Course: Applied Data Analysis and Research Methodology

Module Code: MOD007893

Student ID: 2056338, 2054437, 2063656

Team name: MAY

Project: Sentiment analysis in the Metaverse

Table of Contents

Task Allocation Table 3

Analysis of the problem 4

- Business requirements 4
- Stakeholder requirements 5
- Executive summary 5

Project requirements 7

- Assumptions 7
- Definitions 8
- Functional requirements 8
- Non-functional requirements 12

Outline solution 13

- Flow Chart 13
- Wireframes 15
- Trace Table 17

High-level architecture 18

Justification of technologies 19

Evaluation 20

Reference list 22

Task Allocation Table:

Team Member	Task
2056338	Project requirements, Flow Chart, Wireframe
2054437	Flow Chart, Wireframe, Trace Table
2063656	Literature review, Required Technologies

Analysis of the problem

Business requirements

Two of the biggest issues that any contemporary social medium faces are cyberbullying and internet trolling — both types of cyber abuse. These issues are already highly pernicious on existing platforms for a number of reasons:

- Both cyberbullying and trolling can be subtle or grounded in complex social dynamics, making them hard to detect and/or distinguish from good-faith interactions — especially so for machine learning (Emmery et al., 2020).
- Both issues are very widespread and occur on a regular basis in any social digital environment. (March & Marrington, 2019) (Smith et al., 2008)
- Both issues, especially cyberbullying, can have potentially profound adverse mental health effects (Kwan et al., 2020).

With the upcoming inauguration of the Metaverse — a shared 3D digital environment designed to provide a richer social experience than traditional social media by the addition of sensory inputs such as 3D vision, voice, and touch — cyber abuse is set to be even more significant:

- New forms of cyberbullying and trolling will be enabled, such as bullying through simulated physical contact (Technology Review, 2021).
- The novel means of interaction will create a higher potential for abuse, and hence a higher potential for the severity of mental health effects (NY Times, 2021)
- A closer likeness to real social interaction means a higher likelihood of social stress, and hence higher vulnerability to cyber abuse (verywellminde, 2020)

For these reasons, even a single persistent cyber abuser can turn an entire Metaverse community — to be called *room* from now on in accordance with existing VR terminology — toxic and dysfunctional. Therefore, as a moderator of one of such rooms, one would want to identify, and take appropriate action regarding, all such abusers as soon as possible; not doing so would jeopardise the integrity of the room. There is hence high demand for systems that facilitate this process, as is evident in the success of commercial moderation bots such as MEE6 on existing social platforms such as Discord (SimilarWeb, 2022)

For reasons described above, the demand is expected to be even higher among Metaverse room moderators. Moreover, new measures with unprecedented efficiency are required; measures that may be sufficient for existing social media will not be sufficient for the Metaverse.

Stakeholder requirements

Metaverse identity is expected to become an influential part of the users' personal lives, with the potential to influence employment chances, social status, and other integral facets of life (Sanford, C., 2021). Therefore, while maintaining a low false negative rate in cyber abuse detection is crucial, minimising the false positive rate is necessary as well.

There is growing public frustration with regard to the lack of control that users have in moderation (Telecom, 2021), therefore, allowing users high control over moderation (such as what the appropriate consequences of cyber abuse should be in any given context) is likely to significantly increase user and stakeholder satisfaction.

Finally, the potential severity of singular instances of cyber abuse described in the *Business Requirements* section, combined with the fact that cyber abuse can take place at any moment, means that the final product has to be accessible on-demand at all times.

To summarise, stakeholders are looking for a product that will:

- Help prevent cyber abuse in the Metaverse
- Have a high detection rate
- Have a low false negative rate
- Have a low false positive rate
- Allow users high control over moderation

- Be accessible on-demand at all times

Executive summary

As detailed in the *Stakeholder requirements* section and explained in the *Business requirements* section, there is demand for a product that detects the two most widespread forms of cyber abuse ─ internet trolling and cyberbullying ─ in the Metaverse with unprecedented efficiency and a low error rate. In designing such a product, our team has located what we believe to be the most identifiable and unmistakable (characteristic) features of both forms of cyber abuse, and consequently built a system that efficiently detects these features. We have named this system and the broader application *MAY*, after our team name.

For internet trolling, which is commonly defined as seeking to provoke a negative emotional reaction from others, we have found the characteristic feature to be the disagreeableness ─ or confrontationality ─ of the response to the troll's behaviour. On the other hand, for cyberbullying, which is understood to entail intimidation through the use of chiefly aggressive, derogatory, and other forms of negative vocabulary, we have found the negativity of the sentiment expressed by the cyberbullies to be the characteristic feature.

Based on these features, we have designed two metrics, called Aggregate Response Agreeableness Score, or ARAS, and Aggregate Publication Positivity Score, or APPS, intended to detect internet trolling and cyberbullying, respectively. The ARAS indicates a neural network's evaluation of the general agreeableness of the response that a Metaverse user receives on various platforms; users with an exceptionally low ARAS are deemed likely to be internet trolls. The APPS, on the other hand, indicates a neural network's evaluation of the general positivity of the content generated by the user on these platforms (to be called *publications*); similarly, users with a low APPS are deemed likely to be cyberbullies.

