

Guidelines for Human-AI Interaction

INITIALLY

1
INITIALLY

Make clear what the system can do.

Help the users understand what the AI system is capable of doing.

2
INITIALLY

Make clear how well the system can do what it can do.

Help the user understand how often the AI system may make mistakes.

DURING INTERACTION

3
DURING INTERACTION

Time services based on context.

Time when to act or interrupt based on the user's current task and environment.

4
DURING INTERACTION

Show contextually relevant information.

Display information relevant to the users' current task and environment.

5
DURING INTERACTION

Match relevant social norms.

Ensure the experience is delivered in a way that users would expect, given their social and cultural context.

6
DURING INTERACTION

Mitigate social biases.

Ensure the AI system's language and behaviors do not reinforce undesirable and unfair stereotypes and biases.

WHEN WRONG

7
WHEN WRONG

Support efficient invocation.

Make it easy to invoke or request the AI system's services when needed.

8
WHEN WRONG

Support efficient dismissal.

Make it easy to dismiss or ignore undesired system services.

9
WHEN WRONG

Support efficient correction.

Make it easy to edit, refine, or recover when the AI system is wrong.

10
WHEN WRONG

Scope services when in doubt.

Engage in disambiguation or gracefully degrade the AI system's services when uncertain about a user's goals.

11
WHEN WRONG

Make clear why the system did what it did.

Enable the user to access an explanation of why the AI system behaved as it did.

OVER TIME

12
OVER TIME

Remember recent interactions.

Maintain short-term memory and allow the user to make efficient references to that memory.

13
OVER TIME

Learn from user behavior.

Personalize the user's experience by learning from their actions over time.

14
OVER TIME

Update and adapt cautiously.

Limit disruptive changes when updating and adapting the AI system's behaviors.

15
OVER TIME

Encourage granular feedback.

Enable the user to provide feedback indicating their preferences during regular interaction with the AI system.

16
OVER TIME

Convey the consequences of user actions.

Immediately update or convey how user actions will impact future behaviors of the AI system.

17
OVER TIME

Provide global controls.

Allow the user to globally customize what the AI system monitors and how it behaves.

18
OVER TIME

Notify users about changes.

Inform the user when the AI system adds or updates its capabilities.

The Guidelines for Human-AI Interaction will help you create AI systems and features that are human-centered. We hope you use them throughout your design process – as you evaluate existing ideas, brainstorm new ones, and collaborate with the multiple perspectives involved in creating AI.

These guidelines synthesize more than 20 years of thinking and research in human-AI interaction. Learn more: <https://aka.ms/aiguideelines>.

