Progress Report 1

COSC 4P02 – Software Engineering 2 (Winter 2024)

Professor- Naser Ezzati-Jivan

February 26th, 2024

Project Title: Social Media Post Generator

GitHub: https://github.com/mc16dn/COSC-4P02-Group-Assignment/tree/main

Jira: https://brocku-hh18iq.atlassian.net/jira/software/projects/SMPG/boards/1/backlog

Project Members

| Name | Role | ID | Brock Email | GitHub |
|--------------|---------------|---------|------------------|--------------|
| | | | | Username |
| Basel Abdel | Product Owner | 6826853 | ba19tw@brocku.ca | Hafezberg |
| Hafez | | | | |
| Hevar | Scrum Master | 6701248 | hh18iq@brocku.ca | spiralwind |
| Halabjaee | | | | |
| Rifat | Project | 6830301 | rc19nv@brocku.ca | Rifat- |
| Chowdhury | Manager | | | Chowdhury |
| Raymond Dong | Developer 1 | 6592539 | rd18sx@brocku.ca | raymonddong2 |
| Matthew Cam | Developer 2 | 6195556 | mc16dn@brocku.ca | mc16dn |
| Maisam Anjum | Developer 3 | 6804298 | ma19an@brocku.ca | Koorikdat |

Contents

| Projec | ct Design | | |
|---------|---------------------------------|---|--|
| 1.1 Te | eam Member Contributions | 2 | |
| ı. | Matt: | 2 | |
| II. | Raymond: | 3 | |
| III. | Hevar: | 3 | |
| IV. | Basel: | 3 | |
| V. | Rifat: | 3 | |
| VI. | Maisam: | 3 | |
| 1.2 - I | Project Feasibility and Vision: | 4 | |
| 1.3 – | Progress: | 4 | |
| | Testing: | | |

| 1.5 – Timeline: | 5 |
|------------------------------------|----|
| 2 – Sprints and Iteratives: | 6 |
| 2.1 - Sprint 1 | 6 |
| 2.2 - Sprint 2 | 7 |
| 2.3 - Sprint 3 | 8 |
| 2.4 - Future Sprints/Backlog | 8 |
| 3 – Meetings | 9 |
| Meeting 1 (Saturday January 20th): | 9 |
| Meeting 2: (Monday January 22): | 9 |
| Meeting 3: (Monday January 29th): | 10 |
| Meeting 4: (Monday February 5th): | 10 |
| Meeting 5: (Sunday February 11th): | 11 |
| Meeting 6: (Tuesday February 13): | 11 |
| Meeting 7: (Monday February 26th): | 12 |
| 4 – GitHub | 13 |
| 5 - Issues | 14 |

1 Project Design

This section will cover our design choices and philosophy and a brief description of the project's present and future vision.

1.1 Team Member Contributions

I. Matt

- Researched subs to use for content and created the Git project.
- Created function that grabs URLs of Reddit posts based on sub and amount requested.
- Integrated URLs grabber and content scrapper.
- Researched TTS software and provided code examples.
- Completed User Story research which subreddits are the best to grab stories from
- Completed User Story allow the reddit grabber to grab stories from different reddit.
- Completed User Story fetching from random 10 top posts from a given sub-reddit.
- Completed User Story develop multiply different options for the voice (around 3)
- Completed User Story Research useable voice modules.

II. Raymond:

- researched the maximum duration a video can be.
- ensure a post is of an appropriate length.
- integrated tts code with the reddit grabber.
- Completed User Story test the integration so that there are no errors because of the implementation.
- Completed User Story integrate the voice with the reddit grabber so that it can read the text.

III. Hevar:

- Made the final product and sprint backlog decisions.
- Communicated with Basel to ensure accuracy in the report.
- Contributed to the research material that needs to be looked on to ensure the success of
- the project.
- organization weekly and important meetings and managing time and people.
- Keeping Jira update to date and updating sprint's backlog and times.
- Completed User Story research which subreddits are the best to grab stories from
- Completed User Story make censorship for the voice so that it avoids inappropriate words.
- Completed User Story test the censorship and voice so that there are no errors involved.

