UX DOWNLOADABLE RESOURCE

AI & Agentic UX Checklist

A practical review framework for useful, understandable and controllable AI experiences.

48 review checks	Risk scorecard	Launch test scenarios

CREATED BY

Zahra Ali · UX/UI & Graphic Designer

zahraali.ca

Design AI that users can understand and control

AI changes the interface from a predictable tool into a system that can infer, recommend and sometimes act. This checklist helps teams evaluate the experience before launch, with special attention to expectations, uncertainty, permissions, recovery, accessibility and accountability.

HOW TO USE THIS RESOURCE

Score each question: 0 = missing, 1 = partial, 2 = clearly implemented, N/A = not applicable. Add evidence rather than scoring from memory. Any zero in permissions, safety, privacy or recovery should block launch until reviewed.

Product snapshot

PRODUCT OR FEATURE Name and version	**REVIEW DATE** YYYY-MM-DD
PRIMARY USER Who relies on this experience?	**AI CAPABILITY** Predicts, generates, recommends or acts
HIGHEST-RISK ACTION What is the worst credible mistake?	**REVIEW TEAM** Design, product, engineering, legal or research

Score interpretation

80-96 Strong foundation - verify with users and monitor after launch.	**60-79** Material gaps - prioritize high-risk sections before release.
BELOW 60 Do not treat the experience as launch-ready.	**CRITICAL ZERO** A zero in safety, permission, privacy or recovery overrides the total score.

1 Value and user need

#	Review question	Evidence / score
1	The AI capability solves a specific user problem rather than adding novelty.	0 1 2 N/A
2	The non-AI alternative was considered and the AI option has a clear advantage.	0 1 2 N/A
3	Success is defined in user outcomes, not only model or engagement metrics.	0 1 2 N/A
4	The system does not automate a decision that should remain human-led.	0 1 2 N/A
5	Users can complete the core task when the AI is unavailable.	0 1 2 N/A

2 Expectations and capability disclosure

#	Review question	Evidence / score
1	The interface explains what the AI can do before users depend on it.	0 1 2 N/A
2	Important limitations are visible at the moment they matter.	0 1 2 N/A
3	Output language avoids implying certainty when the system is probabilistic.	0 1 2 N/A
4	Users can distinguish AI-generated content from verified human or source content.	0 1 2 N/A
5	Examples and onboarding set realistic expectations rather than promising perfection.	0 1 2 N/A
6	The experience shows freshness, source or date context where information can expire.	0 1 2 N/A

3 Intent, control and reversibility

#	Review question	Evidence / score
1	Users can review or edit the system's interpretation of their goal.	0 1 2 N/A
2	The AI asks for clarification when intent, recipient, scope or consequences are unclear.	0 1 2 N/A
3	Users can pause, cancel or stop an ongoing agentic action.	0 1 2 N/A
4	High-impact actions require an explicit confirmation with a clear summary.	0 1 2 N/A
5	Users can undo, correct or recover from an AI action whenever technically possible.	0 1 2 N/A
6	The system does not silently expand its task beyond the user's approved scope.	0 1 2 N/A

4 Permissions and action boundaries

#	Review question	Evidence / score
1	Permissions are requested only when needed and explained in plain language.	0 1 2 N/A
2	The interface identifies what data, account or tool the AI will access.	0 1 2 N/A
3	Users can approve actions individually instead of granting unnecessarily broad control.	0 1 2 N/A
4	Recipients, amounts, dates and destinations are shown before consequential actions.	0 1 2 N/A
5	The system prevents duplicate submissions, messages, purchases or destructive actions.	0 1 2 N/A
6	Administrative or irreversible actions require stronger safeguards than routine tasks.	0 1 2 N/A

5 Transparency and explainability

#	Review question	Evidence / score
1	Users can understand why a recommendation or decision appeared.	0 1 2 N/A
2	The explanation is appropriate to the decision's importance and user expertise.	0 1 2 N/A
3	Sources are linked when factual accuracy or verification matters.	0 1 2 N/A
4	The experience separates facts, estimates, inferences and generated suggestions.	0 1 2 N/A
5	Personalization can be inspected, adjusted or disabled.	0 1 2 N/A

