



/ ANITA
B.ORG 2022

GRACE HOPPER
CELEBRATION

NEXT is
NOW



Use Open-Source Tool to Simplify & Automate Video/Image Labeling Process



Samantha Coyle
Software Engineer



Neethu Elizabeth Simon
Senior Software Engineer



Notices and Disclaimers

- Performance varies by use, configuration and other factors. Learn more on the [Performance Index site](#).
- No product or component can be absolutely secure.
- Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.
- Your costs and results may vary.
- Intel technologies may require enabled hardware, software or service activation.
- © Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

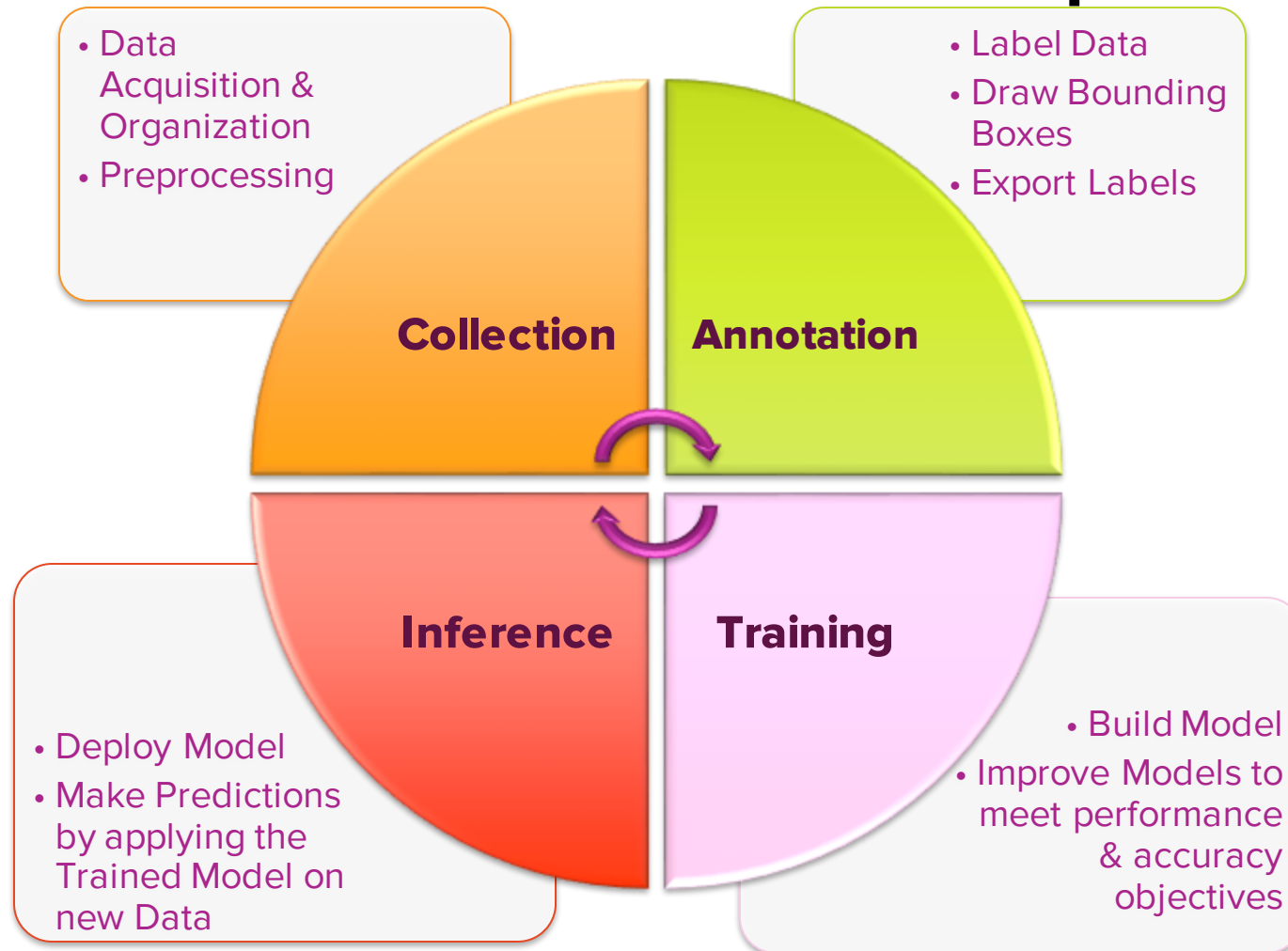
Overview

- Introduction
- Project Setup
- Annotation Specifications
- Annotation Best Practices
- Perform Annotations
- Conclusion



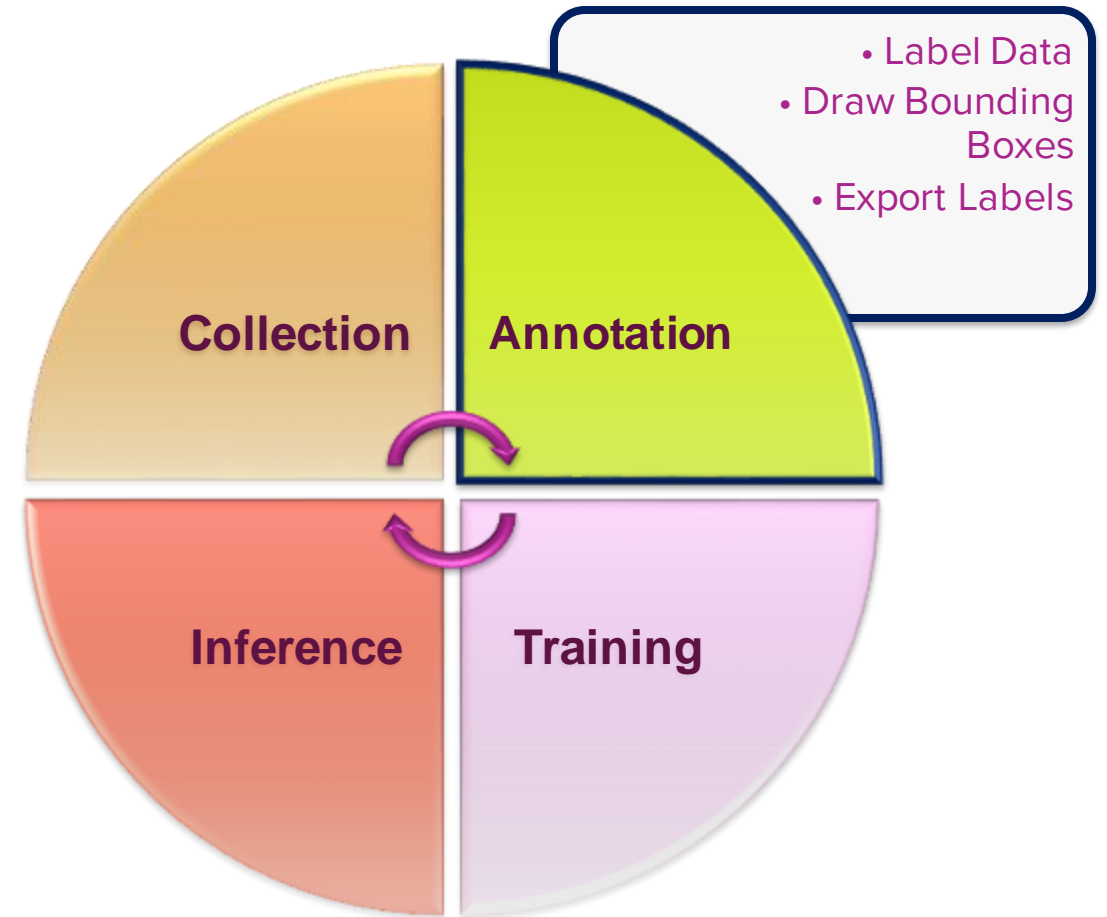
Introduction

AI/ML Solution Development



Data Annotation / Labeling

- Establish ground truth
- Draw bounding boxes around object
- Simplest step
- Challenges
 - Manual vs Automation
 - Time consuming
 - Expensive
 - Difficult to Scale



Data Annotation / Labeling

➤ Commercial Tools

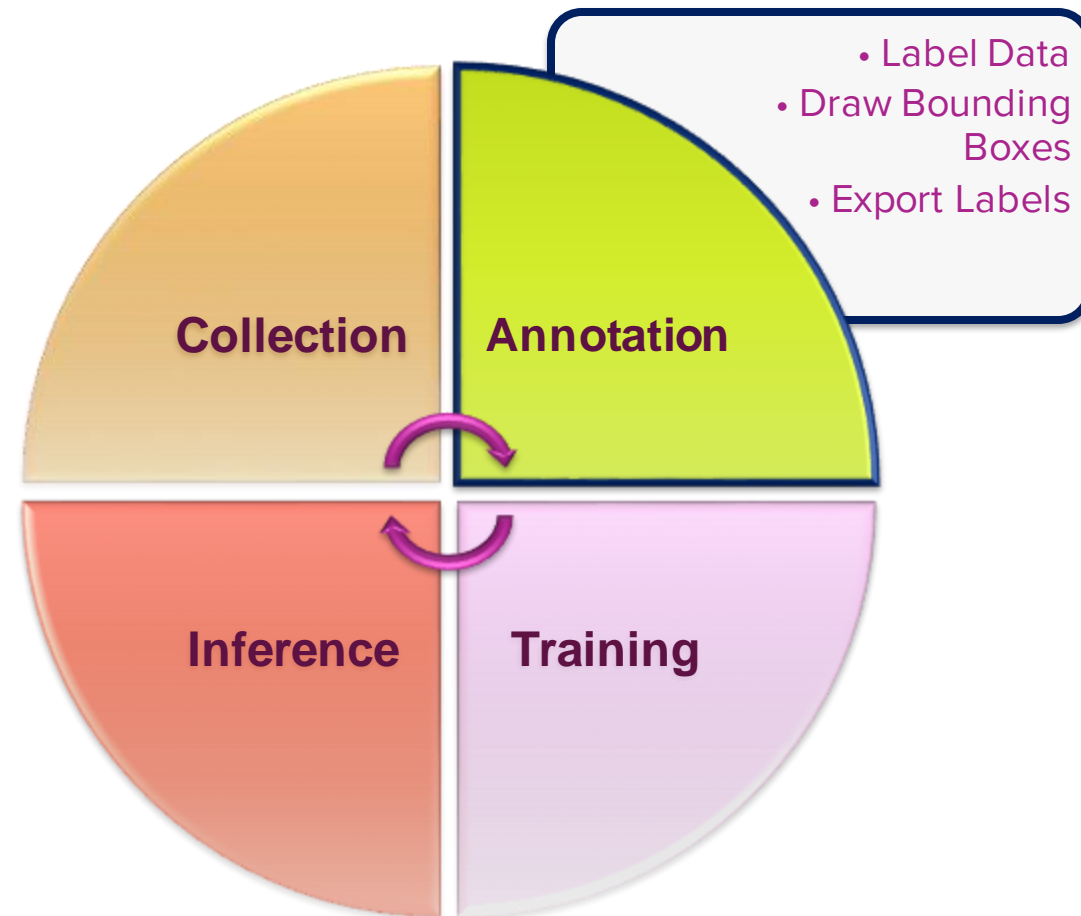
- [Labelbox](#)
- [Dataloop](#)
- [RectLabel](#)

➤ Open-Source Tools

- [Bbox](#)
- [Labellmg](#)
- [VGG Image Annotator \(VIA\)](#)



Computer Vision Annotation Tool
<https://github.com/opencv/cvat>





Project Setup

<https://github.com/sicoyle/ghc-data-annotation-workshop/blob/main/annotationSteps.md>

<https://github.com/sicoyle/ghc-data-annotation-workshop>

<https://app.cvat.ai/auth/login>

Background information

CVAT projects have a hierarchy incorporating tasks and jobs. Annotation projects may have many annotation tasks and jobs. Tasks are where you upload video footage and specify the associated project and labels. Tasks represent the progress for annotation footage and details of the task. They also allow you to upload existing annotations and export annotations. Tasks may have multiple jobs. Jobs allow you to split up tasks by frame to divvy up the annotations among a team.