Since most communication in the Metaverse is expected to be through voice, from which it is difficult to extract data, the APPS is instead calculated from the users' publications on traditional social media (such as Instagram and Twitter) from accounts linked on their public Meta profiles; the ARAS is calculated from the replies to these publications.

MAY will be provided as a web application and might also be available as a separate mobile application. It will allow Metaverse room moderators to view the ARAS and APPS of each room member as well as each user requesting access to the room. Furthermore, detailed analysis, including intuitive explanations of the scores and other useful statistics, of each of these users' behaviours will be available.

MAY is an innovative solution to the problem of cyber abuse that bypasses the most common setbacks of cyber-abuse detectors, such as the elusiveness of cyber abuse to neural networks, by combining machine and human intelligence: humans provide feedback on cyber abusers by replying to their publications, while machines analyse this feedback as well as the cyber abusers' own publications. We believe MAY to be an efficient response to the users' demands.

Project requirements

Assumptions

1. Existence of Meta/FB API to access publicly linked accounts

2. Meta account = FB account

3. Each social network of interest has a public (non-restricted) API to access user's public comments (rationale for existence: everyone can search user's public comments via web interface so it's logical to assume existence of public API to do the same).

4. Each social network of interest has a public API to access all replies to their comments

5. Each social network of interest has a public API to access all comments with the user public handle (assume for simplicity that this is mutually exclusive with replies to comments; in other words, all reaction to a user = {all replies to the user's comments} U {all comments with the user's handle}).

5.1 The user will be notified of all new join requests, as well as whenever the ARAS and APPS of these requests are computed.

6. Meta API allowing subscription to the following event types:

- Room join request (user asks to join a room)
- Room leaves request (user leaves a room)
- Room join request accepted (resulting in user joining the room)
- Room join request declined

Definitions

Publication	A collective term given to all types of user-generated content, including all posts, comments (including replies), stories, etc made by a user.
Aggregate Response Agreeableness Score (ARAS)	A metric that indicates the general agreeableness of replies to a user's account, designed to detect internet trolls. For detailed documentation, see section 1.1.
Aggregate Publication Positivity Score (APPS)	A metric that indicates the average positivity of publications made by a user, designed to detect cyberbullies. For detailed documentation, see 1.2.
Room	A Metaverse community analogous to Facebook Groups.
MU	Metaverse User

Functional requirements

1. Documentation: ARAS and APPS.

1.1. Documentation: ARAS

1.1.1. An MU’s ARAS will be computed as following:

1.1.1.1. A neural network will be trained to evaluate how agreeable the sentiment expressed in a piece of text is ─ or how much the author appears to agree with, support, or commend some other idea or person. The evaluation will be represented as a percentage, with 0% representing maximal confidence in a given piece of text being disagreeable and 100% representing maximal confidence in that piece of text being agreeable. This evaluation will be referred to as the agreeableness score.

1.1.1.2. All the publications to which at least one reply has been posted across all the social accounts publicly linked to the MU’s Meta account will be retrieved; for each such publication, all the corresponding replies will be retrieved.

1.1.1.3. If, across all the publications, more than one reply had been posted by the same user, all the replies posted by that user except the first one retrieved are discarded. This step is necessary to ensure that the sentiments of all users are represented equally; otherwise, the ARAS would be biased towards the views of especially vocal users. For example, if an MU frequently posts on their personal profile, and their friends habitually leave positive comments under these posts ─ a common occurrence ─ then the ARAS would be inflated; hence, even if a large number of users respond negatively to the MU elsewhere due to persistent trolling on their part, that might not be reflected in the ARAS.

1.1.1.4. The average agreeableness score of the remaining replies to each of the retrieved publications will be calculated to give each publication a Response Agreeableness Score. This step is necessary to prevent bias towards an unrepresentative set of highly controversial publications. For example, if an MU usually behaves civilly and gets reasonably agreeable replies but makes one controversial publication which accumulates a large number of negative replies, they might get a deflated ARAS score which isn’t representative of the MU’s typical behaviour.

1.1.1.5. The average Response Agreeableness Score across all publications will be calculated to yield the Aggregate Response Agreeableness Score, or ARAS.

1.2. Documentation: APPS.

1.2.1. A neural network will be trained to evaluate how positive or negative the sentiment expressed in a piece of text is. The evaluation will be represented as a percentage, with 0% representing maximal confidence in a given piece of text being negative and 100% representing maximal confidence in that piece of text being positive. This evaluation will be referred to as the positivity score.

1.2.2. All the publications made by the MU across all the social accounts publicly linked to the MU’s Meta account will be retrieved.

1.2.3. The average positivity score of all the retrieved publications will be calculated to yield the Aggregate Publication Positivity Score, or APPS.

2. Use case: user signs up or signs in to MAY with Meta as the identity provider.

2.1. If the user is already signed into Meta, they will be considered authenticated via Single Sign-On. Otherwise, they will be redirected to Meta’s log-in page and fill in their Meta account credentials. Upon authentication, the user will be redirected to MAY’s main page; without authentication, access to the main page will not be granted.

