KTH ROYAL INSTITUTE
OF TECHNOLOGY

Degree Project in the Field of Technology and the Main Field of Study Computer Science and Engineering

Second cycle, 30 credits

Designing Diverse Features to Reduce the Filter Bubble Effect on Social Media

RAMYA KANDULA

Stockholm, Sweden, 2023

Designing Diverse Features to Reduce the Filter Bubble Effect on Social Media

RAMYA KANDULA

Master's Programme, Interactive Media Technology, 120 credits
Date: September 28, 2023

Supervisor: Jooyoung Park
Examiner: Anders Hedman
School of Electrical Engineering and Computer Science
Host company: COSE
Swedish title: Utformning av Mångsidiga Funktioner för Att Minska Filterbubblans Effekt på Sociala Medier

Designing Diverse Features to Reduce the Filter Bubble Effect on Social Media

RAMYA KANDULA

The filter bubble effect has been an active area of research that has been explored in various contexts within social media. Research on recommender system designs within filter bubbles has received a lot of attention, mainly due to its impact on the phenomena. They have contributed to filter bubbles by personalizing content on social media based on users' behaviours and preferences. This personalization has led to diminished user control over content consumption and a lack of exposure to diverse content and perspectives. This has thus had a significant impact on user agency and polarization across social media. Current research hasn't shown signs of integrating diverse recommender system designs on social media platforms, which is important to understand their effectiveness. Although recommender systems have contributed to filter bubbles largely, the proposition of incorporating them to counteract these bubbles presents a unique opportunity to understand how they function and how they could be optimized for diversity on social media, in a way that improves user autonomy. This project was done with COSE, a social network startup whose focus was on creating and sharing science and technology based content. The project aims to integrate recommender system designs on social media platforms to guide diverse exposure to content and perspectives within the context of science and technology based topics, and thus help in reducing the filter bubble effect. Using a Research through Design (RtD) approach, three design features, Discovery+, Views, and Consumption History were designed and prototyped to allow exposure to diverse content, perspectives and empower users to take control of their content consumption on social media. The evaluation for the three features indicated the possibility of promoting diverse exposure to content and perspectives. Users were exposed to a diverse range of content through the features, which increased their desire to explore new information, promoted transparency in content consumption habits, and provided access to multiple viewpoints within a single post. This urged users to actively seek and engage with a variety of content. A need for intuitive design of the Views feature was highlighted considering its presentation during testing and the intent of the feature itself. Nevertheless, the feature proved to be encouraging of finding diverse perspectives and invoked curiosity. Potential for further research on how the key design principles could be effectively integrated with diverse recommender system designs was also highlighted. The diverse features highlight the importance of fostering ethical social media experiences that support social sustainability, where users have access to empowering information, promote dialogue, help people make informed decisions, and strengthen social cohesion, while also addressing social media perils like polarization, echo chambers, dialogue hindrance, and restrictions on encouraging a global perspective on various topics.

SAMMANFATTNING

Effekten av filterbubblan har varit ett aktivt forskningsområde som har utforskats i olika sammanhang inom sociala medier. Forskning om design av rekommendationssystem inom filterbubblor har fått mycket uppmärksamhet, främst på grund av dess påverkan på fenomenet. De har bidragit till filterbubblor genom att anpassa innehållet på sociala medier baserat på användarnas beteenden och preferenser. Denna anpassning har lett till minskad användarkontroll

Author's address: Ramya Kandula, kandula@kth.se, KTH Royal Institute of Technology, School of Electrical Engineering and Computer Science, Stockholm, Sweden, SE 100 44.

Document date: September 28, 2023

Swedish title: Utformning av Mångsidiga Funktioner för Att Minska Filterbubblans Effekt på Sociala Medier

© 2022 Copyright held by the author.

över innehållskonsumtion och bristande exponering för mångsidigt innehåll och perspektiv. Detta har således haft en betydande inverkan på användarnas handlingsfrihet och polarisering över sociala medier. Nuvarande forskning har inte visat tecken på integration av mångsidiga design av rekommendationssystem på sociala medieplattformar, vilket är viktigt för att förstå deras effektivitet. Trots att rekommendationssystem har bidragit till filterbubblor i stor utsträckning, ger förslaget att inkludera dem för att motverka dessa bubblor en unik möjlighet att förstå hur de fungerar och hur de kan optimeras för mångfald på sociala medier på ett sätt som förbättrar användarnas autonomi. Detta projekt genomfördes med COSE, en startup för sociala nätverk vars fokus var att skapa och dela innehåll baserat på vetenskap och teknik. Projektet syftar till att integrera design av rekommendationssystem på sociala medier för att vägleda mångsidig exponering för innehåll och perspektiv inom ramen för ämnen baserade på vetenskap och teknik, och därmed bidra till att minska filterbubblans effekt. Med hjälp av en forskning genom design (RtD)-metodik utformades och prototyperades tre designfunktioner: Upptäckt+, Visningar och Konsumtionshistorik, för att möjliggöra exponering för mångsidigt innehåll, perspektiv och ge användare befogenhet att ta kontroll över sin innehållskonsumtion på sociala medier. Utvärderingen av de tre funktionerna visade sig främja mångsidig exponering för innehåll och perspektiv. Användarna exponerades framgångsrikt för en mångsidig mängd innehåll genom funktionerna, vilket ökade deras önskan att utforska ny information, främjade transparens i konsumtionsvanor för innehåll och gav tillgång till flera synvinklar inom ett enda inlägg. Detta uppmanade användarna att aktivt söka och engagera sig i olika innehåll. Behovet av en intuitiv design av Visningsfunktionen framhävdes med tanke på dess presentation under testning och funktionens syfte. Trots detta visade sig funktionen uppmuntra till att hitta olika perspektiv och väcka nyfikenhet. Potentialen för ytterligare forskning om hur de viktiga designprinciperna kan integreras effektivt med mångsidiga design av rekommendationssystem lyftes också fram. De varierade funktionerna betonar vikten av att främja etiska sociala medier-upplevelser som stöder social hållbarhet, där användare har tillgång till stärkande information, främjar dialog, hjälper människor att fatta informerade beslut och stärker social sammanhållning, samtidigt som de också tar itu med sociala medier-risker som polarisering, ekkokammare, hinder för dialog och begränsningar för att främja en global perspektiv på olika ämnen.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**.

Keywords: Filter bubbles, recommender systems, diversity, social media

Nyckelord: filter bubblor, rekommenderande system, mångfald, sociala medier

1 INTRODUCTION

The topic of filter bubbles and its effects have been an active area of research largely within the context of social media with respect to polarization and information exposure [21, 47, 51]. The term “filter bubble” was first coined by Eli Pariser in 2011 where he mentioned that “a filter bubble is an environment, created by a personalization algorithm, in which a person only encounters familiar information or opinions” [59]. This has also been studied in topics outside of news-based information such as healthcare [30], scientific research [61] and movie recommendations [57]. As it was explored, recommender systems and algorithms have played a significant role in contributing to the filter bubble effect on social media platforms [73]. Studies have shown that social media users tend to gravitate towards like-minded people and sources of information that reinforce their own opinions [76]. Research also suggests that many interventions aimed towards countering the filter bubble effect that have largely focused on interventions catered toward diversifying audiences in news-based contexts such as Wonkosphere - a new website that displays headlines from different political lenses[4], Balancer - a browser widget that alerts people whether their browsing behaviour is balanced with diverse perspectives or not [52], and CubeThat - a news article recommender that recommends diverse perspectives of news[16]. However, most of these haven’t been integrated into social media platforms, which is crucial

for assessing the effectiveness of them on social media. Filter bubbles on social media can lead to significant effects such as polarization that can impact people on a societal level by dividing them into like-minded groups [53], which makes it all the more necessary to test interventions on social media platforms. Moreover, it is necessary for these kinda of interventions to be tested outside of news-based contexts, such as that of scientific research or happenings as well since content of these topics also affects mindsets of people [13]. Another important gap observed in research was the lack of usage of a user-centric perspective [63] while designing interventions, which considered users' experiences with recommender systems. This was essential to be looked at since this emphasized the users' perspective by analyzing their needs and providing solutions that keep in mind the experience of their interactions with social media platforms.

Therefore, the focus of this thesis has been on exploring ways to improve diversification of content and perspectives keeping the integration of the designed solutions with social media platforms in mind. This thesis focused on using COSE as the social media platform. COSE is a social network startup that aims to build communities on their platform where people can share opinions and interact with each other. The type of content highlighted for this project was written articles within the topic of science and technology. They believe that a social network should be an inclusive space where ideas can be shared, empower creators and users, and be respectful with their mental health. Since depolarization is one of COSE's core values, the organization was keen on understanding how diverse content and perspectives could be provided to users to encourage diversity with positive social outcomes. This was done by exploring the design of recommender systems. The following research question was formulated based on the research topic: How can the design of recommender systems effectively promote diversity to reduce the filter bubble effect on social media?

The methodology adopted a Research through Design (RtD) approach and consisted of 2 studies. Firstly, a focus group session was conducted which included a total of 10 participants spread across 3 sessions, to gain insights into their understanding of filter bubbles, their relationship with social media and how they seek diverse content. Based on the results, 3 design features were designed i.e. Discovery+, Views and Consumption History, which were integrated as part of the COSE prototype on Figma. As the central values of this project focused on achieving diversity and serendipity, the features were designed around them. The final evaluation involved individual user testing sessions with questionnaires before and after the testing, with a total of 6 participants. The questionnaires aimed to understand participants' reading behaviours with respect to science and technology based content and their general experience interacting with the features on the prototype. The results highlighted positive feedback on the design features along with the potential to modify certain elements to enhance their experience and the interpretation of them on users. All of the features provided a novel way of interacting with social media that focused on the importance of providing diverse content and perspectives and user autonomy. The Discovery+ feature allowed people to explore new content and information and gave them control to curate their feed. The Views feature allowed people to explore different viewpoints of related articles and was thought of as a novel way of exploring diverse perspectives. It allowed people access to different perspectives about a specific piece of content. However, the novelty of the feature led to an initial lack of intuitiveness when it came to the viewpoint terms while interacting with the feature for participants. Hence, more research is needed in understanding how viewpoints could be defined to present to users for science and technology based content. The Consumption History feature made people more aware of their consumption patterns, thus allowing them to evaluate their own behaviours on social media.

2 BACKGROUND

The background addresses filter bubbles and different aspects outlining their impact and their connection to recommender systems. It first addresses the impact of filter bubbles on society through social media and then dives deeper into how recommender systems play an active role in contributing to this. It then explores how current interventions in

recommender system designs have contributed towards breaking the filter bubble effect and emphasizes necessary elements to guide the design of future solutions. Finally, COSE's role in this thesis project is elaborated.

2.1 Filter Bubbles and Recommender Systems

Social media has changed the way we consume information. With technological advancements over the years, and the ever-evolving nature of social media, content has become a ubiquitous entity for many where people share their opinions and ideas. This has made social media a space where anyone and everywhere could voice out their views, and reach millions of people. The ease of access towards sending and receiving content has made social media what it is today. However, what used to be considered as a space for open discourse has now become a place that divides people influenced by polarized discussions and extreme filter bubbles [56]. The term "filter bubble", first coined by Pariser suggests that it refers to an environment shaped by a personalization algorithm in which an individual encounters only familiar information or opinions, which leads to a state of intellectual isolation [59]. This phenomenon has the potential to filter out critical issues, resulting in a lack of exposure to different perspectives [75] and restriction of information sources [40]. As opposed to Pariser's initial use of the phrase "filter bubble", which focused on the change in search results online of users due to algorithmic personalization, the phrase is now more widely used to characterize a disturbance in information flows, within social media [15]. This exclusion of important information and viewpoints typically occurs unintentionally because of excessive personalization [59] and usually without user's autonomy [11].

2.1.1 Impact of Filter Bubbles. The issues associated with filter bubbles are numerous. One of the main problems stemming from this is the ability to reinforce echo chambers [11, 71] in which users will be served with information selectively thus, reinforcing their opinions [55] that can lead to opinion polarization [24, 54, 56]. Such an environment has the ability to impact society. This reinforcement of existing views might result in a lack of exposure to other perspectives and crucial information, impeding the development of a well-rounded understanding of many topics. It can increase biases [22], hinder dialogue amongst people [10], expose people to extreme views that could be discriminatory [39], and cause political unrest [77]. Filter bubbles may accelerate societal polarization and fracture societies by segregating people into homogeneous groups. A lack of exposure to different points of view can lead to misunderstandings, disputes, and the deterioration of empathy and social cohesion. Filter bubbles also enable fake news and prevent individuals from being critical towards misinformation [58]. This can worsen the growing lack of trust towards genuinely reliable sources and social media platforms as a whole. Moreover, their ability to critically evaluate information and make informed decisions gets compromised. Filter bubbles can also hinder learning and exposure, which can make it difficult to share new knowledge that can shape communities [29]. As a result, this could represent a threat to scientific research by potentially steering researchers into narrower niches and impeding overall development [61]. Due to the problems outlined above, society's collective knowledge of the world and ability to form opinions is affected significantly.

2.1.2 Role of Recommender Systems in Contributing To Filter Bubbles. The lack of agency provided to users on social media is concerning and requires the need to investigate what interventions could be proposed to combat this. It has been suggested by Wang et al. that the filtering of content due to recommender systems is responsible for creating filter bubbles [73]. Recommender systems refer to the algorithms that are utilized to select, filter and tailor content as per users' interests [45]. They help in suggesting content to users across social media based on various factors such as their past behaviours, preferences and other metrics, and have significantly contributed towards personalization. They have a social and collaborative nature, as they are designed to support user-to-user interactions, and they typically use patterns across users to make recommendations [70]. Personalization, which occurs when content is recommended to users based on their preferences and interests, play an essential role in knowledge management [65]. With increase in

personalization, users' exposure to similar content increases and as a result, leaves them isolated from diverse content. The positive and negative effects of recommender systems have been researched extensively within social media and this matter has been an ongoing debate within academia. On one hand researchers claim that recommender systems help in promoting diversity on social media [31, 57]. It has been suggested that recommendations in the form of social endorsements have led to the exposure of a wider range of perspectives [26] and in addition to that, a vast portion of online connections on social networks are formed between those with contradicting viewpoints that can thus create opportunities for discovering diverse content [37]. This has thus contributed to an enhanced user experience which makes it easier for users to find content they're interested in and saving time while doing so since they wouldn't have to look for relevant content extensively.