Project: Build person tracking annotations for super fun GHC workshop

Task 1: Annotate workshop footage

Job 1

Frames 0-10

Job 2

Frames 10-20

Job 3

Frames 30-40

Job N

Frames 40-N

Task 2: EX: Annotate custom footage

Job 1

Frames A-B

Job 2

Frames B-C

Job 3

Frames C-D

Job N

Frames D-Z

Task N: EX: Annotate custom footage

Job 1

Frames A-B

Job 2

Frames B-C

Job 3

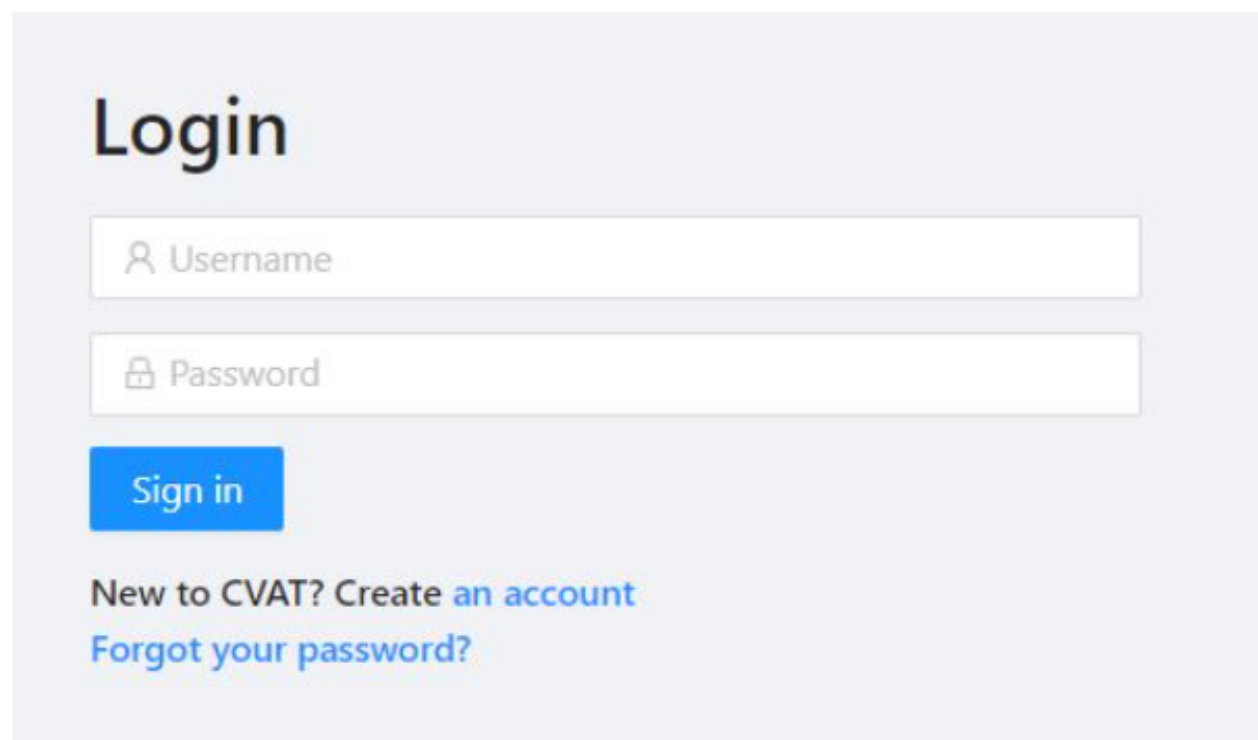
Frames C-D

Job N

Frames D-Z

Steps to follow to create an annotation project

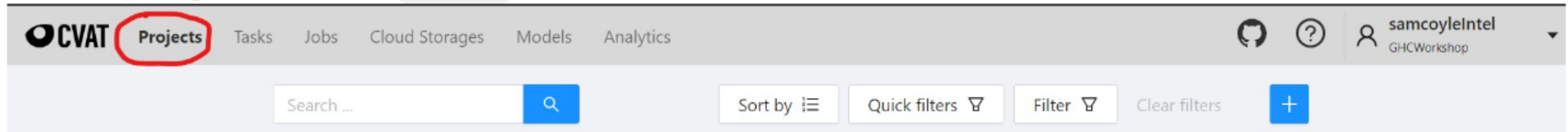
1. Create an account and login at **<https://app.cvat.ai/auth/login>**



The screenshot shows the CVAT login interface. It has a light gray background. At the top left, the word "Login" is written in a large, bold, black font. Below it are two white input fields with gray borders. The first field has a user icon and the text "Username". The second field has a lock icon and the text "Password". Below these fields is a blue rectangular button with the text "Sign in" in white. At the bottom of the form, there are two lines of text: "New to CVAT? Create [an account](#)" and "[Forgot your password?](#)", both with the links in blue.

While creating a new account, a valid email address is required. Tool requires email confirmation by clicking on the link sent to this email address

2. Use the top navigation bar to click **Projects** .



3. Click the blue + sign and then Create a new project .



4. Copy the details from the image below to fill out the project to be created and select `Submit & Open`. For the raw label data, copy/paste the data snipped below the picture. These data represent the labels & their attributes.

Create a new project

* Name

Build person tracking annotations for super fun GHC workshop ✓

Labels:

Raw Constructor

<copy/paste data below this picture here>

Done Reset

> Advanced configuration

Submit & Open Submit & Continue

Raw Label data to copy/paste into the CVAT UI:

```
[
  {
    "name": "Person",
    "id": 1199059,
    "color": "#c06060",
    "attributes": [
      {
        "id": 1198651,
        "name": "Identity",
        "input_type": "number",
        "mutable": false,
        "values": [
          "1",
          "1000",
          "1"
        ]
      }
    ],
  },
  {
    "id": 1198650,
    "name": "Occlusion",
    "input_type": "number",
    "mutable": false,
    "values": [
      "0",
      "2",
      "1"
    ]
  }
]
```

[https://github.com/sicoyle/
ghc-data-annotation-
workshop/blob/main/annot
ationSteps.md](https://github.com/sicoyle/ghc-data-annotation-workshop/blob/main/annotationSteps.md)

5. Click on your newly created project and click on the blue **+** sign at the bottom right of it to create a new **task** for it.

The screenshot displays the CVAT web application interface. At the top, a navigation bar includes the CVAT logo and links for Projects, Tasks, Jobs, Cloud Storages, Models, and Analytics. On the right of the navigation bar, there are icons for GitHub, help, and a user profile for 'samcoyleIntel' with the role 'GHCWorkshop'. Below the navigation bar, the main content area shows a project titled 'Build person tracking annotations for super fun GHC workshop'. It includes a link to 'Back to projects', an 'Actions' menu, and project details such as 'Project #45701 created by samcoyleIntel on August 7th 2022' and 'Assigned to Select a user'. The 'Issue Tracker' section shows 'Not specified'. Below this, there are tabs for 'Raw' and 'Constructor', and a button to 'Add label'. A red circle highlights a blue '+' button at the bottom right of the project card, which is used to create a new task. At the bottom of the interface, there is a search bar, a 'Sort by' dropdown, 'Quick filters' and 'Filter' dropdowns, a 'Clear filters' button, and the highlighted blue '+' button.

6. Create your task copying the details below and click **Submit & Open** on your task:

<https://github.com/sicoyle/ghc-data-annotation-workshop/blob/main/assets/video/workshopFootage.mp4>



Note: It may take a few seconds-minutes to upload the video footage pending the network bandwidth.

Basic configuration

* Name

Annotate workshop footage ✓

Project

Build person tracking annotations for super fun GHC workshop

Subset

Train

Labels

Project labels will be used

* Select files

My computer Connected file share Remote sources Cloud Storage

Click or drag files to this area

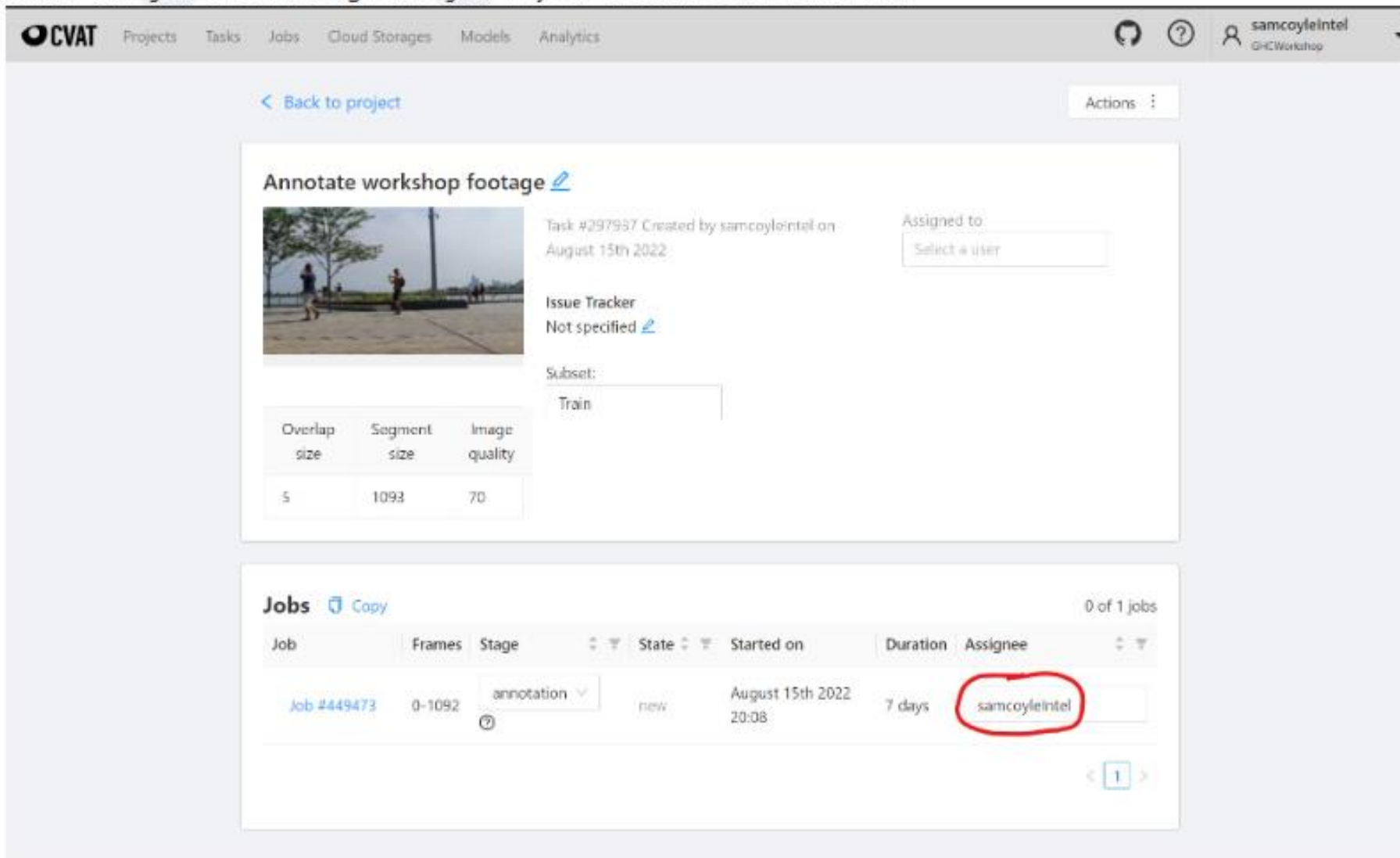
You can upload an archive with images, a video, or multiple images

workshopFootage.mp4

> Advanced configuration

Submit & Open Submit & Continue

7. Under the `job` details, assign the `job` to your username as seen below.



The screenshot shows the CVAT web interface. At the top, there's a navigation bar with links: Projects, Tasks, Jobs, Cloud Storages, Models, and Analytics. The user profile 'samcoyleintel' is visible in the top right corner. Below the navigation bar, there's a 'Back to project' link and an 'Actions' menu. The main content area displays the details for a task titled 'Annotate workshop footage'. This section includes a video thumbnail, task information (Task #297987, created by samcoyleintel on August 15th 2022), an 'Assigned to' dropdown menu with 'Select a user' as the option, an 'Issue Tracker' link, and a 'Subset' dropdown menu with 'Train' selected. Below this, there's a table with columns for Overlap size, Segment size, and Image quality, showing values 5, 1093, and 70 respectively. At the bottom, there's a 'Jobs' section with a 'Copy' button and a table showing job details. The table has columns: Job, Frames, Stage, State, Started on, Duration, and Assignee. The first row shows 'Job #449473', '0-1092', 'annotation', 'new', 'August 15th 2022 20:08', '7 days', and 'samcoyleintel'. The 'Assignee' cell is circled in red. A pagination bar at the bottom right shows '< 1 >'.

CVAT Projects Tasks Jobs Cloud Storages Models Analytics

Back to project Actions

Annotate workshop footage

Task #297987 Created by samcoyleintel on August 15th 2022

Assigned to: Select a user

Issue Tracker Not specified

Subset: Train

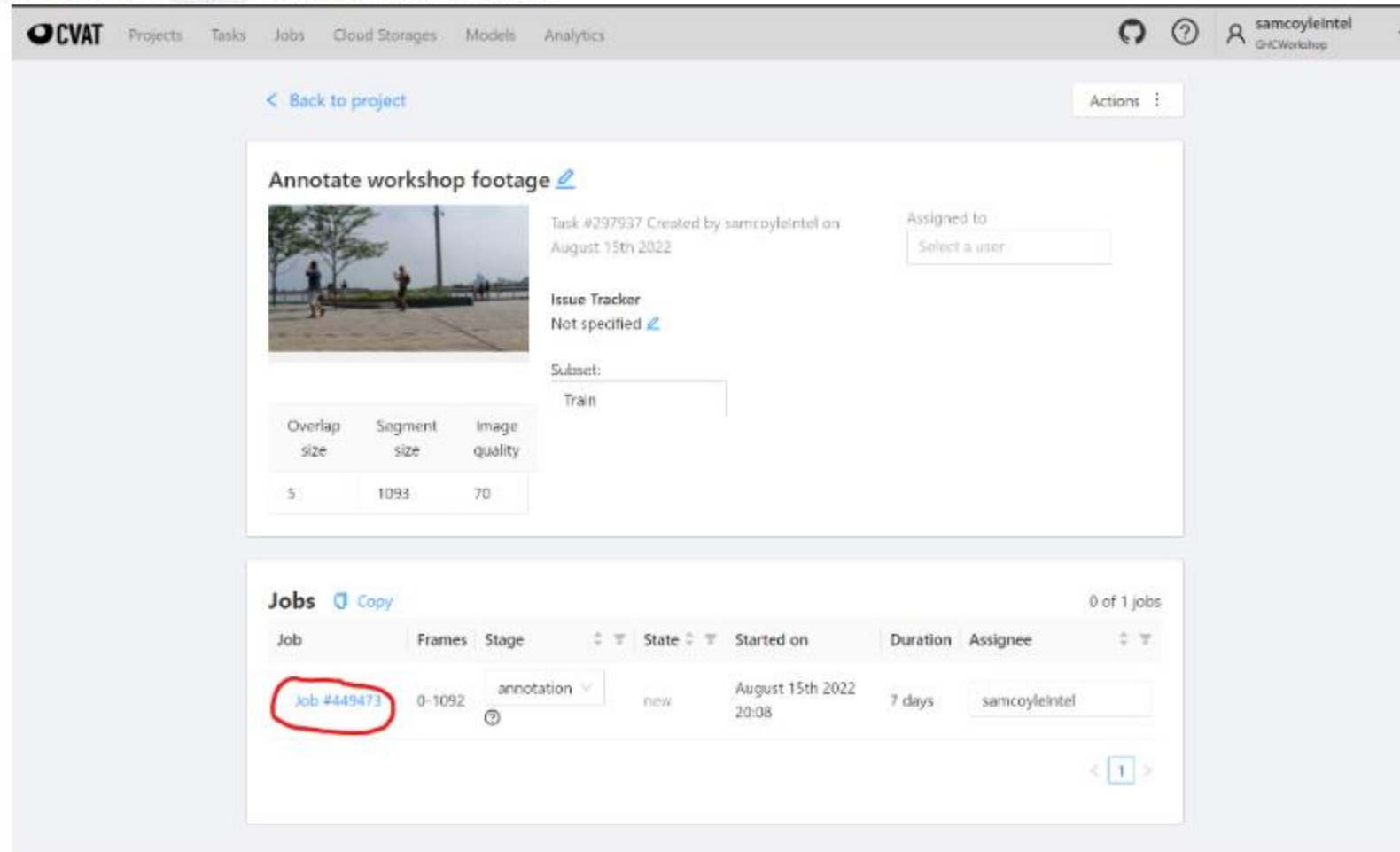
Overlap size	Segment size	Image quality
5	1093	70