2.2. If the user has not been previously authenticated by the system, they will be asked to give the following permissions to MAY:

2.2.1. Receive notifications from the account on all events in rooms in which the account has moderator rights.

2.2.2. Access the public profiles of all the members of these rooms.

2.2.3. View open join requests to these rooms and access the corresponding public profiles.

3. The main page will feature an accordion containing all the rooms in which the user has moderator rights, as well as a search bar allowing the user to search across all such rooms. If no such rooms are found, the text message “No moderated rooms found” will be displayed.

4. Clicking on each room on the main page will reveal:

4.1. A column titled “Entry Requests”, containing a paged list of all open join requests to the room.

4.2. A column titled “Existing Members”, containing a paged list of all existing members of the room.

4.2.1. Both of these columns will be sortable by username, ARAS and APPS.

4.3. A search bar, allowing the user to search all the MUs across both lists.

4.4. A settings icon, which, when clicked, redirects to a page containing the room’s settings. The settings will include the option to auto-accept join requests from members with selected minimum ARAS and APPS.

5. Next to the username of each MU in the above two columns will be displayed:

5.1. The MU’s ARAS and APPS.

5.2. “Accept request” and “Reject request” buttons in the “Entry Requests” column.

5.3. “Remove user” and “Manage Room Rights” buttons in the “Existing Members” columns.

6. Clicking on each username will redirect to a detailed analysis of the corresponding MU’s behaviour, which will include:

6.1. Analysis of the ARAS.

6.1.1. The overall ARAS, as well as an intuitive explanation of that ARAS score means, such as whether or not the user is likely to be a cyber abuser.

6.1.2. A breakdown of the ARAS of the replies to the MU in the form of a pie chart, with slices representing ranges of ARAS (such as 10─20%) and their sizes representing the proportion of replies that have an ARAS that falls within this range.

6.1.3. Recent replies to the MU and the corresponding ARAS of each reply, sortable by criteria such as general topic (such as “politics”), social network, ARAS, date, and others ─ to be called criteria A. The “general topic” filter could be used to get a more accurate prediction of the MU’s behaviour in a context specific to particular room: for example, if an MU has a low politics ARAS but a high chess ARAS, they can be considered unlikely to be problematic in a chess-themed room.

6.1.4. A chart showing the user’s ARAS over time.

6.1.5. The user will have the option to apply a number of filters corresponding to criteria A on the replies to the MU, and re-run the analysis on this filtered set of replies.

6.2. Analysis of the APPS.

6.2.1. The overall APPS, as well as an intuitive explanation of that ARAS score means, such as whether or not the user is likely to be a cyber abuser.

6.2.2. A breakdown of the APPS of the MU’s publications in the form of a pie chart, with slices representing ranges of APPS (such as 10─20%) and their sizes representing the proportion of publications that have an APPS that falls within this range.

6.2.3. Recent publications made by the MU and the corresponding APPS of each reply, sortable by criteria such as general topic, social network, APPS, date, and others ─ to be called criteria B.

6.2.4. A chart showing the MU’s APPS over time, or by criteria B.

6.2.5. The user will have the option to apply a number of filters corresponding to criteria B on the MU’s publications, and re-run the analysis on this filtered set of replies.

6.3. A list of all social accounts publicly linked to the MU’s Meta account.

Graphical User Interface

Non-functional requirements

1. A maximum of 1,000,000 registered MAY users.

2. A maximum of 1,000,000,000 profiles whose publications are collected for ARAS and APPS computation, assuming a potential coverage of every Metaverse user.

3. A maximum of 100,000,000,000 publications collected for ARAS and APPS computation, assuming an average of 100 publications per Metaverse user.