On the other hand, many have argued that these contribute to polarization [18, 59, 71]. Recommender systems have been identified towards isolating users from diverse content and perspectives ideologically which is said to have contributed to group polarization [73], where they get fragmented across social media and those with similar viewpoints get grouped together. This grouping of users would gradually lead to reinforcement of their beliefs where they get exposed to similar content, thus leading to formation of filter bubbles. Various studies have analyzed the impact of these on social media platforms. For example, research has shown that Facebook's news feed algorithm filters posts that are not similar to the user's beliefs, filters out other users that the user isn't interacting with [8]. Another study by Itkonen on link-sharing in Facebook showed that users have friends online who share similar views on topics like climate change and how Facebook's algorithm is designed to display content that aligns with their views and thus groups them [33]. A study on YouTube's recommendation system pointed out that its algorithm promotes channels based on the user's previous history, which could further push them towards echo chambers and ultimately lead to misinformation and polarization [37]. Platforms like Twitter have also played a role in creating ideological divides within online communities as identified in Barbèra et al's findings which indicate that individuals are more likely to share information from ideologically similar sources than those that are dissimilar [9].

Moreover, the nature of discourse happening on social media transcends borders, thus allowing people to engage in discourse on a global scale. This further fuels polarization, misinformation and a lack of diversity on social media. Due to the feedback loop [35], users are constantly presented with suggestions for related content, which can further pique their interests and result in problems with ideological isolation and echo chambers. The consequences of personalization can be felt on an individual level, where it might provide less opportunities for users to make their own decisions, discover different types of content and alternative perspectives [71]. The outcomes can also range from restricting individual creativity, personal growth and reduced ability to build meaningful social connections that could benefit people on a personal and professional level [59]. With such consequences being experienced on a societal level, researchers have attempted to propose interventions that help combat them.

2.2 Redesigning Recommender Systems to Mitigate Filter Bubble Effects

The mitigating of filter bubbles is an open challenge for the way recommender systems function [57], which has been explored in various topics such as healthcare [30], scientific research [61], news [22], movie recommendations [57], and music [31]. Many researchers have proposed redesigns for recommenders that could help in preventing the generation of filter bubbles and improve exposure to diverse content. Moreover, the usage of recommender systems themselves to weaken the filter bubble is an interesting one, since the very factor has been a significant contributor to strengthen filter bubbles. Leveraging the functionalities of recommenders to improve exposure to diverse views and perspectives rather than personalization can help in weakening the filter bubble. The following set of related works will investigate how these have been executed to increase diverse flow of content on social media within news, along with some works that investigate interventions outside of news-based contexts that have aimed towards reducing the filter

bubble effect. These works provide a means to understanding how recommender systems can be redesigned and what functionalities allow them to reduce the filter bubble effect.

2.2.1 Recommender Systems in News-based Contexts. Some of the interventions look at breaking filter bubbles by recommending diverse perspectives of similar content such as AllSides, Flip and Wonkosphere [4, 5, 78]. These look at displaying content (news content in this case) from opposing viewpoints at the same time which can encourage users to be more aware of the way content is conveyed and broaden their horizons. Balancer [52], a recommender intervention designed as a browser widget aims to make users aware of the media they're consuming by showing them their reading behaviors. It does so by displaying a visual graphic symbolising which type of content they're leaning towards more with respect to the political spectrum. This is used as a technique to nudge them towards reading balanced political viewpoints. CubeThat [16] and NewsCube [60] also aim to bring diversity by recommending content similar to the current story being viewed by the user, using mechanisms like clustering and aspect level browsing respectively, to group the related stories. User Controllable Recommendation System [73] addresses the issue of filter bubbles by including a filter bubble indicator that shows the level of recommendation diversity and warns users when their recommended material is getting too limited. Users can also give input on recommended things and choose items of interest to widen their recommendation pool. This way it provides user autonomy in controlling the recommendation process. Social Mirror [42], a web application, provides users with an overview of an ideologically divided social network, and urges them to place themselves within the ideological spectrum that could motivate them to seek out new information and promote diverse sharing of content.

2.2.2 Recommender Systems Outside of News-based Contexts. Although research on the analysis of recommender systems outside of news based contexts are limited, they provide information on how scholars have attempted to improve content and opinion diversity to reduce the filter bubble effect. For example, Nguyen et al. studied how content diversity could be improved at the individual level by analyzing the impacts of MovieLens recommender that recommends movies, on users and whether it led them to a filter bubble [57]. In another study, Huber et al. investigate how users could navigate effectively across podcast recommendations through explanatory labels that could help them in choosing the right podcast to listen to [32]. Another study argues that recommender algorithms filter important information from users when they refer to online resources to manage their own health. They suggest that an unfiltered body of information on the internet among others could help in countering this, in ways that provide unfiltered search where unbiased information would be made available to the public and make them aware of the filter bubble itself. This could thus help in minimizing the filter bubble effect [30]. Portenroy et al. in their study develop a platform that could help users in easy navigation of scholars and their research within the domain of scientific research to counter scientific filter bubbles where isolated bodies of scientific research limit awareness across the literature and hinder innovation [61]. It should also be known that research in communicating science and technology based content on social media that is outside of academia is a relatively new topic but is still relevant to address filter bubbles and improve the spread of diverse content. Moreover, it is necessary to study them further to understand how it influences people's mindsets, their ways of thinking, assumptions and perspectives.

2.2.3 Research Gap Identified in Designing Effective Recommender Systems that Mitigate Filter Bubbles . Although the interventions discussed are varied, they are all unified in their goal of promoting user autonomy and broadening viewers' viewpoints to create a diversified media environment with a variety of content. Existing work on recommendation systems incorporate diversity as an important objective, however they don't prioritize user experience [73]. Moreover, a significant amount of research on recommender systems have focused on algorithmic approaches [36] thus highlighting the need for design focused interventions. The design features of an interface can influence a

user's perceptions on recommended things [49]. Zhao's research suggests that through designing effective interactions, it is possible to communicate diverse viewpoints to users [78]. Hence, for this thesis, the interventions proposed were in the form of design features. It is thus necessary to ensure that the design aspect of a recommender system receives adequate attention. Existing research has also identified that the majority of interventions have not been integrated and evaluated as part of a social media platform. This integration is critical for gaining a better understanding of the effectiveness of the intervention. Anand et al. in their study justified this by stating that their intervention shouldn't be limited to a single dataset or social media platform, whereas testing their interventions within other social media platforms would help in gaining better insights on the filter bubble effect within recommender systems [7]. There is a need to study filter bubbles more within topics outside of news and politics since social media is ubiquitous which consists of a plethora of varied topics of content, that has the ability to influence people's perceptions. Moreover, the research done within filter bubbles outside of news is limited, hence there is a need to expand these areas to gain insights on people's perceptions of other types of content. Interdisciplinary studies provide useful insights into the potential consequences of filter bubbles on scientific information dissemination and perception [13].

2.3 Important Principles to Guide Recommender System Designs

A lot of research has been done in proposing various principles and frameworks to guide the design of recommender systems to mitigate filter bubbles, keeping exposure diversity in mind. Principles like serendipity [38, 64] value sensitive design [23], diversity [27, 47] exposure diversity [28] and transparency [68] have been explored as important design values to guide recommender system designs. Diversity in the context of this project draws from Helberger et al.'s study, where it refers to the idea that in a democratic society, well-informed citizens seek information from a diverse range of sources with different viewpoints, allowing them to make balanced and well-considered decisions [28]. Moreover, it has been a core value to counter filter bubbles. The usage of serendipity has been relevant to the development of recommender systems since it enhances discoverability that thus allows exposure to diverse content. Discoverability empowers users more over content consumption, in a way that doesn't force them to view content they have no interest in. Transparency has also been identified as an important design value to consider to guide the design of recommender systems. This is important to explore since it brings a sense of confidence in users through providing additional information about their actions, which could further instill trust between users and the platform [17].

Nudging and other behavioral change techniques [34, 72] have also been used as design mechanisms to urge users to explore diverse content. It refers to bringing a change in behavior of a user without restricting the freedom and individual choice. Nudge has been termed as an important design aspect to consider while exposing opposing and diverse views to users [34]. Although this principle has been applied more within the context of mitigating echo chambers, it can also be applied to filter bubbles. However, it should also be noted that this mechanism has contributed towards the filter bubble effect since nudges have been used to promote personalization [62]. This also raises ethical concerns on the selection of diverse serendipitous content presented to users, a responsibility which usually rests on the social media platform [59]. Nevertheless, the usage of nudge also shows the possibility of doing the opposite, which is to promote a diverse exposure of content and perspectives. This was especially noticed in Munson et al.'s study where they designed a widget to nudge users to read balanced political viewpoints [52].

The key aspect tying these values together is the need to increase user autonomy by giving them agency in deciding the way they wish to consume and navigate social media along with the ethical concerns associated with the impact of recommender systems on social media users. These aspects are critical in addressing the interventions associated with recommender systems.

2.4 Students as an Essential User Group

Research within the impact of filter bubbles on students is quite limited but growing. For this project, students have been identified as a suitable user group to base this thesis project on. According to research, students are placed as the majority for the highest consumers of social media platforms and are also an influential group as they're exposed to diverse opinions on a regular basis [19]. The student demographic is also included in COSE's target user group. University students specifically make an ideal user group because they have access to education, making it easier for them to acquire information [20]. Moreover, students come from diverse backgrounds and academic disciplines with a wide range of interests and mindsets, thus providing an opportunity for a thorough evaluation of interventions. Wiel's study on educating students about the implications of filter bubbles sheds light on the necessity of bringing awareness amongst them and thus empowering them [74]. This could further lead to a positive impact on their ability to make informed decisions through exposure to diverse content.

2.5 COSE's Vision

COSE is a social network startup that aims to build communities on their platform where people can share opinions and interact with each other [1]. The company believes that a social network should be an inclusive space where ideas can be shared, empower creators and users, and be respectful with their mental health. This is resonant with their platform. COSE's platform allows users to interact with content by accessing regularly updated newsletters that are notified to the user, discovering new content through Discovery and interacting with content and creators on Communities. They can also engage in discussion with creators or communities if they wish to. Diversity [28] is one of their core values and so with this project, COSE aims to understand the impacts of diverse recommender interventions through social media platforms on users and thus contribute to practices aimed towards reducing the filter bubble effect. The current version allows users to explore new content through the Discovery feature. Nevertheless, the Discovery feature alone doesn't allow users to explore content outside of their interests and preferences. Moreover, the content appearing on the Discovery page is based on users' preferred communities to interact with and content related to what they've been consuming. This doesn't provide an opportunity for users to venture out of their bubble and get exposed to new content. The lack of diversity in perspectives hinders users' ability to gain knowledge from varied viewpoints and make informed decisions. Therefore, the following research question arises: How can the design of recommender systems effectively promote diversity to reduce the filter bubble effect on social media? Through this project the company aims to use this research to adapt, design, and provide diverse content in order to encourage diversity with positive social outcomes and ultimately provide diverse social media experiences through their platform.

3 METHODS

This project used the Research through Design (RtD) approach [79] to address the research questions involving the redesign of recommender systems to improve exposure to diverse content and diverse perspectives. This approach involves combining research along with design methods that can lead to design interventions to challenge a problem. Since this method involves usage of the design process, the design artifact produced itself becomes relevant to communicate knowledge [69]. The methods used within this began with user research to understand user needs that could act as insights to develop new solutions. This was done by conducting 3 rounds of focus group sessions with a total of 10 participants. The first part of the session focused on gathering data and insights on students' social media consumption habits along with their understanding and views on filter bubbles. The second part of the workshop focused on brainstorming recommender system solutions using SCAMPER [67] to break the filter bubble within social media platforms. Thematic analysis was used to determine the major themes from the focus group and questionnaire

data by applying the framework proposed by [12]. Three design ideas that were derived from user research insights were incorporated into the COSE mobile platform as features through designing a prototype version of the same. Lastly, the final evaluation involved testing the prototype with 6 participants and gathering qualitative data on the user experience, and then analysing it using thematic analysis.

3.1 Study 1 - Focus Group Study on Effects of Filter Bubbles and Possible Solutions to Tackle Them

3.1.1 Procedure. The first study engaged a focus group where the first part of it involved an open discussion on filter bubbles and recommender systems and the second part consisted of a brainstorming session where participants came up with ideas to improve diverse exposure to content and perspectives. Focus groups, as opposed to interviews, have the advantage of encouraging participant engagement, which facilitates the gathering of more insightful information about participants' experiences [46]. The focus group aimed to gather data on participants' social media habits, their understanding of filter bubbles, and their views on the role of recommender systems, highlighting the importance of enhancing the design of these systems to promote the discovery of new content and diverse perspectives. Getting to know participants' experiences in a group setting allowed the flow of diverse perspectives, interactive discussions and discovering unexpected insights which helped in addressing the research question. The sessions followed a semi-structured list of questions that touched upon these areas.

The second part of the session involved participants in a brainstorming session where they were instructed to use the SCAMPER brainstorming technique to generate ideas on how experiences with recommender systems could be improved. SCAMPER is an acronym for S - Substitute, C - Combine, A - Adapt, M - Modify, P - Put to other use, E-Eliminate and R - Reverse [67]. Due to time constraints in terms of the sessions, participants were made to generate ideas within only two rounds of brainstorming, where they used the prompts S - Substitute and C - Combine respectively. They were given post its to sketch out or describe their ideas as seen in Appendix A.1, which were then grouped according to the prompt, followed by reflections of the ideas after each session. The sessions lasted approximately 90 minutes and took place in KTH's main campus library in group study rooms. The sessions were audio recorded and accordingly transcribed by the researcher using Otter AI [3]. No part of the participants' personal data has been used for this research, while analysing data.

During the sessions, the discussions began with an ice breaker that touched upon participants' social media habits and what each platform meant to them. Subsequently the discussion proceeded towards introducing the research topic and delved into participants' views and experiences with filter bubbles. Finally, in the brainstorming part of the sessions, participants were given the first two prompts from SCAMPER, where they were asked to think of ideas that could improve their exposure to diverse content and perspectives.

3.1.2 Participants. The requirements for the participants were for them to be students enrolled full-time in a program at a university-level. The 10 participants were between the ages 23-27 and were Master's students in various schools at KTH. Out of the 10 participants, 6 came from India, 1 from China, 1 from Germany, 1 from Slovakia and 1 from Russia. Consent from participants was requested via messaging services such as Whatsapp and word-of-mouth before signing up for the workshop using Jotform [2] and written consent was granted before it began. The three sessions were joined by two, five, and three participants each.