Jobs Copy 0 of 1 jobs

Job	Frames	Stage	State	Started on	Duration	Assignee
Job #449473	0-1092	annotation	new	August 15th 2022 20:08	7 days	samcoyleintel

< 1 >

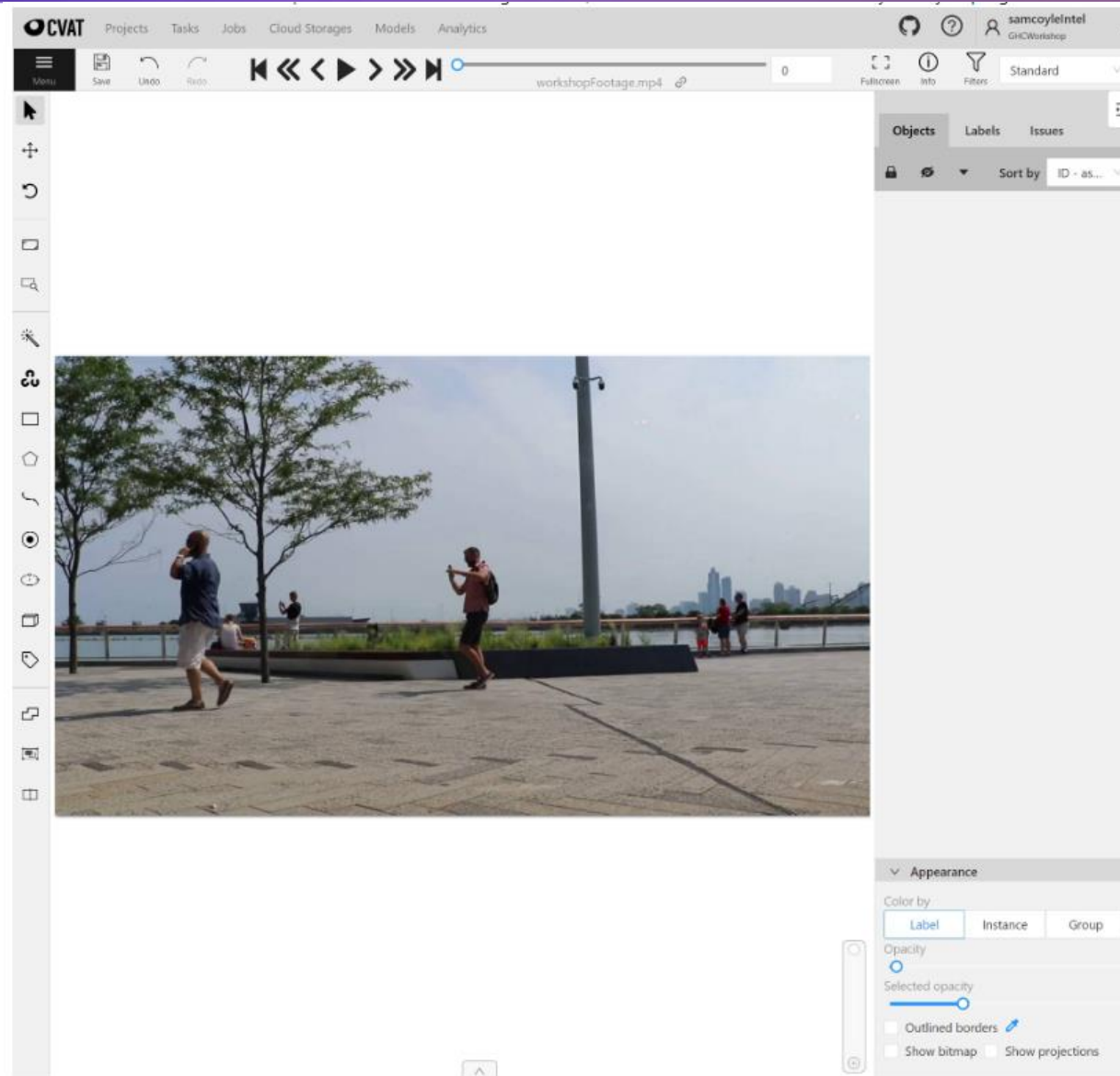
8. Click on the `job #` that was created above.



The screenshot shows the CVAT web interface. At the top, there's a navigation bar with 'CVAT' logo and links for Projects, Tasks, Jobs, Cloud Storages, Models, and Analytics. The user 'samcoyleintel' is logged in. Below the navigation bar, there's a 'Back to project' link and an 'Actions' menu. The main content area displays a task titled 'Annotate workshop footage'. It includes a thumbnail image of a park, task details (Task #297937, Created by samcoyleintel on August 15th 2022), an 'Assigned to' dropdown (Select a user), an 'Issue Tracker' (Not specified), and a 'Subset' dropdown (Train). Below this, there's a table with columns: Overlap size, Segment size, and Image quality. The table has one row with values: 5, 1093, and 70. At the bottom, there's a 'Jobs' section with a 'Copy' link and '0 of 1 jobs' indicator. It contains a table with columns: Job, Frames, Stage, State, Started on, Duration, and Assignee. The table has one row: Job #449473 (circled in red), 0-1092, annotation, new, August 15th 2022 20:08, 7 days, and samcoyleintel.

Job	Frames	Stage	State	Started on	Duration	Assignee
Job #449473	0-1092	annotation	new	August 15th 2022 20:08	7 days	samcoyleintel

This will take you to a screen where you can start data annotations. This is where you will use the supporting documents to perform your data annotations. Feel free to ask questions as needed during this time, and collaborate with those around you as you progress.





Annotation Specifications

<https://github.com/sicoyle/ghc-data-annotation-workshop/blob/main/annotationSpecifications.md>

Data annotation specifications

Specifications to guide annotations for a person tracking project

The information below is meant to guide discussion and provide pointers and considerations that a professional annotation team would have for a person detection annotation project.

General Comments

- Only visible parts of people should be annotated.
- Don't include bags, purses, baby carriages, shopping carts, etc. into a bounding box.
- Individuals should have the same identity if he/she/it/they appear/disappear several times throughout a video.
- Don't annotate small or really blurry people.

Annotation Format

CVAT supports multiple annotation formats that may be found <https://opencv.github.io/cvat/docs/manual/advanced/formats/>

- Annotation Format is to be chosen by the annotator. One example is CVAT XML file schema/metadata.

More information on the XML annotation format specifically may be found https://opencv.github.io/cvat/docs/manual/advanced/xml_format/

The link describes the tags that are present in the XML, what they mean, and demonstrates an annotation example using annotation boxes, polygons, etc.

https://opencv.github.io/cvat/docs/manual/advanced/xml_format/

Person Tracking

The annotation file should contain the following information per frame (from either manual or interpolated annotations):

Annotation	Annotation Type	Encoded by
Person (location)	Rectangular bounding box (x1, y1, x2, y2)	x1: horizontal coordinate of the top left corner y1: vertical coordinate of the top left corner x2: horizontal coordinate of the bottom right corner y2: vertical coordinate of the bottom right corner
Identity	Number	Number indicating the person's identity (maintained over time).
Occlusion	Number	Value: <ul style="list-style-type: none"> • Person is not occluded (0) • Person is partially occluded ($\leq 50\%$) (1) • Person is heavily occluded ($> 50\%$) (2)



Annotation Best Practices

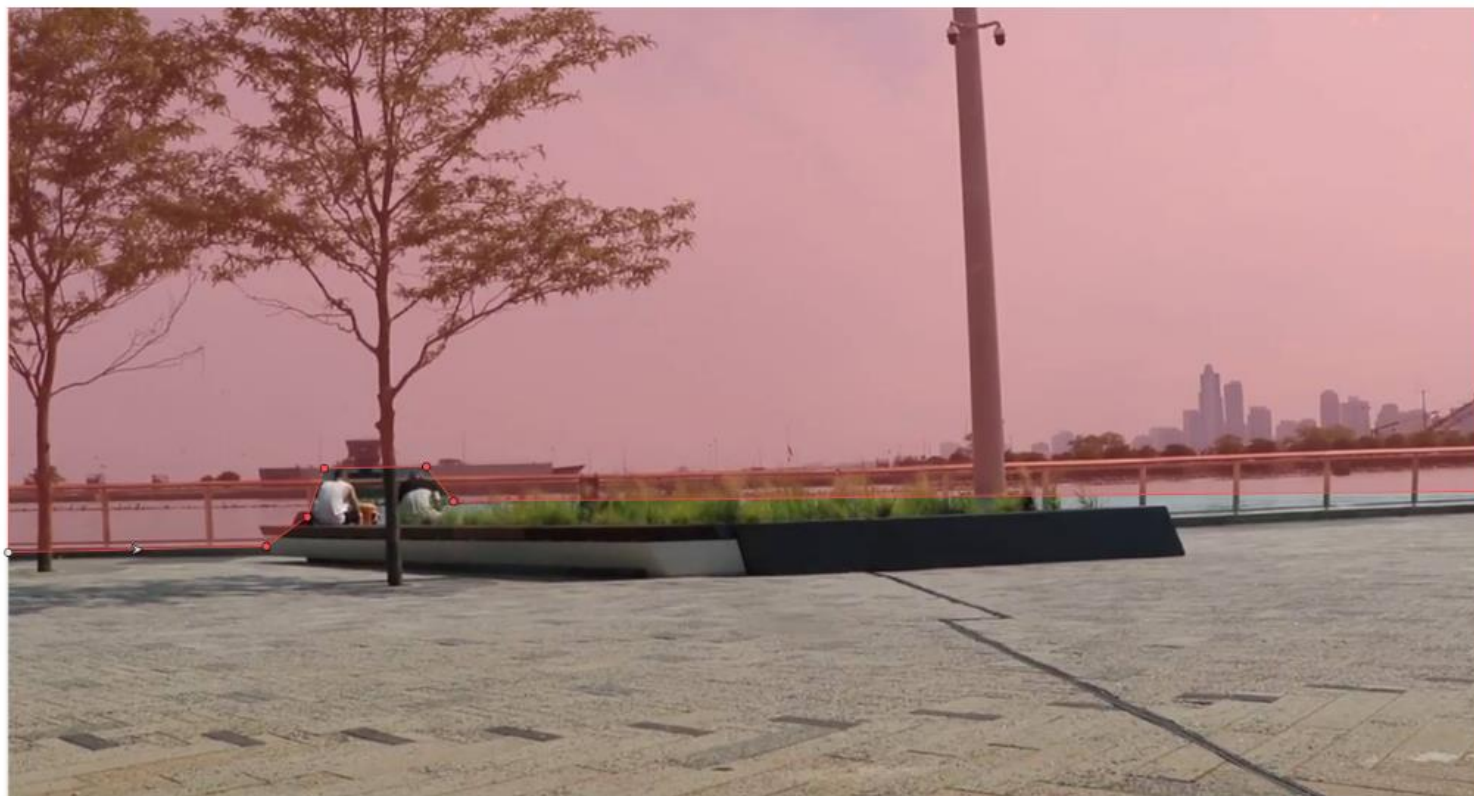
<https://github.com/sicoyle/ghc-data-annotation-workshop/blob/main/annotationBestPractices.md>

Data Annotation Best Practices

1. How big do people have to be to require annotating?

All people should be annotated. However, `ignore regions` may be defined for regions of the image where people are too small. `Ignore regions/zones` are common to have depending on your footage use case.

