Duane to discuss adding performance metrics to this same page. (Week 19)
- Finishing up a video batch inference notebook (Week 19)
- Adding the dockerized C++ tracker to the AWS Elastic Container Registry (ECR) so it can be executed from SageMaker ScriptProcessor. Once this is complete, all of the processing components will be cloud-enabled, and we can scale up processing if we want to. (Week 19)
- Added quick-start documentation for database build. (Week 19)
- Cloud-enabled the video processing pipeline. Only 40 lines of code for a full video processing pipeline in SageMaker. Pretty cool. (Week 22)
- Working with the team on pulling out performance metrics in the VAA SageMaker web app. So far so good. (Week 22)
- 2022 proposal (Week 27)
- Working with Duane Edgington and Kevin Barnard on documentation <http://docs.mbari.org/internal/vaa-docs/> (Week 44)
- Team meeting planned for Friday to work through remaining tasks & assignments for the year (Week 44)
- Mtg planned with AWS Principal Enterprise Strategist on the nonprofits team to discuss alternative funding / support options, models, data formats, and our ML roadmap. (Week 48)
- Deployed changes to deepsea-ai database server with added video process ARN and some refactoring of tables. Docs on queries and mutations updated. (Week 48)
- YOLOv5 model testing on A100 for distributed training. Newest code release needs to be integrated into deepsea-yolov5 for full support. (Week 48)
- Upgraded to OpenCV 4.5, CUDA 11.3.1, and minor tracking bug fixes to handle small objects exceptions in deepsea-video-track pipeline processing. CSRT tracker performs better than KCF for benthic video, but it is slower ~15-20 FPS. (Week 48)
- Team is working on AWS cost/performance benchmarking. Cost per video thus far ranging from \$1.70 to \$4.40 per video for detailed annotation, \$70-170 for training, depending on the size of the model (640x640 or 1280x1280) with varying mAP. Performance plots forthcoming, probably after the holiday break. Thus far, 794673 unique detections and 4846 annotations have been generated during testing. (Week 51)
- Working on Vue app to consolidate all the AWS job/training management (Week 51)
- Updates to deepsea-ai-py code to support latest schema changes. Those are in the branch db1.11.0sync (Week 51)

Out of Office

- March 29-April 2nd. I will be off-the-grid for a good part of this.
- May 28th
-
- Vacation 11/23-11/26
- Vacation Dec 24 - Dec 31

Overhead

- Pulling together a project idea for the Cal Poly Deep Learning and Knowledge Graphs class (Week 1)
- Submitted project description for CSUMB Spring Capstone project (Week 2)
- Picture a Scientist Discussion Panel (Week 2)
- Meeting planned with MBAQ MBA to discuss contributing to the "Into the Deep" virtual exhibit (Week 5)

- Meeting planned with Rich and Karen Friday at 11 for OASIS download/config onboarding. (Week 8)
- OASIS Onboarding meeting with Rich and Karen to discuss mooring turnovers (Week 9)
- OASIS onboarding meeting cancelled this week; will resume next week. (Week 11)
- Our paper was finally reviewed (took a long time due to Covid) “Automated classification of fauna in seabed photographs: the impact of training and validation dataset size, with considerations for the class imbalance,” to be published in the Progress in Oceanography Journal. Responded to feedback along with the other co-authors Jennifer Durden, Henry Ruhl, and William Hosking. (Week 19)
- Our paper was published, finally! Automated classification of fauna in seabed photographs: the impact of training and validation dataset size, with considerations for the class imbalance. Open access here: <https://doi.org/10.1016/j.pocean.2021.102612> (Week 22)
- EMSO tutorial materials completed (Week 44)
- AWS Reinvent conference (Week 48)