3.1.3 Data Collection and Analysis. The audio data from the focus groups were transcribed using OtterAI [3]. The transcribed data was shifted to FigJam using digital post-its and was analysed using a thematic analysis approach [12]. For this, a deductive and inductive approach was used while analysing data from the open discussion. Five codes were present initially, which were based on a general overview of participants' social media habits with respect to

content navigation, experiences of encountering filter bubbles and the role of recommender systems in them. Upon a second iteration of these codes, 11 codes/sub-themes were added. The codes were subsequently organised into 5 main themes: “user control and autonomy”, “issue with recommender systems”, “customised diversity in content curation”, “content discovery” and “designing ethical recommendations”. The table consisting of the codes and themes can be found in Appendix A.2. The ideas from the second part of the focus group i.e. the brainstorming session were also being analysed through thematic analysis. However, this part involved identification codes with a single round alone. Since many ideas overlapped each other or were recurring across sessions and participants, they were categorized into a broader theme of ideas. A total of 9 themes were identified from this as seen in Appendix A.3. After analysing data from the open discussion and the brainstorming session, two personas were created as seen in Appendix A.4, to gain a better understanding on the target user group and how to support them [50].

3.2 Designing the Prototype

3.2.1 Design Process. The prototype design process involved identifying key design values, drawing insights from both user research and literature (see Appendix A.5). These insights, stemming from literature and user research, yielded three design values and one design mechanism that guided the design ideas. These principles and insights informed the hypothesis statements, which in turn shaped the design ideas. Subsequently, these ideas were implemented as features for the COSE mobile platform, namely Discovery+, Views, and Consumption History. Beginning as low-fidelity paper wireframes, these features were refined through feedback sessions with the supervisor. Further iterations involved translating these wireframes into high-fidelity versions using Figma, aligning them with COSE’s design system. This process involved incorporating color palettes, typography, and visual layouts to ensure a seamless flow and connection between the design concepts. This ultimately reflected on the usage of research driven and iterative design process.

3.2.2 Softwares Used. For this project, FigJam and Figma were used extensively for putting up research and prototyping respectively. Apart from these, Jotform was used for inviting participants to take part in the user research and testing phases along with OtterAI for transcribing audio data.

3.3 Study 2 - Evaluation of Diverse Design Features on COSE Mobile Platform

3.3.1 Procedure. The final evaluation consisted of an individual testing session, which included a pre-study interview, interaction with the prototype and a post-study interview. Both the questionnaires were created on Google Forms which consisted of open ended questions to understand participants’ experiences with the prototype. The first questionnaire (pre-study interview) aimed to understand participants’ reading behaviours and personality when it came to interaction with science and technology based content on social media. This was considered necessary to explore since it provided a way to understand how participants perceived content and would further help in justifying their actions. Munson et al. [52] had followed a similar approach where they tried predicting participants’ reading behaviour based on the pre-study questionnaire provided, inspired by the Big Five personality index [25]. The second questionnaire (post-study interview) explored participants’ impressions of the design features implemented in the prototype in terms of functionality and experience. The user testing session took place over 4 days in person, depending on participants’ availability and lasted approximately 40-50 minutes. All sessions included audio and/or video recording and observational note making, and all participants provided written consent. Subsequently, the data from the sessions was transcribed by the researcher.

During the session, the researcher introduced the prototype version of the mobile app to participants. They were instructed to perform 3 tasks where the first task involved using Discovery+ feature to add sub-topics to their Discovery feed, the second task involved reading articles from different viewpoints that were presented in Views, and the third

task involved reading and reviewing the data presented to them about their consumption patterns with the app from the Consumption History feature. In addition to this, the importance of engaging with the content was emphasized during the beginning of the session. The participants who were recruited from the researcher's network, were instructed to interact with the prototype on a mobile device. For the first task alone, participants were instructed briefly on how to interact with the spinning carousel (with respect to locating the topics) in the Discovery+ feature due to the limited functionality of the prototype. The prototype was evaluated in a controlled environment, and participants were advised that the content might not fully align with their actual social media behaviors due to its restricted capability. During testing, participants had ongoing access to instructions to help them complete tasks and could refer to the prototype content while answering the post-study questionnaire.

3.3.2 Participants. The requirements for the participants were same as the ones from the focus group session for user research. A total of 6 students between the ages 24-28, were recruited for this session, who were all pursuing a Master's degree at various schools at KTH. Out of the 6 participants, 5 came from India and 1 from Iran. Only 1 participant from the 6 had attended the focus group session, while the other 5 were not familiar with the topic of the study.

3.3.3 Data Collection and Analysis. The answers to the open ended questions from both the pre-study questionnaire and the post-study questionnaire were shifted from Google Forms to FigJam using digital post-its and were analyzed using thematic analysis [12]. This was done using both a deductive and inductive approach that resulted in 11 codes from the pre-study questionnaire and 12 codes from the post-study questionnaire. A total of 6 themes were established from the pre-study data, based on the codes. The codes along with the themes can be seen in Appendix B. Based on the post-study questionnaire data, 6 themes were established. The codes along with the themes can be seen in Appendix C which have been categorized based on the features.

4 RESULTS

The results from the first study highlighted participants' relationship with social media and revealed their social media consumption habits, along with the challenges of feeling a lack of control while navigating across content. The importance of better access to diverse content and perspectives was highlighted along with the need to balance diversity and personalization in content recommendation. The second study explored participants' reading behaviour with respect to science and technology based content on social media along with emphasis on their reactions to the design features aimed towards reducing the filter bubble effect and improving diversity of content and perspectives on social media.

4.1 Study 1 - Focus Group Results

4.1.1 Findings from Focus Group. The focus group touched upon many issues such as the problems encountered with recommender systems and algorithms on social media to the challenges in designing those that are free from bias and filter bubbles. A lack of control and autonomy was identified when it came to navigating desired content along with an inclination to access a wider range of content. From the thematic analysis, 5 themes emerged: i) User Control and Autonomy, ii) Issues with Recommender Systems, iii) Customized Diversity in Content Curation, iv) Content Discovery, and v) Designing Ethical Recommendations. The 10 participants who took part in the study will be referred to as P1, P2 and so on until P10.

- **User Control and Autonomy** This theme emphasizes the need for platforms to provide more control over their experiences and over the content they see and engage with online communities in ways that are meaningful to

them. As regular users of social media, participants felt that having the control to decide what content to view/not to view was seen as necessary that could have a significant impact on their experience. Based on the discussion 3 sub-themes emerged: "flexibility", "control" and "choice". These sub-themes covered participants' frustrations in social media browsing. This highlighted situations where some participants experienced hindrances in viewing desired content. For instance, the recommendation algorithm would function in a manner where if they viewed a post about a certain topic, their feed would be flooded with more posts on the same topic, which did not pique their interest. This was noticed in statements like "once that I was watching like ski videos, and then after that my feed was full was that" (P8), and "I think I would like it to be more diverse, since sometimes I might just get some random stuff that I just accidentally have opened it and I might not really like it. But then the algorithm thought I like it. So it gives me more. So I think I would prefer to see a lot like more different stuff." (P9).
The discussion also identified some of the needs participants expressed that would improve their experiences on social media by highlighting that personalization shouldn't be mandated, "when you ask for personalization, it should be really a choice thing" (P2). Some mentioned the importance of awareness while consuming content, "what I've noticed is if you read one sort of opinion, and the algorithm sort of understands that, okay, this is how you think they will give you more of that. So instead of getting a perspective, on the other side, you just become more and more and more biased, which is again dangerous" (P3). They also raised important questions against it "where is the line you draw between my content and someone else's content? And how do I know how much of this is really curated for me? And how much of it is kind of general that, you know, all my friends are also seeing? (P7). These quotes from participants emphasized the need for transparency [68] in recommenders and to make users more aware of their content browsing habits, thus giving them more control over them.

- **Issues with Recommender Systems to Ensure Content Diversity** This theme focuses on the issues associated with the implementation and use of recommender systems. Based on this, 2 sub-themes emerged: "algorithmic bias" and "content repetition and filtration". During this part of the discussion, participants were specifically asked about their reactions to content appearing on their feed and the functioning of recommender systems, which revealed instances where they felt that the recommended content did not align with their interests through statements such as "I don't know if the creator has reposted it or if it's just the algorithm showing the same thing, because I don't like seeing those things" (P1), and "it doesn't know what you know what you like anymore, or maybe just tries to dump everything on you" (P2). On the other hand, some participants also acknowledged that they have accepted the way recommender systems function in personalizing content with statements such as "I wouldn't say that I feel overwhelmed by the amount of stuff on my feed. I can't say that because I already trained the algorithm to just give me what I want." (P6), and "when I open it, I see stuff that I want to see, I know what I'm getting into, it's predictable in that sense, and it's kind of comfortable because it's predictable, which is why if I don't want something predictable, then I'll actually actively go seek different ways of doing that" (P10). These statements from participants portray the lack of diversity [27] in content on participants' social media and the lack of control they have in getting access to it. This highlights the need for achieving exposure to diverse content by giving users more autonomy in content curation.
- **Customized Diversity in Content Curation** This was the main theme highlighted during the focus group sessions that emerged from 2 sub-themes: "filtering as a way to seek diverse content" and "balancing personalized and diverse content". The idea of achieving balance between personalized content and diverse content was highlighted in statements such as ""you're stuck in your own you know bubble of content but you're also actively seeking out other stuff, new stuff" (P7) and "I like to get the content I like, but I also like to see new things" (P1). The usage of filtering as a way to seek diverse content was highlighted when participants stated "one way that I used to filter was by using those you know, sometimes when you pick one thing from your Explore page, and

then you can continue scrolling, and it shows you some stuff, I use that as a filter." (P7) and "I want to have some filters with content I want to see now" (P2). This also sheds light on participants' perceptions of the filtering features on social media platforms and associating different purposes with them. This provides an opportunity to explore how nudging users [72] into discovering content outside of their comfort zone while balancing the content personalized to them, could ensure a diverse flow of content. Since nudging looks at influencing users' decisions in a positive manner to promote diversity [27], it proves to be an important design mechanism to consider implementing.

- **Content Discovery** This theme consists of 2 sub-themes, "importance of discoverability" and "discovering diverse content". Participants emphasized the need to account for the discoverability factor and acknowledge its value while seeking a wider range of content through statements such as "I don't know what I like until I see it." (P2) and "I got what I want so I like to stay there rather than see other stuff. I mean, you could say that if something new comes up and I end up liking it instead of not liking it, then it would be nice also but the chances of the stuff that we know would come up is much greater than what we don't" (P10). Participants also expressed that discoverability allows them to encounter or actively search for new information in a way that does not feel imposed by the platform. This natural approach piques their interest in learning more about new content and also highlights the importance of serendipity or discoverability [64] as a design value to consider while designing diverse features.
- **Designing Ethical Recommendations** When introduced to the topic of re-designing the current recommender systems, 2 sub-themes emerged: "challenges with designing solutions" and "reforms to avoid reinforcing biases". The challenges mainly stressed on the practicalities of delivering a recommender system that is truly diverse and might be too idealistic to achieve. This was seen in statements such as "I agree that it's important to be exposed to diversity on social media but in the real world I don't think it's possible where you have the algorithm designed in such a way that it gets you hooked to the platform" (P4). However, P4 also added that "people may not post as much but might be interested in where content is coming from, on what basis is a post coming on your feed". The matter of subjectivity in people's opinions about diverse perspectives arose when P3 mentioned "it's very subjective. For some people, they don't mind staying in their own bias because it gives them the comfort to live in their biases. But for some people, they would like to have more new perspective.", which questions the idea of considering reform in recommender systems. P7 raised an important concern about how oblivious people are to filter bubbles because of the way recommender systems function by saying "even if we do give them an option to like choose whether to support a bubble or not, if you don't even know you're in a bubble, and how are you going to make that choice, right, inside or out.", and further added that "that would require like a lot more education to the general public, and they don't always want to be educated". P7's quote here highlights the lack of awareness about the filter bubble effect amongst users of social media and educating them about it poses another challenge in designing interventions.

4.1.2 Categorizing Ideas and Persona Generation. Using SCAMPER method, participants brainstormed numerous design ideas for alternative recommender systems. These ideas focused on improving the experience of recommender systems to improve diversity in content and viewpoints on social media platforms. A total of 9 themes of ideas were identified within all the ideas as seen in Appendix A.3 out of which 3 of them were chosen to be taken forward since they aligned with the design values and these themes contained the most number of ideas that participants came up with: "customizable filters to allow users to curate their own feed", "web of related content within each post from different creators or perspectives" and "spread of new content at regular intervals on user feed to get out of the bubble".

Based on the findings, two user personas were created as seen in Appendix A.4. They were based on the participants who attended the focus group studies, most of whom were university students, with brief bios and descriptions about their needs, behaviours, goals and pain points. Creating personas ensured that the user is at the center of the design process and they also acted as references to turn to while making design decisions that are user focused.

4.1.3 Insights and Hypotheses. The findings from the focus group and literature study ultimately helped in gaining insights that served as starting points towards designing solutions for diverse spread of content and perspectives through redesign of recommender systems. Three design values and one design mechanism emerged as critical elements that guided the development of the design ideas during the research phase. The design values were serendipity [1], diversity [20] and transparency [65]. Along with these values, a design mechanism, nudge [19] was also seen as necessary to shape the ideas. The insights from research were based on five important elements discovered during literature study and user research from Appendix A.5. The insights are as follows:

- **I1.** Users desire the flexibility to seamlessly switch between content aligned with their interests and explore new content.
- **I2.** Improved discoverability encourages individuals to actively seek out a wider range of diverse content.
- **I3.** By displaying users' content consumption patterns, platforms can enhance transparency and increase users' awareness of their own behaviour.

These insights helped in shaping the hypothesis statements as well, that acted as guides to shape the design ideas. The hypothesis statements are as follows:

- **H1.** Exposing consumers to a varied range of new information on their feed on a regular basis can positively influence their inclination to actively explore and find new content.
- **H2.** Incorporating discoverability features that allows users to access multiple viewpoints within a post will significantly influence user behaviour, leading to a greater inclination for actively seeking and engaging with diverse content.
- **H3.** By providing users with information about their content consumption habits, platforms can enhance transparency and promote users' awareness regarding their own behaviour.

Using these hypothesis statements, the design ideas were generated and designed to be integrated into the COSE mobile platform.

4.1.4 Final Design Ideas. The design ideas were iterated as features that could be part of a single userflow (Figure 1) for the prototype where each idea was supported by each hypothesis statement. The 3 design ideas were: i) Discovery+ feature ii) Views feature and iii) Consumption History feature. The prototype was designed to accommodate two user flows that were similar in terms of information architecture but different in terms of the content itself i.e. the topics that the user would pick to explore in the Discovery+ feature. The user flows differed based on the topics of artificial intelligence (AI) and sustainability for the prototype. These topics were chosen as they were considered relevant areas within science and technology. Moreover, COSE has also actively contributed towards content on these topics. Since the emphasis of this prototype had been on the content, it differed based on what the user chose to view in the first design idea (Discovery+), which influenced the content appearing within the other 2 design ideas (Views and Consumption History) as well. This would thus dictate the flow of the prototype for the user interacting with it.