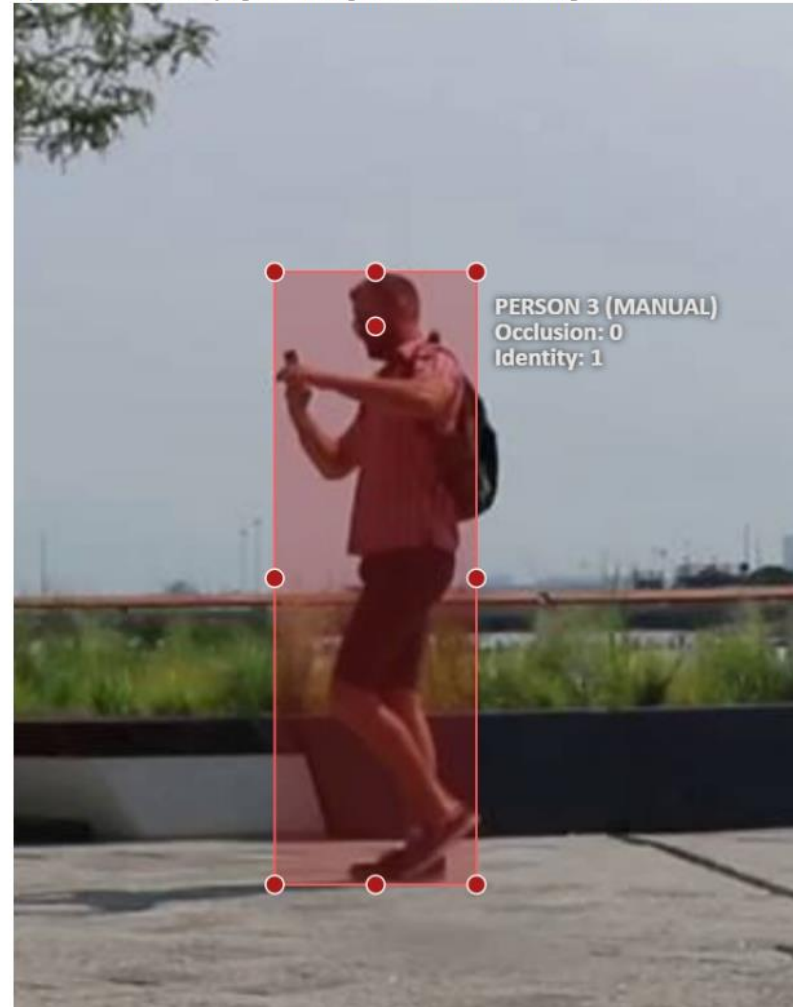
For example, in the image below, the red region can be described as our `ignore region`. This region will be excluded from annotations. In the case that ships pass by in the water with passengers on deck, then we do not want them counted towards our data annotations.



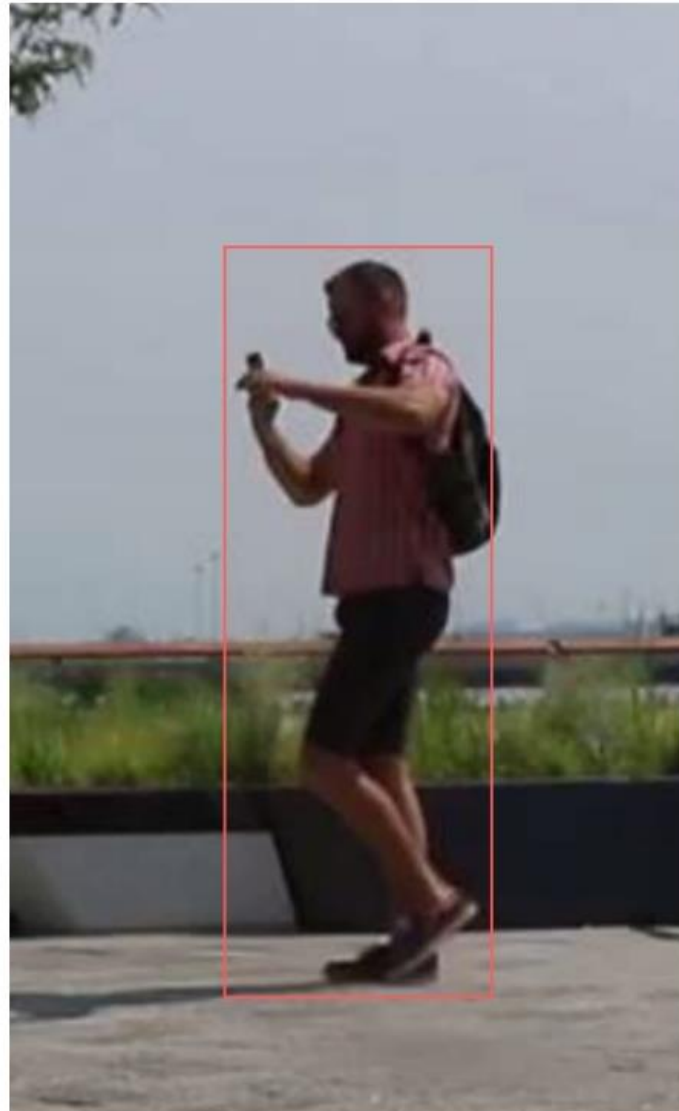
2. How precise should annotations be?

Annotations may be made with very tight bounding boxes, with 1-5% margin, or with >5% margin of boundaries.

Option 1: Below is a very tight bounding box with little to no margin of error:



Option 2: Below is a looser bounding box with a 1-5% margin of error:



Option 3: Below is the loosest bounding box with a $>5\%$ margin of error:



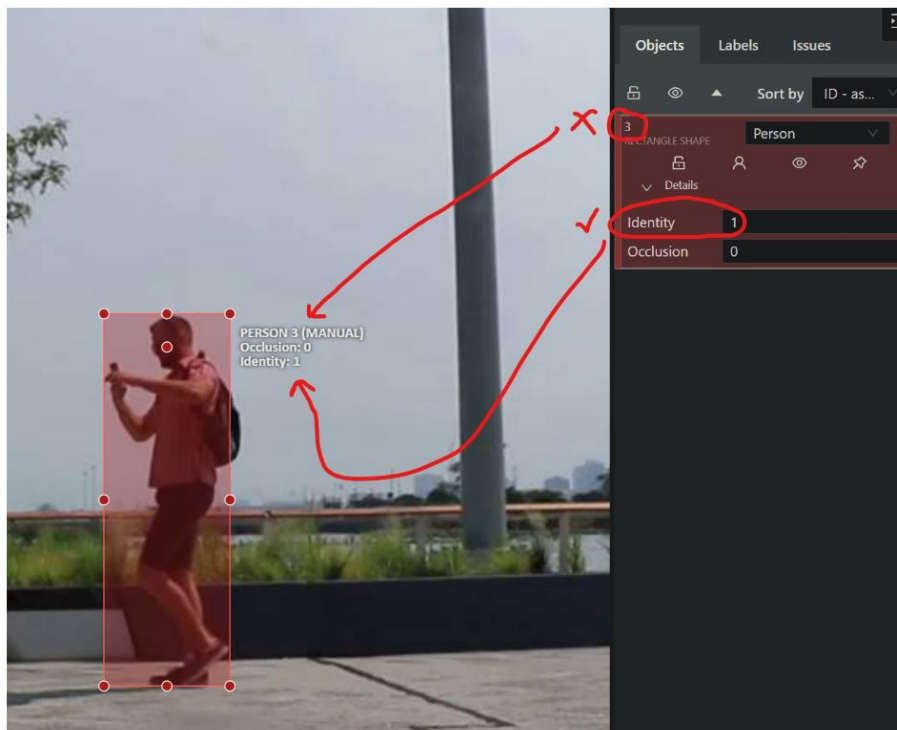
For the purposes of this workshop, Option 2 is suitable for our use case. Please utilize Option 2 for reference when performing your data annotations.

3. What is the `Identity` attribute?

Annotation	Annotation Type	Encoded by
Identity	Number	Number indicating the person's identify (maintained over time).

`Identity` is not personally identifiable information in the way one might initially think. It does **not** mean name, DOB, address, etc. `Identity` in this sense refers to a means to track the same person across different frames.

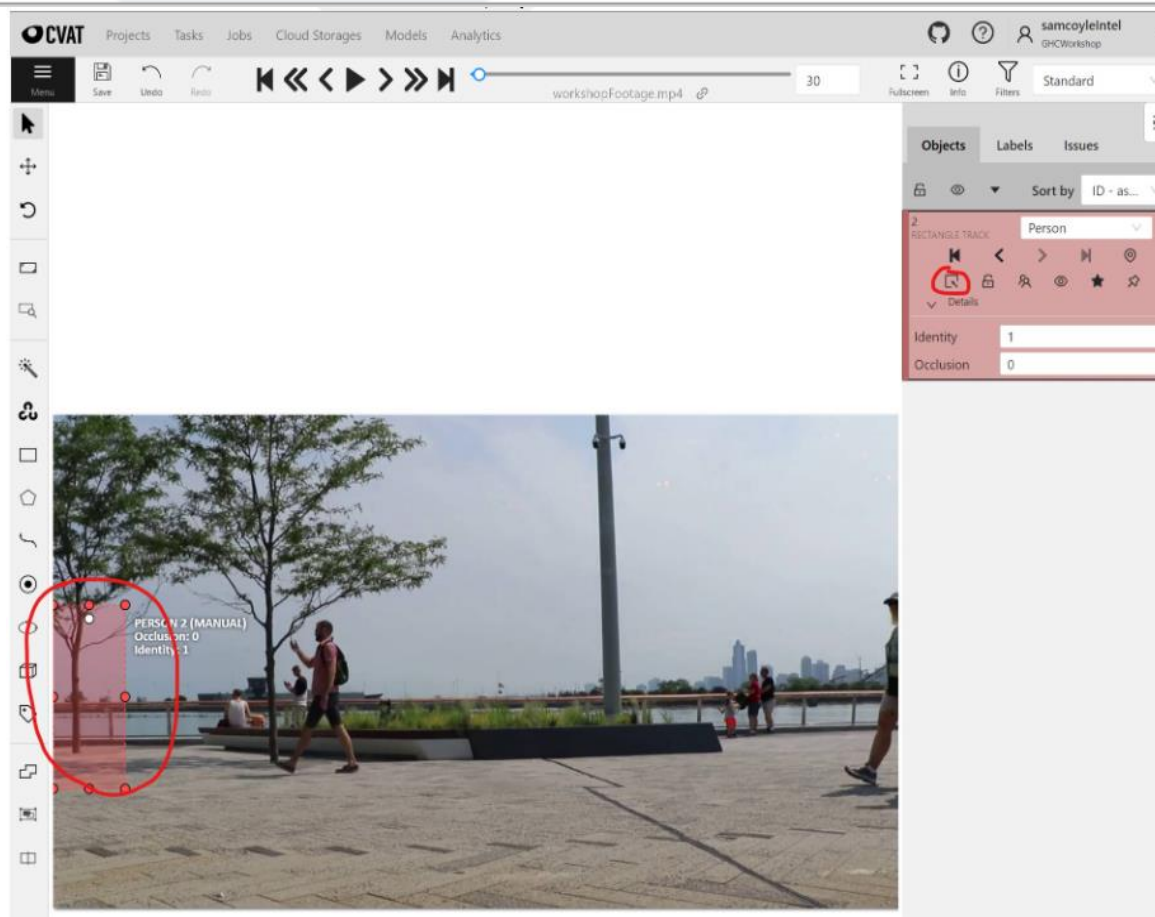
CVAT creates automatic `IDs` which cannot be changed. For example, the image below has the automatic `ID` of `PERSON 3`. However, we refer to the person with the `Identity` label of `1`. This is the `Identity` with which we will refer to this person throughout the duration of the video. For these purposes, you may disregard the automatic `IDs` given to each of your annotations. In other words, please disregard the numbers by the red `x` in the image.



4. How to annotate partial people?

For cases where people are present only partially in the frame, annotate the parts of the person that you can. Be sure to mark the `Occlusion` field accordingly.

5. How to annotate people who left the frame? For cases where someone walked to where they are no longer visible, mark the person as outside the frame by clicking on the **Switch Outside Property** field circled in the image below.



