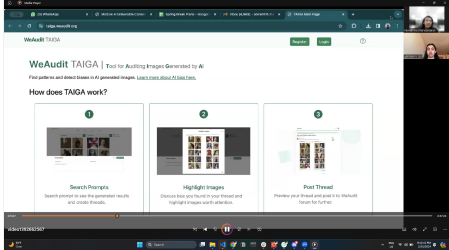
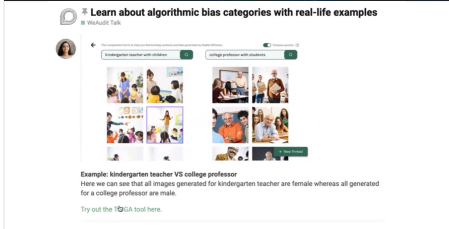
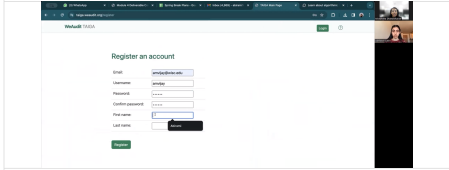
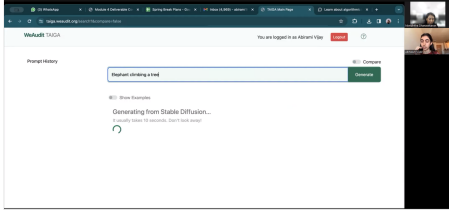
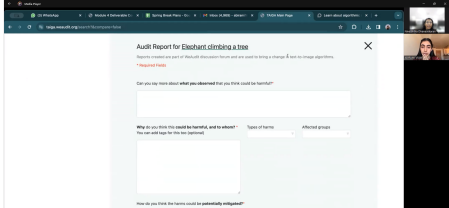
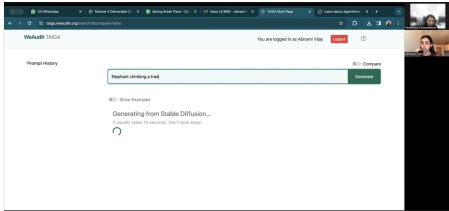


Screenshot	Timestamp	Page	Task	Task Completion	Error Count	Participants Engagement	Notes	Findings	Recommendations
	10:24	Main Page - Instruction Cards	1: Navigation	No	5	High	<p>Reads the cards. Thinks it is for comparing images. Tries to click on cards. Doesn't understand what to do next. Points out that 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.</p>	The wording on the cards is confusing.	Simplify the wording on the cards and make them links since the user attempted to click on them to be re-directed 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	<p>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. Navigates 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.</p>	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 for the user 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	<p>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.</p>	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	<p>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.</p>	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	<p>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 harms. 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.</p>	Doesn't spend too much time analyzing the generated images. Tries to click around.	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.

Snapshot	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 posting a gif that shows how users should interact with the system instead of the three cards since users frequently skip over it.</p>
							<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>		
	20:24	Compare Prompt Page	2: Prompt 3: Identify Patterns & Biases	Yes	0	High	<p>Issues several single prompts.</p> <p>single prompt: 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 report button, and clicks.</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 that it is overall not that harmful.</p> <p>Notifies some things, doesn't want to comment or report.</p> <p>single prompt: foreigner on new york metro</p> <p>Notifies 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 get 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 better. The meaning of compare was also 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>
	21:23	Report	4: Feedback	No	3	High	<p>When asked if he would report it - says wouldn't give detailed textual responses, but would anonymously report it.</p> <p>Clicks the report button.</p> <p>Fills None in the required fields and only elaborates in the optional text field.</p> <p>Barely selects the harmful tags or grouping - complains that the dropdown visual cue is nearly invisible.</p> <p>Clicks submit.</p> <p>Redirected to image tagging.</p> <p>Zoom bar blocks the Post and accidentally comes back to the form.</p> <p>Goes back to the tagging.</p> <p>Does it again and posts it this time.</p> <p>Feels like it is controversial and wouldn't want to post publicly outside the study.</p>	<p>Users find it hard to report the issues as they find it hard to remember the different types of biases they should look for.</p> <p>They also don't report due the text fields on the form - heavy workload.</p> <p>They find some things controversial to talk about and post publicly.</p> <p>The submit button on the report re-directs to the tagging followed by a post - which is an unclear workflow.</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. Prevent the re-direction to another pop-up for image tagging.</p>
LINK to Recording:							<p>https://drive.google.com/drive/folders/1YPRmUjG9HHZbNvwthd8awK2NvWkK2?usp=sharing</p>		

Links to the Notes: <https://docs.google.com/document/d/1WDS-TGiMnIQmnTMcM7wcDn2qIKNpSuq5U-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

Said that it guided the direction in which they framed their report

<https://forum.weaudit.org/t/an-electrical-engineer-by-alec-chen/980>