- **Discovery+ feature** This feature was designed as an extension of the existing Discovery feature of COSE platform. While the existing Discovery feature allowed users to discover personalized content, the enhanced Discovery+ feature aimed to expose users to content that was outside of their interests and broaden their horizons. The goal of this design was to foster a wide range of content by allowing users to carefully add topics of their

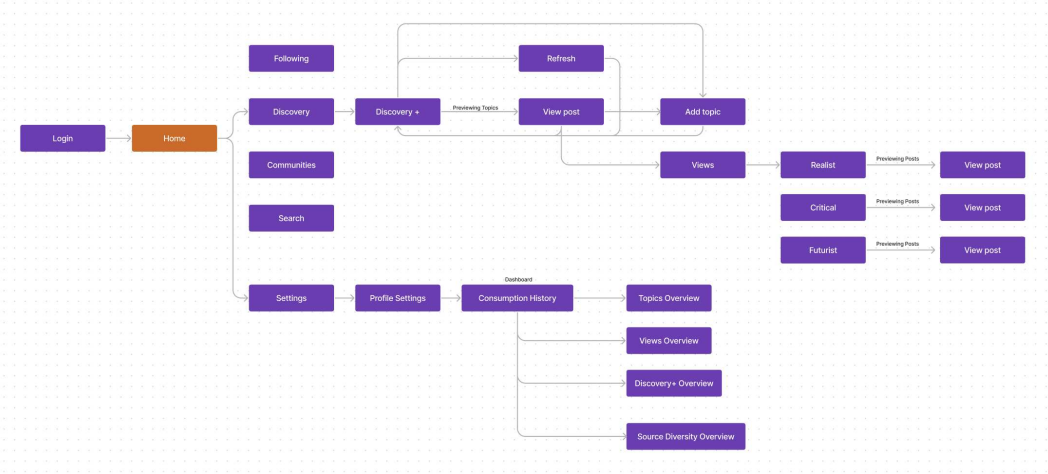


Figure 1: Prototype Userflow tying the 3 design features

choice into their feed. Users would gain more control over their social media experience as a result, becoming active participants in curating their own feed. The Discovery+ feature would get activated when users tapped on the button from Discovery, upon which a spinning carousel menu appeared as an overlay that contained different umbrella topics to choose from (Figure 2). In the prototype, the topics of artificial intelligence (AI) and sustainability were taken as examples. The topic of AI had been placed as a default selection where the overlay would consist of posts with their titles and sub-topic tags, which the user could choose to add to their feed. They could either add the sub-topics directly to their feed by clicking on the icon for "More" or they could read the post first and then add the sub-topic tags to their feed. Similarly, users could also choose sub-topic tags from the topic of sustainability to their feed.

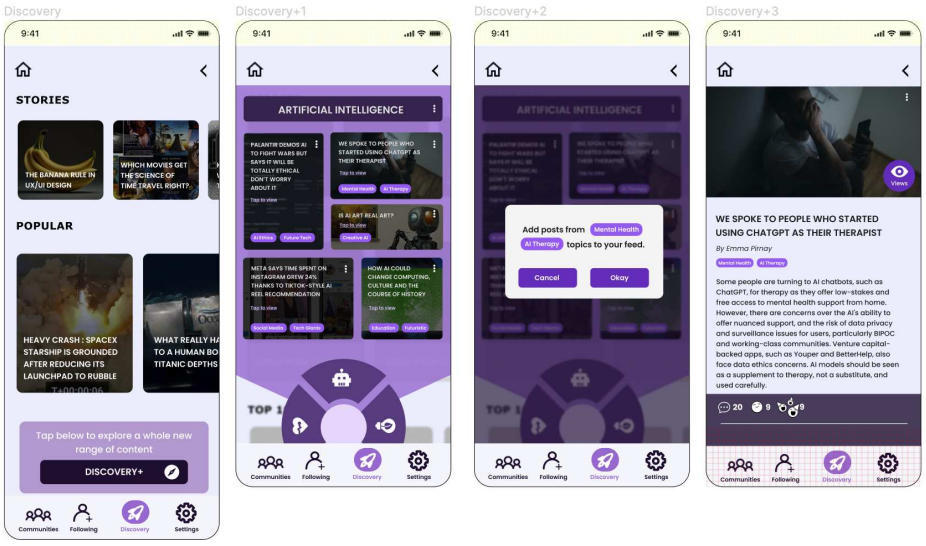


Figure 2: Wireframe prototype of Discovery+ feature

- **Views feature** This feature was designed in an attempt to expose users to diverse perspectives. Users could choose to access Views from any post, where they would be able to view related posts based on different viewpoints. Since COSE mainly dealt with content based on Science and Technology, the viewpoints used to categorize posts into were “Realist”, “Critical” and “Futurist” (Figure 3). Under each viewpoint, a body of posts with their titles would appear, where the user can view each post individually. Within the study, the "Realist" perspective demonstrated practicality by presenting related posts based on current circumstances or available resources, adopting a grounded approach. The "Critical" perspective, on the other hand, provided an analytical and objective approach, challenging assumptions and emphasizing ethical considerations. Furthermore, the "Futurist" viewpoint took a long-term perspective, highlighting posts that probed the potential impact of the discussed context on the future and imagined various possibilities. The inclusion of these perspectives was justified by the inherent presence of such perspectives within science and technology-based content [43, 48]. Moreover, Brossard argues the need to further evaluate the way scientific content is being perceived by audiences on the internet. She believes this could help scientists, educators, and policymakers communicate scientific content more effectively to the public [14]. Brossard also argues that scientific information consumed on social media is always accompanied by cues such as likes, comments and reactions that add to the reader’s interpretation of it [13]. This could mean that the usage of these cues influence users’ perceptions of the content presented to them. Perhaps, the introduction of various viewpoints in this context could also act as a cue to add to the user’s interpretation of the post.

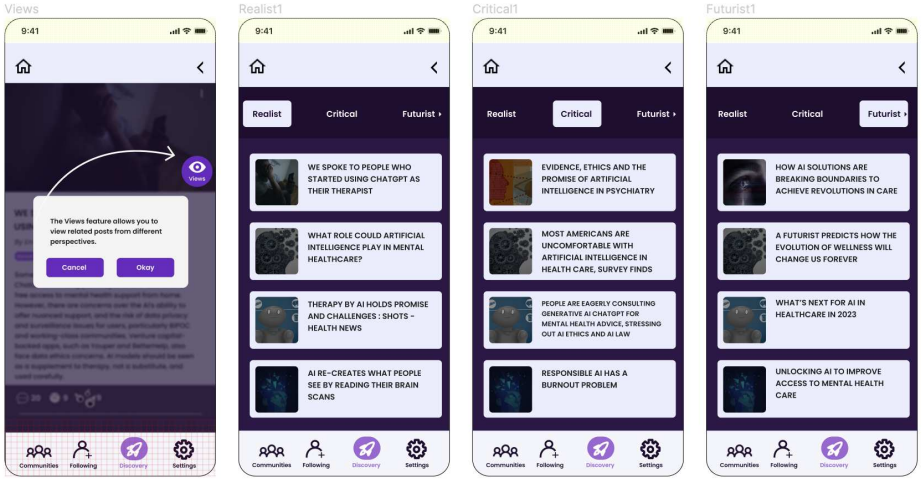


Figure 3: Wireframe prototype of Views feature

- **Consumption History feature Consumption History feature** The overarching goal of incorporating this feature was to establish a transparent environment that fostered trust between the platform and its users. The objective was to provide users with insights on the topics that they engaged with the most and the associated time spent on each by showing the user’s consumption history through graphical representations. The feature was divided into 4 parts that focus on 4 different aspects of content consumption which are Topics Overview, Discovery+ Overview, Views Overview and Source Diversity Overview (Figure 4). The Topics Overview allowed users to gain insights on their interests and whether they were spending too much/too little time on a specific topic. The Discovery+ Overview allowed users to gain insights on the frequency at which they wanted to refresh

their main feed and explore new topics, and what new topics they were showing interest towards. The Views Overview allowed users to gain insights on what kind of perspectives they were spending too much time on that could inform them of their own biases or if they were leaning towards a specific view. The Source Diversity Overview allowed users to gain insights on which communities and creators they were consuming a lot of.

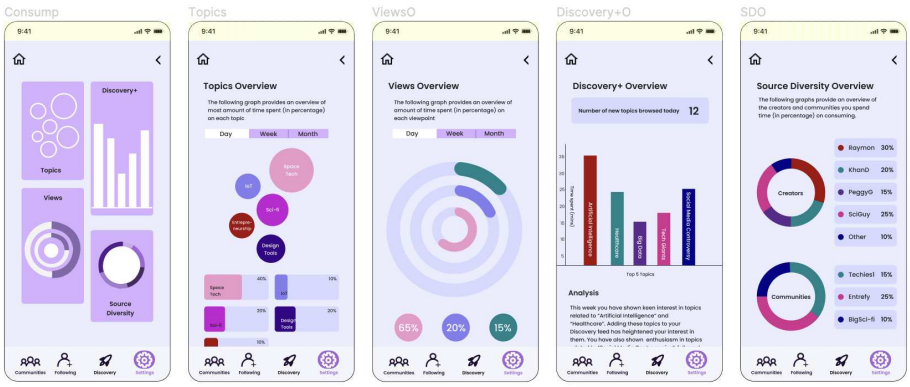


Figure 4: Wireframe prototype of Consumption History feature

4.2 Study 2 - Prototype Evaluation Results

4.2.1 Overview of Findings from Evaluation. The user evaluation explored the following research questions: 1. How does improved discoverability impact users' inclination to actively seek new content? 2. Does exposure to different views contribute to users' curiosity on the viewpoints itself on social media? 3. How much does users' access to their consumption history influence social media consumption behaviour? Based on these research questions, the evaluation procedure was accordingly planned and executed. Several themes emerged from the participants' responses on their browsing habits on science and technology content on social media in the pre-study questionnaire. These themes included how they interact with such content, the difficulties they have when seeking varied perspectives, shifts in their perspectives, and their perceptions of the impact of social media on their understanding of these subjects. The post-study questionnaire from the evaluation highlighted participants' views on interacting with the Discovery+ feature, Views feature and Consumption History feature on the COSE prototype. It particularly brought out their impressions on how these features could impact social media habits and their integration into social media platforms. The six participants who took part in the evaluation will be referred to as P1, P2, P3, P4, P5 and P6.

4.2.2 Findings from pre-study. The initial questionnaire highlighted many aspects of participants' views on interacting with science and technology-based content on social media such as the kind of platforms they seek such content from, what factors influence them to discover more content, challenges with seeking diverse content and perspectives and the overall influence of social media in their browsing habits. The following themes emerged: i) Variations in Platforms and Content on Social Media, ii) Active Discussion Outside of Social Media, iii) Influence of Compelling Headlines, Thumbnails, and Credible Sources, iv) Discovering Novel Content and Different Types of It, v) Lack of Authenticity, Trust and Diversity in Social Media and vi) Positive Influences on Social Media Habits. The themes identified included participants' content preferences, offline conversations on science and technology, the importance of catchy content, shifting perspectives through in-depth exploration, difficulties in discovering diverse

views, and positive impacts on social media usage. These themes helped in gaining a general overview of participants' reading behaviours, and better understanding of what inclines them towards science and tech content on social media.

4.2.3 Findings from post-study - Discovery+ feature. This section captured participants' impressions of the Discovery+ feature on the COSE prototype under which 2 themes emerged: i) User Control and ii) Content Diversity.

- **User Control** This theme emerged from 2 sub-themes: "Ease of navigation" and "Control over content curation". The level of intuitiveness and ease of navigating across this feature was seen in comments like "Yes it was clear. both the written instruction and the hints in the design guided me" (P1) and "The discovery+ feature was quite intuitive. It felt quite similar to Explore pages on social media apps" (P2). Participants also felt a sense of control while adding topics to their Discovery feed. This was noticed in comments like "Saving the tags into my home feed is also helpful to track the topics or an industry of interest" (P4) and "I feel this tagging system would make it easier for me to filter out as i can easily remember it if I am intentionally avoiding or seeking that sort of information" (P5).
- **Content Diversity** This theme emerged from 2 sub-themes: "Access to variety in content" and "Access to new information". The ability to find variety in topics of the content displayed was highlighted in participants' statements such as "It was a good point to see different topics and be able to choose from them" (P1) and "Without the feature it's always a hustle to explore the results obtained, ending up with the similar information in the post thus losing huge amount of time" (P3). Discoverability of new information was emphasized through statements such as "Since it acts as a central hub of all the new information on the platform, I believe it made it easier for me to engage with the app" (P2).

4.2.4 Findings From Post-study - Views feature. This section captured participants' impressions of the Views feature on the COSE prototype under which 2 themes emerged: i) Novel approach to communicating diverse viewpoints and ii) Exploring diverse perspectives and positive impacts.

- **Novel Approach to Communicating Diverse Viewpoints** This theme emerged from 2 sub-themes: "Initial misinterpretation of feature" and "Better ways of communicating Views". It summarizes participants' initial perceptions of the feature and how this could be communicated in a way that could have been more intuitive. This was seen in statements such as "Purpose was clear, but the definitions of the terms of the viewpoints was not clear" (P4), "Initially I thought it was the article view count because of its name but once I started using it, it become quite clear what was the intention of the feature" (P2), and "The viewpoints presented are good enough but could possibly have a well rounded approach in selecting or presenting the articles like having the contrasting article very next to it to help in understanding the topic in depth" (P3).
- **Exploring Diverse Perspectives and Positive Impacts** This theme emerged from 2 sub-themes: "Importance of exploring diverse views" and "Positive influence of the feature". It conveys participants' impressions that the Views feature has a beneficial impact and there is significant value in considering diverse perspectives of content through it. Participants thought that this feature offered them insights into the content they encountered, providing greater clarity on considering different perspectives, as conveyed through statements such as "I particularly enjoyed toggling between the three different directions, it gave me a perspective" (P2), "Here, it was nice to be given the opportunity to see all sides, whether I wanted to pick a side or not" (P6) and "It influences my way of understanding the topic, and helps me to speak about it in daily conversations with more agency" (P1).

4.2.5 Findings from Post-study - Consumption History feature. This section captured participants' impressions of the Consumption History feature on the COSE prototype under which 2 themes emerged: i) User autonomy and ii) Novel approach to content consumption.