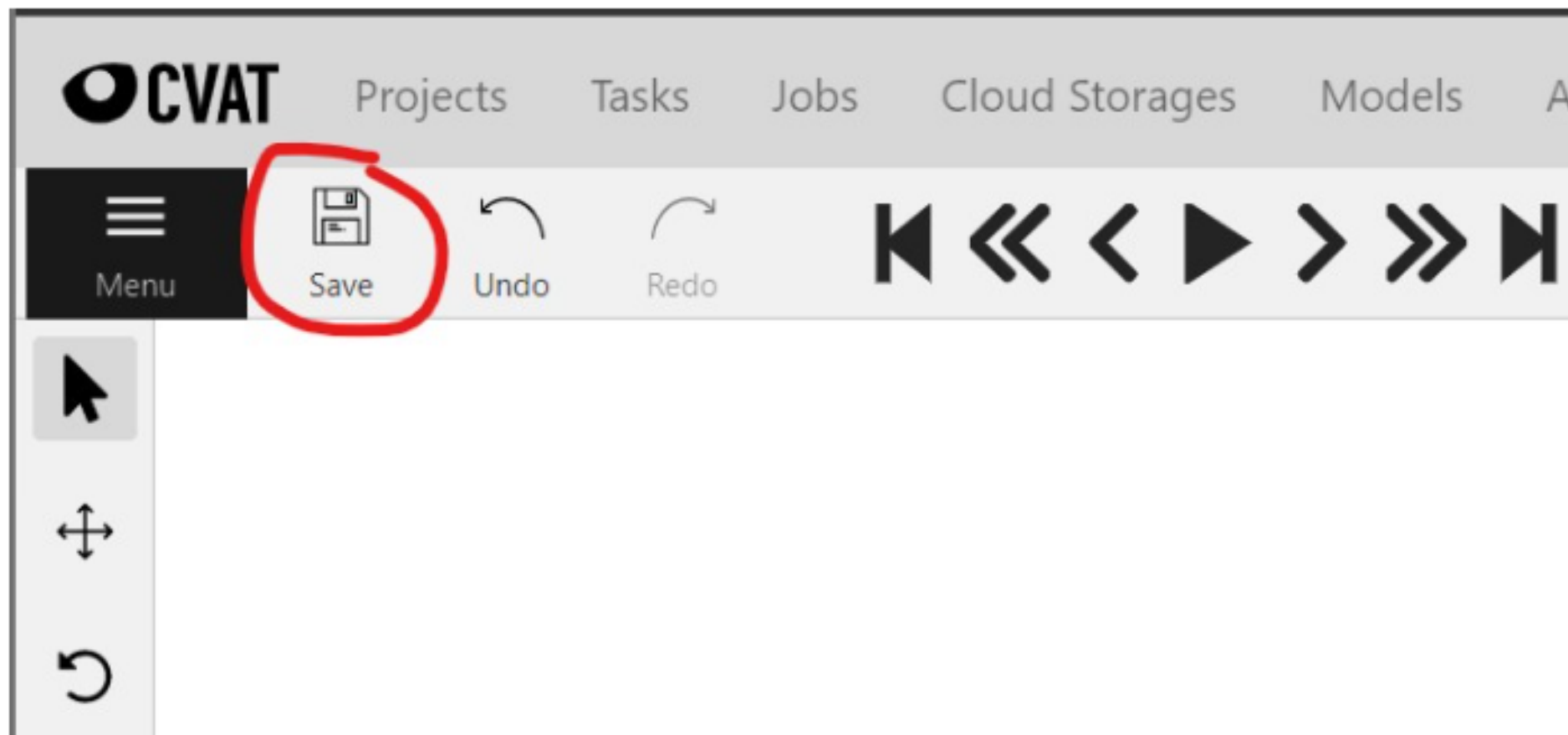
In the example above, the person who was being tracked by `Identity = 1` has now walked outside the frame. By switching the `Outside Property` circled under the `objects` tab on the right, CVAT understands that the person is no longer visible within the frame. This field will ensure that the data annotations properly account for people as they walk outside the frame.

6. If annotating every 10 frames, then how to account for the frame count not being evenly divisible by 10?

The workshop footage contains 1093 frames. Being that it is a video and people are being tracked, then it is acceptable to annotate every 10 frames. This is because CVAT knows to interpolate the frames between that are unannotated to automatically annotate them for us. If annotating every 10 frames, you will end up on frame 1090. Go ahead and annotate the last frame numbered 1093.

7. When to save annotation progress?

Click **Save** after each frame has been annotated.





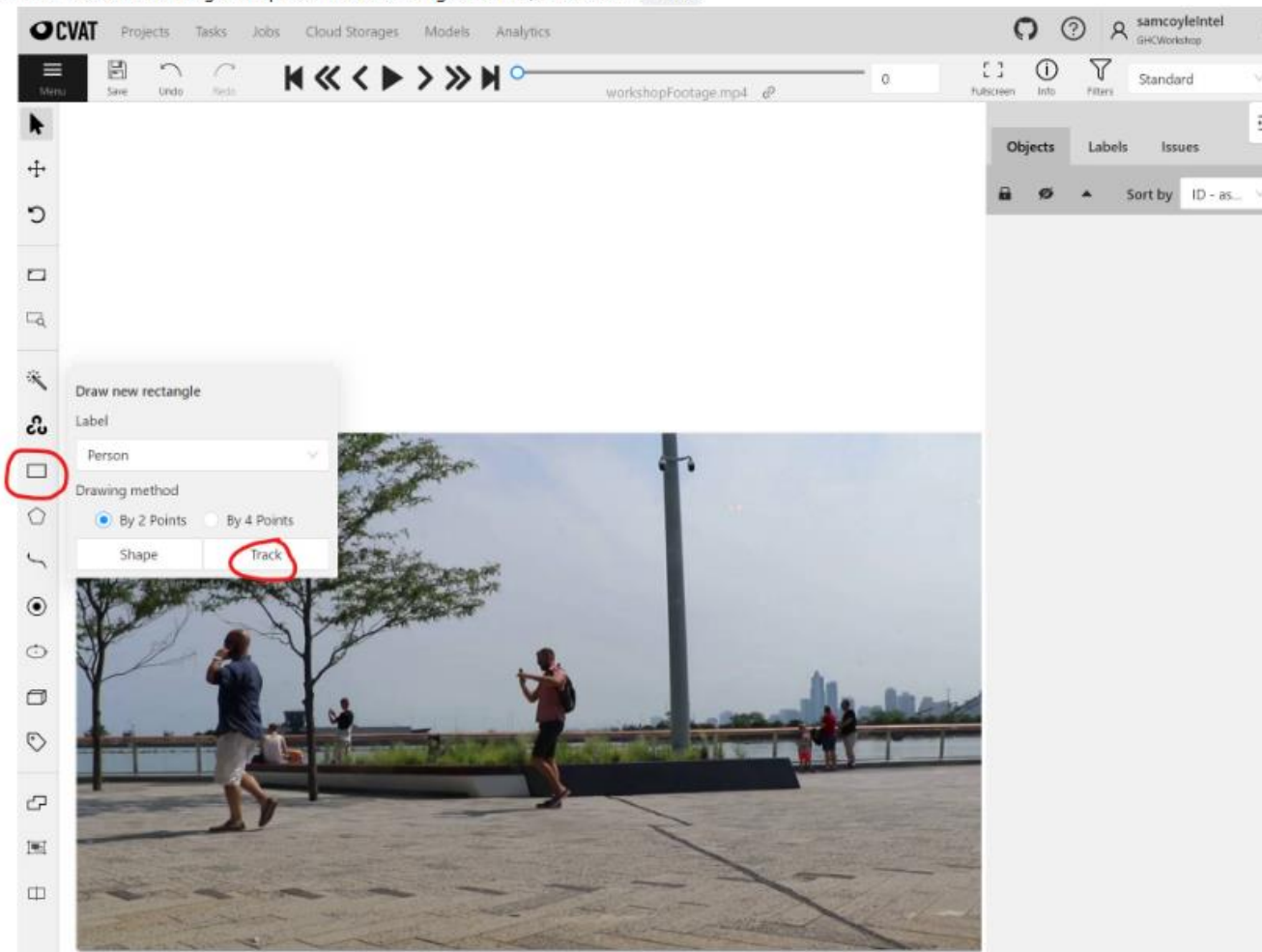
Perform Annotations

<https://github.com/sicoyle/ghc-data-annotation-workshop/blob/main/annotationSteps.md>

<https://github.com/sicoyle/ghc-data-annotation-workshop>
<https://app.cvat.ai/auth/login>

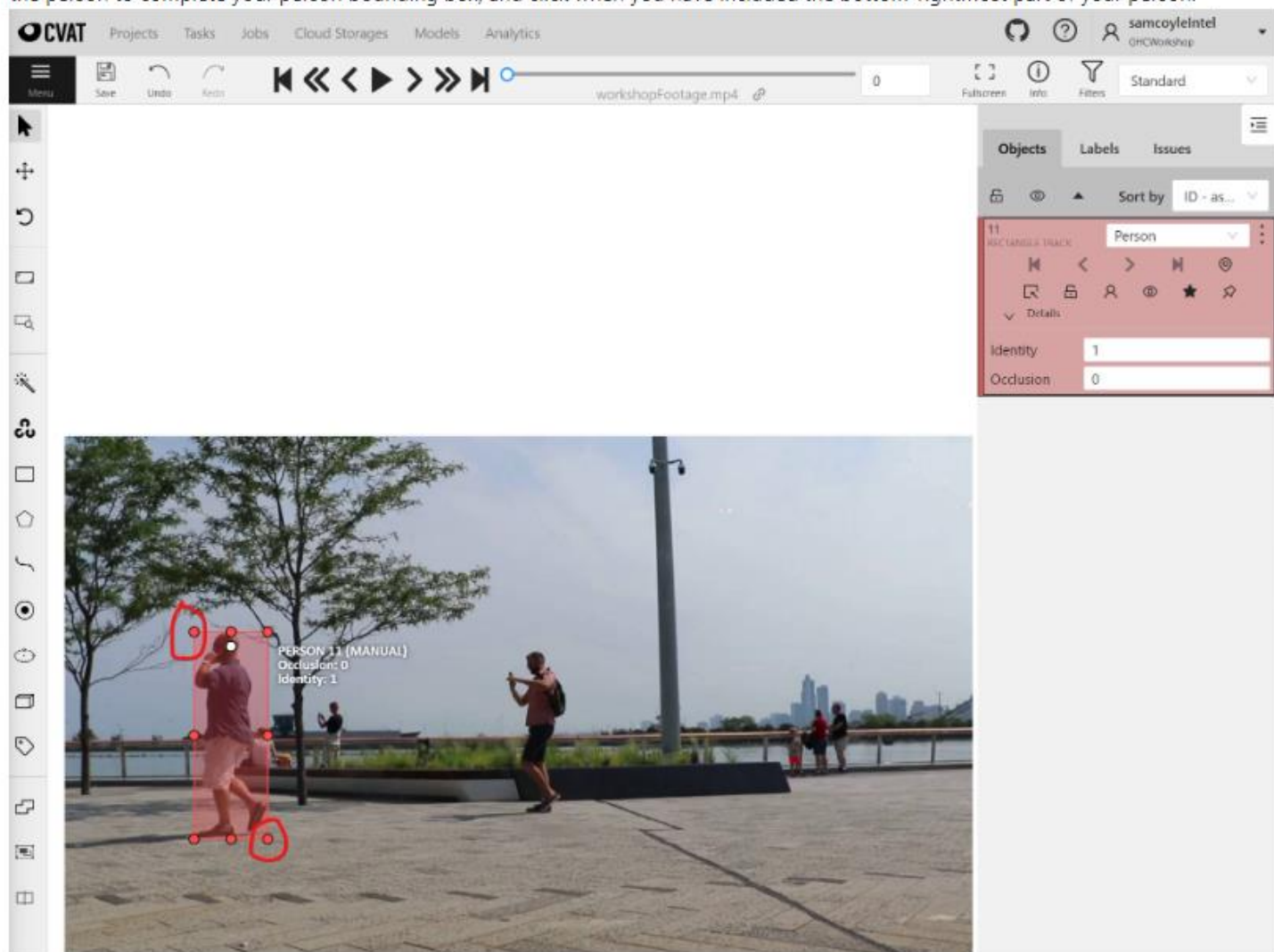
Steps to follow to perform data annotations

1. Hover over the rectangle shape on the left navigation bar, and select **Track**.

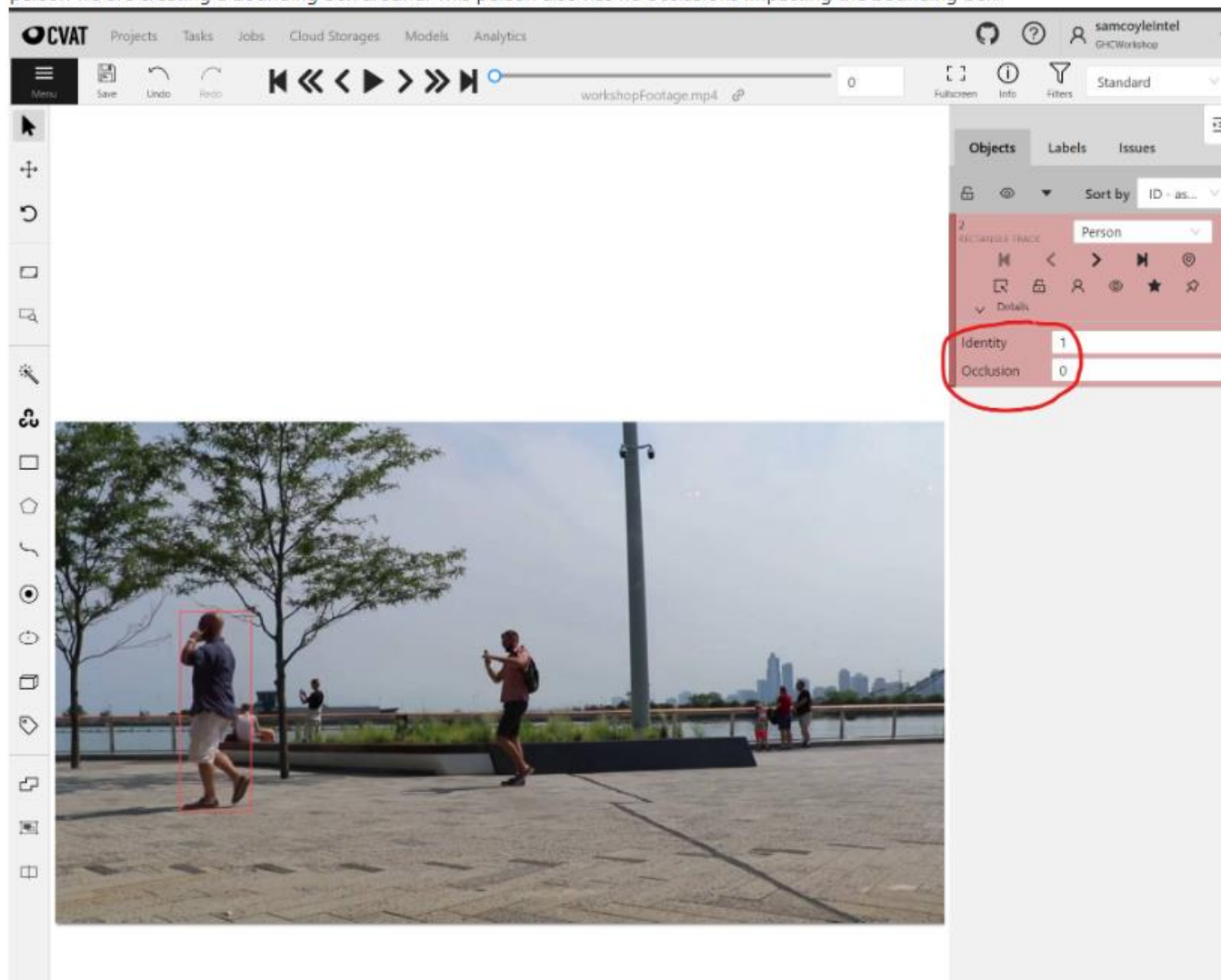


Note: We are annotating a video file. This means we will want to track people as they move across the frame throughout the video.

