JESS - Fostering Empathy through Virtual Communities in Augmented Reality

Jack Rivera
esr@mit.edu

Jessica Yin
yinj@mit.edu

Harry Yu
harryyhw@mit.edu

Abstract

With the greater, growing social distance in today's society, it is relatively easy to feel a sense of isolation and unfamiliarity with the people we pass by each day. We find that many social applications are designed to promote communication between known friends, yet neglects bonding of general communities. We present Just Everyday Shared Stories (JESS), a mobile social application in which individuals can join communities to share augmented reality posts for others in the same location as them later that day to discover. In this manner the app aims to facilitate the spontaneous casual sharing of ideas and experiences between communities of physically proximal like-minded individuals, and through so encourage a more empathetic and hopeful view of the real-life communities that users reside in.

Author Keywords

Augmented reality; virtual communities; google cloud anchors; virtual presence; social application; mobile application.

Introduction

A strong sense of community helps us build connections and empathize with others around us, and improves mental well-being by helping us feel more safe when out in public. However, in recent years, the degree of trust between Americans has been declining, to a level where most Amer-

icans believe it is imperative that increasing social confidence be addressed [1]; in recent times face masks and social distancing measures have further impeded interpersonal trust by obscuring facial expressions and decreasing interpersonal interactions. We believe that camaraderie within local communities is sorely lacking, and this forms our motivation for creating Just Everyday Shared Stories (JESS).

In the mobile application JESS, users generate posts composed of AR digital content, then tag a virtual community that the creator wishes to share the post with. These posts are of limited duration (generally one day) and are tied to the physical location at which the post was created. When a member of said tagged community is physically near the post, they are notified via an in-app map that a member in the shared community made a post at a nearby location and can travel there to view the post. Users are invited to express their everyday stories, experiences and emotions through 3D model templates and attached text/audio, and discover new communities to share their posts with.

Design

Our goal is for people to feel a sense of companionship and mutual understanding with other people in their local communities as users connect with other users in their local communities that they may not know in the physical world by sharing their day to day experiences. This is done by leveraging AR content to achieve spatially local and temporal experiences that magnify the intimacy and extend the reach of users' virtual footprint, while maintaining anonymity and centering the content around communities, thus emphasizing the profile of a local community rather than the individual.

To this end, it was critical that posts be only instantiated

and viewable at the physical location of the placed AR object; when viewing a post, the viewer is thus assured that the post was created by another wanderer passing by that same location, and thus could equally well be the story of any one of many nameless passersby. The transient one-day nature of posts encourages expectation-free creation on the fly (and incidentally solves the issue of old posts taking up too much virtual real estate due to the space-occupying nature of AR content); and from the perspective of the viewer lends more immediacy to the content, while being long-lived enough to ensure that a daily commuter along their commute should see all content other commuters have posted. Put another way, JESS slightly relaxes the constraint on communication from being at the same place at the same time, to being at the same place within a day – allowing users to leave a virtual echo behind them, and speak through that echo. We believe that hearing these echoes induces a much stronger sense of the presence of the creator than hearing the experiences of creators possibly miles away, days after those experiences have left their mind. Therefore, the enforced temporality and spatial locality is key to our design.

We reinforce this vision using AR by encouraging users to leave a message or digital footprint in the form of a 3D model with other attached media such as text or audio (e.g. a raincloud to signify a somber mood with a short message expression some associated sentiment), which marks the occasion of their previous occupancy of the physical space. Mapped onto real physical space, the digital content feels more real and intimate as is the affordance of AR.