4. System availability at least 99.99999% of the time.

5. Maximum response time:

5.1. For ARAS and APPS computation of new join requests: 15 seconds

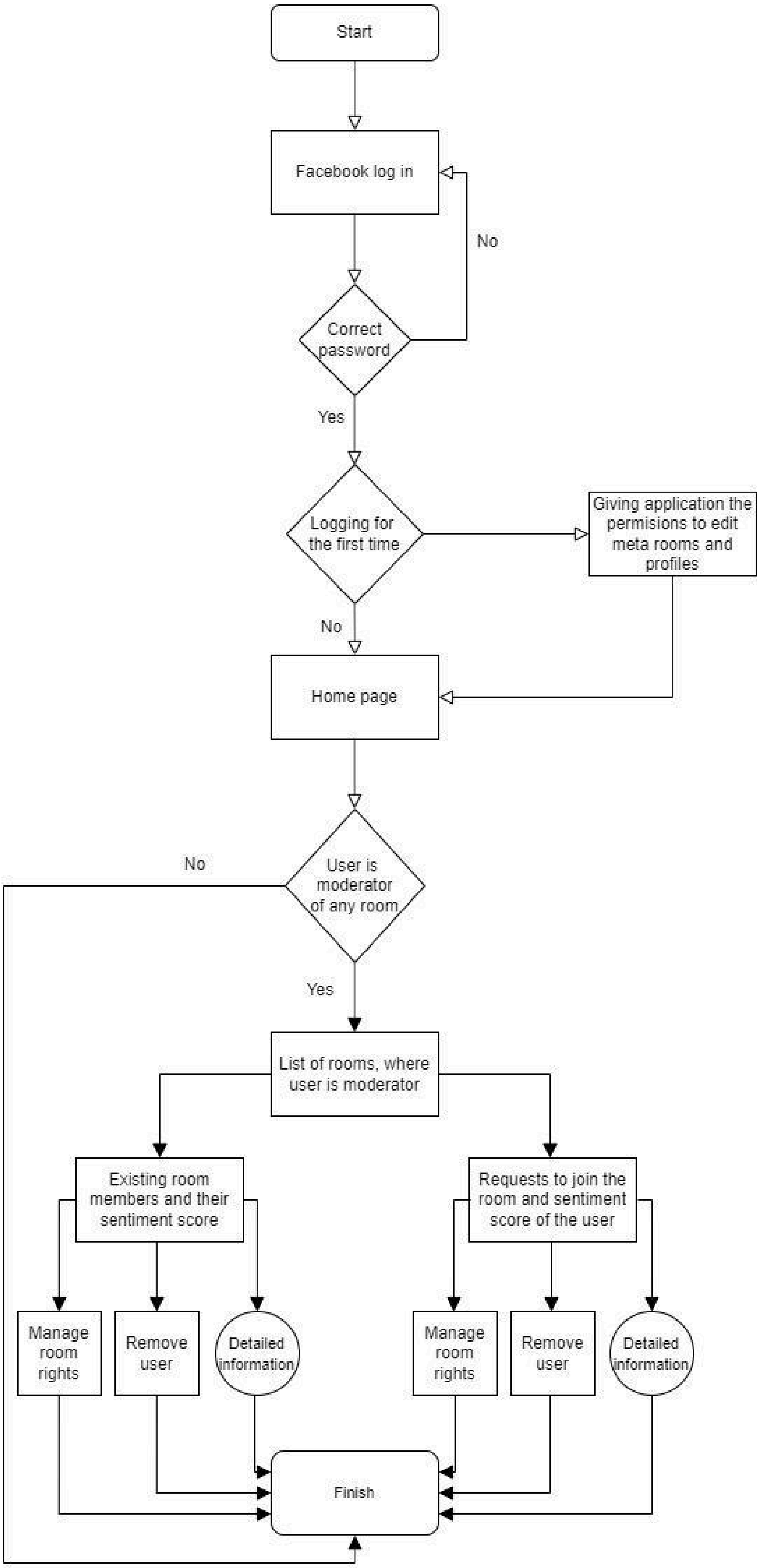
5.2. System-wide: 0.5 seconds

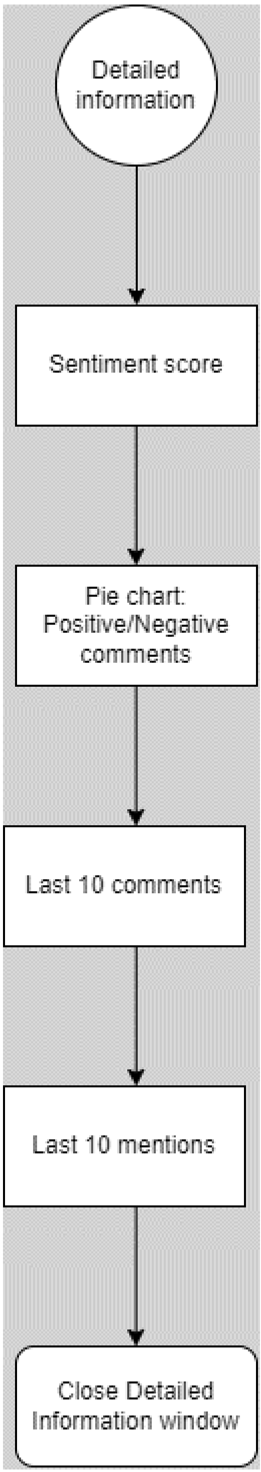
6. Security: private information provided by the user must be protected.

