

AI Tool Evaluation Checklist

Save all prompts used for testing and share them at the end of this document.

Target tool:

Intended use(s) for tool:

Documentation consulted (provide links if possible):

Overall thoughts and comments on this tool:

Usability

Ease of use

- ☐ The purpose of the tool is clear.
- ☐ The user interface is intuitive.
- ☐ There are instructions for using the tool.
- ☐ The tools features and settings are findable and understandable.

Efficiency

- ☐ The tool provides answers in a reasonable time frame.
- ☐ The tool performs well when receiving a high volume of requests.

Limitations

- ☐ The tool has hard limitations on use, such as the number of prompts (tokens) allowed within a given period or throttling of response times after a certain number of uses.
- ☐ Limitations (or lack thereof) are clearly communicated to users before they are encountered.

Model Parameter Customization

- ☐ Temperature can be adjusted.
- ☐ Top-p can be adjusted.
- ☐ Top-k can be adjusted.
- ☐ Presence penalty can be adjusted.

User Satisfaction

- ☐ I am satisfied with the responses provided by this tool.
- ☐ I am satisfied by the experience of using this tool.
- ☐ The tool reliably assists with library-related tasks.

Comments on Usability

Accuracy and Reliability

Accuracy of Responses

- ☐ The tool understands queries.
- ☐ It consistently provides accurate and relevant information.
- ☐ The tool catches misunderstandings and errors and offers helpful suggestions.
- ☐ It provides sources or citations for its answers that are accurate and reliable.
- ☐ The information provided is current.
- ☐ Responses to similar prompts are consistent.
- ☐ The tool provides explanations for its reasoning and decision-making when requested.

Contextual Awareness

- ☐ The tool maintains context across multiple prompts (for example, if the user asks follow-up questions the tool remembers the previous answers)
- ☐ The tool seems able to interpret user intent.
- ☐ The tool provides information that aligns with my specific needs.

Accuracy and Reliability Comments

Accessibility and Inclusivity

WCAG 2.1 AA Compliance

- ☐ The user interface adheres to WCAG 2.1 AA guidelines according to an evaluation tool such as the WAVE browser plugin.
- ☐ The user interface can be navigated by keyboard effectively.

Multilingual Support

- ☐ The tool understand queries in languages commonly spoken by our user base.
- ☐ Translations provided by the tool are accurate.
- ☐ Translations provided by the tool sound natural and authentic.

Accessibility and Inclusivity Comments

Privacy, Consent, and Security

Data Protection

- ☐ The tool provides clear information on its privacy practices.
- ☐ The tool provides information on its data retention practices.
- ☐ The tool uses data encryption.
- ☐ The tool is FERPA compliant.
- ☐ The tool is HIPAA compliant.

Data Usage Transparency

- ☐ The tool provides clear information on its data collection and usage practices.
- ☐ The tool uses data collected to train the model.

Privacy, Consent, and Security Comments

Ethical and Environmental Considerations

Cultural Sensitivity and Bias

- ☐ The tool avoids biases and stereotypes.
- ☐ The tool handles sensitive topics responsibly.
- ☐ The tool provides responses that are inclusive for users from diverse cultural backgrounds.

Accountability

- ☐ There is a clear procedure for addressing errors, biases, or inappropriate content generated.
- ☐ The provider has a public statement or report addressing the tool's environmental impact.
- ☐ The provider has a public statement or report addressing their sustainability efforts.
- ☐ The tool offers a dashboard for assessing the tool's environmental impact.

Model Transparency

- ☐ There is clear access to a model card, which provides detailed information about the model.
- ☐ The model card provides details about the data it was trained on.
- ☐ The model card describes the tool's intended uses and potential limitations.
- ☐ The model card includes a section detailing the energy consumption and/or estimated carbon emissions associated with training the model.

Ethical and Environmental Considerations Comments

Integration and Customization

Integration with Existing Library Systems

- ☐ The tool can be easily embedded in library websites using embed code or a plugin.
- ☐ The tool provides API access to submit queries.
- ☐ The tool provides API access to retrieve responses.
- ☐ Any API rate limits are clearly communicated.

Authentication and Authorization

- ☐ The tool integrates with the library's existing authentication system.

Interface Customization

- ☐ The tool's UI can be customized to match the library's branding.
- ☐ The layout or presentation of information can be adjusted.

Functionality Customization

- ☐ The tool's core functionality can be modified or extended through configuration options, plugins, or custom scripts.
- ☐ The tool can be configured to prioritize certain data sources or types of information.

Integration and Customization Comments

Documentation of Testing Prompts & Responses

Copy & paste text or screenshots of all exchanges used in testing.