



Module 6

Deliverable 1

Research Report

Esther Lin, Sylvia Shi, Kester Tan, Yousif Alnajjar

Table of Contents

Does interacting with other users benefit the overall experience?

01 Executive
Summary

02 Research
Methods Review

03 Insights +
Supporting
Evidence

04 Appendix

Executive Summary

We conducted contextual interviews with four participants experienced with AI tools to understand how interacting with others on an AI bias reporting platform impacts their engagement and likelihood to return. Key findings revealed that users feel more validated and encouraged to report AI biases when sharing with a community rather than alone. Viewing others' posts allows users to gain new perspectives on types of biases and strategies to identify them.

However, users desire more in-depth educational content, enhanced interaction features like upvoting and following expert users, and robust moderation to keep discussions constructive. Inefficient navigation and sorting suggest a broader need for an intuitive UX, improved content discovery, regular educational updates, and incentives for active users to compel people to visit the platform regularly as the AI landscape evolves.

Research Methods Review

For our contextual interviews, we found **four participants** who have experience with using AI tools. Out of the four participants, three experienced AI bias before and one reported AI biases before.

Our goals for the interview were to understand how interacting with other users on an AI reporting platform would impact their engagement on the platform and whether they would use the platform again.

For our contextual interview, we started with a **directed storytelling** activity to find out more about their experience generating AI bias reports if they have experience doing so. Next, we performed an in-depth **contextual inquiry** activity by asking a user to generate an AI bias report while finding ways to interact with other users' posts. This allowed us to see how the participant navigated through the platform to find posts and whether the action of interacting with other users' posts benefited their experience.

Directed Storytelling

Before we begin, I'd love to understand more about your previous experience about generating a report about AI bias. Would you be keen to share more?

Contextual Inquiry

Instead of just going through the process of creating an audit of an AI bias, we would like you to try and look for other user's posts when you can.

Affinity Diagram



Insights



Insight 1

Users are more likely to use TAIGA when there is a sense of community

People value not feeling like they alone when participating in GenAI bias auditing. User research revealed that people feel more validated when reporting findings when it is posted to a community of people who are also reporting their own findings.

It was also found that people approach GenAI bias auditing in a more positive manner when posting to a community. Users felt as though community building encouraged them to report more findings from the platform as compared to if they were unable to share their posts to others.

Insight 2

When posting to a social platform, users are able to learn new perspectives and methodologies in GenAI bias auditing.

User research showed that in many cases, by being able to view other people's posts and findings, users are able to learn new perspectives of what kinds of AI biases there are. Users reporting that they saw posts on certain types of biases that they had never seen before, and was only able to learn about them by seeing other people's posts. Users were able also see what strategies and methodologies other people used, and were able to note them and implement them as well going forward.

Insight 3

Users expressed demands for quality contents that are more in-depth and educational.

Users indicated preference for educational contents over generic contents, and there's a noted dissatisfaction with the surface-level discussions, which points to a demand for more substantive posts that will lead to more comprehensive and deeper understanding of GenAI biases. Possible implications will be implementing new features that gives random chosen words to generate images and let users identify if there exhibit any bias, therefore enable new insights outside of the box.

Insight 4

Users are looking for enhanced user interaction and moderation for the platform.

There's a clear call for TAIGA to have features that support stronger user interaction and engagement. Preferred implication includes developing and incorporating features such as upvoting, downvoting, and following expert users to increase interactive learning. Alongside this users also raised the need for strong moderation and reporting tools to ensure that the environment remains respectful, inclusive, therefore conducive to their sense of safety and foster constructive users' discourse.

Insight 5

Users expressed need for sustained engagement through features and updates

While users indicated that they encountered inefficient navigation and sorting issues, these appeared as signs of a larger concern regarding the platform's ability to maintain long-term user engagement. Users seek a compelling reason to return to the platform regularly, highlighting the necessity for a more intuitive user experience, better content discoverability, continuous educational insights and updates that keep pace with the evolving landscape of AI and user interests, as well as possible incentives for returning users and active users.