- **User Autonomy** This theme emerged from 2 sub-themes: "Awareness over content consumption" and "Control over consumption". It reveals participants' reflections on the awareness that develops from their exposure to consumption patterns, and the level of control it affords them in shaping their subsequent content preferences. Through this feature participants' felt more in control over their content consumption patterns and a heightened awareness of their personal interests as seen in statements such as "It definitely works as a visual reminder or report of my consumption from which I can reflect on my actions" (P2), "It helped me to understand better the inclination i have towards receiving the information thus hinting me of the topics which actually keeps me engaged" (P3) and "Whenever the graph felt too high on one aspect, there was a tendency to learn more on the other side of it, to make sure there is a balance in content consumption about the topic" (P5).
- **Novel Approach to Content Consumption** This theme emerged from 2 sub-themes: i) Impact on future browsing patterns and ii) Visual data as informative elements. It conveys the impact of the Consumption History feature on future browsing patterns and participants' reactions to the usage of graphic data to convey information on their consumption. Participants reflected on this feature by highlighting its advantages in providing insights into their browsing behaviour that could enhance mindful browsing in the future, along with the clarity achieved through the usage of graphic elements which was noticed in statements such as "I think it would encourage more reading on different topics, views on the same subjects to balance my perspective and understanding others' view on each regard would be insightful to understand my consumption journey" (P5), "After exploring through the different quadrants and the very graphical representation of the browsing history it was super interesting specially with the screen time aware population" (P3) and "I was anticipating just a browsing history in the form of lists or something similar but to my surprise it was presented in a visual info-graphics format, which I think definitely added to the experience" (P2).

4.2.6 Overall User Experience with Design Features. From the participants' reactions and feedback with the design features in the prototype, it was evident that they provided a novel way of interacting with social media, in a way that allowed participants more control and agency over their decisions with content consumption and exposure to diversity in content and perspectives. The feedback on the Discovery+ feature was positive because it gave participants more control over the way of consuming content in terms of providing access to a variety of content with P1 mentioning that "saving the topic tags into my home feed was helpful to track the topics". The feature also received positive response in terms of providing content diversity with P6 mentioning that "it had a lot of topics so variety was evident". It also brought up the notion of providing access to new information with P2 suggesting that "it acted as a central hub of new information on the platform so it made it easier for me to engage with the app". Another reason was due to its ease of navigation where P4 mentioned that "it was easy to navigate and makes sense".

The feedback received on the Views feature was mixed, since participants initially found it difficult to understand what they viewpoints meant where P4 mentioned that "the purpose was clear but the definitions of the terms of the viewpoints was not clear" and P3 mentioning that "it was a bit difficult to figure out what it was leading to". In addition to this, alternative ways of communicating this feature was also brought up with P1 mentioned that "too many conflicting views may not be useful in some cases to make up your mind and decide" and P4 suggesting that "maybe the definition of the viewpoints could be changed to something that is easily understandable by everyone". However, some positive aspects of the feature were also highlighted with P6 mentioning "it was nice to be given the opportunity to see all sides". The importance of exploring diverse perspectives was also emphasized where P2 mentioned "its important for me to hear all sides of a story and then form my perspective".

The feedback received on the Consumption History was positive since it provided participants with awareness and control over their content consumption. Participants' inclination towards content and their interests were made clearer

to them where P6 mentioned “it was super cool to see what invokes my interest and how much time I spent on them” and P3 mentioned “It helped me to understand better the inclination I have towards receiving the information”. The sense of control it brought to content consumption was highlighted when P2 mentioned “I feel like I was in control of my consumption”. This feature also brought up the usage of visual data as a tool to convey information where P5 mentioned that “it felt like a data oriented journal” and P3 mentioned “it was super interesting with the screen time data”. Finally the impact of this feature on future browsing patterns was emphasized with P5 mentioning that “it would encourage more reading on different topics and views on the same subjects to balance my views”.

5 DISCUSSION

In the results mentioned above, a focus group study to understand users’ relationship with recommender systems and diversity on social media was conducted. After understanding their views, an evaluation of diverse recommender systems, in the form of design features on COSE’s social media platform was conducted. The results indicate that the three design features, Discovery+, Views and Consumption History had a positive impact on the exposure to diverse content and perspectives. This section will highlight the 3 hypothesis statements that answer the research question for the project, while reflecting on the results. This section reflects on the findings from the evaluation by drawing parallels from the background section. The reflections will be stated in the following themes:

- Contributions of design features for diverse social media experiences
- Ethical aspects of designing features to reduce the filter bubble effect

5.1 Contributions of design features for diverse social media experiences

The aim of the design features was to assess the impact of the affordances of viewing a diverse range of content and viewpoints on users and how that influenced their social media behaviours. Thus, this was tested by qualitatively understanding their user experiences and the potential future impacts on their behaviours.

For the Discovery+ feature, results indicate that participants received this feature as diverse and encouraging, fostering exploration of variety in content. They perceived it as empowering, allowing them to control content curation and as a means to discover content outside their personalized interests. Based on the first hypothesis (H1), this feature showed positive traits in exposing users to a varied range of new information on their feed on a regular basis and can positively influence their inclination to actively explore and find new content. The Discovery+ feature corresponds with users’ needs to actively explore new content while maintaining access to personalized content based on their needs as revealed during the focus group sessions. The goal of this strategy was to find a balance between personalisation and content diversity which was identified in the focus group sessions. Therefore, the findings from testing this feature contribute to discoverability in recommender systems, building on Reviglio’s idea of a diverse social media to face filter bubbles, that, creating an environment that allows users to encounter diverse content naturally, instead of focusing only on curation of diverse content, can enhance discoverability [64]. Portenroy et al’s study revealed that participants had found some newly discovered content that was broken down into shorter bits of information to complement topics related to a known author’s research [61]. This meant that serendipitous content recommended to participants in smaller chunks made them more interesting for participants to explore. In this project, serendipitous content in the form of new topics, along with sub-topic tags under each post made it easier for participants to explore new content and understand their own interests. It also allowed them more control to keep track of their newfound interests by adding those sub-topic tags to their feed. This means that categorizing new content through tags would allow users to gain a comprehensive understanding of that topic, giving them more control over content consumption and curation in their feed. However, it should also be noted that discoverability or “serendipity” as stated in Kotkov et al.’s research suggests

that studying this concept through recommender systems can be quite challenging due to its emotional dimension that can be difficult to capture and it is a complex concept that consists of various other definitions [44]. Further research could look into understanding a clearer definition and how that could reflect on recommender system designs.

Coming to the Views feature, participants found the feature to lack intuitiveness while first interacting with it, and their interactions highlighted their impressions of its novelty. However, upon further interaction, their reactions emphasized the importance of exploring diverse viewpoints and reflected positively on it. Users felt that the feature allowed them to gain a comprehensive overview of the post they were reading, motivating them to learn more about the post. Thus, based on the second hypothesis (H2), this feature showed positive traits of allowing users to access multiple viewpoints within a post which significantly influences user behaviour, leading to a greater inclination for actively seeking and engaging with diverse content, which supports the second hypothesis. In line with Chhabra et al. [16] and Park et al.'s [60] research, users were motivated to explore diverse perspectives of a single topic when stories were presented in meaningful clusters. In this feature, stories were clustered based on the viewpoint which acted as a line of thought and thus provided users with more information. This allowed them to explore diverse views. Nevertheless, participants did not find the feature to be intuitive initially and needed some guidance to carry out the task, after which they provided suggestions for better ways of communicating the feature that could be easier to comprehend. It should also be noted that the viewpoints presented in this study were based on content catering to topics related to science and technology, and not based on the political spectrum as seen in numerous studies. The novelty of the viewpoint terms alone weren't intuitive as these terms haven't been explored enough to be comprehended easily by the general public. This led to a significant lack of intuitiveness initially when participants first encountered the View feature. Nonetheless, this does contribute to research in diverse recommender system designs. Further research is required in understanding and evaluating effective ways of communicating this feature, along with exploring viewpoint terms in this context that would be easily comprehensible by users.

For the Consumption History feature, results indicate that participants expressed the sense of awareness and control they felt over their content consumption when exposed to it. They felt empowered with the knowledge received on the way they were consuming content through visual data. By being transparent about users' content consumption patterns, the feature aims to develop a sense of trust between users and the platform, as supported with Sinha et al.'s research [68]. However, it should be noted that through this feature, transparency is being brought about by informing users of their consumption patterns and not through directly recommending content as cited in Sinha et al.'s research. Nevertheless, it contributes to transparency in reducing the filter bubble effect. Further research is needed in investigating the implications of a feature that displays users' consumption history and its relationship to transparency in recommender systems. The usage of graphical data to communicate users' consumption patterns proved to be engaging. The feature ultimately helped participants understand their own consumption journey and further encouraged reading diverse content and perspectives. Based on the third hypothesis (H3), this feature shows positive inclinations in providing users with information about their content consumption habits, which could thus allow platforms to enhance transparency and promote users' awareness regarding their own behaviour. This feature corresponds with the need to inform users of a balanced mix of stories from different viewpoints as studied in Munson et al.'s research [52]. Hence, the findings of this feature build on existing evidence of usage of visualization to communicate content exposure on social media.

5.2 Ethical Aspects of Designing Features to Reduce the Filter Bubble Effect

The features designed for reducing the filter bubble effect were addressed by incorporating key design principles such as diversity, serendipity, transparency and nudge as an important design mechanism, as mentioned in the background. These were considered as necessary to help reduce the filter bubble effect. In this project, autonomy [41] was seen

as a core value that could contribute to diverse social media experiences. The design principles mentioned aimed to contribute to it by allowing users to take control of their social media experiences with respect to content navigation, exposure to diverse content, perspectives, and consumption patterns. Their objective was to promote individual empowerment, responsible engagement, user satisfaction and privacy and security. The features designed aim to also address the negative impacts of social media on autonomy that have been emphasized in academia [66]. However, there are ethical aspects to consider while using the design principles to promote autonomy. Combining diversity with serendipity in the design features could lead to users being drawn towards serendipitous content that meets their preferences rather than engaging with content which is genuinely diverse that presents opposing views [65]. This could further lead to users' selective exposure of content and the formation of echo chambers that could impact societal mindsets. This might alter the extent of control users possess over their social media experiences. Due to these concerns, the design features were designed and evaluated to ensure a seamless blend of personalization and diversity in content consumption on the design prototype. Transparency was also identified as a crucial principle to guide the Consumption History feature. Since transparency is necessary to develop trust and confidence amongst users and social media platforms [68], this feature was designed to inform users of their consumption patterns, giving them more control over their experiences on social media. This involves handling of users' data, which must be executed in a manner that respects their privacy and refrains from exploiting it. Participants reacted positively to the feature, where they felt in control of their experience, curious to understand their interests and explore diverse content. The usage of nudge as a design mechanism to reduce the filter bubble effect has received mixed reactions within the research community. On one hand, it has contributed towards the filter bubble by promoting personalization of content on social media [62], while on the other it has shown to promote a diverse exposure of content [52]. Acquisti et al. discuss nudging through information and presentation that could help users gain awareness over content consumption [6], which in this project was addressed through the design features. Since nudging offers a means to influence user behaviour positively, it is crucial to recognize the importance of identifying those with the authority to do so, and enabling users to opt out of it if they wish to so that their autonomy isn't compromised. In this project, the mechanism brought about a positive response from participants, where revealing their consumption patterns appeared to nudged them towards seeking diverse content and the presentation of diverse viewpoints in the Views feature nudged them to explore diverse perspectives. The Consumption History feature served as a nudging mechanism in the way it conveyed information about users' consumption habits, which encouraged mindfulness and self-awareness about their interests and empowered them to choose the information they consume. The Discovery+ and Views feature, through their manner of conveying new information, by allowing accessibility to new topics and different viewpoints also nudged users to explore diverse content and perspectives, outside of their bubble without imposing it on them.

6 LIMITATIONS FUTURE WORK

The limitations of this project mainly deal with the methodology of the evaluation phase and some with the designing of the prototype. The number of participants who took part in the evaluation were limited, hence the insights generated from this are based on the sessions of the 6 participants who took part in it. All the 6 participants were students from KTH, most of whom were Indian (with 1 international participant), hence a further exploration of experiences of participants from other countries might provide more learnings that are based on a global worldview of social media. This could also shed light on the social and cultural differences that impact people's perspectives on social media.

Another important limitation was identified during the designing of the prototype that involved the regulation of the flow of the prototype and participants' evaluation tasks. The participants couldn't enter their preferences to the prototype due to the limiting functionalities of Figma that was used to design it. Participants were thus instructed to perform the tasks as though they were native users of the app, interacting with the design features for the first time through the

prototype. Future work could focus on gaining a better understanding of how the design principles could be applied into reducing the filter bubble effect and how the design features could be improved for a better user experience. As mentioned in the results, it took participants some time to understand it, since they found the Views feature unfamiliar and difficult to understand its intention. Therefore, future work should look at investigating what viewpoint terms would be suitable for segregating content that is science and technology based in social media. This is to improve the way people perceive such information and further understand how the viewpoints act as cues that influence people's perception of content. Moreover, participants also suggested better ways of communicating this feature, that could be more user friendly as mentioned in Appendix C from the thematic analysis of the evaluation, but these aren't critical changes. Further research is required into the broader and long-term implications of the Consumption History feature on users, to gain a comprehensive understanding of its potential future effects on users' social media behaviours. Further investigation on serendipity, transparency and diversity is required to understand how these principles could be effectively communicated through the design features, and the role they play in reducing the filter bubble effect.

7 CONCLUSION

The aim of this thesis was to investigate the following research question: How can the design of recommender systems effectively promote diversity to reduce the filter bubble effect on social media? This involved usage of a Research through Design (RtD) approach [69], which first led to studying the impacts of filter bubbles, the role of recommender systems in contributing to them and how interventions in recommender system designs aimed to reduce/break them. After gaining further insights from focus groups sessions with target users, it resulted in the design of 3 features on the COSE platform's prototype, Discovery+, Views and Consumption History. These features aimed towards promoting diverse social media experiences and thus reduce the filter bubble effect. The Discovery+ feature showed positive traits in exposing users to a varied range of new information on their feed on a regular basis that positively influenced their inclination to actively explore and find new content, thus possibly suggesting the usage of diversity [27] and serendipity [64] as important design values. The Consumption History feature showed positive characteristics of transparency [68] where it appeared to nudge users [72] with information about their content consumption habits which helped in improving their sense of awareness of their own behaviours. The Views feature exhibited favourable characteristics by allowing users to access multiple viewpoints within a post which appreared to significantly influence user behaviour, leading to a greater inclination for actively seeking and engaging with diverse content. However, the Views feature lacked intuitiveness which proved difficult for users to comprehend the feature initially. Overall, this thesis contributed to the field of filter bubbles and recommender systems designs on social media by proposing design ideas to guide the diverse exploration of content and perspectives on social media and improve user experiences. The diverse features also have implications for social sustainability, as they highlight the need for diverse and ethical social media experiences, where users have access to information that can empower them, enhance dialogue amongst individuals and enable them to make informed decisions that could ensure social cohesion. They address the perils of social media such as polarization, including issues like polarization, the reinforcement of echo chambers, the obstruction of constructive dialogue, and the impediment of fostering a global perspective on various topics.