Since the idea is for users to share their everyday moments instead of more curated or premeditated content, it was also important that the application be easy to use. More specifically, the app must have a low barrier of entry to cre-

JESS relaxes the constraint on communication from being at the same place at the same time, to being at the same place within a day – allowing users to leave a virtual echo behind them, and speak through that echo.
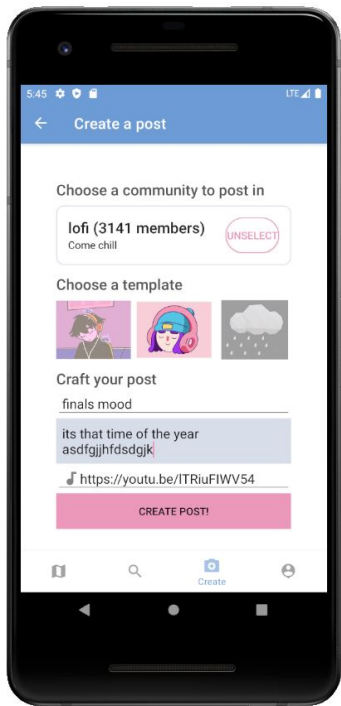
ate and join communities, as well as tools that allow quick creation of virtual content. We accomplish this by using a mobile app with built-in templates for post creation, so that users with little time or experience with creating digital content may simply choose a predesigned 3d model (e.g. avatar or 3d object like a cloud) for the AR content of their post while attaching various forms of media such as text by filling out text-boxes or music by providing links. Additionally, the AR sample app we adapted from Google [1] provided many text cues guiding the user to select a location and place their AR content which helped facilitate ease of use.

Finally, as mentioned earlier, we emphasize the community over the individual in our design of the application. Users subscribe to communities as opposed to individual users, and we avoid offering direct user-to-user messaging or 'friend-ing' functionality. Moreover, as described above, this community-centric focus is crucial to ensuring the transient and local nature of posts in JESS induce the feeling that the content comes from just another person the viewer might have passed by on the street. Through this we hope to cultivate lively local communities around the casual sharing of content and alleviate interpersonal distrust.
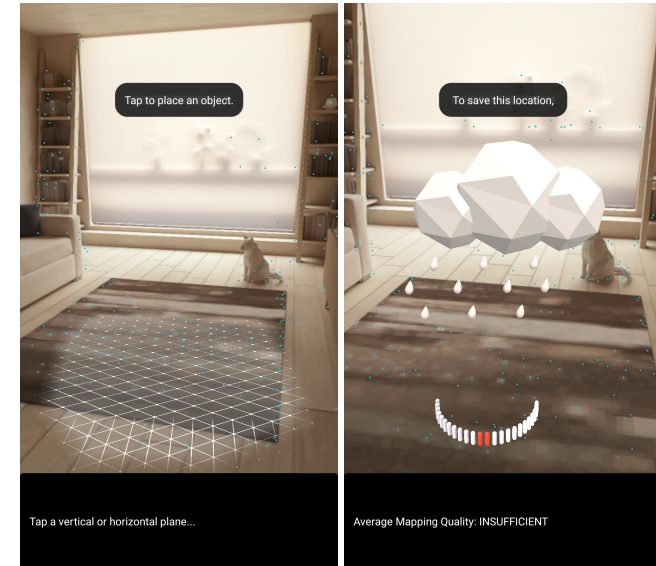
Implementation

Our implementation of JESS is composed of three main parts: creating posts, viewing posts, and supporting functionality.

Creating posts

Posts in JESS are multimedia content, at minimum composed of one piece of three-dimensional content serving as the base of the AR content, a title and a text description. Posts are tied to a geographical location and have a fixed

[1] https://developers.google.com/ar/develop/java/cloud-anchors/developer-guide-android

Figure 1: A sample template on the post creation screen. The user has selected the lofi community, a 3D model, and filled in a title, description and link to some music. A button saying "Create post!" is lit up under the template entries.

Figure 2: A user is scanning for a location to place their post. The camera display lights up with a grid showing where object placement is possible, and once placed, a wheel showing the progress of scanning the surface from different angles. Text popups instruct the user as to how to scan, place and orient the object.

orientation, and persist for by default one day (the technology supports up to a year; more on duration in censorship section of future features) upon creation. This storage and anchoring onto physical locations is enabled by Persistent Cloud Anchors from Google's ARCore Software Development Kit which offer high accuracy AR object placement and storage. Using Cloud Anchors, 3D objects can be precisely placed in a specific orientation, supporting (e.g. a human sitting and leaning against a tree), which enhance

the realism of the AR experience.

