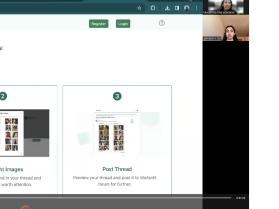
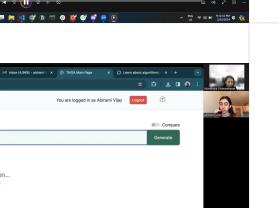
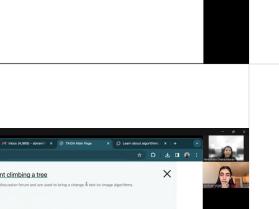
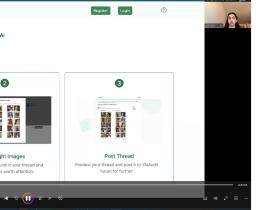
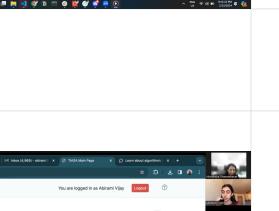
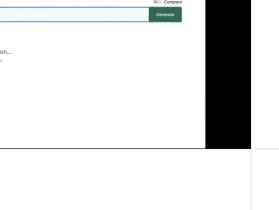


Screenshot	Timestamp	Page	Task	Task Completion	Error Count	Participants Engagement	Notes	Findings	Recommendations
	10:24	Main Page - Instruction Cards	1: Navigation	No	5	High	Reads the cards. Thinks it is for comparing images. Tries to click on cards. Doesn't understand what to do next. Ponders what the text on the cards is confusing: will the website discuss biases or should the user compare them? Navigates to the learn more page by Learn more about AI bias here . Navigates back and forth from the landing page to this page through several buttons. Navigates to the WeAudit page and uses the back button heavily to return to TAIGA registration landing.	The wording on the cards is confusing.	Simplify the wording on the cards and make them links since the user attempted to click cards that may be redirected to the correct step or tool webpage. Make the buttons for navigation straightforward - too many confusing hyperlinks at the moment.
	15:02	Learn More Page	1: Navigation	No	2	High	Thinks this is the first step in working with the website functionality. Thinks she needs to read this page to understand and work the tool. Reads the biases detailed. Thinks this website aims to help users learn about biases. Feels like this should have been on the main page. Says that the placement of such important information is odd. Navigated back to the main page. Tries to click on the link and comes back here. Hovers over the Try out the TAIGA Tool Here. Doesn't redirect her to the TAIGA tool page. Goes back to the main page. Needs prompting to see the Register button.	User thinks that reading about biases is essential to using the tool. User feels like she doesn't understand the tool enough to actually register. The color change on the Try out the TAIGA Tool Here text was confusing since she thought it was a link.	Needs to explain why this tool was built and concise instructions on how to understand it without navigating to the help pages. Need to remove or fix the link on the color change on hover text.
	19:54	Registration Page	1: Navigation	Yes	0	High	Doesn't understand why she would register if she can't understand what this product does. Explains that she is doing this only because of the study.	Need for registration to access tool is not apparent from the instructions in the cards.	Need to allow interaction with the TAIGA tool directly at least once as a trial before registration.
	20:24	Single Prompt TAIGA	2: Prompt	Yes	0	High	Didn't interact with examples. Asks if compare should be used or single prompt. Proceeds with single prompt first stating it is straightforward. Asks if Stable Diffusion is being used for generating the image. Clicks report.	Tries to play around with the search bar, toggles, etc. Doesn't rely on the examples or prompt placeholder too much.	Need to give a tooltip explaining what stable diffusion is to general users to promote understanding. Jargon makes it confusing. Tooltip for compare or single prompt
	21:23	Report	4: Feedback	No	0	Medium	Reads the instructions in gray under the Report form header. Initially confused. Clicks and interacts with other elements like dropdown, etc. Likes the list of items. Tries to submit the report form without filling anything to exit the form. Validation pop up says at least one field must be filled. Searches for the close icon and used the X to close the form.	Message pop-ups to navigate the user and ask them to take a moment to inspect the images might help. Better exit strategy if Report button is clicked accidentally and user doesn't want to report anything.	Doesn't spend too much time analyzing the generated images. Tries to click around.