8 ACKNOWLEDGMENTS

Thank you to my supervisors, Joo Young Park and Carlos Tirado for their valuable guidance and my examiner, Anders Hedman for valuable advice and support throughout the project. Thank you to Manuel Fraile from COSE for this incredible opportunity. I am also grateful to all the participants who took part in my studies, as their insights and feedback helped shape the project. My biggest thanks to my family and friends for being my biggest source of strength and cheering me on throughout. Final thanks to myself for all the dedication and hard work invested in this journey.

REFERENCES

[1] [n. d.]. COSE. https://www.cose.community/

[2] [n. d.]. Free Online Form Builder & Form Creator | Jotform. https://www.jotform.com/

[3] [n. d.]. Otter.ai - Voice Meeting Notes & Real-time Transcription. https://otter.ai/

[4] [n. d.]. Wonkosphere. https://wonkosphere.com/

[5] 2019. AllSides | Balanced news via media bias ratings for an unbiased news perspective. https://www.allsides.com/unbiased-balanced-news

[6] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, Yang Wang, and Shomir Wilson. 2017. Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online. *Comput. Surveys* 50, 3 (Aug. 2017), 44:1–44:41. https://doi.org/10.1145/3054926

[7] Vivek Anand, Matthew Yang, and Zhanzhan Zhao. 2022. Mitigating Filter Bubbles within Deep Recommender Systems. https://doi.org/10.48550/arXiv.2209.08180 arXiv:2209.08180 [cs].

[8] Eytan Bakshy, Solomon Messing, and Lada A. Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (June 2015), 1130–1132. https://doi.org/10.1126/science.aaa1160 Publisher: American Association for the Advancement of Science.

[9] Pabloe Barberá, John T. Jost, Jonathan Nagler, Joshua A. Tucker, and Richard Bonneau. 2015. Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? - Pablo Barberá, John T. Jost, Jonathan Nagler, Joshua A. Tucker, Richard Bonneau, 2015. https://journals.sagepub.com/doi/full/10.1177/0956797615594620?casa_token=V6sk6HxXKXEAAAAA%3AqAqlCXLGWj4BHJPoOemgKtkifwSEIgsQ-M8Wda78kcWTxlUx9A6HrmySCQJSwHXuKmZ6yJGpYzoX3A

[10] V. E. Bozdag. 2015. Bursting the Filter Bubble: Democracy, Design, and Ethics. (2015). https://repository.tudelft.nl/islandora/object/uuid%3A87bde0a2-c391-4c77-8457-97cba93abf45

[11] V. E. Bozdag and J. F. C. Timmermans. 2001. Values in the filter bubble Ethics of Personalization Algorithms in Cloud Computing. *1st International Workshop on Values in Design – Building Bridges between RE, HCI and Ethics, Lisbon, Portugal, 6 September 2011* (2001). https://repository.tudelft.nl/islandora/object/uuid%3A5988617e-de91-4afa-9bc6-a820c41a47d1

[12] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (Jan. 2006), 77–101. https://doi.org/10.1191/1478088706qp063oa Publisher: Routledge _eprint: https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa.

[13] Dominique Brossard. 2013. New media landscapes and the science information consumer. *Proceedings of the National Academy of Sciences* 110, supplement_3 (Aug. 2013), 14096–14101. https://doi.org/10.1073/pnas.1212744110 Publisher: Proceedings of the National Academy of Sciences.

[14] Dominique Brossard and Bruce Lewenstein. 2010. A Critical Appraisal of Models of Public Understanding of Science: Using Practice to Inform Theory. In *Communicating Science: New Agendas in Communication*. 11–39. https://doi.org/10.4324/9780203867631 Journal Abbreviation: Communicating Science: New Agendas in Communication.

[15] Axel Bruns. 2019. Filter bubble. *Internet Policy Review* 8, 4 (Nov. 2019). https://policyreview.info/concepts/filter-bubble

[16] Sidharth Chhabra and Paul Resnick. 2012. CubeThat: news article recommender. In *Proceedings of the sixth ACM conference on Recommender systems*. ACM, Dublin Ireland, 295–296. 978-1-4503-1270-7 https://doi.org/10.1145/2365952.2366020

[17] Henriette Cramer, Vanessa Evers, Satyan Ramlal, Maarten van Someren, Lloyd Rutledge, Natalia Stash, Lora Aroyo, and Bob Wielinga. 2008. The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction* 18, 5 (Nov. 2008), 455–496. https://doi.org/10.1007/s11257-008-9051-3

[18] Pranav Dandekar, Ashish Goel, and David T. Lee. 2013. Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences* 110, 15 (April 2013), 5791–5796. https://doi.org/10.1073/pnas.1217220110 Publisher: Proceedings of the National Academy of Sciences.

[19] Theresa Davidson and Lee Farquhar. 2020. Prejudice and Social Media: Attitudes Toward Illegal Immigrants, Refugees, and Transgender People. In *Gender, Sexuality and Race in the Digital Age*, D. Nicole Farris, D'Lane R. Compton, and Andrea P. Herrera (Eds.). Springer International Publishing, Cham, 151–167. 978-3-030-29855-5 https://doi.org/10.1007/978-3-030-29855-5_9

[20] Michael X. Delli Carpini. 2000. In search of the informed citizen: What Americans know about politics and why it matters. *The Communication Review* 4, 1 (Jan. 2000), 129–164. https://doi.org/10.1080/10714420009359466 Publisher: Routledge _eprint: https://doi.org/10.1080/10714420009359466.

[21] Gali Einav, Ofir Allen, Tamar Gur, Yossi Maaravi, and Daniel Ravner. 2022. Bursting filter bubbles in a digital age: Opening minds and reducing opinion polarization through digital platforms. *Technology in Society* 71 (Nov. 2022), 102136. https://doi.org/10.1016/j.techsoc.2022.102136

[22] Seth Flaxman, Sharad Goel, and Justin M. Rao. 2016. Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly* 80, S1 (Jan. 2016), 298–320. https://doi.org/10.1093/poq/nfw006

[23] Batya Friedman. 1996. Value-sensitive design. *Interactions* 3, 6 (Dec. 1996), 16–23. https://doi.org/10.1145/242485.242493

[24] Nabeel Gillani, Ann Yuan, Martin Saveski, Soroush Vosoughi, and Deb Roy. 2018. Me, My Echo Chamber, and I: Introspection on Social Media Polarization. In *Proceedings of the 2018 World Wide Web Conference (WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 823–831. 978-1-4503-5639-8 https://doi.org/10.1145/3178876.3186130

[25] Lewis R. Goldberg. 1990. An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology* 59, 6 (1990), 1216–1229. https://doi.org/10.1037/0022-3514.59.6.1216 Place: US Publisher: American Psychological Association.

[26] Thomas Gottron and Felix Schwagereit. 2016. The Impact of the Filter Bubble – A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities. https://doi.org/10.48550/arXiv.1612.06551 arXiv:1612.06551 [cs].

[27] Natali Helberger. 2011. Diversity by Design. *Journal of Information Policy* 1 (Jan. 2011), 441–469. https://doi.org/10.5325/jinfopoli.1.2011.0441

[28] Natali Helberger, Kari Karppinen, and Lucia D'Acunto. 2018. Exposure diversity as a design principle for recommender systems. *Information, Communication & Society* 21, 2 (Feb. 2018), 191–207. https://doi.org/10.1080/1369118X.2016.1271900

[29] Benjamin Herm-Morris. 2022. Education and the dislike society: The impossibility of learning in filter bubbles. *Educational Philosophy and Theory* 54, 5 (April 2022), 502–511. https://doi.org/10.1080/00131857.2021.1935233 Publisher: Routledge _eprint: https://doi.org/10.1080/00131857.2021.1935233.

[30] Harald Holone. 2016. The filter bubble and its effect on online personal health information. *Croatian Medical Journal* 57, 3 (June 2016), 298–301. https://doi.org/10.3325/cmj.2016.57.298

[31] Kartik Hosanagar, Daniel Fleder, Dokyun Lee, and Andreas Buja. 2014. Will the Global Village Fracture Into Tribes? Recommender Systems and Their Effects on Consumer Fragmentation. *Management Science* 60, 4 (April 2014), 805–823. https://doi.org/10.1287/mnsc.2013.1808 Publisher: INFORMS.

[32] Bernd Huber, Yixue Wang, Jean Garcia-Gathright, and Jenn Thom. 2022. Explaining Podcast Recommendations To Users with Content Diversity Labels. (2022).

[33] Juha V. A. Itkonen. 2015. Social ties and concern for global warming. *Climatic Change* 132, 2 (Sept. 2015), 173–192. https://doi.org/10.1007/s10584-015-1424-0

[34] Youngseung Jeon, Bogoan Kim, Aiping Xiong, DONGWON LEE, and Kyungsik Han. 2021. ChamberBreaker: Mitigating the Echo Chamber Effect and Supporting Information Hygiene through a Gamified Inoculation System. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 472:1–472:26. https://doi.org/10.1145/3479859

[35] Ray Jiang, Silvia Chiappa, Tor Lattimore, András György, and Pushmeet Kohli. 2019. Degenerate Feedback Loops in Recommender Systems. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. 383–390. https://doi.org/10.1145/3306618.3314288 arXiv:1902.10730 [cs, stat].

[36] Michael Jugovac and Dietmar Jannach. 2017. Interacting with Recommenders—Overview and Research Directions. *ACM Transactions on Interactive Intelligent Systems* 7, 3 (Sept. 2017), 10:1–10:46. https://doi.org/10.1145/3001837

[37] Jonas Kaiser and Adrian Rauchfleisch. 2020. Birds of a Feather Get Recommended Together: Algorithmic Homophily in YouTube's Channel Recommendations in the United States and Germany. *Social Media + Society* 6, 4 (Oct. 2020), 205630512096991. https://doi.org/10.1177/2056305120969914

[38] Marius Kaminskas and Derek Bridge. 2016. Diversity, Serendipity, Novelty, and Coverage: A Survey and Empirical Analysis of Beyond-Accuracy Objectives in Recommender Systems. *ACM Transactions on Interactive Intelligent Systems* 7, 1 (Dec. 2016), 2:1–2:42. https://doi.org/10.1145/2926720

[39] Daniel Kilvington. 2021. The virtual stages of hate: Using Goffman's work to conceptualise the motivations for online hate. *Media, Culture & Society* 43, 2 (March 2021), 256–272. https://doi.org/10.1177/0163443720972318 Publisher: SAGE Publications Ltd.

[40] Brent Kitchens, Steve L. Johnson, and Peter Gray. 2020. Understanding Echo Chambers and Filter Bubbles: The Impact of Social Media on Diversification and Partisan Shifts in News Consumption. *MIS Quarterly* 44, 4 (Dec. 2020), 1619–1649. https://doi.org/10.25300/MISQ/2020/16371

[41] Michael Klenk and Jeff Hancock. 2019. Autonomy and Online Manipulation. *Internet Policy Review* 1 (2019), 1–11. https://philarchive.org/rec/KLEAAO-3

[42] Joel Klinger, Juan Mateos-Garcia, and Konstantinos Stathoulopoulos. 2022. A narrowing of AI research? https://doi.org/10.48550/arXiv.2009.10385 arXiv:2009.10385 [cs].

[43] Sabine Koch. 2010. Healthy ageing supported by technology – a cross-disciplinary research challenge. *Informatics for Health and Social Care* 35, 3-4 (Sept. 2010), 81–91. https://doi.org/10.3109/17538157.2010.528646 Publisher: Taylor & Francis _eprint: https://doi.org/10.3109/17538157.2010.528646.

[44] Denis Kotkov, Jari Veijalainen, and Shuaiqiang Wang. 2016. Challenges of Serendipity in Recommender Systems. SCITEPRESS. https://doi.org/10.5220/0005879802510256 Accepted: 2016-05-13T06:18:57Z.

[45] Akos Lada, Meihong Wang, and Tak Yan. 2021. How machine learning powers Facebook's News Feed ranking algorithm. https://engineering.fb.com/2021/01/26/ml-applications/news-feed-ranking/

[46] Jonathan Lazar, Jinjuan Feng, and Harry Hochheiser. 2017. Interviews and focus groups. 187–228. 978-0-12-805390-4 https://doi.org/10.1016/B978-0-12-805390-4.00008-X

[47] Q. Vera Liao and Wai-Tat Fu. 2013. Beyond the filter bubble: interactive effects of perceived threat and topic involvement on selective exposure to information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2359–2368. 978-1-4503-1899-0 https://doi.org/10.1145/2470654.2481326

[48] Anna MacLeod, Paula Cameron, Victoria Luong, Robin Parker, Vinson Li, and Cora-Lynn Munroe-Lynds. 2023. Questions of Well-Being and Inclusion in Online Undergraduate Medical Education During COVID-19: A 2-Phased Focused Scoping Review. *Academic Medicine* 98, 4 (April 2023), 521. https://doi.org/10.1097/ACM.0000000000005119

[49] Kirsten Medhurst and Rashmi Sinha. 2002. Interaction Design for Recommender Systems. *Presentation at the International Conference on Designing Interactive Systems, London, June* (March 2002).

[50] Tomasz Miaskiewicz and Kenneth A. Kozar. 2011. Personas and user-centered design: How can personas benefit product design processes? *Design Studies* 32, 5 (Sept. 2011), 417–430. https://doi.org/10.1016/j.destud.2011.03.003

[51] Lien Michiels, Jens Leysen, Annelien Smets, and Bart Goethals. 2022. What Are Filter Bubbles Really? A Review of the Conceptual and Empirical Work. In *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '22 Adjunct)*. Association

for Computing Machinery, New York, NY, USA, 274–279. 978-1-4503-9232-7 https://doi.org/10.1145/3511047.3538028

[52] Sean Munson, Stephanie Lee, and Paul Resnick. 2013. Encouraging Reading of Diverse Political Viewpoints with a Browser Widget. *Proceedings of the International AAAI Conference on Web and Social Media* 7, 1 (2013), 419–428. https://doi.org/10.1609/icwsm.v7i1.14429 Number: 1.

[53] Sean A Munson and Paul Resnick. 2010. Presenting diverse political opinions: how and how much. (2010).

