## **Week 1 (05/30-06/2)**

**Day 1 (Tuesday, May 30th, 2023):** Today was my first day working at Argonne. I was pretty nervous to start but after going through my day I calmed down. The day started off with attending orientation from 9-12, and then lunch from 12-1. It was then time to report to my workstation and meet with my mentors and supervisor. It was really great to meet them, and they gave me a rundown of my project and what to expect from it. I didn’t do much actual work, but I spent the rest of the day completing all my required training, and also set up my computer to be able to code.

**Day 2 (Wednesday, May 31st, 2023):** Today started off by once again attending orientation from 9-12. I went to lunch yet again, and at 1 reported back to my workstation. I had to complete the rest of my training as they had added some for me to do. I then began to start working on my project. My first task was to convert all of the xml files from the given data into one large csv file, with information such as the filename, the folder, and the coordinates of the smoke. I worked on this all day, but got stuck on figuring out how to get the coordinates into their correct spots. The reason for this was because the direct child of the object in the xml file was not multiple x’s and y’s, but one large text which I would have to split up in order to get the respective x’s and y’s.

**Day 3 (Thursday, June 1st, 2023):** Today was my first full day at work. I got to my station at 8:30, and began by checking all my emails and making sure I'm up to date with everything. I then created my personal introduction chart, and then I began trying to crack down on the problem. I managed to figure out how to successfully get the maximum and minimum of the x’s and y’s into their respective dataframe locations. I did this by looping through the child object and comparing them to preset variables in order to get the absolute max and absolute min. After I completed this, I went over to the weekly meeting at 10:00, in which we all introduced ourselves, and asked any question that we may have had. I created a different function, which would iterate through every folder and file of xml’s from the data, run each of those xml files through the single xml to csv file function which I created. In the code it would create each individual dataframe, and then concat or combine them all into one large dataframe. I was very proud after I completed this step, because it felt like it took forever. After this I went to lunch until 1:00. Once I came back, I started trying to figure out how to crop each image. My next step was to take every instance of smoke from the csv file I created, and crop it at the given points to only have pictures of the actual smoke. Bhupendra told me to first only crop one individual image, in order to see how it crops. I used the Pillow Image functions in order to do this, but once I cropped the image I realized that something was wrong. It didn’t crop only the smoke and instead cropped the image randomly. I thought about why this may be and after looking through my code I realized that when I actually converted the xml to the df, I was grabbing the points of the image, and not the points of the boundary boxes of the smoke. I went back to fix my mistake, and once I had gone through my other code to fix and other errors that arose, the single image crop worked perfectly. I then wrote a function which would take the csv file of size 9,550, and crop each image in the dataset to their smoke coordinates. This function was much easier to create, but when I was running it (which took a very long time because I was saving 9,500 cropped images into a separate folder on my computer), the function kept crashing at one given moment. After looking at the error message which said the given directory was not found, I double checked and found that there was actually a typo in the xml file for the folder name, where it contained a dash and not an underscore. I added a condition in the code where if that specific folder was found, to change the name to its correct name, which worked. This error happened a few times while running the code, but I managed to correct all of the discrepancies and the function finished running. I now have a complete file with all images of the smoke which were cropped.