To create a post, a user selects "Create a post" in the "Create" tab. The user then selects a community to share the post with, and fills in a predefined template associated with the community to provide the content of the post; our demo prototype currently supports a template for the "lofi" community, a hypothetical community of aficionados of low-fidelity (lofi) music, which is typically listened to while relaxing or focusing on a task. This template includes a choice of a 3D model for the AR component, simple text boxes for the title and description, and a box for users to paste in a link to some lo-fi music relevant to the post. The template takes a link to a song rather than providing a selection of songs to choose from as with the 3d models because we envisioned this community around users who are at least somewhat familiar with lofi. Attaching images and other media is possible with the technology but is currently omitted from this lofi template for demo purposes.

Once a template is filled in, the user is prompted to scan for a nearby surface with their phone camera to place the post on. On-screen displayed instructions guide the user through the process of scanning for the surface and the placement of the object. Following this, the user must walk around the object or capture different angles of the object at a fixed distance to create a robust feature map of the object in the environment. The user is progressively notified of their progress with small markers that go from blank to red to yellow to green with increasing completion (note: an object does not need all markers to be green to have a sufficient feature map and the device notifies the user automatically when a sufficient map is achieved). Once the feature map is successfully created, this information along with the digital content and pose location is uploaded to Google Cloud using oauth 2.0 authentication. This completes the

post creation process.

Viewing posts
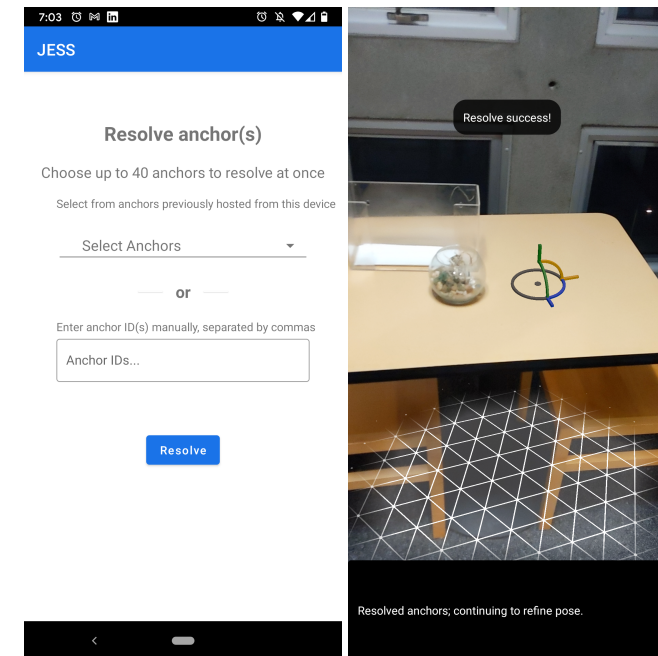
Figure 3: Left: Here is the current UI for choosing a post but it is currently using google cloud jargon. Right: When a user chooses an anchor and clicks resolve, the camera will try and match the current environment to what it expects to find according to the stored information in the anchor; This screenshot shows a successful resolution

To begin viewing a post you must first choose a post from the posts available to you. The posts available to you include those posts belonging to communities that you have joined and that are spatially local to you (within a defined/fixed radius of your current geographic location), thus achieving

the temporality and spatial locality described earlier (note: Currently, there is no filtering on the MVP; more on this in filtering in future features section). To choose a post, a user can select an available post by name from a drop down menu or search for it by name. Admittedly, this not an ideal design because you would have to remember the name of the post you are trying to find rather than the app automatically keeping a record of what post you are currently trying to find (more on this in future features section).

Once a post/anchor is chosen, clicking the resolve button (displayed in figure 3 UI) will open the camera. The camera will then compare the current environment to the environment stored previously at the chosen post/anchor. More specifically, the camera will compare the current feature map being calculated by the current camera input to the feature map stored on the chosen post/anchor previously calculated when the post was made. If there is a match the post with all of its digital content will appear (right image in Figure 3 shows successful match).

Currently, in this MVP, the anchors created are unable to be stored in Google Cloud and thus are unable to be resolved/found because keyless authentication (oauth 2.0) used by the Google Cloud Anchor API is incompatible with keyed authentication used by the Google Maps API (more on this in future features section). The reason the successful resolution (right Figure 3) was able to be captured was because the Google Cloud Anchor API functionality was developed/tested separately to the following supporting functionality.