Appendix



Interpretation Notes 1

Note Number	Note		User Profile
SH1-1	Search bar results were not entirely relevant and not many posts showed up, perhaps due to not enough users on the platform yet	SH1-1: Search bar results were not entirely relevant and not many posts showed up, perhaps due to not enough users on the platform yet	SH1 - Stakeholder One
SH1-2	Showed dissatisfaction with the lack of effectiveness from using key words in search feature	SH1-2: Showed dissatisfaction with the lack of effectiveness from using key words in search feature	Has used GenAI before
SH1-3	Encountered racial bias and gender bias when searching through posts, mentioned there were also some types of bias they had never thought about before	SH1-3: Encountered racial bias and gender bias when searching through posts, mentioned there were also some types of bias they had never thought about before	Has experienced AI bias before
SH1-4	Made a note that most of the people who were posting would be people who are interested in AI bias	SH1-4: Made a note that most of the people who were posting would be people who are interested in AI bias	Has never reported AI bias
SH1-5	Stakeholder recommended upvoting feature if a user agrees to what the person is saying in the post and wants to promote it in the algorithm	SH1-5: Stakeholder recommended upvoting feature if a user agrees to what the person is saying in the post and wants to promote it in the algorithm	
SH1-6	Stakeholder recommended downvoting feature if the content is inaccurate	SH1-6: Stakeholder recommended downvoting feature if the content is inaccurate	
SH1-7	Stakeholder recommended against adding additional reaction features as there are already a sufficient number of features	SH1-7: Stakeholder recommended against adding additional reaction features as there are already a sufficient number of features	
SH1-8	Need to think about the diminishing utility of adding more features to the platform	SH1-8: Need to think about the diminishing utility of adding more features to the platform	
SH1-9	Would the ability to downvote posts impact the community aspect and inclusiveness of the platform	SH1-9: Would the ability to downvote posts impact the community aspect and inclusiveness of the platform	
SH1-10	Stakeholder said they would be interested in following specific users on the platform	SH1-10: Stakeholder said they would be interested in following specific users on the platform	
SH1-11	Stakeholder said they would be interested in following specific areas of AI bias on the platform	SH1-11: Stakeholder said they would be interested in following specific areas of AI bias on the platform	
SH1-12	Stakeholder said they would feel more comfortable reporting AI bias on the platform if other people did as well	SH1-12: Stakeholder said they would feel more comfortable reporting AI bias on the platform if other people did as well	
SH1-13	Need to think about if a community of reporting AI bias could be overall positive or negative	SH1-13: Need to think about if a community of reporting AI bias could be overall positive or negative	
SH1-14	Encountered types of biases they had not previously seen	SH1-14: Encountered types of biases they had not previously seen	
SH1-15	Stakeholder reported feeling a sense of community when using the post features	SH1-15: Stakeholder reported feeling a sense of community when using the post features	
SH1-16	Stakeholder did not encounter any dismissive, offensive, or unhelpful posts	SH1-16: Stakeholder did not encounter any dismissive, offensive, or unhelpful posts	
SH1-17	Stakeholder thought a reporting feature would be unnecessary, biased since they did not encounter any harmful posts	SH1-17: Stakeholder thought a reporting feature would be unnecessary, biased since they did not encounter any harmful posts	
SH1-18	Would need to gauge a reporting feature according to how many posts and subsequently inappropriate posts are uploaded	SH1-18: Would need to gauge a reporting feature according to how many posts and subsequently inappropriate posts are uploaded	
SH1-19	Stakeholder reported feeling respected and included when using the post features	SH1-19: Stakeholder reported feeling respected and included when using the post features	
SH1-20	Need to think about the audience of who would be using the platform and how that impacts the type of posts and communication on the platform	SH1-20: Need to think about the audience of who would be using the platform and how that impacts the type of posts and communication on the platform	