IV. Basel:

- Created User epic.
- Contributed to user stories.
- Collected the team's thoughts.
- Contributed to the Release Planning doc editing.
- Formatting and updated the Progress report 1 constantly.
- Keeping track of the meetings by meeting logs

V. Rifat:

- Contributed to user stories.
- Formatting the Release Planning doc.
- Managing Miro.
- Updating readme file.
- Created the License file.
- Assisting Basel with User Epic decisions.
- Created a timeline of the different sprints with Basel.
- Formatting and updating the Progress report 1 constantly.
- Arranged meeting's and keeping track of due dates.
- Assisting Basel with the meeting logs.

VI. Maisam:

• Completed User Story – register with reddit API.

- Completed User Story use reddit API to pull from given URL.
- Completed User Story research how to grab and read stories from reddit.
- Completed User Story test the censorship and voice so that there are no errors involved.
- Completed User Story allow easy switching between voices and test that it doesn't cause errors.
- Completed User Story make the voice of the bot read any text.
- Assisted Hevar with research.
- Assisted with making decisions regarding the program's implementation.
- Assisted in discussing the risk management of the project.

1.2- Project Feasibility and Vision:

The end goal of this project is a video generator that will upload videos to YouTube Shorts autonomously. These videos will consist of a TTS (text-to-speech) voice reading a Reddit post to the viewer. It seems simple on paper but an overwhelming number of possibilities on these videos' design can heavily influence a given video's success. Considering this fact, the program will allow the user to edit the video in a multitude of ways, we will refer to these edits as permutations and they will be a critical feature of the final product. Some examples of permutations include music, sound effects, the type of TTS voice used, layout, length, type of background video, and subtitle size. The program will mix and match different combinations of these permutations and analyze the metrics, outputting the relevant information directly to the user after a set data-collection period. The decision tree can analyze these metrics to spot a pattern to videos' success, and then output that correlation to the user for future reference. A video to demonstrate this project's premise can be found here:

https://www.tiktok.com/@tiktok.readings/video/7341460576436112670?q=reddit&t=1706570724915

1.3 – Progress:

As of this report's writing, we have successfully implemented the project's foundation. Namely, interfacing with Reddit's API, setting up a reliable grabber, and censorship. Configuring the text-to-speech functionality is our current objective. With these obstacles behind us, the results of future sprints will be much more tangible. Starting with the next sprint, videos will begin to be generated; with layout and permutation testing to follow soon after. Below is an example execution that demonstrates the work done so far.

URL: https://www.reddit.com/r/stories/comments/lavriue/i_never_realized_how_disgusting_a_lot_of_men_were/.json Subreddit: stories

Upvotes: 1575

Precensored Text: I'm 25F. Now let me start off by saying that I know not all men are disgusting. There are so many really good and kind men out there and I truly do appreciate the men that I and other women can feel safe with. When I was a teenager, I was very dumb and naive. At the time, I really didn't see anything wrong with an 18 year old or 20 something year old trying to talk to me, flirting with me and calling me pretty. I felt so cool because an older guy was talking to me. I remember being 17 and men in their 38 asking me when I was turning 18. I look pretty young for a 25 year old and I've noticed when I tell men older than me, my age, they seem to lose interest. It's creepy honestly. Not t hat long ago, I was talking to a guy that I met at Nork. He's 36. He was a very good looking guy, very respectful to me in the workplace, etc. But he was very creepy about our age difference and would say things like "you were in element ary school when I was in my late teens or early 20s, but I would've still f'eked you' immediately, I was like "HELL NO!" . I stopped talking to him. It's especially more concerning to me because I have a daughter. No one I seen has ever met her and I never let them meet my kids unless I feel safe around them. They know I have kids but there is absolutely no way, I'm letting randoms meet them. I honestly don't think I even want to date until my kids are adults themselves . I know not all men are like this, but the men I've encountered make me feel like it's litteraly all men. Fuck