[54] Sean A. Munson and Paul Resnick. 2010. Presenting diverse political opinions: how and how much. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. Association for Computing Machinery, New York, NY, USA, 1457–1466. 978-1-60558-929-9 https://doi.org/10.1145/1753326.1753543

[55] Sayooran Nagulendra and Julita Vassileva. 2014. Understanding and controlling the filter bubble through interactive visualization: a user study. In *Proceedings of the 25th ACM conference on Hypertext and social media*. ACM, Santiago Chile, 107–115. 978-1-4503-2954-5 https://doi.org/10.1145/2631775.2631811

[56] Matti Nelimarkka, Salla-Maaria Laaksonen, and Bryan Semaan. 2018. Social Media Is Polarized, Social Media Is Polarized: Towards a New Design Agenda for Mitigating Polarization. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, New York, NY, USA, 957–970. 978-1-4503-5198-0 https://doi.org/10.1145/3196709.3196764

[57] Tien T. Nguyen, Pik-Mai Hui, F. Maxwell Harper, Loren Terveen, and Joseph A. Konstan. 2014. Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of the 23rd international conference on World wide web*. ACM, Seoul Korea, 677–686. 978-1-4503-2744-2 https://doi.org/10.1145/2566486.2568012

[58] Raymond S. Nickerson. 1998. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology* 2, 2 (June 1998), 175–220. https://doi.org/10.1037/1089-2680.2.2.175 Publisher: SAGE Publications Inc.

[59] Eli Pariser. 2011. *The Filter Bubble: What The Internet Is Hiding From You*. Penguin UK. 978-0-14-196992-3

[60] Souneil Park, Seungwoo Kang, Sangyoung Chung, and Junehwa Song. 2009. NewsCube: delivering multiple aspects of news to mitigate media bias. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Boston MA USA, 443–452. 978-1-60558-246-7 https://doi.org/10.1145/1518701.1518772

[61] Jason Portenoy, Marissa Radensky, Jevin D West, Eric Horvitz, Daniel S Weld, and Tom Hope. 2022. Bursting Scientific Filter Bubbles: Boosting Innovation via Novel Author Discovery. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–13. 978-1-4503-9157-3 https://doi.org/10.1145/3491102.3501905

[62] Markus Prior. 2013. Media and Political Polarization. *Annual Review of Political Science* 16, 1 (2013), 101–127. https://doi.org/10.1146/annurev-polisci-100711-135242 _eprint: https://doi.org/10.1146/annurev-polisci-100711-135242.

[63] Pearl Pu, Li Chen, and Rong Hu. 2011. A user-centric evaluation framework for recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems (RecSys '11)*. Association for Computing Machinery, New York, NY, USA, 157–164. 978-1-4503-0683-6 https://doi.org/10.1145/2043932.2043962

[64] Urbano Reviglio. 2017. Serendipity by Design? How to Turn from Diversity Exposure to Diversity Experience to Face Filter Bubbles in Social Media. In *Internet Science (Lecture Notes in Computer Science)*, Ioannis Kompatsiaris, Jonathan Cave, Anna Satsiou, Georg Carle, Antonella Passani, Efstratios Kontopoulos, Sotiris Diplaris, and Donald McMillan (Eds.). Springer International Publishing, Cham, 281–300. 978-3-319-70284-1 https://doi.org/10.1007/978-3-319-70284-1_22

[65] Urbano Reviglio. 2019. Serendipity as an emerging design principle of the infosphere: challenges and opportunities. *Ethics and Information Technology* 21, 2 (June 2019), 151–166. https://doi.org/10.1007/s10676-018-9496-y

[66] Siavosh Sahebi and Paul Formosa. 2022. Social Media and its Negative Impacts on Autonomy. *Philosophy & Technology* 35, 3 (July 2022), 70. https://doi.org/10.1007/s13347-022-00567-7

[67] Olivier Serrat. 2017. The SCAMPER Technique. In *Knowledge Solutions: Tools, Methods, and Approaches to Drive Organizational Performance*, Olivier Serrat (Ed.). Springer, Singapore, 311–314. 978-981-10-0983-9 https://doi.org/10.1007/978-981-10-0983-9_33

[68] Rashmi Sinha and Kirsten Swearingen. 2002. The role of transparency in recommender systems. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems (CHI EA '02)*. Association for Computing Machinery, New York, NY, USA, 830–831. 978-1-58113-454-4 https://doi.org/10.1145/506443.506619

[69] Pieter Jan Stappers and Elisa Giaccardi. 2017. Research through Design. In *The Encyclopedia of Human-Computer Interaction*, M. Soegaard and R. Friis-Dam (Eds.). The Interaction Design Foundation, 1–94. https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/research-through-design

[70] Jonathan Stray, Alon Halevy, Parisa Assar, Dylan Hadfield-Menell, Craig Boutilier, Amar Ashar, Lex Beattie, Michael Ekstrand, Claire Leibowicz, Connie Moon Sehat, Sara Johansen, Lianne Kerlin, David Vickrey, Spandana Singh, Sanne Vrijenhoek, Amy Zhang, McKane Andrus, Natali Helberger, Polina Proutskova, Tanushree Mitra, and Nina Vasan. 2022. Building Human Values into Recommender Systems: An Interdisciplinary Synthesis. https://doi.org/10.48550/arXiv.2207.10192 arXiv:2207.10192 [cs].

[71] Cass R. Sunstein. 2009. *Going to Extremes: How Like Minds Unite and Divide*. Oxford University Press. 978-0-19-979314-3 Google-Books-ID: jEWplxVkEEEC.

[72] Calum Thornhill, Quentin Meeus, Jeroen Peperkamp, and Bettina Berendt. 2019. A Digital Nudge to Counter Confirmation Bias. *Frontiers in Big Data* 2 (2019). https://www.frontiersin.org/articles/10.3389/fdata.2019.00011

[73] Wenjie Wang, Fuli Feng, Liqiang Nie, and Tat-Seng Chua. 2022. User-controllable Recommendation Against Filter Bubbles. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, Madrid Spain, 1251–1261. 978-1-4503-8732-3 https://doi.org/10.1145/3477495.3532075

[74] T. van de Wiel. 2021. *Sorry to burst your bubble! Developing an Educational Application to Raise Awareness About Filter Bubbles Among Teenagers*. Master's thesis. https://studenttheses.uu.nl/handle/20.500.12932/41324 Accepted: 2021-08-29T18:00:11Z.

[75] Yunfei Xing, Xiwei Wang, Chengcheng Qiu, Yueqi Li, and Wu He. 2022. Research on opinion polarization by big data analytics capabilities in online social networks. *Technology in Society* 68 (Feb. 2022), 101902. https://doi.org/10.1016/j.techsoc.2022.101902

[76] Sarita Yardi and danah boyd. 2010. Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter. *Bulletin of Science, Technology & Society* 30 (Oct. 2010), 316–327. https://doi.org/10.1177/0270467610380011

[77] Xinzhi Zhang, Wan-Ying Lin, and William H. Dutton. 2022. The Political Consequences of Online Disagreement: The Filtering of Communication Networks in a Polarized Political Context. *Social Media + Society* 8, 3 (July 2022), 20563051221114391. https://doi.org/10.1177/20563051221114391 Publisher: SAGE Publications Ltd.

[78] Rong Zhao. 2021. Flip: An Interaction Design Project for Helping People to Break Echo Chamber - ProQuest. https://www.proquest.com/openview/ba089d480c9e078f19670a681a7562d0/1?pq-origsite=gscholar&cbl=18750&diss=y

[79] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. Association for Computing Machinery, New York, NY, USA, 493–502. 978-1-59593-593-9 https://doi.org/10.1145/1240624.1240704

A USER RESEARCH ELEMENTS

A.1 Ideas from Focus Group Session

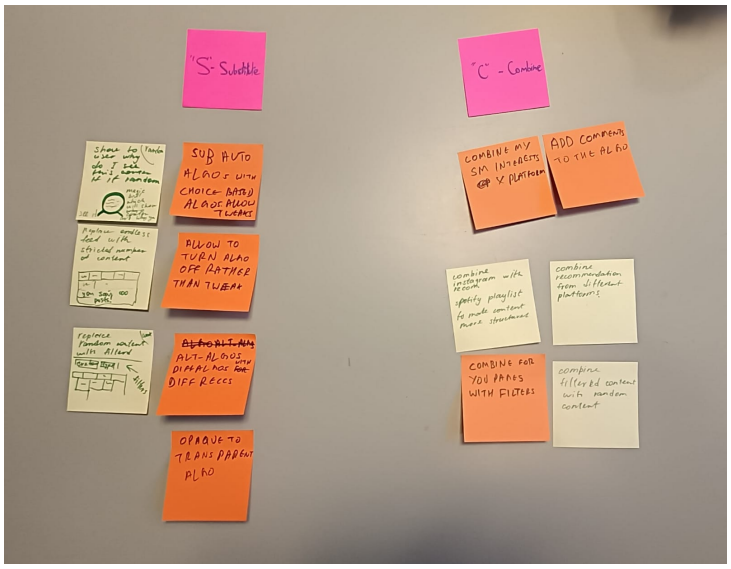


Figure 5: Ideas from Session 1

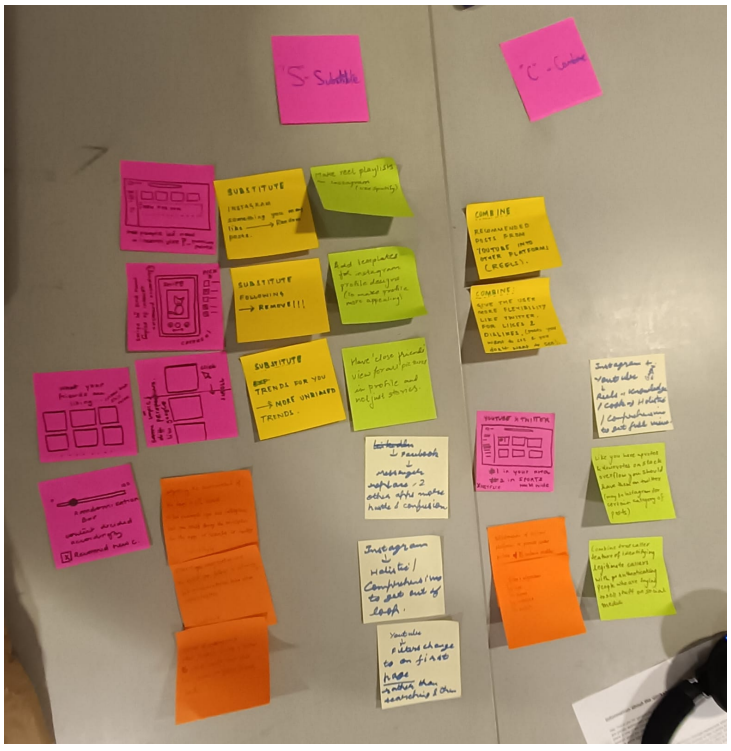


Figure 6: Ideas from Session 2

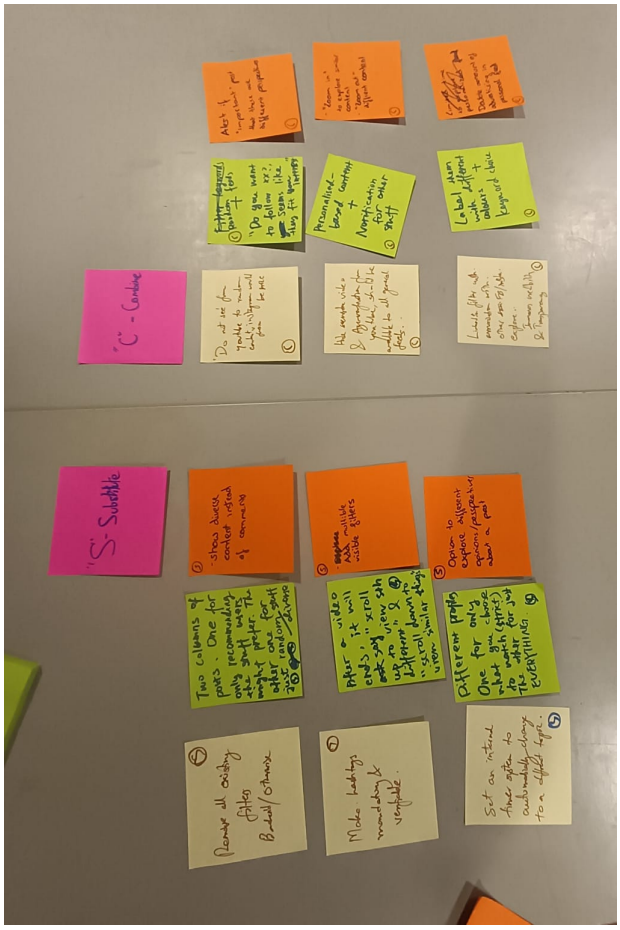


Figure 7: Ideas from Session 3

A.2 Thematic Analysis from Focus Group Session

Themes	Codes	Comments
User Control and Autonomy	Control Choice	P6: I wouldn't say that I feel overwhelmed by the amount of stuff on my feed. I can't say that because I already trained the algorithm to just give me **what I want**. P2: I get to see the stuff from this zoomed out angle, I get to see what is happening. So like, **I get to be an observer**. P7: **I have to be given a choice**. Rather than algorithm thinking I want my choice to be reflected. P9: I want to have some filters with **content I want to see** now. P2: I feel **more controlled** when using Reddit. Maybe it's because I **don't have to join** channels that **I don't like**. P10: I think there should be "**I do not want to see this kind of thing**" in which you can just add keywords, which you don't really want to see at all.
Issues with Recommender Systems	Algorithmic bias Content repetition and filtration	P2: I'm assuming that **depending on which post you stop at**, the algorithms are located like this, and it's **probably filtering that**. P4: Your **personal data value gets amplified** which might be **creating this bias** P1: I don't know if the creator has reposted it or if it's just the algorithm showing the same thing, because **I don't like seeing those things.** P8: Once I was watching like ski videos, and then after that **my feed was full of that**. P10: There has been a case in which I was talking with my friend about a lock or something like that. And then she searched for the lock, and then **I started getting those posts** about the lock. P9: I was watching something with my partner on TV. And then he mentioned that, this actor is Jack Black. So we were talking about it. But then the next day, I saw him on my Instagram. I was like, **I didn't search for that so it bothered me very much.**
Customized Diversity in Content Curation	Filtering as a way to seek diverse content Balancing personalized and diverse content	P7: One way that I used to filter was by **picking one thing from my Explore page**, and then I can continue scrolling, and it shows me some stuff, **I use that as a filter.** P5: I want to get the **content I like**, but I also **like to see new things**. P7: One thing that I think is social media has become important to people is because **it's a very personal thing**. Like you see what you know, the **feed and all is curated to you**, **your social media will be very different from somebody else's social media**. P3: For example, on Twitter, I get a lot of bias content. So when I read something on Twitter, which I think is biased, I move on to Google, as you said, check out what's what's happening and **what's going on with the same topic on other other news articles or something like that**.
Content Discovery	Exposure to new ideas Diversity of content	P4: I don't know what I like **until I see it**. P1: I like to **discover** content. P2: You can't realistically label, like what label works for you might not work for me. I don't think it's realistic for me to think that I should get all sides of this discourse or have **an algorithm that does that**. It's not going to be able to do that. Realistically, you won't. But yeah, **I like discoverability**. That's what I like. P7: You're **stuck in your own bubble of content** but you're **also actively seeking out other new stuff**. P7: I'll **find all the articles** that are talking about or videos or podcasts, and then I kind of like **cross check it and get a complete idea**. That is **the way I kind of challenge what I already know**.
Designing Ethical Recommendations	Challenges with designing solutions Reforms to avoid reinforcing bubbles	P3: it's very **subjective**. For some people, they **don't mind staying in their own bias** because it gives them the comfort to live in their biases. But for **some people**, **they would like to have more new perspective**. P6: I really agree with the news bit that you need to get angles from everywhere. But at least when you ask for **personalization**, **it should be really a choice thing.** P10: **Choice** and **variety** are key to avoiding a **conflict of interest** in personalized content.