Supporting functionality
Map.   The Map tab on the navigation bar leads to an in-app display of Google Maps. The location of each post from communities the user has joined and that are within the defined radius for locality will be visible on the map. The

map is interactive and helps the user locate where nearby posts are. Each available post can be clicked on to view the name, community, author, time since creation, and a short preview of the description to entice users to visit the post.

Communities.   The Communities tab on the navigation bar contains a searchable list of available communities in the app, with a short description of each community (the current MVP does not allow for creation of communities). The user can search for communities aligned to their interests, and join or leave communities from this tab.

Profile.   The Profile tab on the navigation bar allows the user to view (and leave) their currently joined communities, as well as view a list of posts they have created. Sometimes if the post has a specific media, it may be previewed. For example, in the MVP a sample post is created with a music link and clicking the music icon will go to YouTube to play the song.

Usage
We intend for JESS to be an effective means for users to express their immediate emotions and experiences on the go, and to hear more stories from other users who have recently shared the same space. We have described how the temporal and spatial locality, and the AR content of JESS's posts promotes a more intimate connection with the community and the creator; and discussed how the ease of use and community focus of the app promotes casual, expectation-free exchanges between community members and helps them feel less isolated and more connected with their local communities as a whole. In our example of the lofi community, users view a 3d model and listen to linked music in the specific location where the creator was inspired to create said experience, as though listening through a shared pair of earphones with the creator
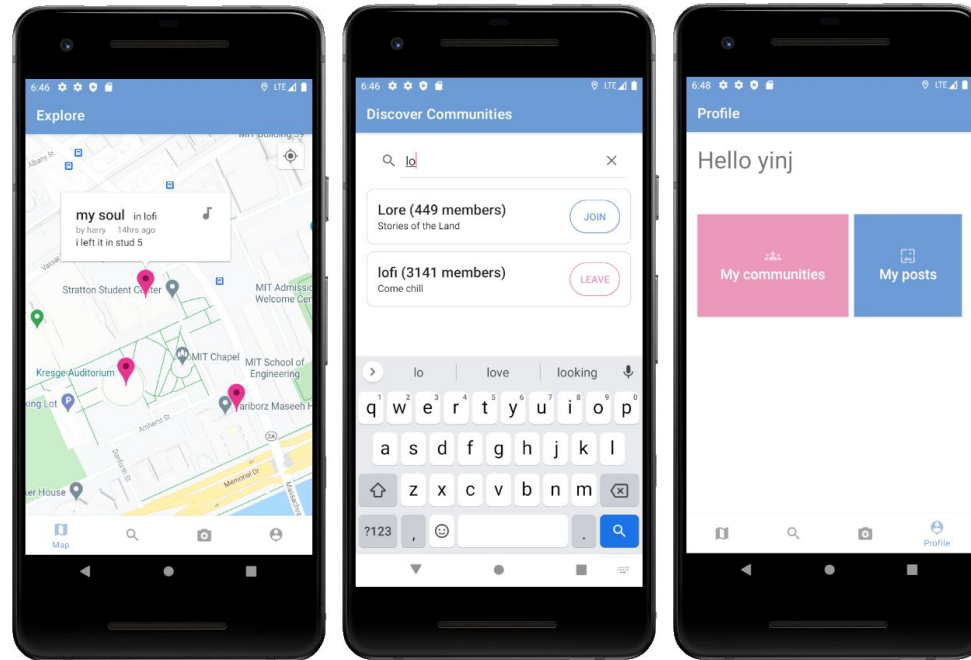
Figure 4: Left: the **Map** tab, showing an interactive map of nearby posts in joined communities and a short description of the post. Middle: the **Communities** tab showing a searchable list of all communities that the user can join or leave. Right: the **Profile** tab. The buttons saying "My Communities" and "My Posts" bring the user to a tab showing a list of those communities and posts when clicked with various functionality.

and looking through the same pair of eyes, which is a showcase of how the multimedia aspect of JESS can lead to stronger resonance within the community. Here we detail an additional few specific ways JESS can tell stories that are unique to its design.