Postcensored Text: I'm 25F. Now let me start off by saying that I know not all men are disgusting. There are so many really good and kind men out there and I truly do appreciate the men that I and other women can feel safe with. When I was a teenager, I was very dumb and naive. At the time, I really didn't see anything wrong with an 18 year old or 20 something year old trying to talk to me, flirting with me and calling me pretty. I felt so cool because an older guy was talking to me. I remember being 17 and men in their 36s asking me when I was turning 18. I look pretty young for a 29 year old and I've noticed when I tell men older than me, my age, they seem to lose interest. It's creepy honestly. Not that I ong ago, I was talking to a guy that I met at work. He's 36. He was a very good looking guy, very respect 12 by men to me in the workplace, etc. But he was very creepy about on age difference and would say things like "you were in elemen tary school when I was in my late teens or early 20s, but I would've still f'cked you' immediately, I was like "HELL NO!". I stopped talking to him. It's especially more concerning to me because I have a daughter. No one I seen has ever met her and I never let them meet my kids unless I feel safe around them. They know I have kids but there is absolutely no way, I'm letting randoms meet them. I honestly don't think I even want to date until my kids are adults themselve s. I know not all men are like this, but the men I've encountered make me feel like it's litteraly all men. ****

I never realized how disgusting a lot of men were until I became an adult

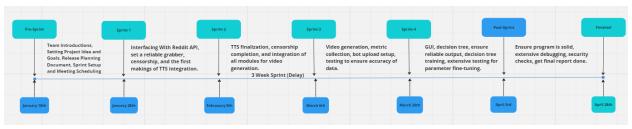
Venting

I'm 25F. Now let me start off by saying that I know not all men are disgusting. There are so many really good and kind men out there and I truly do appreciate the men that I and other women can feel safe with. When I was a teenager, I was very dumb and naive. At the time, I really didn't see anything wrong with an 18 year old or 20 something year old trying to talk to me, flirting with me and calling me pretty. I felt so cool because an older guy was talking to me. I remember being 17 and men in their 30s asking me when I was turning 18. I look pretty young for a 25 year old and I've noticed when I tell men older than me, my age, they seem to lose interest. It's creepy honestly. Not that long ago, I was talking to a guy that I met at work. He's 36. He was a very good looking guy, very respectful to me in the workplace, etc. But he was very creepy about our age difference and would say things like "you were in elementary school when I was in my late teens or early 20s, but I would've still f*cked you" immediately, I was like "HELL NO!" . I stopped talking to him. It's especially more concerning to me because I have a daughter. No one I seen has ever met her and I never let them meet my kids unless I feel safe around them. They know I have kids but there is absolutely no way, I'm letting randoms meet them. I honestly don't think I even want to date until my kids are adults themselves. I know not all men are like this, but the men I've encountered make me feel like it's litteraly all men.

1.4 - Testing:

As of this report, no extensive testing has been conducted. Though, a lot of trial and error was applied to find a correct, efficient, and free method of interfacing with Reddit. We eventually found a way to turn a post into a JSON file using a Scraper, which worked very well. When the next sprint starts, testing will play a huge role in video generation and parameter fine-tuning. The makings of a test framework has already been made, and can be accessed through this link: https://github.com/mc16dn/COSC-4P02-Group-Assignment/blob/main/testing.py

1.5 – Timeline:

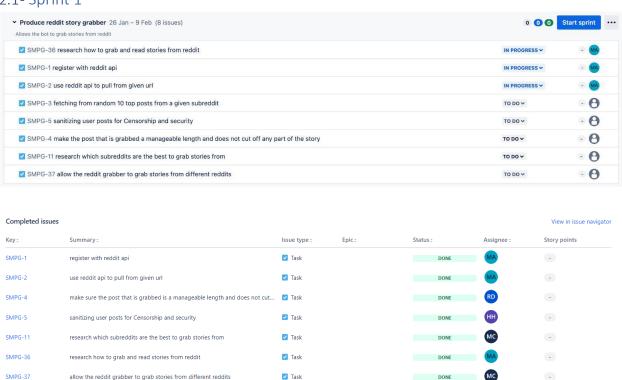


2 – Sprints and Iteratives:

Please refer to the timeline for a brief overview of each sprint's goals.

Our project completely supports the Agile philosophy, which emphasizes flexibility, cooperation, and responsiveness to change. This approach is represented in our project model, which supports a variety of meetings, including weekly sprint reviews. These reviews provide a forum for the team to address difficulties and implement adjustments based on meeting outcomes, ensuring that the project moves forward effectively toward its ultimate state. A prominent example came during the problems of the second sprint, when more work was necessary during the reading week. We were able to increase the second sprint's duration by one week using the Agile methodology. This modification gave us more time to appropriately handle possible difficulties. Due to such a process approach, it is possible to make certain changes with ease. Besides it allows the development team to make fast fixes and solutions by breaking it down to manageable units for either the current or the future sprints.