Interpretation Notes 2

	Note	User Profile
SH2-1	the navigation on main page is not user-friendly and needs improvement. (SH1 has no idea what to do when navigation to the main page, and found the login button after given hints)	SH2-1: the navigation on main page is not user-friendly and needs improvement. (SH1 has no idea what to do when navigation to the main page, and found the login button after given hints) SH1 - Stakeholder One
SH2-2	a more educative search function is necessary to guide users in understanding and finding biases. (SH1 feels frustrated what to search for while navigating through the posts)	SH2-2: a more educative search function is necessary to guide users in understanding and finding biases. (SH1 feels frustrated what to search for while navigating through the posts) Have participated in a research about GenAI websites
SH2-3	filtering functions seems limited, with categories and contents primarily revolving around people-related biases, suggesting a narrow generating/search scope.	SH2-3: filtering functions seems limited, with categories and contents primarily revolving around people-related biases, suggesting a narrow generating/search scope. Have experiences with GenAI tools
SH2-4	posts are generally lack depth, leading to user disinterest and a need for more insightful and diverse content.	SH2-4: posts are generally lack depth, leading to user disinterest and a need for more insightful and diverse content.
SH2-5	standardization in prompts and interactions fails to encourage deep exploration or engagement with AI bias topics.	SH2-5: standardization in prompts and interactions fails to encourage deep exploration or engagement with AI bias topics.
SH2-6	emotional engagement with the platform is low, with posts described as standardized, not fostering a sense of community or shared purpose.	SH2-6: emotional engagement with the platform is low, with posts described as standardized, not fostering a sense of community or shared purpose.
SH2-7	participants see no negative content, suggests the platform is not harmful	SH2-7: participants see no negative content, suggests the platform is not harmful
SH2-8	a respectful and inclusive environment is essential for users to feel safe in expressing viewpoints and reporting biases.	SH2-8: a respectful and inclusive environment is essential for users to feel safe in expressing viewpoints and reporting biases.
SH2-9	platform is criticized for missing engaging functions and does not provide incentives for long-term user engagement.	SH2-9: platform is criticized for missing engaging functions and does not provide incentives for long-term user engagement.
SH2-10	user interaction is considered unhelpful, pointing to a lack of features that support collaborative learning and reporting.	SH2-10: user interaction is considered unhelpful, pointing to a lack of features that support collaborative learning and reporting.
SH2-11	participant suggest alternative methods, such as using random word combinations, to uncover AI biases, indicating a gap in current platform capabilities.	SH2-11: participant suggest alternative methods, such as using random word combinations, to uncover AI biases, indicating a gap in current platform capabilities.
SH2-12	there's a disconnect between the users' expectation of content depth and what is currently available.	SH2-12: there's a disconnect between the users' expectation of content depth and what is currently available.
SH2-13	participant suggest word checking to report or flag inappropriate or harmful behavior	SH2-13: participant suggest word checking to report or flag inappropriate or harmful behavior
SH2-14	the need for incentives and enhanced features is emphasized to maintain user interest and frequent platform interaction	SH2-14: the need for incentives and enhanced features is emphasized to maintain user interest and frequent platform interaction
SH2-15	platform lacks mechanisms to spotlight insightful content or encourage in-depth examination of AI biases.	SH2-15: platform lacks mechanisms to spotlight insightful content or encourage in-depth examination of AI biases.
SH2-16	transparency, trust, and community feel are highlighted as crucial factors for a productive and engaging AI bias reporting experience.	SH2-16: transparency, trust, and community feel are highlighted as crucial factors for a productive and engaging AI bias reporting experience.
SH2-17	There's an expressed need for a wider array of moderation and reporting tools to ensure constructive and respectful discourse.	SH2-17: There's an expressed need for a wider array of moderation and reporting tools to ensure constructive and respectful discourse.

Interpretation Notes 3

	Note		User Profile
SH2-18	Upvotes are motivated by agreement or new insights; downvotes are due to harmful content or lack of insight, highlighting the need for varied engagement options.	SH2-18: Upvotes are motivated by agreement or new insights; downvotes are due to harmful content or lack of insight, highlighting the need for varied engagement options.	
SH2-19	Participants are hesitant to follow experts or join communities on the platform, preferring more innovative and research-oriented methods.	SH2-19: Participants are hesitant to follow experts or join communities on the platform, preferring more innovative and research-oriented methods.	