7. Usability: MAY should be easy to use and navigate.

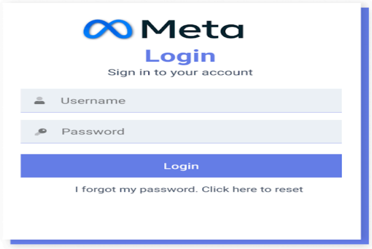
Outline solution

Flow Chart

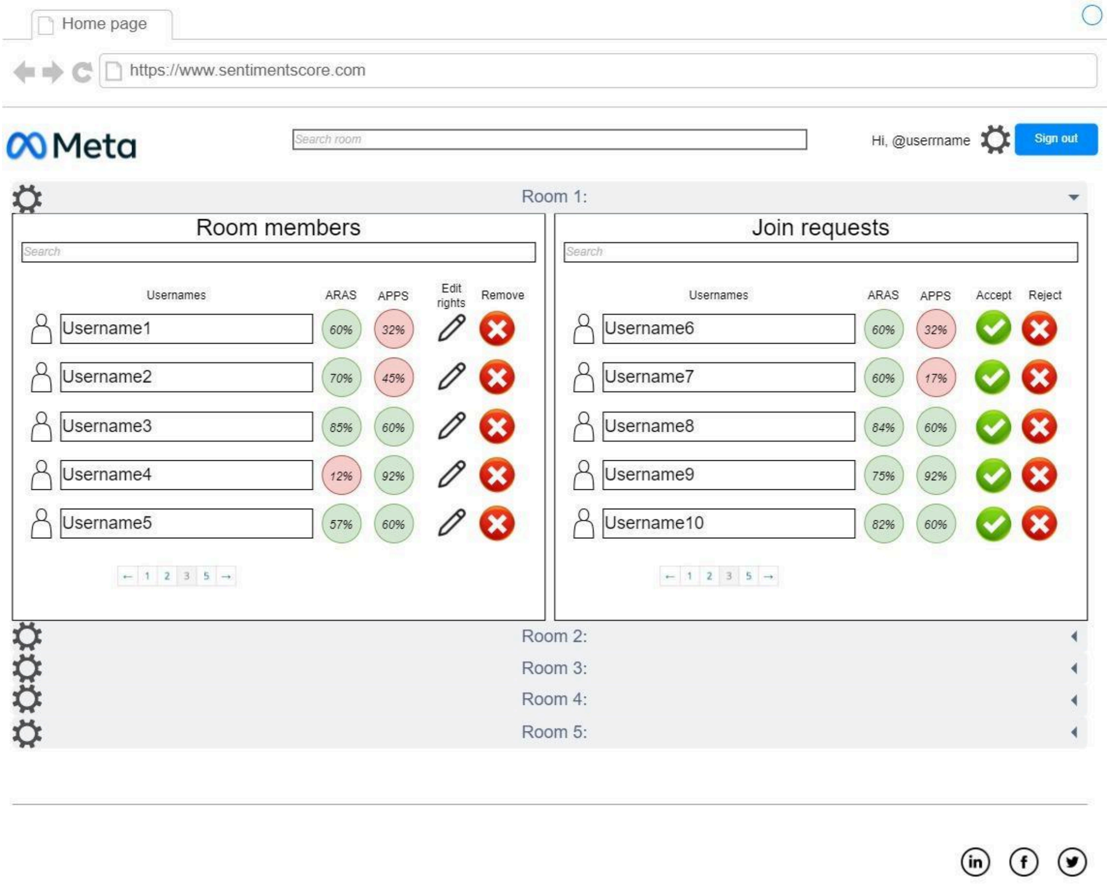

Start
Facebook log in
No
Correct password
Yes
Logging for the first time
Giving application the permisions to edit meta rooms and profiles
No
Home page
No
User is moderator of any room
Yes
List of rooms, where user is moderator
Existing room members and their sentiment score
Requests to join the room and sentiment score of the user
Manage room rights
Remove user
Detailed information
Manage room rights
Remove user
Detailed information
Finish


Detailed information
Sentiment score
Pie chart: Positive/Negative comments
Last 10 comments
Last 10 mentions
Close Detailed Information window

Wireframes



This is the login page of the application with Meta logo, having a username and password for the user to login and options for resetting the password just in case the user forgets his password.



Room members

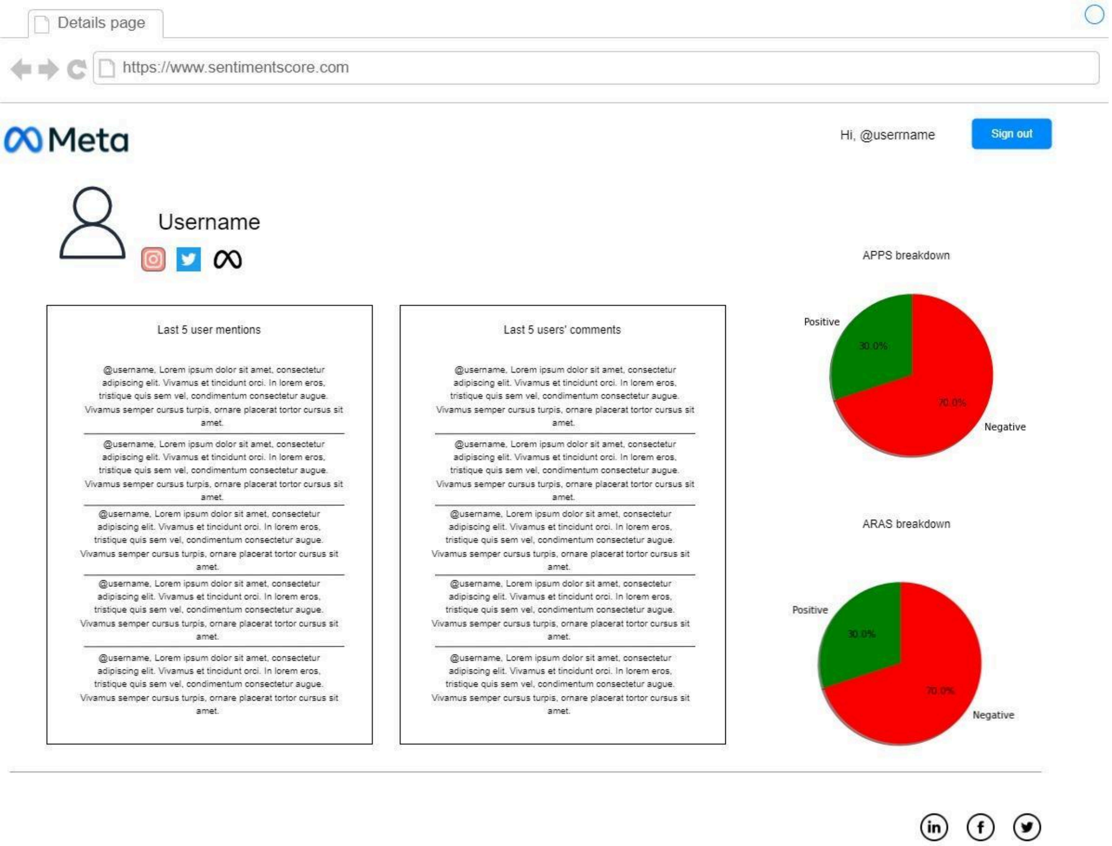
Usernames	ARAS	APPS	Edit rights	Remove
Username1	60%	32%	✎	✗
Username2	70%	45%	✎	✗
Username3	85%	60%	✎	✗
Username4	12%	92%	✎	✗
Username5	57%	60%	✎	✗