- *Pen pals* - a commuter in the "self improvement" community posts a recording of her learning piano with

her son at a bus stop on her commute. Another commuter posts the next day at the same bus stop about her progress in a weight loss program. The two commuters regularly exchange updates on their morning commutes, inspiring more people to take inspiration and join in. The location-based aspect of the posts allow the temporally-separated commuters to effectively share a chat between acquaintances each day at the bus stop about their progress, with some of the

comfort and casual intimacy of an actual bus stop conversation.

- *Tour guide* - a coffee shop regular posts a dozing boy avatar in his seat in the "cozy places" community, and attaches a photo of his go-to studying drink and mentions the way the door chimes are causing a bird roosting opposite the street to caw back. The recommendations of what drink to order and what to listen for are appropriate and relevant to the audience who must be in that same community and visiting the shop already. We note here that coffee shops are an ideal location for posting, as users can relax and learn about other patrons' day as though they were each still seated in the shop and open to conversation.

- *Lore of the land* - a couple posts in the "memory lane" community on an anniversary a virtual heart in front of an office building. Attached to the post is a photo of their proposal in the bar that used to exist in the same space before it was bought and rebuilt into the new building. The location-specific AR nature of posts in JESS makes it much more impactful and evocative of nostalgia and loss to the office workers who happen upon the post, by superimposing the bar lost to time on top of the present building with which they have had physical interactions.

We also argue here that JESS is uniquely designed to accomplish its design goals by comparing it against three existing applications it shares some commonalities with – Snapchat, Reddit, and Mark, which we took inspiration from when designing JESS to account for the shortcomings we see with each application.

Snapchat [3] is a multimedia mobile social application that specializes in transient posts much like those of JESS. Snapchat succeeds in encouraging unpremeditated content where users share small stories and casual thoughts "in the moment". However, even aside from the benefits of location-based posts and AR we mentioned, we find Snapchat's interaction model of sharing between close friends to be unhelpful to our goal of promoting camaraderie and interpersonal trust in local communities. We designed JESS in part to share this type of casual content to a wider audience, while neatly avoiding the resulting issue of excessive spam often found on various social media paltforms, including Snapchat, through our stringent limitations on where and when one can view posts.

Reddit [4] is a forum website that revolves around sharing of posts in communities that users can join freely. We appreciate the community spirit it fosters; however, being a forum leads to a "popularity contest" problem, wherein the majority of posts are never seen and only the massively upvoted threads rise to the top, a problem JESS avoids with spatial locality and temporality. Moreover, a forum format encourages extra forethought put into each post (in part because of the need to garner votes), which runs contrary to our vision of users sharing non-deliberate snapshots of everyday life. Finally, we specifically wish to encourage more warmth towards the faces we brush by each day, and the digital platform of an online site we feel does not allow for effective translation of online togetherness to the physical world. We believe JESS through the power of spatial locality and temporality much more effectively generates the impression that some passerby you might have encountered had

---

[3]https://www.snapchat.com
[4]https://reddit.com
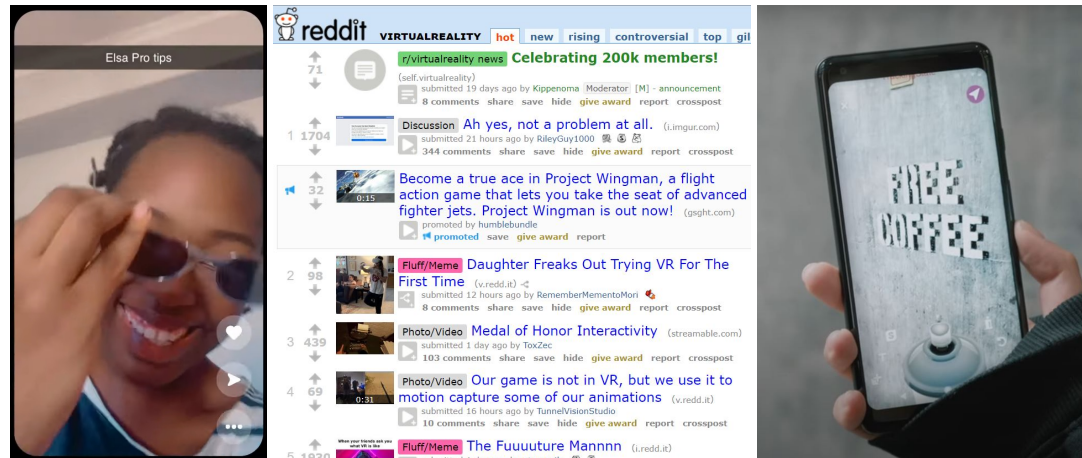
Figure 5: Left: an image of **Snapchat**, showing a captioned "snap". Middle: an image of the **Reddit** website, showing a list of the most popular posts in a community. Right: an image of the **Mark** app, showing virtual graffiti of "free coffee" on a blank wall. [2]