Interpretation Notes 4

Note Number	Note	User Profile
SH3-1	The stakeholder found the redirection after creating a post to be faulty as it did not redirect a user to the forum.	SH3-1: The stakeholder found the redirection after creating a post to be faulty as it did not redirect a user to the forum.
SH3-2	They observed posts related to bias in certain prompts, indicating a focus on stereotypical biases prevalent in Western cultures.	SH3-2: They observed posts related to bias in certain prompts, indicating a focus on stereotypical biases prevalent in Western cultures.
SH3-3	The stakeholder appreciated insightful posts that provided detailed explanations and thoughtful perspectives on AI bias, indicating a preference for well-reasoned and informative content.	SH3-3: The stakeholder appreciated insightful posts that provided detailed explanations and thoughtful perspectives on AI bias, indicating a preference for well-reasoned and informative content.
SH3-4	They highlighted criteria for upvoting or downvoting posts, emphasizing the importance of accuracy, avoidance of harmful stereotypes, and offering unique insights.	SH3-4: They highlighted criteria for upvoting or downvoting posts, emphasizing the importance of accuracy, avoidance of harmful stereotypes, and offering unique insights.
SH3-5	They suggested that reaction faces could provide quicker feedback, but detailed options might encourage more thoughtful engagement.	SH3-5: They suggested that reaction faces could provide quicker feedback, but detailed options might encourage more thoughtful engagement.
SH3-6	The stakeholder expressed interest in following expert users and joining communities focused on specific types of bias, indicating a desire for collaboration and mutual learning on AI bias platforms.	SH3-6: The stakeholder expressed interest in following expert users and joining communities focused on specific types of bias, indicating a desire for collaboration and mutual learning on AI bias platforms.
SH3-7	Interacting with other users on the platform motivated the stakeholder to report AI biases.	SH3-7: Interacting with other users on the platform motivated the stakeholder to report AI biases.
SH3-8	They encountered dismissive comments that downplayed the importance of AI bias, this did not discourage them from reporting biases.	SH3-8: They encountered dismissive comments that downplayed the importance of AI bias, this did not discourage them from reporting biases.
SH3-9	They suggested the need for moderation tools to address inappropriate behavior.	SH3-9: They suggested the need for moderation tools to address inappropriate behavior.
SH3-10	The stakeholder emphasized the importance of fostering a respectful and inclusive environment for reporting AI bias to encourage open and honest discussions.	SH3-10: The stakeholder emphasized the importance of fostering a respectful and inclusive environment for reporting AI bias to encourage open and honest discussions.
SH3-11	They highlighted features that would encourage long-term engagement, such as regular updates and interaction with other users. This highlights the importance of ongoing engagement strategies and user feedback mechanisms to sustain user interest over time.	SH3-11: They highlighted features that would encourage long-term engagement, such as regular updates and interaction with other users. This highlights the importance of ongoing engagement strategies and user feedback mechanisms to sustain user interest over time.
SH3-12	They emphasized the value of community engagement in fostering a sense of shared purpose and collective action towards addressing AI bias.	SH3-12: They emphasized the value of community engagement in fostering a sense of shared purpose and collective action towards addressing AI bias.
SH3-13	The stakeholder acknowledged learning new strategies and approaches to identifying and reporting AI bias from interacting with other users.	SH3-13: The stakeholder acknowledged learning new strategies and approaches to identifying and reporting AI bias from interacting with other users.
SH3-14	The stakeholder acknowledged learning new perspectives and tolerance for AI bias from interacting with other users.	SH3-14: The stakeholder acknowledged learning new perspectives and tolerance for AI bias from interacting with other users.
SH3-15	They expressed feeling optimistic about addressing AI bias after reading thoughtful posts from other users.	SH3-15: They expressed feeling optimistic about addressing AI bias after reading thoughtful posts from other users.