2.1- Sprint 1



Issue type :

Task

View in issue navigator

Story points

Assignee :

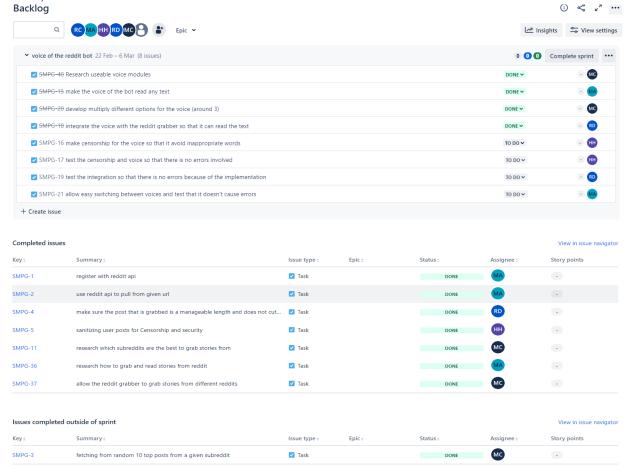
МС

SMPG-3

Issues completed outside of sprint

| Date | Event | Issue | Completed | Scope |
|-----------------------------|------------------|---|-----------|-------|
| Sun, Jan 28 2024, 8:00pm | Sprint started | SMPG-1 register with reddit api SMPG-2 use reddit api to pull from given url SMPG-3 fetching from random 10 top posts from a given subreddit SMPG-4 make sure the post that is grabbed is a manageable length and does not cut off any part of the story SMPG-5 sanitizing user posts for Censorship and security SMPG-11 research which subreddits are the best to grab stories from SMPG-37 essearch how to grab and read stories from reddit SMPG-37 allow the reddit grabber to grab stories from different reddits | 0 | 0 |
| Sun, Feb 04 2024, 4:40pm | Issue completed | SMPG-11 research which subreddits are the best to grab stories from | 0 | 0 |
| Mon, Feb 05 2024, 6:03pm | Issue completed | $\label{eq:SMPG-4} \textbf{SMPG-4} \ make sure the post that is grabbed is a manageable length and does not cut off any part of the story$ | 0 | 0 |
| Mon, Feb 05 2024, 6:03pm | Issue completed | SMPG-2 use reddit api to pull from given url | 0 | 0 |
| Mon, Feb 05 2024, 6:03pm | Issue completed | SMPG-1 register with reddit api | 0 | 0 |
| Mon, Feb 05 2024, 6:03pm | Issue completed | SMPG-36 research how to grab and read stories from reddit | 0 | 0 |
| Tue, Feb 06 2024, 6:28pm | Issue completed | SMPG-37 allow the reddit grabber to grab stories from different reddits | 0 | 0 |
| Tue, Feb 06 2024, 6:28pm | Issue completed | SMPG-5 sanitizing user posts for Censorship and security | 0 | 0 |
| Thu, Feb 22 2024, 2:27pm | Sprint completed | SMPG-1 register with reddit api SMPG-2 use reddit api to pull from given url SMPG-3 fetching from random 10 top posts from a given subreddit SMPG-4 make sure the post that is grabbed is a manageable length and does not cut off any part of the story SMPG-5 sanitizing user posts for Censorship and security SMPG-11 research which subreddits are the best to grab stories from SMPG-36 research how to grab and read stories from reddit SMPG-37 allow the reddit grabber to grab stories from different reddits | 0 | 0 |