Join requests

Usernames	ARAS	APPS	Accept	Reject
Username6	60%	32%	✓	✗
Username7	60%	17%	✓	✗
Username8	84%	60%	✓	✗
Username9	75%	92%	✓	✗
Username10	82%	60%	✓	✗

On this homepage, after the user has been able to login, all rooms where the user is a moderator is displayed, showing all the members of the room with their sentiment scores so far, with the option of the moderator editing rights or removing users based on their sentiment scores in the group.

It also displays new members who intend to join the room. The moderator has access to their sentiment scores and decides to accept or deny them access to the group based on their sentiment scores.



Last 5 user mentions

@username, Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus et tincidunt orci. In lorem eros, tristique quis sem vel, condimentum consectetur augue. Vivamus semper cursus turpis, ornare placerat tortor cursus sit amet.

@username, Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus et tincidunt orci. In lorem eros, tristique quis sem vel, condimentum consectetur augue. Vivamus semper cursus turpis, ornare placerat tortor cursus sit amet.

@username, Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus et tincidunt orci. In lorem eros, tristique quis sem vel, condimentum consectetur augue. Vivamus semper cursus turpis, ornare placerat tortor cursus sit amet.

@username, Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus et tincidunt orci. In lorem eros, tristique quis sem vel, condimentum consectetur augue. Vivamus semper cursus turpis, ornare placerat tortor cursus sit amet.

@username, Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus et tincidunt orci. In lorem eros, tristique quis sem vel, condimentum consectetur augue. Vivamus semper cursus turpis, ornare placerat tortor cursus sit amet.

Last 5 users' comments

@username, Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus et tincidunt orci. In lorem eros, tristique quis sem vel, condimentum consectetur augue. Vivamus semper cursus turpis, ornare placerat tortor cursus sit amet.

@username, Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus et tincidunt orci. In lorem eros, tristique quis sem vel, condimentum consectetur augue. Vivamus semper cursus turpis, ornare placerat tortor cursus sit amet.

@username, Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus et tincidunt orci. In lorem eros, tristique quis sem vel, condimentum consectetur augue. Vivamus semper cursus turpis, ornare placerat tortor cursus sit amet.

@username, Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus et tincidunt orci. In lorem eros, tristique quis sem vel, condimentum consectetur augue. Vivamus semper cursus turpis, ornare placerat tortor cursus sit amet.

@username, Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus et tincidunt orci. In lorem eros, tristique quis sem vel, condimentum consectetur augue. Vivamus semper cursus turpis, ornare placerat tortor cursus sit amet.

This homepage displays when the moderator clicks on each user, be it existing members or new users. It redirects the moderator to the last 5 mentions and the 5 last comments of the users with a detailed pie chart breakdown and also displays all other social media platforms linked to the users' account.

Trace Table

Log in

Username	Password	Output
Y	N	0
N	Y	0
N	N	0
Y	Y	1

MU belongs to the room	MU is moderator	Output
Y	N	0
N	Y	0
N	N	0
Y	Y	1

Showing the list of rooms

Showing the list of existing users in the group

MU belongs to the room	MU is moderator	Number of users is room>0	Output
Y	N	N	0
Y	Y	N	0
Y	N	Y	0
N	Y	Y	0
N	N	Y	0
N	Y	N	0
Y	Y	Y	1

Showing the MU requests in the group

MU belongs to the room	MU is moderator	Number of users is room>0	Output
Y	N	N	0
Y	Y	N	0
Y	N	Y	0
N	Y	Y	0
N	N	Y	0
N	Y	N	0