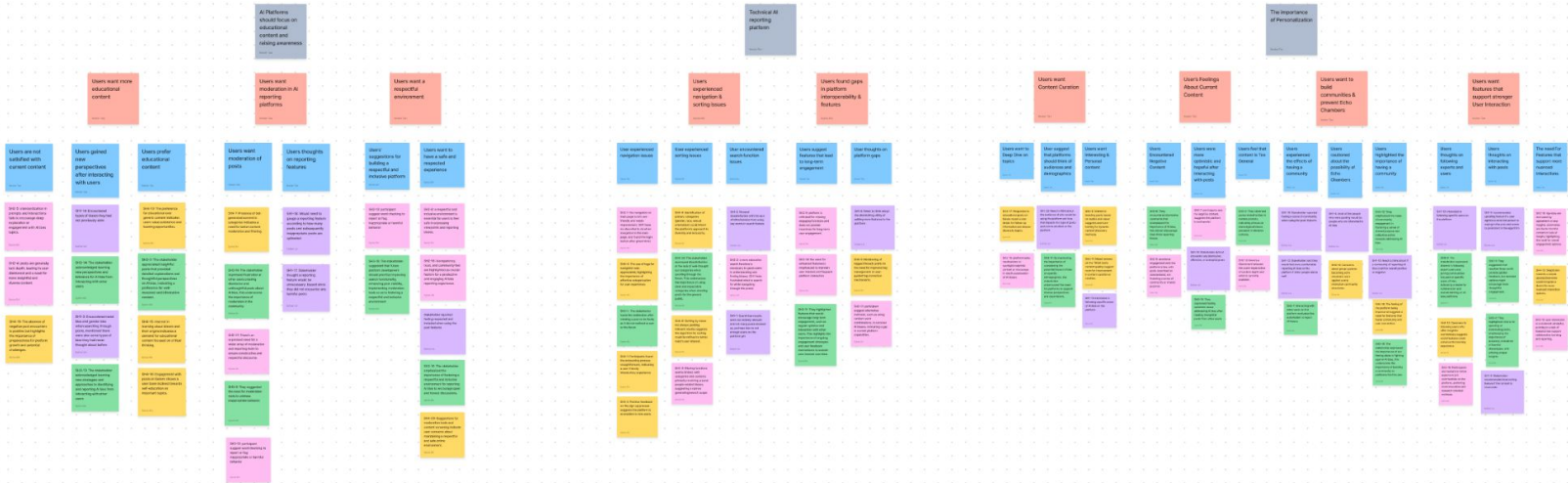
Interpretation Notes 5

Note Number	Note	User Profile
SH3-16	The stakeholder suggested that future platform development should prioritize improving search functionality, enhancing post visibility, implementing moderation tools so as to fostering a respectful and inclusive environment	SH3-16: The stakeholder suggested that future platform development should prioritize improving search functionality, enhancing post visibility, implementing moderation tools so as to fostering a respectful and inclusive environment
SH3-17	By mentioning the importance of considering the potential impact of bias on specific demographics, the stakeholder underscored the need for platforms to support diverse perspectives and experiences.	SH3-17: By mentioning the importance of considering the potential impact of bias on specific demographics, the stakeholder underscored the need for platforms to support diverse perspectives and experiences.
SH3-18	The stakeholder expressed the importance of not feeling alone in fighting against AI bias, this underscores the importance of building a community on platforms like this one.	SH3-18: The stakeholder expressed the importance of not feeling alone in fighting against AI bias, this underscores the importance of building a community on platforms like this one.
SH3-19	The stakeholder expressed frustration at other users creating dismissive and unthoughtful posts about AI bias, this underscores the importance of moderation in the community.	SH3-19: The stakeholder expressed frustration at other users creating dismissive and unthoughtful posts about AI bias, this underscores the importance of moderation in the community.
SH3-20	The stakeholder expressed dissatisfaction at the lack of well-thought out categories when scrolling through the forum. This underscores the importance of using clear and expectable categories when showing posts for the general public.	SH3-20: The stakeholder expressed dissatisfaction at the lack of well-thought out categories when scrolling through the forum. This underscores the importance of using clear and expectable categories when showing posts for the general public.