2.2- Sprint 2



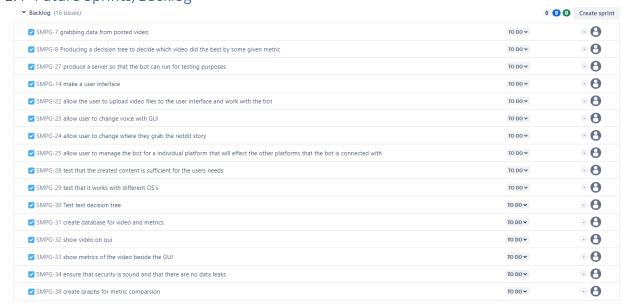
| Date | Event | Issue | Completed | Scope |
|-----------------------------|---------------------|--|-----------|-------|
| Thu, Feb 22 2024, 2:27pm | Sprint started | SMPG-15 make the voice of the bot read any text SMPG-16 make censorship for the voice so that it avoid inappropriate words SMPG-17 test the censorship and voice so that there is no errors involved SMPG-18 integrate the voice with the reddit grabber so that it can read the text SMPG-19 test the integration so that there is no errors because of the implementation SMPG-20 develop multiply different options for the voice (around 3) SMPG-21 allow easy switching between voices and test that it doesn't cause errors SMPG-40 Research useable voice modules SMPG-26 combine the video with the text from reddit and the voice to produce content that can be uploaded | 0 | 0 |
| Mon, Feb 26 2024, 6:17pm | Removed from sprint | $\ensuremath{SMPG-}\xspace 26$ combine the video with the text from reddit and the voice to produce content that can be uploaded | 0 | 0 |
| Mon, Feb 26 2024, 6:23pm | Issue completed | SMPG-18 integrate the voice with the reddit grabber so that it can read the text | 0 | 0 |
| Mon, Feb 26 2024, 6:23pm | Issue completed | SMPG-15 make the voice of the bot read any text | 0 | 0 |

2.3- Sprint 3

Projects / Social Media Post Generator

Backlog Q RC MA HH RD MC P Epic V Insights 🗢 View settings 0 0 0 Start sprint ··· ··· 🖰 ··· ☐ ✓ SMPG-6 producing 18 video templates ✓ TO DO ~ · 0 ✓ SMPG-41 allow user to upload file TO DO **▽** · 0 ☑ SMPG-43 create generic background for text and make it readable TO DO 🕶 **8** ☑ SMPG-39 create subtitles for reddit post on video automatically SMPG-26 combine the video with the text from reddit and the voice to produce content that can be uploaded TO DO 🗸 • **(9 · 0** TO DO 🗸 **• 0** ☑ SMPG-35 Test that the metrics given returned from the video are sound and accurate TO DO 🗸 **· (9** ✓ SMPG-9 grab video metrics from a given channel (watch time, total clocks, total view and total shares) TO DO 🗸 **8** ✓ SMPG-44 test that the given metrics are similar to the ones stated in the creator hub for the video TO DO 🗸 + Create issue

2.4- Future Sprints/Backlog



3 – Meetings

Meeting 1 (Saturday January 20th):

Agenda: Release Planning & Backlog Meeting

- 1 License.md: GPL or MIT.
- 2 We decided to use Miro for recording our progress and scheduling our meetings. Rifat will make the board and invite us later.
- 3 Sprint Planning: Biweekly for now, weekly later if necessary.
- 4- We made the decision to use Jira to record our sprints and tasks.
- 5- Product Owner and Project Manager will handle all the reporting. We will reference our Jira often to stay updated on sprint progress.
- 6 Talked about the rough idea of the project. Hevar mentioned an app that can pull from Reddit very easily and we can use that in the program.

Questions to ask the TA (Will be updated on Monday as well):

- 1 How often do we need to report to you? Will it be after every sprint/meeting, or will it simply be at the assignment's pace? Do we need to record our meetings in detail?
- 2 How will the program itself be graded? Will the complexity of it be accounted for? Will the amount of testing be graded? Etcetera.
- 3 How detailed do the sprint reports need to be? How general can we be when it comes to the tasks we record on Jira?
- 4 How should we handle pairing? Will it be better for 2 dev's to work on the same thing for the duration of the sprint, or is it better to have each dev work on one aspect instead?
- 5 What exactly is expected of the Product Owner and Project Manager? We already made the decision to handle all the reporting, but what else can we do to even the workload?

Meeting 2: (Monday January 22):

Agenda: Sprint Review & Retrospective

- 1. We confirmed the license GPL v3.0 as our license format.
- 2. For the Ta's meeting we decided to create 4 user stories:

User Statistic analyse, User entrepreneurs, Users develop, User youtuber.

