Logo-istics has progressed pretty smoothly over the last two weeks as significant progress has been made in preprocessing images to be passed into the model. The implementation of BLIP has worked pretty well being able to describe the abstract nature of logos. Currently, only a subset of the main dataset has been run through blip and shows pretty promising results. This subset can be found at: <https://huggingface.co/datasets/samp3209/blip2logos>

The next steps are to blip the entire logo dataset and then once we get ahold of Champlain’s creative assets, blip them as well. The next step would be to create a training configuration setup and find a cloud computing service to run the training on. This would be the most difficult part of the project and most likely will consume the most amount of time. My strategy for this would be to create a new GitHub repo for the sake of cleanliness/privacy and then upload all the required JSON and YAML files needed to run the training for the model. The expected cost for this training should be around ten dollars depending on the current rates for prices and how long it will take for our image-text captions to be trained on. The Pokemon Stable Diffusion project, which was a massive inspiration for this project, took about 6 hours to train on a dataset of image-text that is a similar size to what we have. The cost for that run was about 10$ so assuming that our blip captions are more lengthy we could expect a larger bill and longer run time.