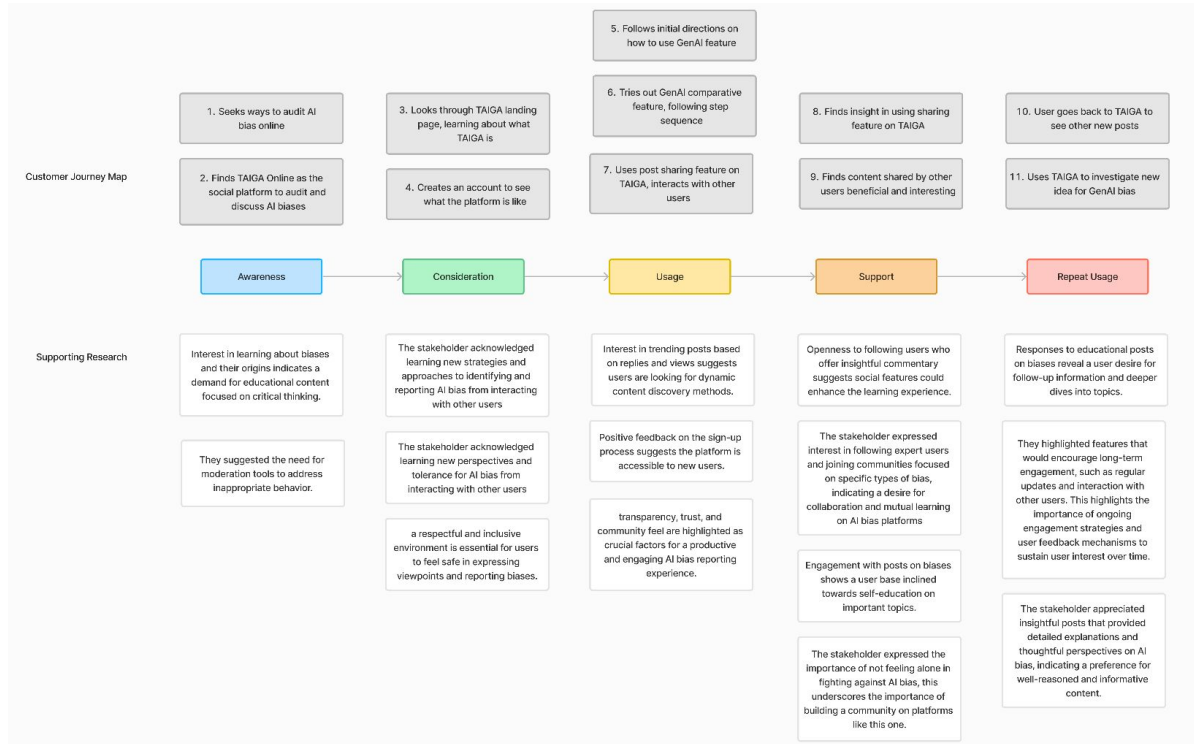
Interpretation Notes 6

	Note		User Profile
SH4-1	Participants found the onboarding process straightforward, indicating a user-friendly introductory experience.	SH4-1: Participants found the onboarding process straightforward, indicating a user-friendly introductory experience.	SH4 - Stakeholder Four
SH4-2	Positive feedback on the sign-up process suggests the platform is accessible to new users.	SH4-2: Positive feedback on the sign-up process suggests the platform is accessible to new users.	Has used GenAI before
SH4-3	Active engagement with posts and discussions reflects a level of user interest and comfort with navigating the platform.	SH4-3: Active engagement with posts and discussions reflects a level of user interest and comfort with navigating the platform.	Has experienced AI bias before
SH4-4	The use of tags for navigation was appreciated, highlighting the importance of effective categorization for user experience.	SH4-4: The use of tags for navigation was appreciated, highlighting the importance of effective categorization for user experience.	Has never reported AI bias
SH4-5	Mislabeling of tagged threads points to the need for improved tag management or user-guided tag correction mechanisms.	SH4-5: Mislabeling of tagged threads points to the need for improved tag management or user-guided tag correction mechanisms.	
SH4-6	Sorting by views not always yielding relevant results suggests the algorithm for sorting could be refined to better match user interest.	SH4-6: Sorting by views not always yielding relevant results suggests the algorithm for sorting could be refined to better match user interest.	
SH4-7	Presence of bot-generated content in categories indicates a need for better content moderation and filtering.	SH4-7: Presence of bot-generated content in categories indicates a need for better content moderation and filtering.	
SH4-8	Interest in trending posts based on replies and views suggests users are looking for dynamic content discovery methods.	SH4-8: Interest in trending posts based on replies and views suggests users are looking for dynamic content discovery methods.	
SH4-9	Identification of primary categories (gender, race, sexual orientation) could inform the platform's approach to diversity and inclusivity.	SH4-9: Identification of primary categories (gender, race, sexual orientation) could inform the platform's approach to diversity and inclusivity.	
SH4-10	The preference for educational over generic content indicates users value substance and learning opportunities.	SH4-10: The preference for educational over generic content indicates users value substance and learning opportunities.	
SH4-11	Mixed reviews on the TAIGA tool's content quality suggest room for improvement in content curation or creation.	SH4-11: Mixed reviews on the TAIGA tool's content quality suggest room for improvement in content curation or creation.	
SH4-12	Skepticism towards a simple upvote/downvote system implies a desire for more nuanced interaction options.	SH4-12: Skepticism towards a simple upvote/downvote system implies a desire for more nuanced interaction options.	
SH4-13	Openness to following users who offer insightful commentary suggests social features could enhance the learning experience.	SH4-13: Openness to following users who offer insightful commentary suggests social features could enhance the learning experience.	
SH4-14	Concerns about group systems becoming echo chambers warn against overly restrictive community structures.	SH4-14: Concerns about group systems becoming echo chambers warn against overly restrictive community structures.	
SH4-15	Interest in learning about biases and their origins indicates a demand for educational content focused on critical thinking.	SH4-15: Interest in learning about biases and their origins indicates a demand for educational content focused on critical thinking.	
SH4-16	Engagement with posts on biases shows a user base inclined towards self-education on important topics.	SH4-16: Engagement with posts on biases shows a user base inclined towards self-education on important topics.	
SH4-17	Responses to educational posts on biases reveal a user desire for follow-up information and deeper dives into topics.	SH4-17: Responses to educational posts on biases reveal a user desire for follow-up information and deeper dives into topics.	
SH4-18	The feeling of the platform being impersonal suggests a need for features that foster community and user connection.	SH4-18: The feeling of the platform being impersonal suggests a need for features that foster community and user connection.	
SH4-19	The absence of negative post encounters is positive but highlights the importance of preparedness for platform growth and potential challenges.	SH4-19: The absence of negative post encounters is positive but highlights the importance of preparedness for platform growth and potential challenges.	
SH4-20	Suggestions for moderation tools and content screening indicate user concerns about maintaining a respectful and safe online environment.	SH4-20: Suggestions for moderation tools and content screening indicate user concerns about maintaining a respectful and safe online environment.	