- Place your cursor at the top left of the person and click. This should create the start of your bounding box. Now go to the bottom right of the person to complete your person bounding box, and click when you have included the bottom-rightmost part of your person.



3. In the right most part of the CVAT UI, under the `objects` tab, you will leave `Identity` as `1`, and `occlusion` as `0` as this is the first person we are creating a bounding box around. This person also has no occlusions impacting the bounding box.



4. Once you have completed annotations for one person, then move onto the next person in the frame.
5. Repeat steps 1-3 for each person going left to right in the frame.

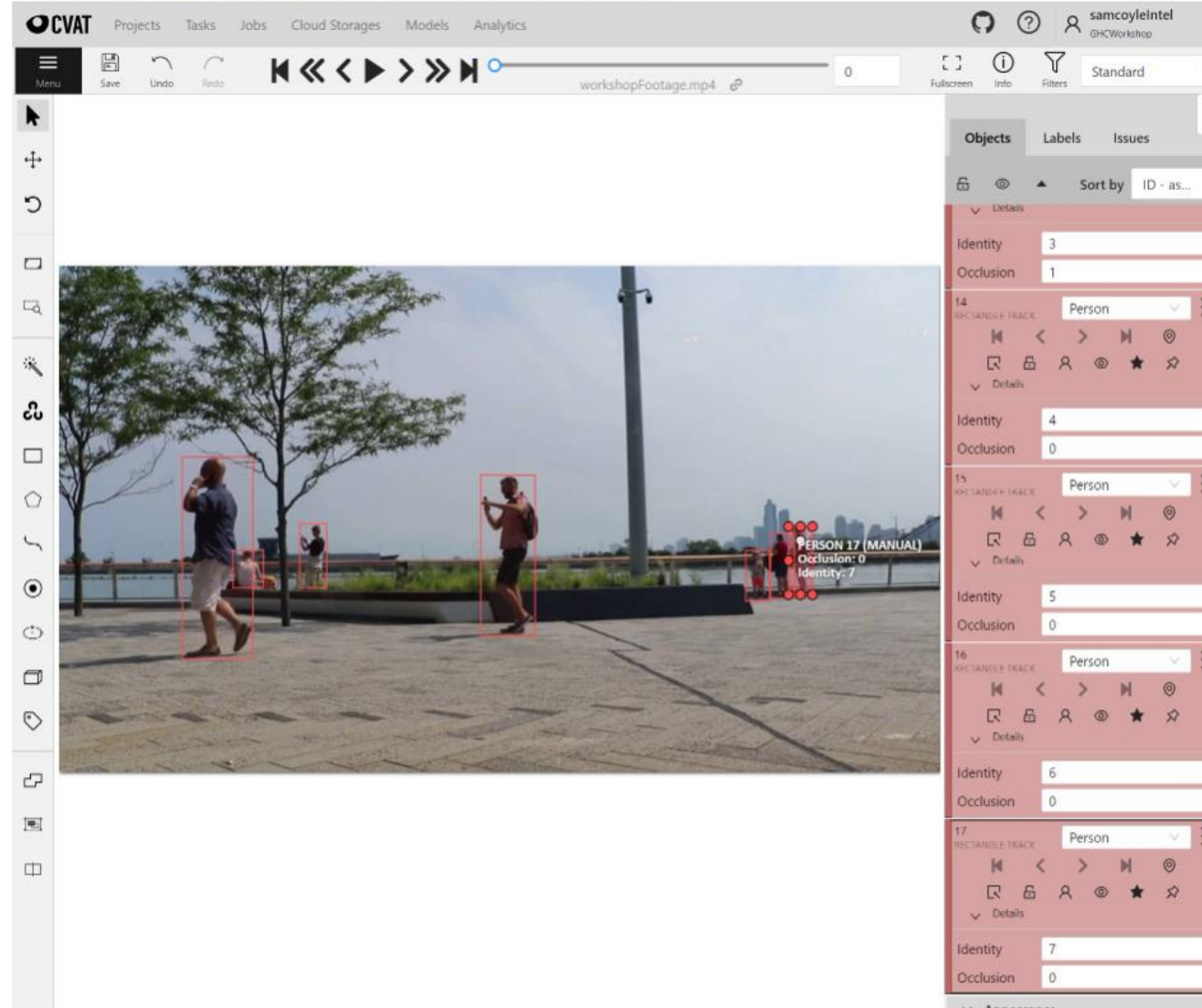
Note: Since Person 2 is seated and we only see ~50% of their body, we will mark their **occlusion** field as 1.

Also make sure to increment the number in **Identity** field for each new person. Refer to the point 3 in [Annotation Best Practices](#) for more details on this **Identity** field

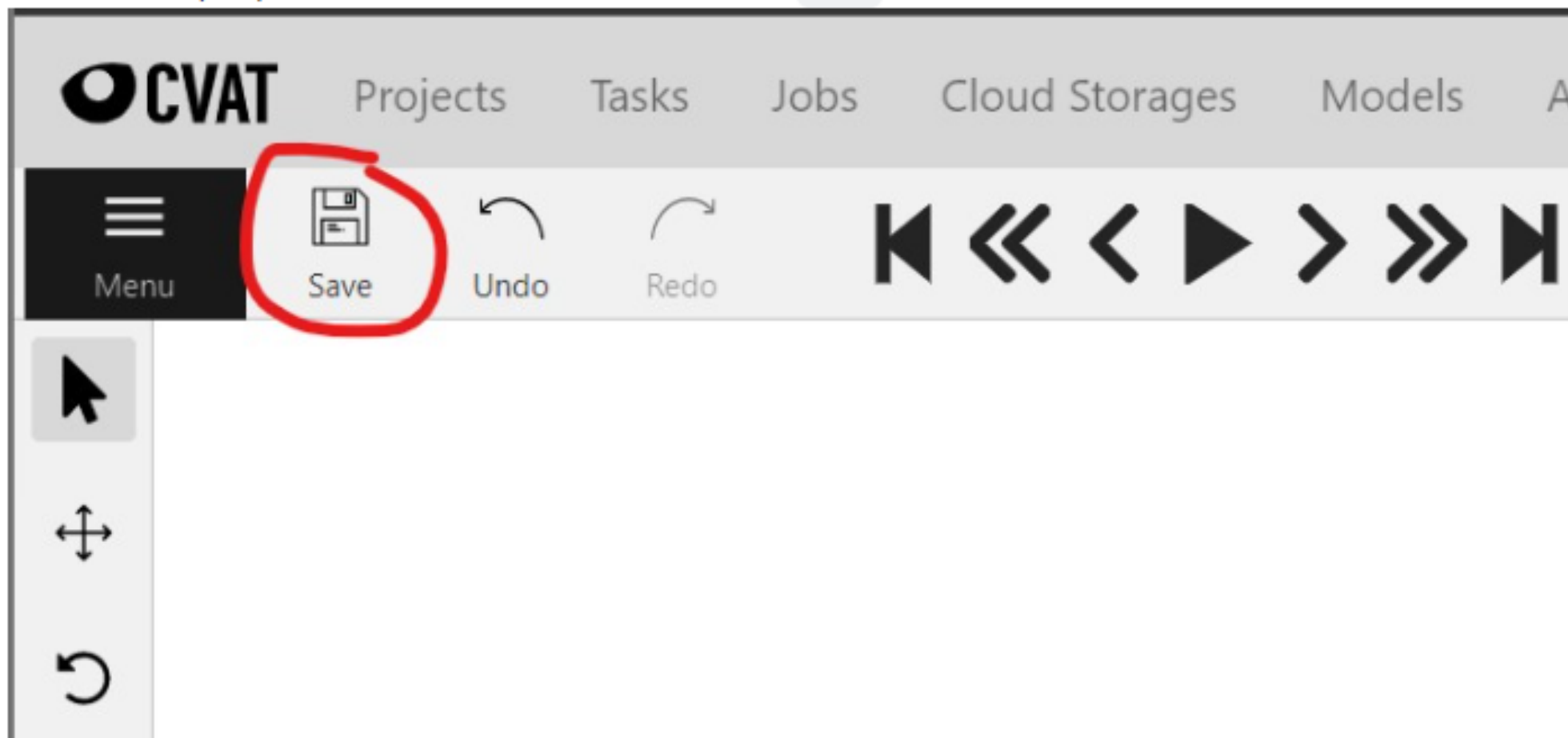
Remember to refer to the [Annotation Specifications](#) as needed to adjust your annotations based on occlusions.

The screenshot displays the CVAT web interface. At the top, there's a navigation bar with tabs for Projects, Tasks, Jobs, Cloud Storages, Models, and Analytics. Below this is a toolbar with icons for menu, save, undo, redo, and video navigation controls. The main area shows a video frame of two people walking on a path. A red bounding box is drawn around the person on the left. The right sidebar has tabs for Objects, Labels, and Issues. The 'Objects' tab is active, showing a list of objects. Two objects are listed: '11 PERSON' and '12 PERSON'. The '12 PERSON' entry has 'Identity' set to 2 and 'Occlusion' set to 1. The video player at the bottom shows the current frame with the bounding box and a label 'PERSON 12 (MANUAL)' with 'Occlusion: 1' and 'Identity: 2'.

When you are done annotating frame 0, it should look like the below picture.



6. Now that all people in the frame are annotated, click `Save` to save the current annotations.



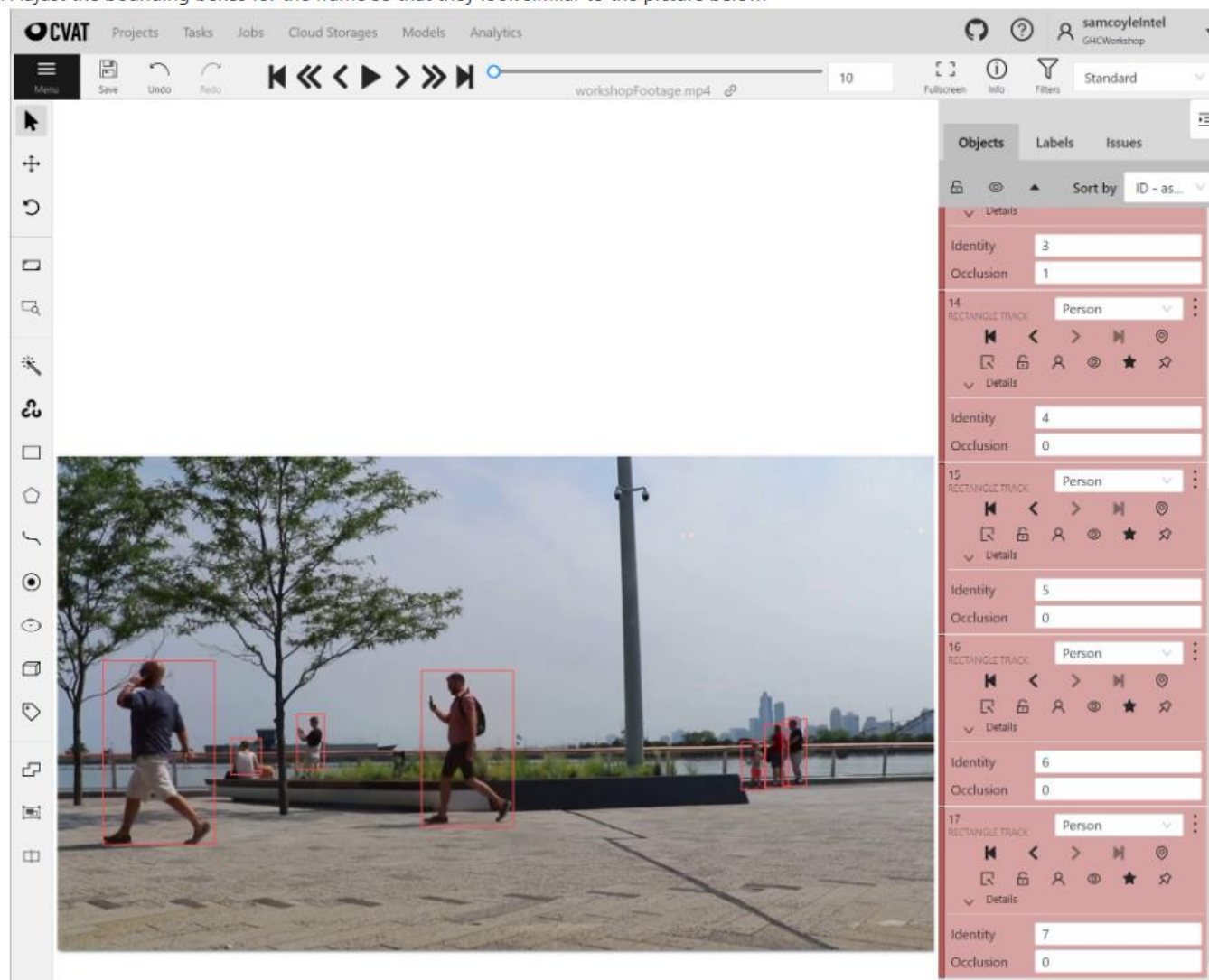
7. The track bar (marked in red below) can be used to scroll through the frames one after the other.