6 Uncertainty, errors and recovery

#	Review question	Evidence / score
1	The UI communicates uncertainty without overwhelming the user.	0 1 2 N/A
2	The system gracefully narrows its service when it lacks enough information.	0 1 2 N/A
3	Error messages explain what happened, what may have changed and the next safe step.	0 1 2 N/A
4	Users are not blamed for model failures or ambiguous output.	0 1 2 N/A
5	A human support or escalation path exists for important failures.	0 1 2 N/A
6	The system preserves useful work when an action or generation fails.	0 1 2 N/A

7 Trust, privacy and safety

#	Review question	Evidence / score
1	The experience clearly states how prompts, files and feedback may be used.	0 1 2 N/A
2	Sensitive information is minimized, masked or excluded when it is not required.	0 1 2 N/A
3	The UI warns users before sharing private data with another person or system.	0 1 2 N/A
4	Abuse, bias and harmful-output risks were tested with realistic scenarios.	0 1 2 N/A
5	Users can report harmful, incorrect or inappropriate behavior.	0 1 2 N/A
6	Logs and audit trails exist for consequential agentic actions.	0 1 2 N/A

8 Accessibility and inclusion

#	Review question	Evidence / score
1	All AI controls work with keyboard navigation and assistive technology.	0 1 2 N/A
2	Status changes, progress and generated results are announced accessibly.	0 1 2 N/A
3	The interface does not rely on color alone to communicate confidence or risk.	0 1 2 N/A
4	Generated content can be reviewed at a comfortable pace without disappearing.	0 1 2 N/A
5	Plain-language summaries are available for complex output.	0 1 2 N/A
6	Testing includes people with disabilities and users with varied language or digital confidence.	0 1 2 N/A

9 Feedback, learning and monitoring

#	Review question	Evidence / score
1	Feedback controls ask for useful context rather than only thumbs up or down.	0 1 2 N/A
2	Users know whether feedback changes their experience, the product or both.	0 1 2 N/A
3	Teams monitor task success, corrections, reversals, complaints and harmful failures.	0 1 2 N/A
4	Model or behavior changes are communicated when they materially affect users.	0 1 2 N/A
5	The experience supports resetting or correcting learned preferences.	0 1 2 N/A

Critical launch scenarios

Run these scenarios with the actual interface. Record what the user sees, what the system does and whether recovery is possible.

#	Review question	Owner	Evidence / score
1	The user's request is ambiguous and could affect the wrong person or account.	_____	0 1 2 N/A
2	The AI is confidently wrong about an important fact.	_____	0 1 2 N/A
3	The user changes their mind during a multi-step agentic action.	_____	0 1 2 N/A
4	A connected service becomes unavailable halfway through the task.	_____	0 1 2 N/A
5	The AI attempts to request or expose unnecessary sensitive information.	_____	0 1 2 N/A
6	A keyboard or screen-reader user needs to review and correct the output.	_____	0 1 2 N/A
7	The same action is submitted twice because the interface appears unresponsive.	_____	0 1 2 N/A
8	The product behavior changes after a model update.	_____	0 1 2 N/A

Launch decision

DECISION

Launch, launch with conditions, or do not launch. Identify the evidence behind the decision.

_____

_____

_____

_____

TOP THREE RISKS

Name the risk, the owner and the mitigation deadline.

_____

_____

_____

_____

References and responsible use

This checklist is an original practical synthesis. It does not replace security, privacy, legal, accessibility or domain-specific review.

Microsoft HAX Toolkit

Evidence-based guidance, design patterns and planning tools for human-AI experiences.
https://www.microsoft.com/en-us/haxtoolkit/

Microsoft Guidelines for Human-AI Interaction

Eighteen guidelines covering initial interaction, regular use, failure and long-term behavior.
https://www.microsoft.com/en-us/haxtoolkit/ai-guidelines/

Google People + AI Guidebook

UX and machine-learning guidance for designing helpful AI products. https://pair.withgoogle.com/guidebook/

W3C Accessibility, Usability and Inclusion

Clarifies how accessibility and usability overlap and why inclusive evaluation matters.
https://www.w3.org/WAI/fundamentals/accessibility-usability-inclusion/

Created by Zahra Ali · UX/UI & Graphic Designer · zahraali.ca