A.3 Categorised Ideas

A.4 Personas

Persona 1

Chetan

Bio

Chetan is a 20-year-old university student who likes to keep things comfortable and predictable.

He prefers to have control over the content he sees on his social media feeds and doesn't want the algorithm to dictate what he sees. However, he also understands the importance of being exposed to diverse viewpoints and new ideas.

He is willing to actively seek out new content but doesn't want to be bombarded with too much of it.

He sometimes struggles to identify what he likes until he sees it.

Needs

Control over the content he sees on social media, predictability, exposure to diverse viewpoints and new ideas without being overwhelmed.

Behaviors

Actively seeks out new content but is hesitant to engage with it if it's too different from what he's used to.

Goals

To have a comfortable social media experience while also being exposed to new and diverse content.

Pain points

Feeling overwhelmed by too much diverse content, not being able to identify what he likes until he sees it.

Persona 2

Ananya

Bio

Ananya is a 22-year-old university student who enjoys exploring new content and ideas

She values discoverability and seeks out diverse viewpoints to challenge what she already knows.

She understands that it's impossible to get all sides of every discourse but still wants to have access to a variety of content.

She is skeptical of how much of the content she sees is curated for her and wants to make sure she's getting a complete idea.

Needs

Discoverability, exposure to diverse viewpoints and new ideas, access to multiple sources of information.

Behaviors

Actively seeks out new content, checks multiple sources to get a complete idea.

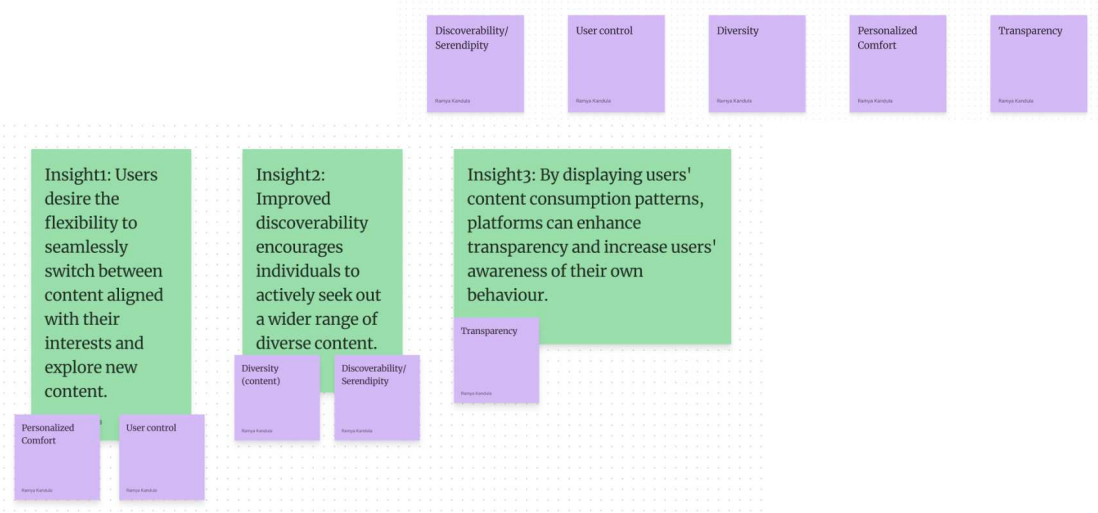
Goals

To explore and discover new content, to challenge what she already knows and broaden her perspectives.

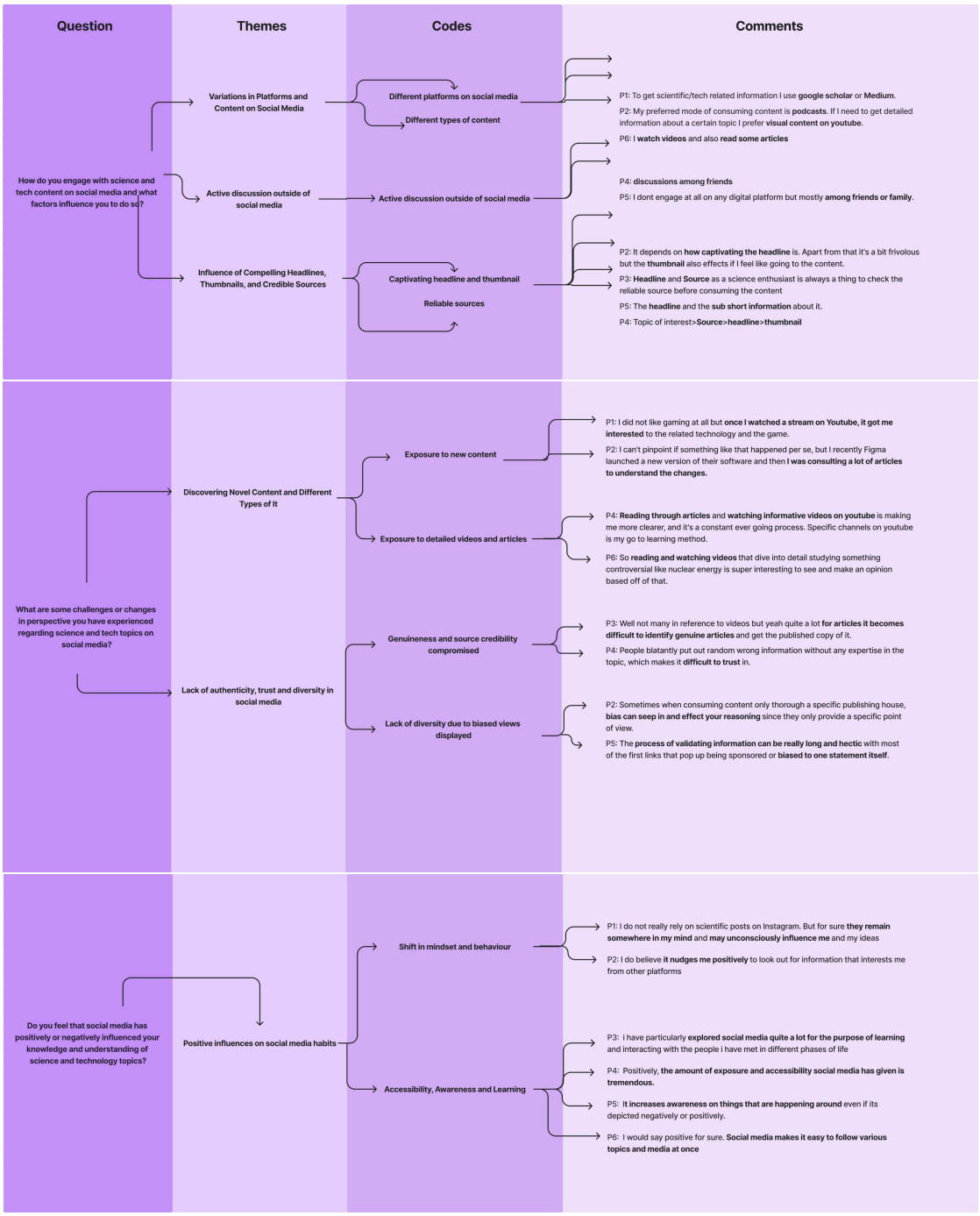
Pain points

Feeling like she's only seeing curated content, not having access to enough diverse viewpoints.

A.5 Insights and Key Principles Driving Them

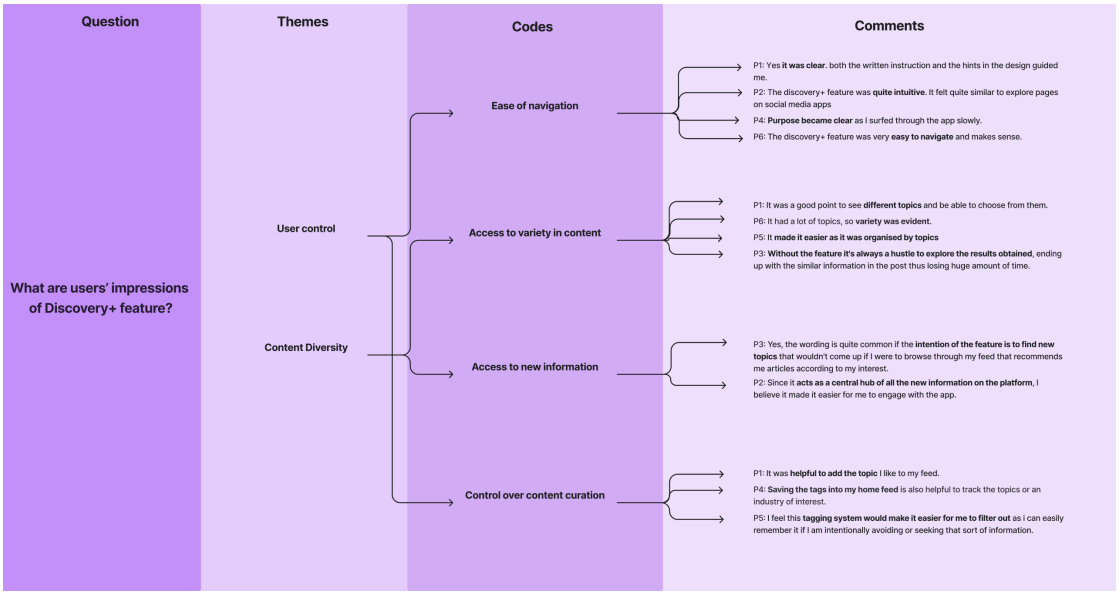


B THEMATIC ANALYSIS OF PRE-STUDY

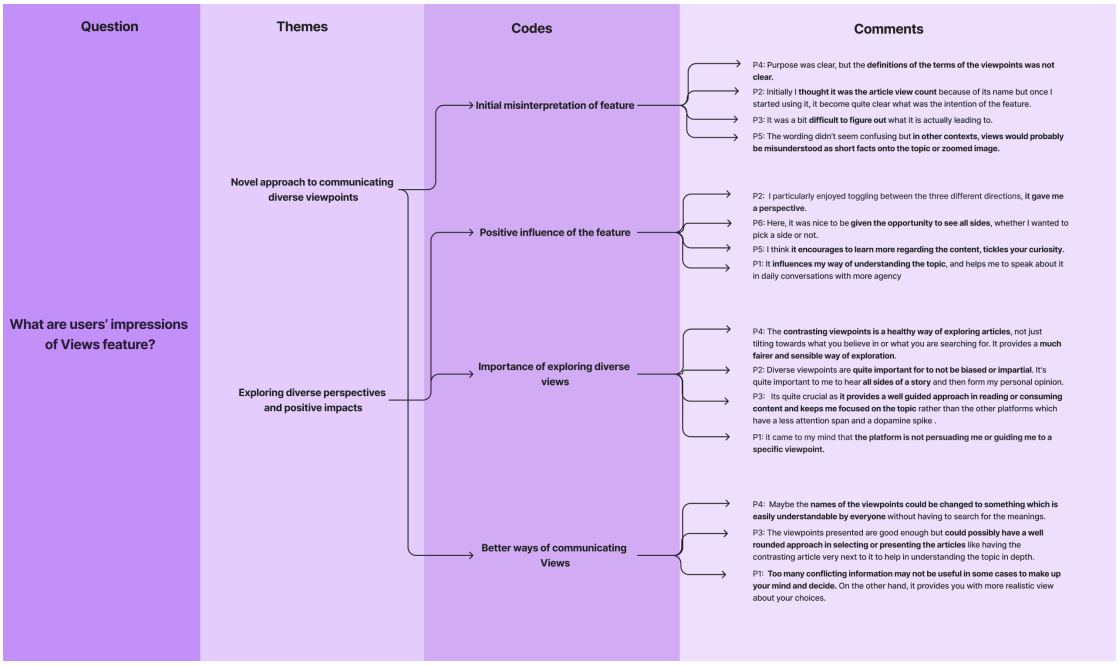


C THEMATIC ANALYSIS OF POST-STUDY

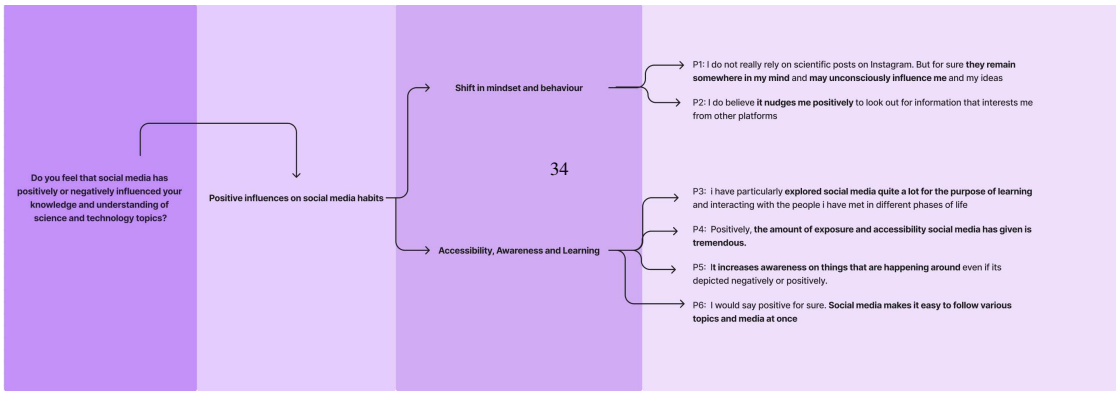
C.1 Discovery+ Feature



C.2 Views Feature



C.3 Consumption History Feature



TRITA-EECS-EX-2023:778