Click on the double right arrow to increment the frame count by 10.

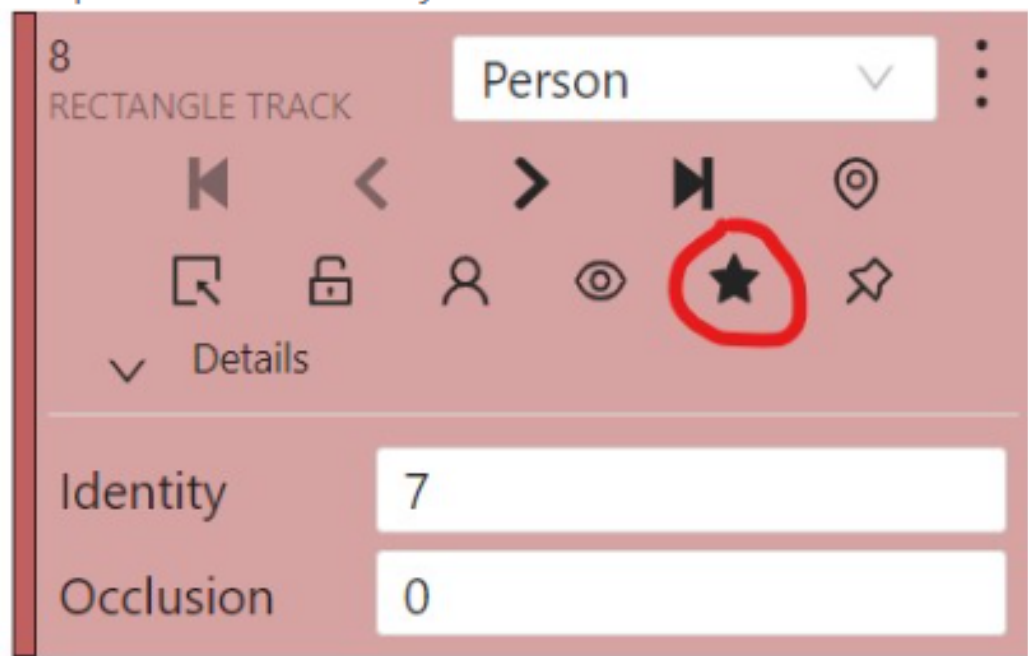
The screenshot displays the CVAT web interface for video annotation. The top navigation bar includes links for Projects, Tasks, Jobs, Cloud Storages, Models, and Analytics. The user is logged in as 'samcoyleIntel' in the 'G4CWorkshop' environment. The main toolbar contains icons for Menu, Save, Undo, Redo, and navigation controls. The navigation controls include a double right arrow (highlighted with a red circle) to increment the frame count by 10. The video file 'workshopFootage.mp4' is loaded, and the frame count is currently at 10 (also highlighted with a red circle). The video player shows a scene with several people walking. Red bounding boxes are drawn around some of the people. A tooltip for 'PERSON 14 (MANUAL)' is visible, showing 'Occlusion: 0' and 'Identity: 4'. On the right side, the 'Objects' panel is open, showing a list of tracked objects. The first object is 'Person' with Identity 3 and Occlusion 1. The second object is 'Person' with Identity 4 and Occlusion 0. The third object is 'Person' with Identity 5 and Occlusion 0. The fourth object is 'Person' with Identity 6 and Occlusion 0.

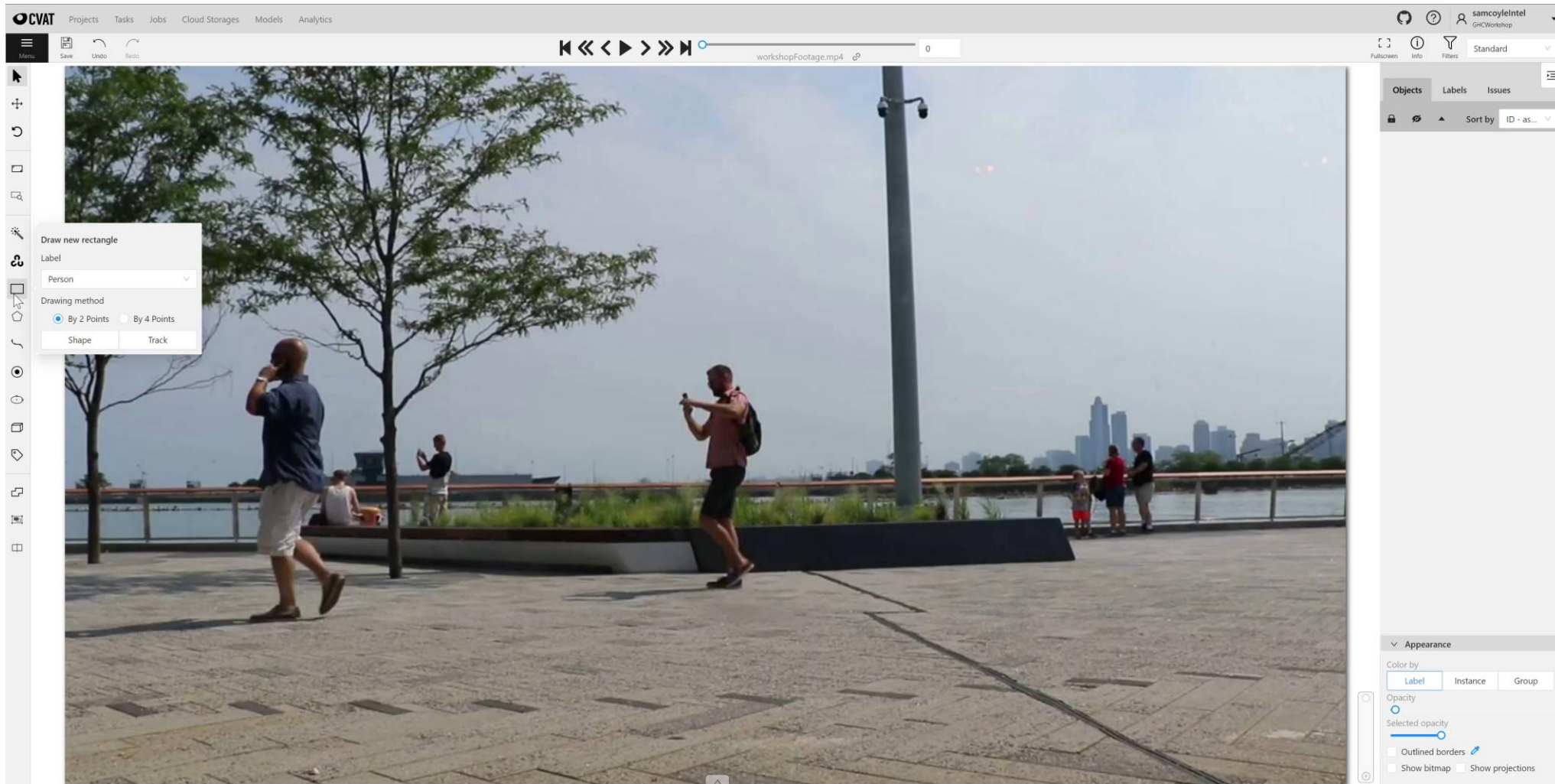
8. Adjust the bounding boxes for the frame so that they look similar to the picture below.



9. Continue steps 1-7 for every 10 frames of the video annotated each person in the frame.

Note: You will see that the initial bounding boxes are no longer properly placed. The bounding boxes size and placement will need to be adjusted since this is a different frame, and people have moved. Also, by annotating every 10 frames, we are leveraging CVAT's interpolation feature. The interpolation feature allows annotators to annotate key frames, and then the frames between them will be interpolated automatically. Each frame annotated is automatically denoted as a key frame with the star shown below.

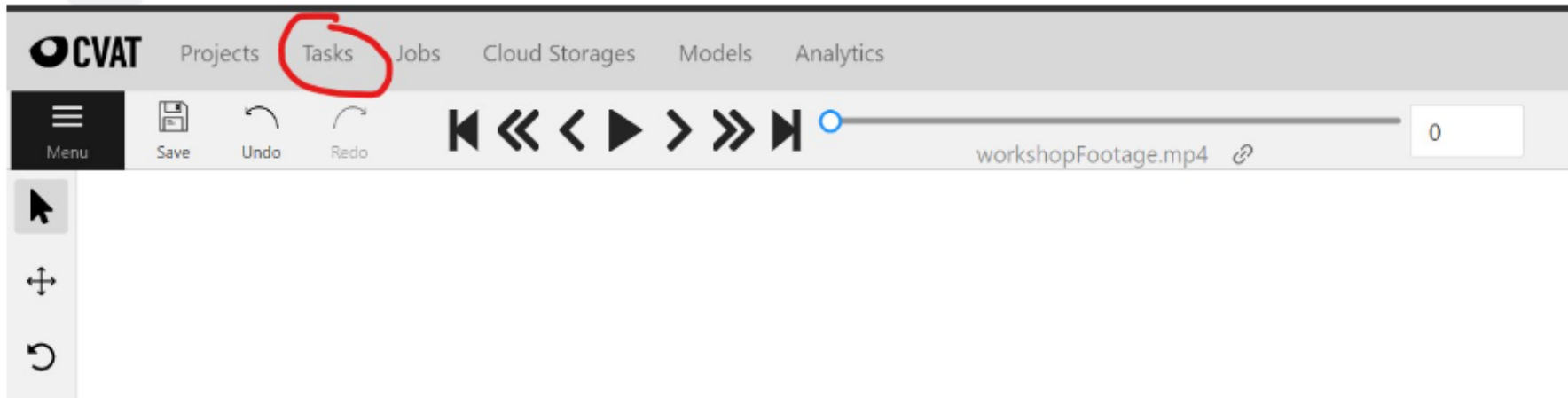




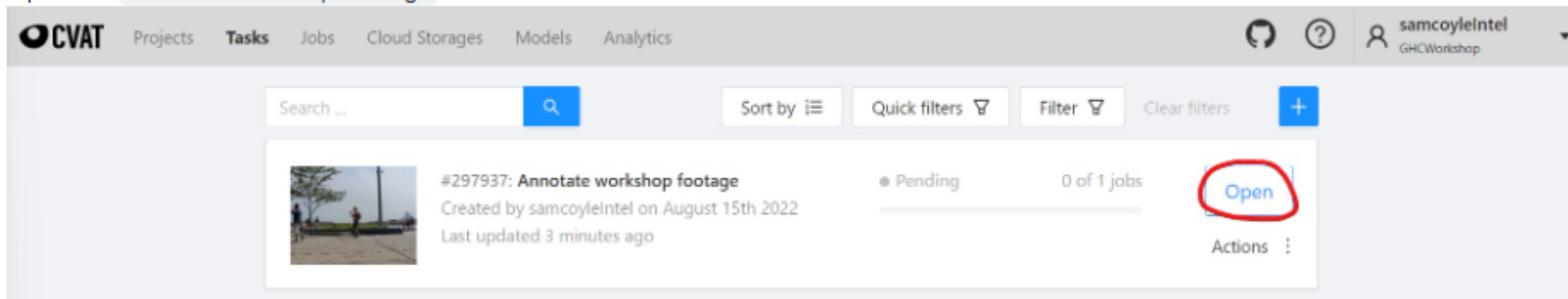
Steps to follow once frames have been annotated

The following steps are meant to demonstrate to workshop attendees that data annotations go beyond a one-person effort. Data annotation efforts may be reviewed by other team members, and even exported upon completion.

1. Click **Tasks**.



2. Open the **Annotate workshop footage** task.



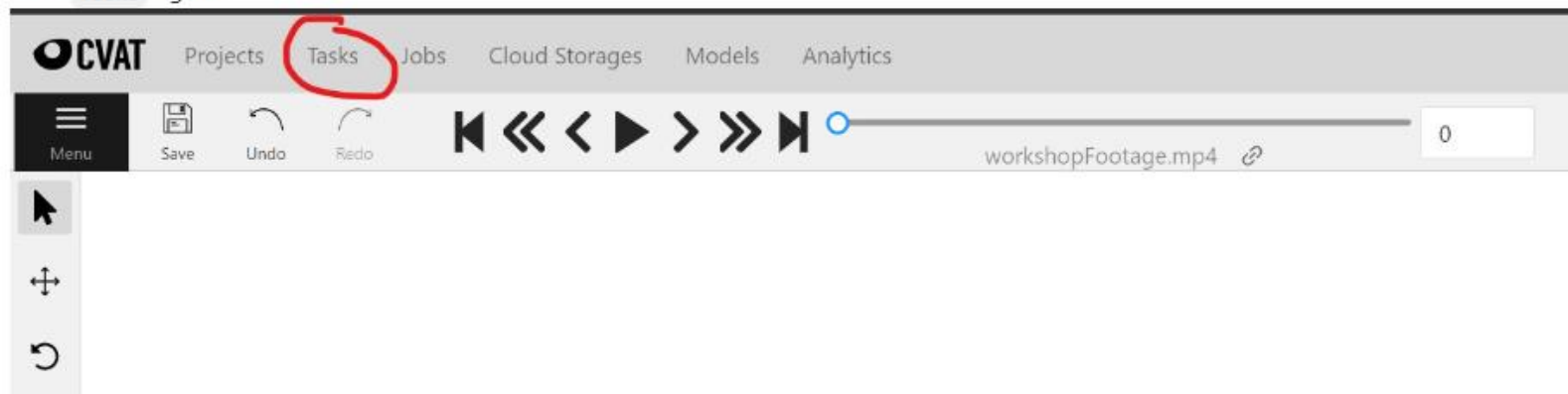
3. Update the Job Stage to validation .