3. For the Ta's meeting -15 mins project basic discussion & 5 mins Ta answering question.

Questions to ask the TA (added with the Saturday's meeting questions):

6. Is there any minimum push requirements for everyone for the project?

7. How'd you suggest approaching the 5% bonus marks for the assessment part (slide 17 in the course outline)?

Meeting 3: (Monday January 29th):

Agenda: Sprint Planning

- 1 Maisam created multiple branches to maintain version control.
- 2 Maisam did some research on Reddit's API and found out that simply adding. Json at the end of the URL creates a file that contains all the relevant post's information. This will be our method for pulling reddit posts.
- 3 Matthew researched subreddits and it prompted us to discuss comment-dependent versus post-dependent content. This will have to be decided a little later, but the issue has been discovered. We decided to leave this decision to later down the line.
- 4 Matthew brought up censorship and how we should tackle it. Firstly, we need to define what censorship entails. So far, swear words, sensitive words, strong words (kill, murder, etc...) can be replaced with words like "unalive" etcetera.
- 5 We discussed security in the program and what it entails, we decided it mainly means we need to try to ensure that metrics and the like are not easily leaked or stolen.
- 6 We discussed code pacing. And decided our current pace is fine for now. But there will likely be a crunch period which we will try to avoid.
- 7 Basel brought up the point of measuring data obtained both from the program and through user metrics. It was decided that concessions will be made by the dev team to be able to measure this data.
- 8 Basel brought up the issue of pulling metrics from our posts to serve in the decision tree. We agreed that it wouldn't pose a big challenge if we use websites like social blade. Detail will have to wait for its relevant sprint, but we have a good idea on how to go about doing it.
- 9 Raymond brought up a question about what qualifies as a "manageable" post length. Hevar suggested measuring the text-to-speech's time through google and the decision should be made based on that. Raymond decided he would do this through measuring WPM (words-per-minute). But further detail will have to wait for the creation of the voice-over module. This will warrant extensive testing.
- 10 We decided to research and see if Reddit has a system of collecting screenshots. This brought up a permutation possibility which is either having subtitles on the screen or a screenshot of the post itself. We can upload a different version of this to two channels and seeing which one generates more attention.

Meeting 4: (Monday February 5th):

This meeting is taking place just before the sprint ends. Not much was discussed but we ensured everything was going to plan.

Agenda: Sprint Review

- 1 6 out of the 8 sprints were marked as done
- 2 -Discussion was done regarding one of the sprint backlogs sanitizing user posts for Censorship and security.
- 3- We discussed when the TA meeting was taking place.
- 4- We planned the Sprint Retrospective Meeting and set plans for the next sprint's Sprint Planning Meeting.

Meeting 5: (Sunday February 11th):

This meeting will serve as a Sprint Retrospective Meeting, we will tie up any loose ends and reflect on what we could have done better.

Agenda: Sprint Retrospective

- 1- It's come to our attention that integrating the swear word database for censorship along with our reddit post grabber is proving more difficult than anticipated. This task will now be moved to the next sprint.
- 2- We cited a database we used for swearing words and decided on a reliable way to cite our sources from now on.
- 3- What everyone learned from the first sprint proved to be too little compared to the detail that's expected of us. So, Basel and Rifat will take measures to make some stuff up. As well as ask questions to the dev team to get some more context.
- 4- We decided to compile our work and what we learned. The Reddit API proved to be difficult to work with, so we had to come up with a roundabout way of interfacing with Reddit. Also, unfortunately it is not possible to run the program through GitHub directly which made our work more difficult. We decided that a more reliable and readable way of formatting our code to allow the other developers to grab what they need from another member's commits into their own branch will prove helpful. An extended looping issue in our pulling algorithm was also discovered during one of our tests which will need to be addressed very soon before building on top of the foundation we made this sprint, as execution is taking much longer than desired.

Meeting 6: (Tuesday February 13):

This meeting will serve as the sprint planning session of the second sprint.

Agenda: Sprint Planning

1. Old Sprint Review: - The team initiated the meeting by discussing the outcomes of the previous sprint. Any challenges or unfinished tasks were reviewed, and decisions were made on whether they should be included in the upcoming sprint.