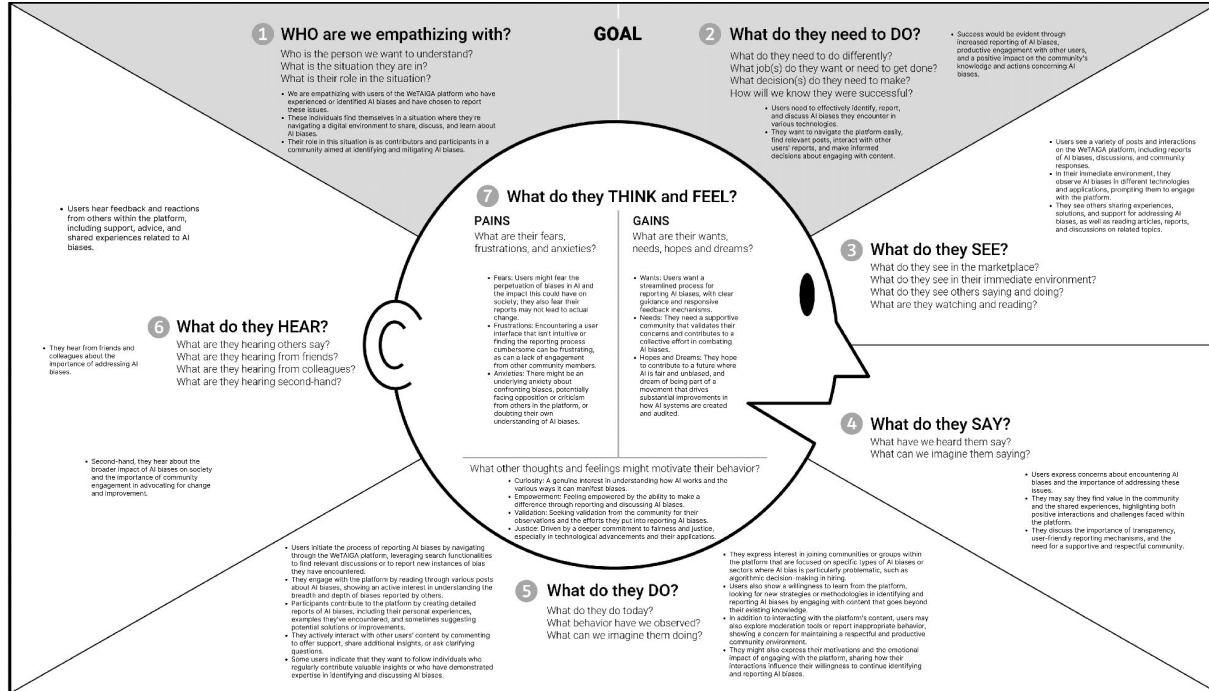
Affinity Clusters



Model 1: Customer Journey Map



Model 2:



Thanks