Y	Y	Y	1

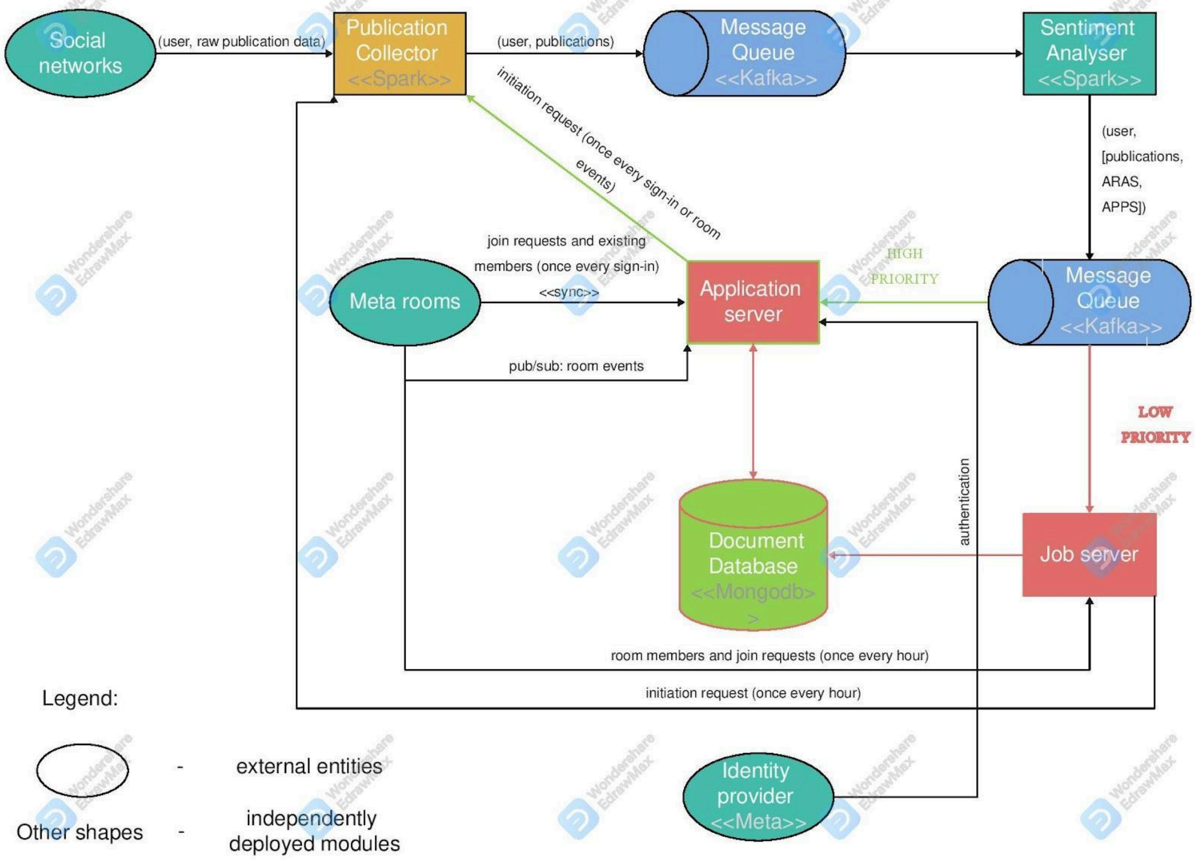
Computing ARAS

MU profile is public	Number of MUs' replies>0	Outcome
N	N	0
Y	N	0
N	Y	0
Y	Y	1

Computing APPS

MU profile is open	Number of MUs' comments>0	Outcome
N	N	0
Y	N	0
N	Y	0
Y	Y	1

High level architecture


Social networks
(user, raw publication data)
Publication Collector <<Spark>>
(user, publications)
Message Queue <<Kafka>>
Sentiment Analyser <<Spark>>
(user, [publications, ARAS, APPS])
initiation request (once every sign-in or room events)
join requests and existing members (once every sign-in)
<<sync>>
Meta rooms
Application server
HIGH PRIORITY
Message Queue <<Kafka>>
pub/sub: room events
LOW PRIORITY
authentication
Document Database <<Mongodb>>
Job server
room members and join requests (once every hour)
initiation request (once every hour)
Legend:
- external entities
Other shapes - independently deployed modules
Identity provider <<Meta>>

1. Machine resource allocation. Every module in the diagram will be deployed to multiple machines due to the large number of data that cannot be processed by a single machine.

2. High- and low-priority paths. The singular schematic path in the diagram from social networks to the Sentiment Analyser represents two actual paths: a low-priority path to the job server and a high-priority path to the application server.

3. Low-priority path explained. The job server batch jobs run every hour and update the ARAS and APPS of all members and join requests for all rooms, and store them in the Document Database. Since these jobs are computationally heavy, batch processing is required to reduce the response time to below the maximum value established in the Non-functional requirements section. Ton increase efficiency, if an MU is already stored in the Document Database, data about this MU will be updated incrementally based on the MU’s publications in the last hour.

4. High-priority path explained.

4.1. Pub/Sub. Whenever a user signs in, the application server subscribes to receive all Meta room events from the user's rooms. These events include the room leave event, the new join request event, etc. This is necessary for the information displayed to the user to be up to date.

4.2. Scenario A. User signs into MAY. The sign-in triggers a synchronous transfer of all the existing members' and join requests' IDs across all rooms from Meta to the application server. It then requests the ARAS and APPS for each of these IDs from the Document Database and displays these scores on-screen; if no such scores are available (if the join requests are new), then a request is sent to the Publication Collector to compute these scores. These scores will be computed and displayed asynchronously, and the user will be notified when the computation is complete. The process is asynchronous so as to not keep the user waiting.

4.3. Scenario B. User is already signed into MAY and receives a new join request. This triggers a request to the Publication Collector to compute these ARAS and APPS for these requests. Like in scenario A, these scores will be computed and displayed asynchronously, and the user will be notified when the computation is complete.