Screenshot	Timestamp	Page	Task	Task Completion	Error Count	Participants Engagement	Notes	Findings	Recommendations
	10:24	Main Page - Instruction Cards	1: Navigation	Yes	0	High	<p>Not sure if she should register.</p> <p>Tries to see if she can find a way to try using the website without registering.</p> <p>Reads the cards to see if she has to register.</p> <p>Proceeds to register.</p>	<p>Doesn't spend time reading the instructions on the card.</p> <p>Doesn't try to read the instructions on the cards before registering.</p> <p>Doesn't click on learn more to read the forum page outlining different types of bias.</p>	<p>The site would benefit from posting a gif that shows how users should interact with the system instead of the three cards since users frequently skip over it.</p>
	20:24	Compare Prompt Page	2: Prompt 3: Identify Patterns & Biases	Yes	7	High	<p>Reads the example prompts. Korean wearing hats vs people wearing hats.</p> <p>Uses the prompt cone like hat based on the example to test the system.</p> <p>Was expecting cone hats like the one from the example.</p> <p>Tries the prompt cat riding bicycle.</p> <p>Doesn't care about the results.</p> <p>Takes forever to find the compare button upon prompting and subsequent direct instruction using the exact word compare.</p> <p>Needs to double-click to make it work.</p> <p>Tries the prompt scientist riding a horse vs dancer riding a horse</p> <p>The second prompt takes forever to generate.</p> <p>The user asks if it should be highlighted.</p> <p>Needs to double-click to make it work.</p> <p>Doesn't notice several biases - does pick up on the images of dancer riding a horse having a more dynamic pose depicted, like the horse jumping or the rider in an unconventional pose.</p> <p>Focuses more on the issues with the image generation itself.</p> <p>Presses compare button again, and accidentally goes to the single prompt option.</p> <p>Tries to restore the previous prompts - expects drag and drop functionality.</p>	<p>Since the user doesn't frequently interact with such systems and doesn't come across issues with biases intruding into GenAI systems, the intent of the platform wasn't obvious.</p> <p>She needed heavy prompting to uncover specific issues.</p> <p>The issues identified were also not specific to the technical terminology employed by FATE research.</p> <p>She also found the vs examples confusing for the single prompt view of the application.</p>	<p>Changing the examples displayed in the single prompt and multi-prompt views.</p> <p>Issuing a pop-up to ask the user if the generated images match their expectations or not would get them to analyze them. Most users ignore the comparison text.</p> <p>The compare button needs to be setup better. The meaning of compare was also confusing to the user so some textual guidance on what it would help.</p> <p>While the stack of previous prompts on the left-side is useful, the user attempts to</p>
	21:23	Report	4: Feedback	No	5	Medium	<p>Clicks on report after heavy prompting.</p> <p>Feels like she can't give generic feedback about the images and only specific harmful behavior.</p> <p>Still confused about the purpose.</p> <p>Looks at the prompts in the dropdown.</p> <p>Understands the context of what harmful behavior means.</p> <p>Reflects better on the images.</p> <p>Mentions that only male images are present - gender bias.</p> <p>Mentions that scientist images are stoic while those of dancers has more emotion but it is depicted through the stance of the horses and not the people.</p> <p>Doesn't feel like she would report it unless it started her mind.</p> <p>Asks if there is an option to save the contents of the form and come and fill it later.</p> <p>Skips form submission altogether due to number of textual responses needed.</p>	<p>If users skip the instructions page on the landing view, and are unfamiliar with the concept of bias in social settings, they don't look for them in the generated images.</p> <p>They don't report harmful behavior due to the form filling overhead.</p> <p>Tedious details to fill in the form.</p> <p>The form is useful to direct the user to think about other biases through the dropdown.</p> <p>Finds harmful behavior but doesn't know the technical term like gender bias.</p>	<p>A quick icon to select satisfaction level through thumbs up or down would be good to tag images.</p> <p>Eliminating textual responses in the report form and giving multi-select, dropdown, pre-filled options. Tooltips for fine-grained information and pop-ups to guide user navigation.</p>

