Quality & Reliability Report:

Introduction:

This report evaluates the quality and reliability of ChatGPT Web, focusing on its functional aspects, LLM/chat capabilities, UI/UX design, and potential areas for improvement. Through test cases, UI evidence, and proposed solutions, the report aims to provide a comprehensive assessment of the platform's current state.a

1. Functional Testing: [High Level Test]

- Landing Page (If not logged in)
    - Signup
        - Test: User is able to signup with valid email address. ✅
        - Test: User is able to signup with Third Party accounts. ✅
            - Google account ✅
            - Mircrosoft account ✅
            - Apple account ✅
    - Login
        - Test: User is able to login with valid credentials. ✅
        - Test: User is able to login with Third Party accounts. ✅
            - Google account ✅
            - Mircrosoft account ✅
            - Apple account ✅
    - Navigations:
        - Privacy Policy
            - Test: Link is working fine. ✅
        - Terms of Use
            - Test: Link is working fine. ✅
- Landing Page (If logged in)

- New Chat (Request)
    - Test: User is able to send chatgpt query and a response against that is being provided. ✅
    - Test: User is able to keep on asking the queries on the provided response. ✅
    - Test: User is able to select a prewritten query and a response to that has been provided. ✅
    - Test: User is able to edit the query submitted and a new response is generated for the same. ✅
    - Test: User is able to regenerate the response by invoking regenerate button. ✅
    - Test: User is able to provide feedback to the chatgpt response. ⚠️
    - Test: User is able to copy the provided response. ✅
    - Test: User is able to regenerate a new response to the existing query. ✅
    - Test: User is able to navigate between the responses. ✅
    - Test: User is able to share the chat with public link. ✅
    - Test: User is able to delete the shareable link. ✅
- Profile Details
    - Test: User is able to set custom interaction. ✅
    - Test: User is able to interact with the settings. ✅
    - Test: User is able to log out. ✅
- Chat History
    - Test: User is able to view the chat history. ✅
    - Test: User is able to interact with the previous chat from the chat history. ✅
    - Test: User is able to delete/archive/rename the chat from the chat history. ✅

2. LLM/Chat result testing:

- Factual Accuracy of Responses:

    - Test: User interacting with app provides accurate and relevant responses. ✅
        - Approach: Ask a set of diverse questions and analyze the responses.
    - Test: User interacting with app responds correctly to a variety of questions. ✅

- Contextual Understanding:
    - Scenarios: Assess the model's ability to maintain context in a conversation.
        - Test: User interacting with app upon asking a query model engage in a multi-turn conversation and check if the model understands context. ✅
        - Test: User interacting with app upon retrying the query stays within the relevant context. ⚠️
    - Scenario: Understanding of a situation and maintain the thread
        - Test: Once a specific context is set and user start building queries on top on it than application maintains the context and do respond wiith considering all the possible situations defined from beginning. ⚠️
- Bias detection:
    - Test Case: User interacting with app upon asking questions prone to social or cultural biases should not be biased. ✅
- Creativity and humor:
    - Test case: User interacting with app upon asking query related to creative writing, jokes, or different writing styles respond within the context with expected creativity. ✅

3. UI/UX & Accessibility Testing:

- Design Consistency:
    - Test Case: User experience should be consistent in design elements across the web application platform. ✅
    - Test Case: User should not be experience any design flaws that may affect user experience. 🔴
    - Test Case: User should feel comfortable while using the application and ease of access in the application should be provided. ✅

    - Test case: User experiences should be adaptive and consistent across different devices and screen sizes. ✅

- Accessibility
    - Test Case: User with special abilities should be able to access the web application platform screen reader and keyboard navigation to access all features. ✅
    - Test Case: Web application platform should be complaint and developed considering WCAG guidelines for visual, auditory, and cognitive disabilities. ✅
- GDPR
    - Test Case: Web application platform is following the GDPR guideline. ✅
    - Test Case: User consent should be taken before proceeding any user data. ✅

4. Performance Testing

- Web application Performance:
    - Test Case: Measure the web application performance with the hep of google lighthouse tool for the following information
        - First contentful paint ✅
        - First meaningful paint ✅
        - Time to interactive ✅
        - Estimated input latency ✅
    - Test Case: Application performance under the slow network condition ✅
    - Test Case: Application performance under the heavy load of users accessing the application on same it. ✅

5. Security Testing:

- Security Testing:
    - Identify and address potential security vulnerabilities. ✅
    - Review data security practices and user privacy policies. ✅
    - Assess the platform's transparency about LLM operation and potential limitations. ✅

Future Recommendations:

- Continuously gather user feedback to inform future improvements.
- Integrate advanced testing tools and automation into the development process.
- Establish clear quality and reliability metrics for consistent evaluation.
- Prioritize transparency and ethical considerations in AI development.

Conclusion:

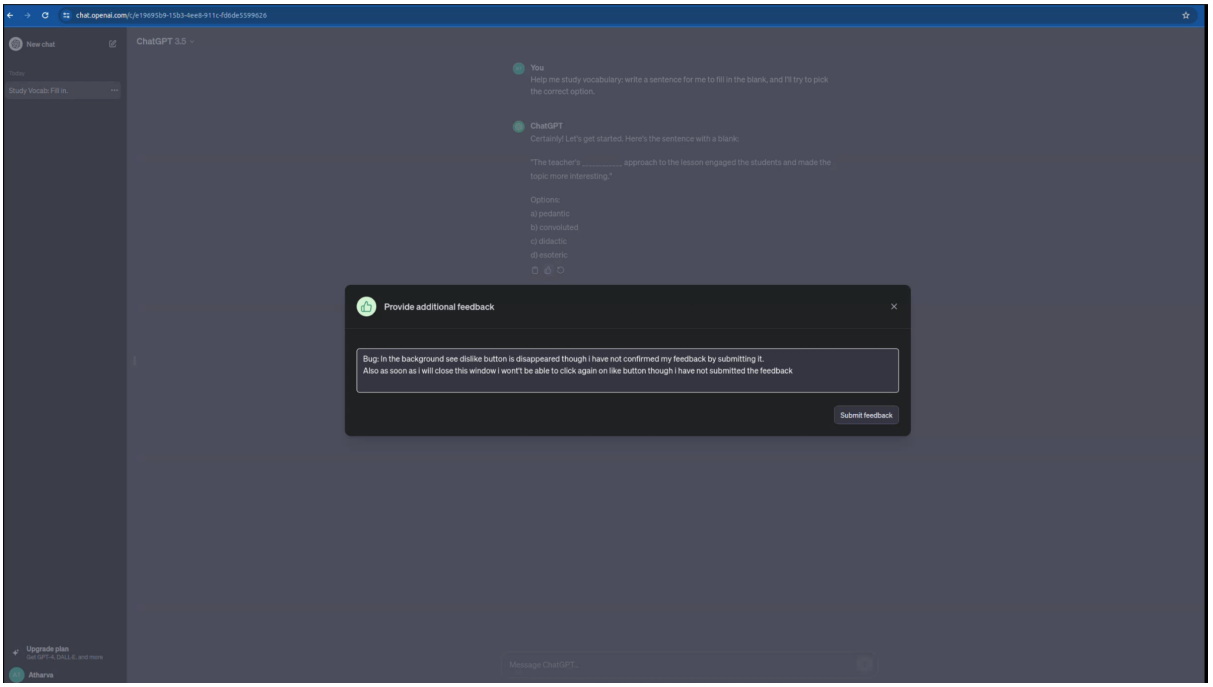
ChatGPT Web demonstrates impressive capabilities in LLM/chat interactions, but requires improvement in functional aspects, UI/UX design, and transparency. By addressing the identified issues through the proposed solutions and continuous testing, ChatGPT can enhance its quality, reliability, and user experience.

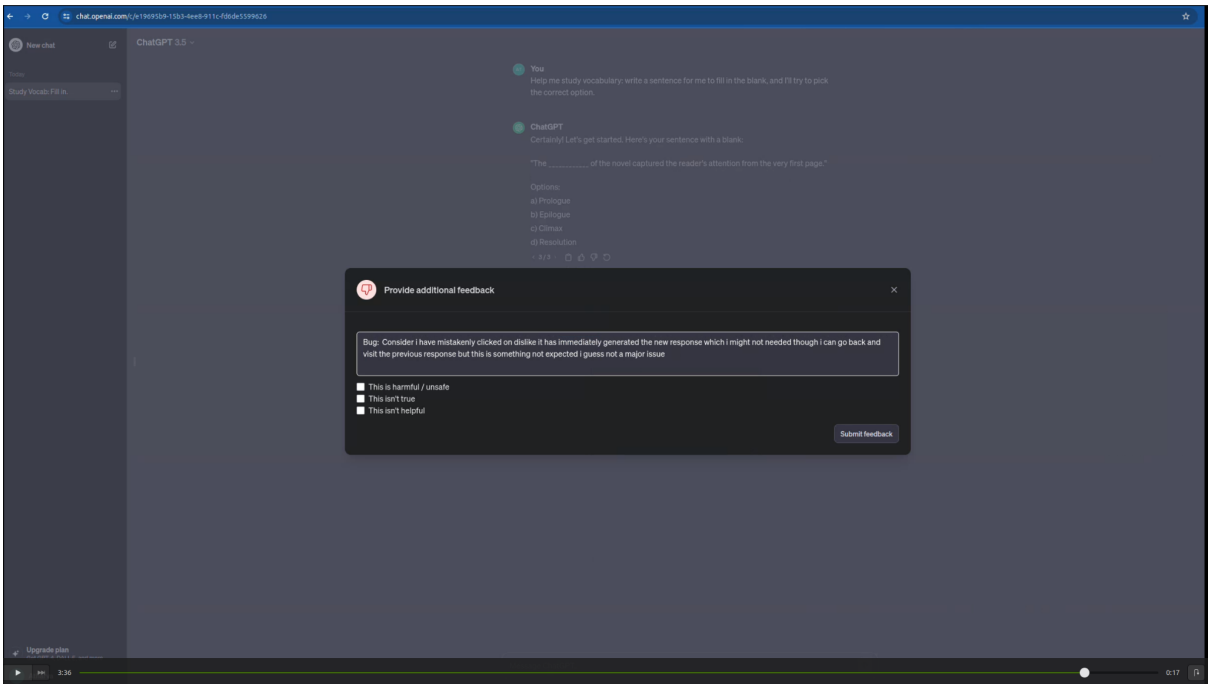
This report provides a starting point for further evaluation of ChatGPT Web. Regular testing and user feedback are crucial to ensure its continued improvement and responsible use.

Type of Testing	Output
Functional Testing	⚠️ [Could be more improvised]
LLM/Chat results testing	⚠️ [Could be more improvised]
UI/UX & Accessibility Testing	🔴 [Some scenarios failed]
Performance Testing	✅
Security Testing	✅

Bugs/Improvements:

1. Once a response is generated and user clicks on 👍 button to provide the feedback an immediate popup window appears on which user can provide additional feedback but it disappears the choice of deselcting the 👎option in background which means once user has provided one type of feedback it cannot be reverted considering it might be done with a mistake.
   Also user cannot click back to the 👍/ 👎button again if they have clicked it once.





2. LLM Issues:
   Some times the chat looses the context if theread is long running and user do not keep on providing the enough context considering it already set on the top of the chat.