Justification of technologies

1. *Spark (Publication Collector).* Spark is a highly efficient (Agrawal, Butt, Doshi and Larriba-Pey, 2022) MapReduce framework capable of running multiple MapReduce jobs on multiple machines simultaneously (Overview - Spark 3.2.1 Documentation, 2022). This is necessary due to the heavy computational requirements of our system. MapReduce is suitable for the tasks performed by the module because:
    1. The map job is (user ID, social network) →(user, all publications by user with this ID on this social network).
    2. The reduce job yields the publications per user from all social networks from the map job.
2. *Kafka Message Queue.* Kafka is a potentially high-availability message queue capable of running on multiple machines, which is suitable for our system's heavy computational requirements. In our system, it allows the Publication Collector, Sentiment Analyser, job server and application server to function independently of each other and not have to wait for each other (High availability with Apache Kafka - Azure HDInsight, 2022) (Solace PubSub+ vs Kafka: High Availability - Solace, 2022).

3. *Spark (Sentiment Analyser).* Similarly to the Publication Collector module, this module is computationally tasking, which Spark’s MapReduce multi-machine capabilities are suitable for. MapReduce is suitable for the tasks performed by the module because:
    1. The map job is publication →ARAS, APPS. This map job will be achieved via a **pre-templated sentiment analysis engine**.
    2. The reduce job yields a list of (publication, ARAS, APPS) tuples for each user.
4. *Mongodb.* MongoDB is a potentially high-performance and high-availability database capable of storing a key-to-JSON value mapping (JSON And BSON, 2022) Advanced MongoDB Performance Tuning, 2022). It is suitable for the Document Database module because the mapping is user →all data related to user (including ARAS and APPS for each publication, publication data, Meta data, etc).

Evaluation

Nowadays cyber-abusing, including cyber-bullying and trolling, is the most significant problem in the digital world. Our team came up with a solution to resolve the cyber-abusing problem in the Metaverse. We have created an idea of an application for the moderators of the rooms in the Meta community, which aim is to connect people with the same interests.

Our application allows moderators to check the prospective and existing digital personas of the metacommunity on the potential cyber-abuse based on the sentiment analysis. Our sentiment decisions are based on the comments and the mentions of people across the web.

The system has a user-friendly interface and meets all the security requirements. Also used technologies are making the system highly responsive to our users.

Our system is built for improving the modern community and will help to avoid any cyber-bullying in the metaverse.

Reference list

A global look at YouTube and its censorship policies - Telecoms.com. (2022). Retrieved 28 February 2022, from https://telecoms.com/opinion/a-global-look-at-youtube-and-its-censorship-policies/

Agrawal, D., Butt, A., Doshi, K. and Larriba-Pey, J., 2022. [online] Available at: <https://www.researchgate.net/publication/308901386_SparkBench_-_A_Spark_Performance_Testing_Suite> [Accessed 28 February 2022].

Docs.microsoft.com. 2022. High availability with Apache Kafka - Azure HDInsight. [online] Available at: <https://docs.microsoft.com/en-us/azure/hdinsight/kafka/apache-kafka-high-availability> [Accessed 28 February 2022].

Emmery, C., Verhoeven, B., De Pauw, G., Jacobs, G., Van Hee, C., Lefever, E., Desmet, B., Hoste, V. and Daelemans, W., 2020. Current limitations in cyberbullying detection: On evaluation criteria, reproducibility, and data scarcity. *Language Resources and Evaluation*, 55(3), pp.597-633.

How You Can Help Your Bullied Child With Social Anxiety. (2020). Retrieved 28 February 2022, from https://www.verywellmind.com/bullying-effects-social-anxiety-child-3024250

March, E., & Marrington, J. (2019). A Qualitative Analysis of Internet Trolling. Cyberpsychology, Behavior, And Social Networking, 22(3), 192-197. doi: 10.1089/cyber.2018.0210

Medium. 2022. Advanced MongoDB Performance Tuning. [online] Available at: <https://medium.com/idealo-tech-blog/advanced-mongodb-performance-tuning-2ddcd01a27d2> [Accessed 28 February 2022].

MongoDB. 2022. JSON And BSON. [online] Available at: <https://www.mongodb.com/json-and-bson> [Accessed 28 February 2022].

Sanford, C. (2021). Meta (Facebook) Connect 2021 Metaverse Event Transcript. Retrieved 28 February 2022, from https://www.rev.com/blog/transcripts/meta-facebook-connect-2021-metaverse-event-transcript

SimilarWeb (2022). Retrieved 28 February 2022, from https://www.similarweb.com/website/mee6.xyz/#overview

Smith, P., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: its nature and impact in secondary school pupils. *Journal Of Child Psychology And Psychiatry*, *49*(4), 376-385. doi: 10.1111/j.1469-7610.2007.01846.x

Spark.apache.org. 2022. Overview - Spark 3.2.1 Documentation. [online] Available at: <https://spark.apache.org/docs/latest/> [Accessed 28 February 2022].

Spark.apache.org. 2022. Overview - Spark 3.2.1 Documentation. [online] Available at: <https://spark.apache.org/docs/latest/> [Accessed 28 February 2022].

The metaverse has a groping problem already. (2022). Retrieved 28 February 2022, from https://www.technologyreview.com/2021/12/16/1042516/the-metaverse-has-a-groping-problem/

The Metaverse’s Dark Side: Here Come Harassment and Assaults. (2021). Retrieved 28 February 2022, from https://www.nytimes.com/2021/12/30/technology/metaverse-harassment-assaults.html