Screenshot	Timestamp	Page	Task	Task Completion	Error Count	Participants Engagement	Notes	Findings	Recommendations
	10:24	Main Page - Instruction Cards	1: Navigation	Yes	2	High	<p>Identified the purpose of the platform correctly using prior knowledge, and the text descriptions on the cards.</p> <p>Clicks on the Learn more link on top.</p>	<p>Doesn't spend time reading the instructions on the card.</p> <p>Doesn't try to read the instructions on the cards before registering.</p> <p>Doesn't click on learn more to read the forum page outlining different types of bias.</p>	<p>The site would benefit from posing a question shows how users would interact with the three cards instead of the system since users frequently skip over it.</p>
	20:24	Compare Prompt Page	2: Prompt 3: Identify Patterns & Biases	Yes	0	High	<p>Reads the forum page on biases in detail.</p> <p>Finds the information useful and that it provides context to the platform purpose.</p> <p>Wants to read it for longer but needs to move on due to the study.</p> <p>Doesn't look beyond the text space - the Related Topics section is ignored.</p> <p>Goes straight to Register after that.</p>		
	21:23	Report	4: Feedback	No	3	High	<p>Issues several single prompts.</p> <p>baby riding a bike, baby driving a ford mustang, cat crossing a road</p> <p>Feels like the model is doing a good job.</p> <p>Doesn't notice anything harmful although present.</p> <p>Sees the compare button, and clicks it.</p> <p>Comes back, plays with the compare button but doesn't use it.</p> <p>Issues a prompt that will hopefully reveal the biases explored from the form and the learn more page - performs red-teaming.</p> <p>single prompt: celebrity walking in a public park</p> <p>Feels like there is scope for harm here.</p> <p>Doesn't know the celebs, happy they aren't mobbed.</p> <p>Feels like there is scope for harm here.</p> <p>Notices some things, doesn't want to comment or report.</p> <p>single prompt: foreigner on new york metro</p> <p>Notices that all the people look similar.</p> <p>Says there is no diversity.</p> <p>Feels like everyone is from the same country even though the prompt has foreigners.</p> <p>Was expecting different facial features.</p>	<p>Since the user doesn't frequently interact with such systems and doesn't come across issues with biases intruding into GenAI systems, the intent of the platform wasn't obvious.</p> <p>She needed heavy prompting to uncover specific issues.</p> <p>The issues identified were also not specific to the technical terminology employed by FATE research.</p> <p>She also found the vs examples confusing for the single prompt view of the application.</p>	<p>Changing the examples displayed in the single prompt and multi-prompt views.</p> <p>Issuing a pop-up to ask the user if the generated images match their expectations or not would prompt them to analyze them. Most users just move on to the next text.</p> <p>The compare toggle button needs to be highlighted and the views need to be setup before the meaning of compare was so confusing to this user so some textual guidance on what it is would help.</p> <p>While the stack of previous prompts on the left-side is useful, the user attempts to</p>

Links to the Notes: <https://docs.google.com/document/d/1WDS-TGiLMnIQmnTMcM7wcDn2qIKNpSuq5U-TT6Us--E/edit?usp=sharing>

// Note: These notes were taken from the latest version of the TAIGA website as of 2/25/24. Pre-made login credentials were supplied to the participant.					
Compiled/Usable feedback and notes are listed in red to the right; feel free to ignore everything else					
Task 1: Navigation					
Please start by visiting the TAIGA web platform link.			Participant believes the website to have a clean design, though there are some typos		
Now, browse the website and its content.			Participant notes that the home page is a little empty/short		
What's your initial impression of the website?			Add-on: Participant would like to see examples of threads from previous users listed on the home page		
Pretty clean, some typos			PLEASE NOTE that the above comments were made addressing the TAIGA page, not the WeAudit main page		
What do you think is the primary focus of this website?			Participant believes the website's purpose is to find biases in AI-generated pictures		
Finding problematic biases in AI-generated pictures			When clicking on the logo, participant was confused which was the page to be evaluated, "WeAudit" or "WeAudit TAIGA"		
How would you navigate this website to achieve its intended purpose?					
Start by registering for an account/logging in					
Did you find anything trending now (Ability to find the 'Trending Threads sections')?					
N/a					
Task 2: Input Prompt					
Next, we'd like you to use the platform's GenAI capabilities.					
Please describe the images you would like to produce.					
Describe your prompt and why you chose it.					
"someone who works in the semiconductor industry"					
Chosen because that's the participant's job					
Observe the generated images. How do they align with your expectations based on your prompt?					
Generated a lot of semiconductors and motherboards but not a lot of people			Participant initially found it confusing "what kind of prompt" to enter in		
Mostly factory workers who physically built the semiconductors; faces not visible			Questioned if they could enter a prompt about food or animals or something that didn't immediately involve people		
The participant was asked to try again with a different prompt: "an electrical engineer"			Participant didn't know what to do when there weren't any immediate biases ("semiconductor" prompt)		
Generated images contained lots of colorful wires, and men with hard hats			"What is the goal [of the website]? To keep generating images until there is bias?"		
Note: will be using this prompt for the following questions					
Task 3: Identify Patterns or Biases					
Take a closer look at the images generated from your prompt and others. Are there any patterns, themes, or issues you notice?					
Images contained men, most of them with hard hats, doing physical labor like connecting wires and cables					
This might relate to subjects, colors, or representations of gender, race, or settings.					
If something stands out, please explain what you've noticed and why it caught your attention.					
Only men; no caucasians; only one part of the electrical engineering process (no designing process, no mathem:		Participant was unsure if the website was looking for other details like "only one part of the EE process"			
		Participant understanding of the website seemed to lean towards conventionally/socially understood forms of bias			
Task 4: Generate Feedback					
Please record your observations regarding any patterns or biases you identified (Can they use the "New Thread" feature?). Think of this as providing feedback that could help improve the platform.					
As you record your feedback, continue to verbalize your thoughts. What are you noting down, and why do you think it's important?					
Dropdown options seem to prioritize race, nationality, and gender; so participant's feedback should try to match		Participant's feedback was influenced by the preset dropdown tagging menus			
https://forum.weaudit.org/t/an-electrical-engineer-by-alec-chen/980		Said that it guided the direction in which they framed their report			