The screenshot shows the CVAT web interface. At the top, there's a navigation bar with 'CVAT' logo and links for Projects, Tasks, Jobs, Cloud Storages, Models, and Analytics. The user 'samcoyleIntel' is logged in. Below the navigation bar, there's a 'Back to project' link and an 'Actions' menu. The main content area displays a task titled 'Annotate workshop footage'. It includes a video thumbnail, task details (Task #297937, Created by samcoyleIntel on August 15th 2022), an 'Assigned to' dropdown, an 'Issue Tracker' link, and a 'Subset' dropdown set to 'Train'. Below this is a table with columns 'Overlap size', 'Segment size', and 'Image quality', containing the values 5, 1093, and 70 respectively. At the bottom, there's a 'Jobs' section with a 'Copy' link and a table showing 0 of 1 jobs. The table has columns: Job, Frames, Stage, State, Started on, Duration, and Assignee. The first row is for Job #449473, with Frames 0-1092, Stage 'annotation', State 'in progress', Started on August 15th 2022 20:08, Duration 6 days, and Assignee samcoyleIntel. A dropdown menu is open for the 'Stage' column, showing options: 'annotation', 'validation' (circled in red), and 'acceptance'. There is also a 'validation' button below the dropdown.

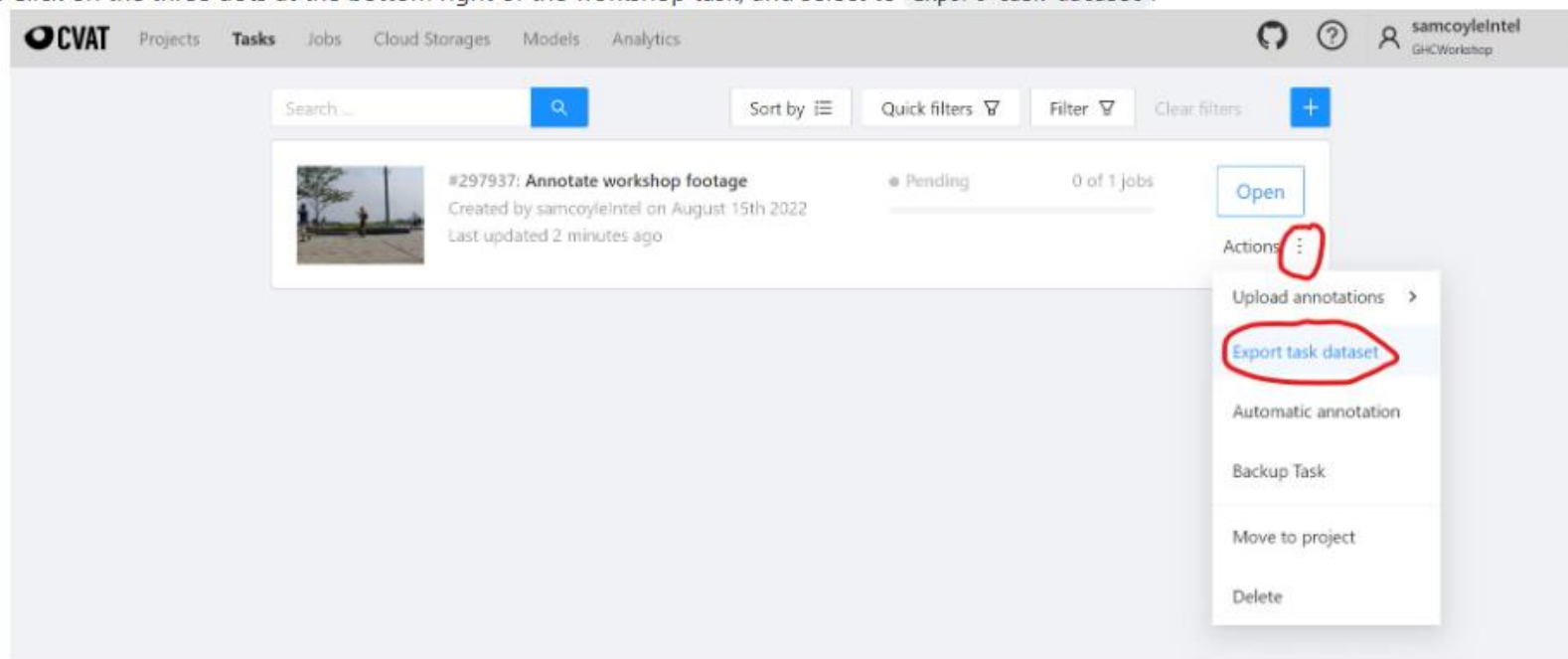
Job	Frames	Stage	State	Started on	Duration	Assignee
Job #449473	0-1092	annotation	in progress	August 15th 2022 20:08	6 days	samcoyleIntel

Note: This is an important step when participating in data annotations within a team environment. Updating the stage from `annotation` to `validation` signals to others that the current annotations are ready for others to review. Once they have been reviewed and adjusted as necessary, then the annotations may be marked `accepted` .

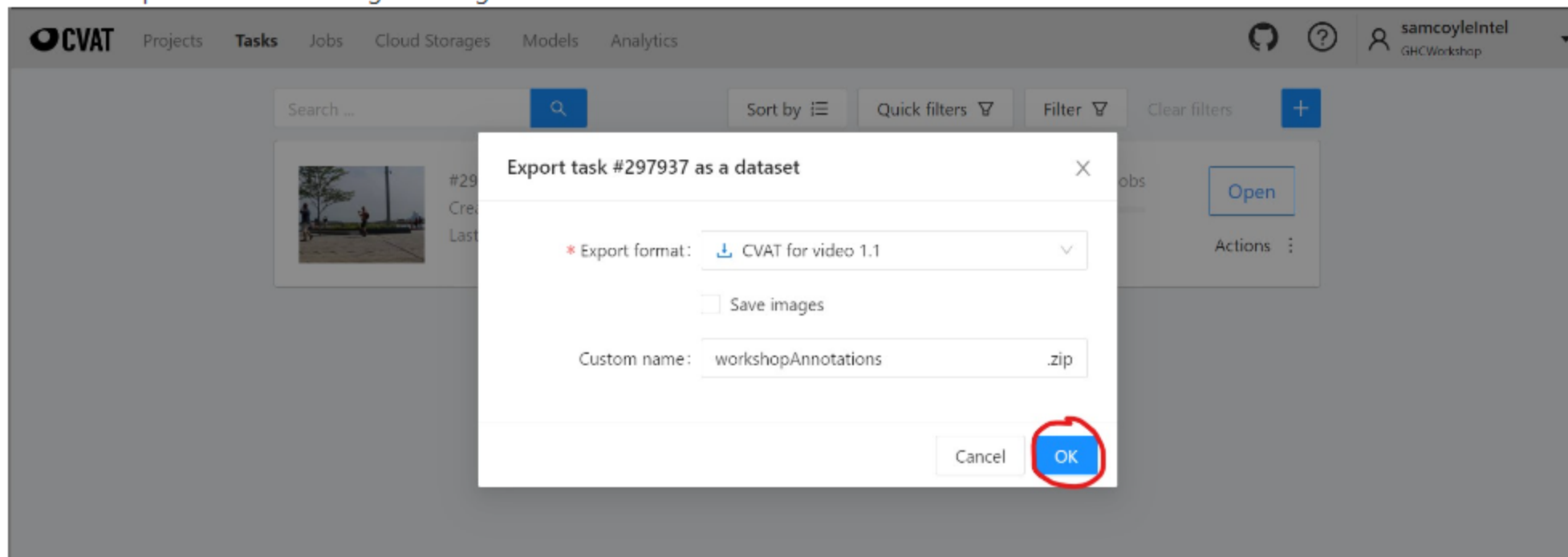
4. Click **Tasks** again.



5. Click on the three dots at the bottom right of the workshop task, and select to **Export task dataset**.



6. Fill in the export details following the image below.

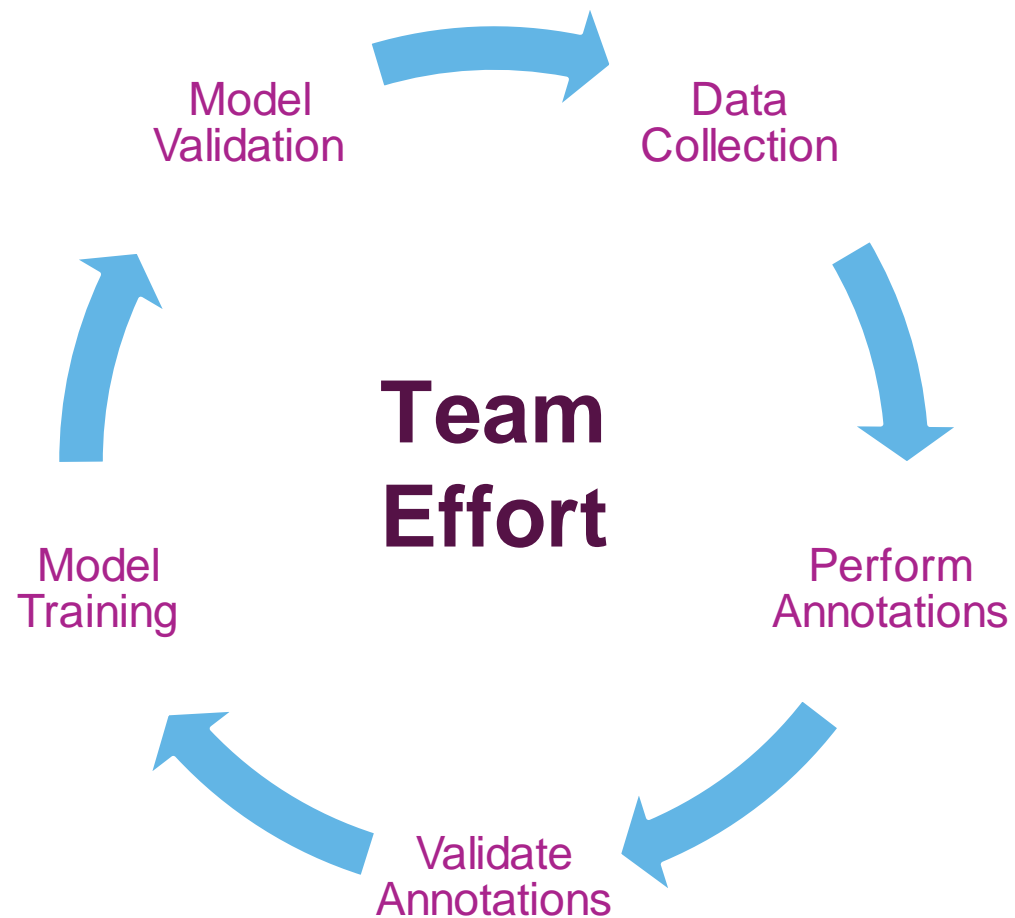


Note: You will now see the annotations downloaded to a zip folder in the Downloads folder. The zip folder contains an annotations XML formatted file that encodes the annotated data. Different formats may be selected pending the format needed to train the specified model of choice. More information may be found in the section on [Annotation Format](#). The file contains tags corresponding to the project pieces created for this workshop (ie `<task>`, `<labels>`, `<track>`, `<box>`). Those tags correspond to the task, label, and track ID, bounding box of the people annotated, etc. This exported data is what would be used to train a model.



Conclusion

Annotation Workflow



CVAT for Teams

CVAT Projects Tasks Jobs Cloud Storages Models Analytics

Back to project Actions

Annotate workshop footage

Task #297937 Created by samcoyleIntel on August 15th 2022

Assigned to
Select a user

Issue Tracker
Not specified

Subset:
Train

Overlap size	Segment size	Image quality
5	1093	70

Jobs

Copy 0 of 1 jobs

Job	Frames	Stage	State	Started on	Duration	Assignee
Job #449473	0-1092	annotation	in progress	August 15th 2022 20:08	6 days	samcoyleIntel

annotation
validation
acceptance validation

- Local installation
- Simple setup
- Open-Source solution
- Better support professional data annotation teams

CVAT Roles

Permissions

☒ Active
Designates whether this user should be treated as active. Unselect this instead of deleting accounts.

☐ Staff status
Designates whether the user can log into this admin site.

☐ Superuser status
Designates that this user has all permissions without explicitly assigning them.

Groups:

Available groups ⓘ

admin
annotator

Chosen groups ⓘ

observer
user

Choose all ⓘ

Remove all ⓘ

The groups this user belongs to. A user will get all permissions granted to each of their groups. Hold down "Control", or "Command" on a Mac, to select more than one.

Thank You



Samantha Coyle
Software Engineer



Neethu Elizabeth Simon
Senior Software Engineer





NEXT IS
NOW

Use Open-Source Tool to Simplify & Automate Video/Image Labeling Process



Samantha Coyle



Neethu Elizabeth Simon



Overview

- Introduction – 5 min (N)
- Project Setup – 10 min (S)
- Annotation Specifications – 2 min (N)
- Annotation Best Practices – 3 min (S)
- Perform Annotations – 30 min (S & N slides 42-45)
- Conclusion – 5 min (S)
- QA / wiggle room – 5 min



Thank You

