2021-12-17 Clowder Dev Meeting Notes

Date

17 Dec 2021

Attendees

- Luigi Marini
- Maxwell Burnette
- Sara Lambert
- Todd Nicholson
- Bing Zhang absent
- Michael Bobak
- Rob Kooper
- Sandeep Puthanveetil Satheesan absent
- Chen Wang
- Kastan Day
- Elizabeth Yanello
- Kathryn Naum absent

Agenda:

- V2
- Hackathon turned into talkathon
- Poster for ACSW - http://acsw.core.edu.au/2022-call-for-posters
- Quick way to take meeting notes: https://hackmd.io/?nav=overview
- Use the google doc to note meetings that you have for V2 and mark any changes, issues that you want to share with the V2 team https://docs.google.com/document/d/1GnQKMyM4FMaEKm9UqIw3QQzI_DiiyqpXtf0k5cPoAaw/edit#

Discussion:

Nightingale - singularity is not an option. Conda can be used on Radiant; turned images into a Conda container. The Pod that they gave to us cannot work, so is working with Maria Jaromin on a work around. Rob requests that Chen document what works and what doesn't, which can then be put into official documentation. (Docker, Kubernetes, Singularity, etc.)

- V2: Max fixed some bugs.
  - Downloading still needs some work, perhaps because I was running multiple instances. We can provide a signature to a model and will convert Json for you. Tested front end in docker compose - seemed to be working!
  - Adding notes to read.me
  - We are using FastAPI for Swagger
  - We need to create datasets and have been tested
  - Name Alpha 1 as a temporary name 2.0-Alpha.1 to distinguish between V1 and V2
  - Mike L notes that PR does not have good support for downloading file blobs - client makes the assumption that we are working with Json or Txt and does not return binary data. Options are to look at a different generator or use non-generated code for now
  - Should we use cookies instead of auth? Since we are fairly early in dev of v2, this might be a good time to change these patterns.
  - Using a generated client, it's returning binary blobs.
  - Working on how to sync a client and have put together a GitHub action and will check out the current FastAPI and run the latest docker image and generates the code.
  - Can we watch other GitHub actions? For extractors, it runs once a day/once an hour, etc. to check for merges OR make an access token can commit to another repository. However, if you're not careful, the secret token could be hacked. It would be best to run a cron job in the front end. GitHub actions can run a cron job. https://github.com/clowder-framework/extractors-core/blob/master/.github/workflows/pyclowder.yml. If you run a complex job on Friday night, it will be updated over the weekend and on Monday there will be a completely updated version.
  - Todd going to be 25% on clowder after the new year. The PI should fill in the form themselves, including the CFOP, but the developer should help with how much space etc you will need in Clowder and Radiant usage.
  - Releasing V2 2.0-Alpha.1 this afternoon. Congrats to all! There will be difference in the backend

Who	Notes
Luigi	absent

Max	absent
Mike L.	WIP GitHub action to sync OpenAPI spec changes with generated client code (still testing / trying to think of better patterns) - Rob says look into turning this into a cron job action
Todd	pull requests open edit and increment file upload multiple files
Bing	
Mike B.	
Rob	
Kastan	- [x] Get a reply from PDG machine learning folks - Scheduled for Monday 12/20 - [ ] Ensure Dask worker.distributed works on HAL (workers getting killed by garbage collector) - [ ] Talk to HAL admins about Ray support (for IBM Codeflare) - [ ] Does Codeflare support current SOTA models (like MASK R-CNN) used by PDG? - [ ] Begin collaboration with IN-CORE grad student Tu Nguyen on building-type classification workflow.
Sandeep	
Chen	▪ Conda podman run Clowder stack (clowder-scala, rabbitmq and mongodb) works on radiant machine; still figuring out on nightingale ▪ V2 Alpha 1 release finalized
Katie	
Lisa	

Action Items/To Dos:

- [ ] Workbench Webinar in 2022 ending CDDR?