created the post.

Finally, Mark [5] is an AR social application wherein users tag walls and other public flat surfaces with their choice of preset virtual graffiti patterns. We enjoy the expectation-free creation and spontaneous self-expression it encourages but users have complained it is hard to find existing graffiti. Although our MVP currently does not address this geographic issue directly (more on navigation in future features), JESS is our attempt at expanding the range of AR self-expression with customizable, description-equipped posts with all the ease of use that selecting from predefined AR content affords, while providing users greater ease in finding content

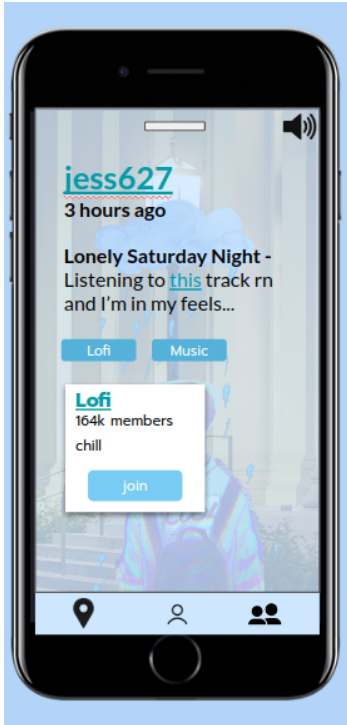aligned to their interests via showing nearby and relevant posts in a Map tab.

Future Features
Given the associated time constraints, technical on-ramp with Android development, and various integration issues in regards to Google Cloud ARCore, certain features were excluded from our MVP. Additionally, only one of us had a physical Android phone to test the AR features on, which further impeded our development. Below is a list of future improvements, starting with expected features that were not fulfilled. This does not intend to be an exhaustive list, but merely an attempt to catalogue the more important missing features.

Features planned for inclusion
Better navigation. Currently navigating to a post requires keeping the Map tab open and reading the map carefully,

---

[4]Snapchat screenshot taken from Snapchat Spotlight's trailer at https://www.youtube.com/watch?v=aDfzNGg_byU. Mark screenshot taken from the Mark app official trailer at https://www.youtube.com/watch?v=EQ0v_LoLS9g
[5]https://www.mark.app

Figure 6: A frame taken from our paper prototype showcasing some of the overlay functionality that were hoping to achieve in our MVP

which makes viewing (and even discovering) posts difficult when on the streets. The original plan included sending users to Google Maps when clicking on a post in the Map tab to plan a route to the post. In the following future improvements section, we expand upon this to propose a newer technology to resolve this issue of lacking a route as well as improves upon it by not diverting the user to outside the app.