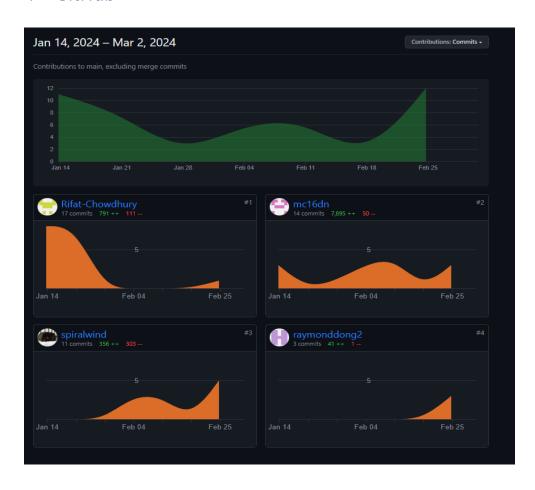
- 2. API Integration: The team engaged in a comprehensive discussion on the integration of multiple open-source solutions for the API. Various options were explored, and decisions were made to enhance the functionality and efficiency of the system.
- 3. Sprint 2 Kickoff: Confirmation was given that Sprint 2 is scheduled to commence tomorrow. The team expressed readiness and commitment to achieving the sprint goals.
- 4. Potential Additional Sprints: Considering the complexity of the data, there was a proposal for two more sprints, with the possibility of a third if required. The decision was not finalized, and further assessment will be made during the ongoing sprints.
- 5. User Backlogs for Sprint 3: The team discussed and added user backlogs for Sprint 3, demonstrating a proactive approach to ensure a well-prepared backlog for future sprints.
- 6. Time Constraints Discussion: Deliberations were held on time constraints, acknowledging the importance of managing time effectively to meet project milestones. The team committed to staying vigilant and proactive in addressing any potential challenges related to time.
- 7. Next TA meeting discussion Discussion regarding the next meeting with the TA.
- 8. Assigning tasks to the dev team The Scrum master (Hevar) assigned individual tasks to the dev team (Maisam, Matt & Raymond). More roles might be added depending on the need.
- 9. Generating Video Templates Produce 18 video templates for the coming sprint 3.
- 10. Possible Challenges of Sprint 3 Discussion was made on the censorship of the voice.

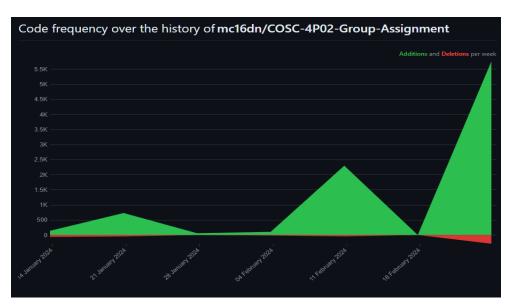
Meeting 7: (Monday February 26th):

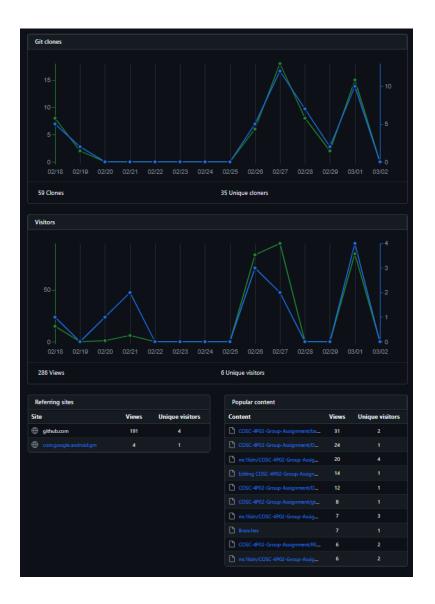
Agenda: Sprint Review

- 1 Hevar made the decision to extend this sprint's duration by one week. Our current topic of discussion is analyzing how to increase our efficiency with completing these tasks by breaking down the current issues using Jira.
- 2- We discussed tomorrow's TA meeting to get in the correct mindset. As a result of this discussion, we are going to take measures to compile everything we recorded so far to show the TA.
- 3- The developer team discussed how to handle text-to-speech (voices, censorship, and timing).

4 – GitHub







5-Issues

Currently, the biggest looming problem is the fine-tuning of TTS integration with the grabber method, timing, censorship, and voice must all be correct. Though, we predict that we will overcome this issue in a timely manner. We are not accustomed to working with such big groups, which resulted in some version control problems through GitHub, but this problem will be temporary, and we are taking measures to end it by improving how we communicate. The project's time constraint is intimidating at times, but looking at our steady workflow and task completion rate gives us confidence on getting the project done in time.