

Team Name: StandAI

Group #3

Arya Miryala USCID: 9567678120

Grace Wu USCID: 1980098805

Hui Xie USCID: 7956658480

Meeting Minutes

Date	Attendance + Summary
2/11 Meeting duration: 1 hour	All members present <ul style="list-style-type: none">- Reviewed the assignment and delegated tasks- Decided on what subreddit to scrape- Hui - Build the scraper and preprocess the text. Connect scraper to mysql- Arya - Look over scraper and preprocessing and assist with clustering. Build OCR- Grace - Assist with clustering and build automation Next steps: Begin implementing the scraper and setting up the database schema.
2/12 Meeting duration: 3 hours	All members present <ul style="list-style-type: none">- We discovered the Reddit API requires access permissions due to new updates from a few months ago. Arya and Hui submitted requests, but they were not automatically approved.- Hui used JSON endpoint instead for data scraping- Set up README so everyone has the same set up on Linux Next steps: Continue preprocessing implementation and begin testing OCR on subreddit images.
2/13 Meeting duration: 3 hours	All members present <ul style="list-style-type: none">- Continued to work on each part simultaneously- The subreddit we used, Hardwareswap didn't have many images, but we used /r memes to check and see if the OCR is working- Switched to testing with /r investing Next steps: Integrate scraper, preprocessing, OCR, and clustering into one pipeline.
2/14 Happy Valentines~ Meeting duration: 2 hours	All members present <ul style="list-style-type: none">- Everyone tested the whole pipeline in Linux

	<p>environment to ensure it is replicable</p> <ul style="list-style-type: none">- Updated the README for submission- Deligated who to submit the files.
--	--