Notifications/filtering anchors.   Originally, we planned on sending a notification to users if a post in a community they have joined was in a close enough proximity to their current location. In other words, we wanted to filter the anchors/posts according to location and community to achieve the spatial locality and temporality and intimate communities outlined above. Filtering anchors based off of community and location leading to notifications is not done for the following reasons. Firstly, as mentioned, tthere is a conflict between the Google Maps API (keyed authentication) and Google Cloud Anchors API (OAuth 2.0/keyless authentication), which prevents both from running at the same time on the app currently. Given this, there was no way to test the filtering anchors functionality even if implemented 2) Again, due to time constraints, we could not set up a function that efficiently and continuously filtered the anchors. For example, given all retrievable anchors from database, which are in x user's radius according to x's location and are part of x's current communities. This process needs to be continuous because user x is moving and possibly updating their joined communities list 3) Notifications is another layer on top of all this that requires keeping track of previously recorded posts that past the filter and were notified to x user. To have a record of these notifications, would require a cloud database (e.g. Firebase) because despite various attempts, we could not write to a non-cloud service on the android app - all files local to the app created before

or after the build were considered local to the device and thus unable to be written to.

Post/author description overlay.   In our paper prototype we promised a UI similar to the one depicted in Figure 6 where a window can be dragged over the post as you are viewing it yielding the author name, post description and various community tags. Clicking a community tag would reveal more buttons that allow you to join the community or lead to the community description page. Clicking the author name would go to their profile page consisting of a short description and their joined communities with similar extended buttons outlined here. This functionality was ommitted due to time constraints.

Feedback.   Positive feedback on posts can facilitate the creation of good posts and can provide motivation to continue using the application. However, the current design does not allow for feedback on posts you have made. For instance, adding a views counter and allowing viewers to upvote or 'like' a post while enabling the creator to see the number of views or likes they have garnered in the "My Posts" would provide this motivation.

Other support functionality.   We focused on creating and refining the core functionality of JESS - creating an AR post via a template that can be then viewed even after closing the app and returning to the location. Other support functionality including switching accounts, making communities, tagging multiple communities and choosing a community template to use, and more are to be added in a final product.

Future improvements
Earth Anchors.   Instead of sending the user outside the app to Google Maps or using the Google Maps API to mon-

itor your location relative to a post, a better alternative for navigation would have been to use Google ARCore's Earth Anchors which provide built-in navigation to anchors (turn which direction at which intersection, etc) at higher fidelity than GPS. This new functionality is currently unavailable to the public, but is soon to be integrated into ARCore.

Adding content to existing anchors.   Currently creating a successful feature map is difficult as it requires scanning a location from multiple angles and can take upwards of several minutes to resolve. Being able to attach content to previously scanned and used feature maps would make the process of making posts more efficient, particularly in the "pen pals" example detailed above.

Ability for users to add templates.   As denoted in our MVP, we believe templates help making posts easier and more efficient, but as the app scales, it isn't feasible for the development team to develop templates for each emerging community. As a result, the ability for users to create templates and add them to communities within the app would be crucial.

Audio playing.   Since we already support attaching music to a post, playing that music at a volume inversely proportional to the distance of a viewer in the 'Find a Post' mode would create an effective illusion of that music emanating from the post's AR object.

Censorship.   The unmoderated, anonymous, and broadcasting (as opposed to messaging) functionality of JESS make it quite vulnerable to profane posts from malicious users. While most posts only last a day(there may be some future metric for determining eligibility for making posts that last longer than a day like some threshold of cumulative number of likes), some automatic vulgarity detection or flagging by moderators (akin to reddit moderators) will be needed if this app is to be released to the public.

Multiple objects tied to a single anchor.   This is no doubt a major technical challenge because the feature map would need to take into account multiple objects. The reason for proposing this idea is to increase the range of posts. For example, choosing a combination of models from the various templates yields higher possibilities than a set of objects with single models or models conjoined in a fixed way.

Texture mixing.   This is a conceptual idea that arose from when we used a random png as the texture for a cloud 3d model. We got a crazy looking, but interestingly, deeply evocative creation that sparked the idea of having a section for choosing textures in the template choosing section in addition to choosing a model and choosing a song (song in the case of the lofi community).

Editing posts after posting them.   I think the ability to revise posts after posting them would be nice as it is a commonly available feature on various social media platforms. Some of the predicted edits include editing text, community tags, deleting before expiration data, and more.

Acknowledgements

REFERENCES

[1] Pew Research Center. 2019. Trust and Distrust in America. (July 2019). Retrieved December 7, 2020 from `https://www.pewresearch.org/politics/2019/07/22/the-state-of-personal-trust/`.