



Designing with Purpose: Building Human-Centered Copilots Using the Microsoft HAX Toolkit

Building secure, trustworthy AI experiences that empower users and strengthen assurance.



Padmini Soni

Founder & CEO, Rezonance AI

Specialize in AI Literacy, AI Ethics & Strategy,
and AI-driven business transformation.

Cofounder of a global community - Asian
Women Advancing AI

Co-author of an upcoming book on OpenAI GPT
Builders

25+ years bridging technology, business, and
risk management





 YouTube



Microsoft AI Commercial ft Common (*Microsoft & Common are Register...

Please Get Back to The Person Who Referred You, to This Video. or Explore The World's Most Profitable,... "A.I. - Powered" FOREX Trading Software !!! ☆ Previously...

Responsible AI *is* Human-Centered

"It's about ensuring that what we build benefits people in society and how we build it begins and ends with people in mind. "

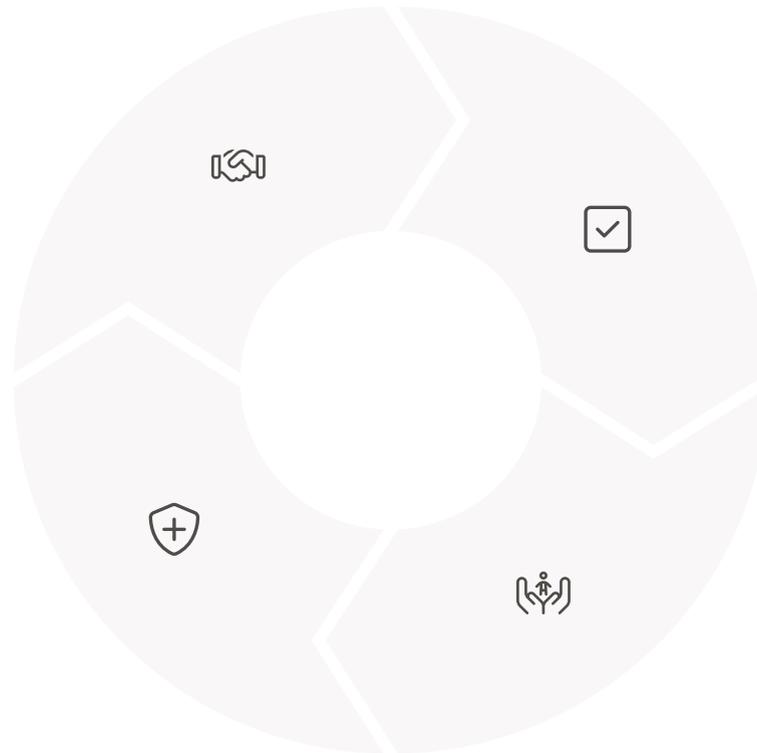
Human-Centered AI Cornerstones

Trust and Assurance

Human-centered AI builds trust by being transparent, explainable, and predictable in its behavior.

Safety and Risk Management

Keeps humans in the loop thereby providing critical safeguards against AI system failures, unexpected behaviors, and misuse.



Ethical Decision-Making and Bias Mitigation

Prioritizes human values and perspectives in AI design, which can help better identify and address harmful biases.

Enhanced Human Capabilities Rather Than Replacement

Augments human intelligence and abilities rather than replacing human judgment entirely.

What is HAX Toolkit?

The Microsoft HAX (Human-AI eXperience) Toolkit is a **resource designed to help teams create human-centered AI products and experiences.**



Guidelines

Best practices to ensure AI behavior is predictable, transparent, and resilient.



Workbook

Work together with your team to prioritize which Guidelines to implement in your product.



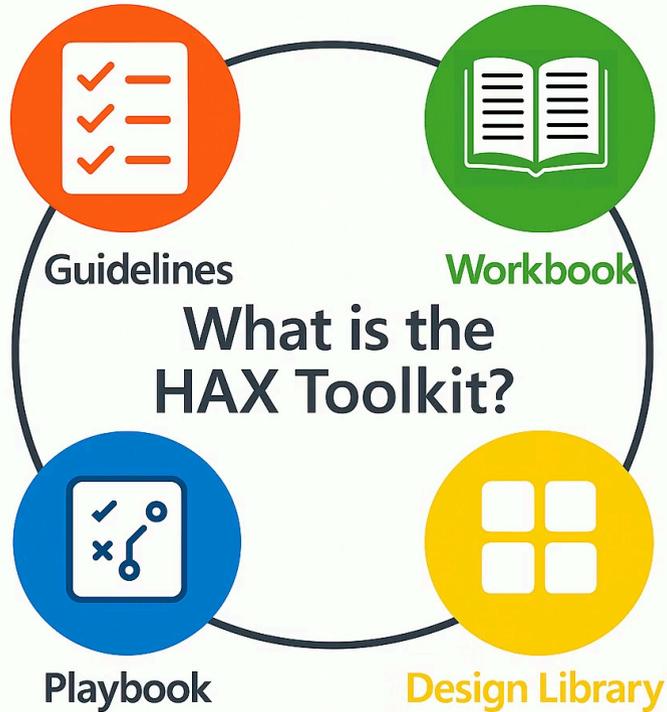
Playbook

For applications using NLP, identify common failures so you can plan for mitigating them.



Design Library

Learn how to apply the guidelines, using patterns and examples.



Guidelines for Human AI Interaction



Initially

Help users understand what the AI can do and how to start using it effectively.



During interaction

Keep users informed, in control, and supported while interacting with the AI.



When wrong

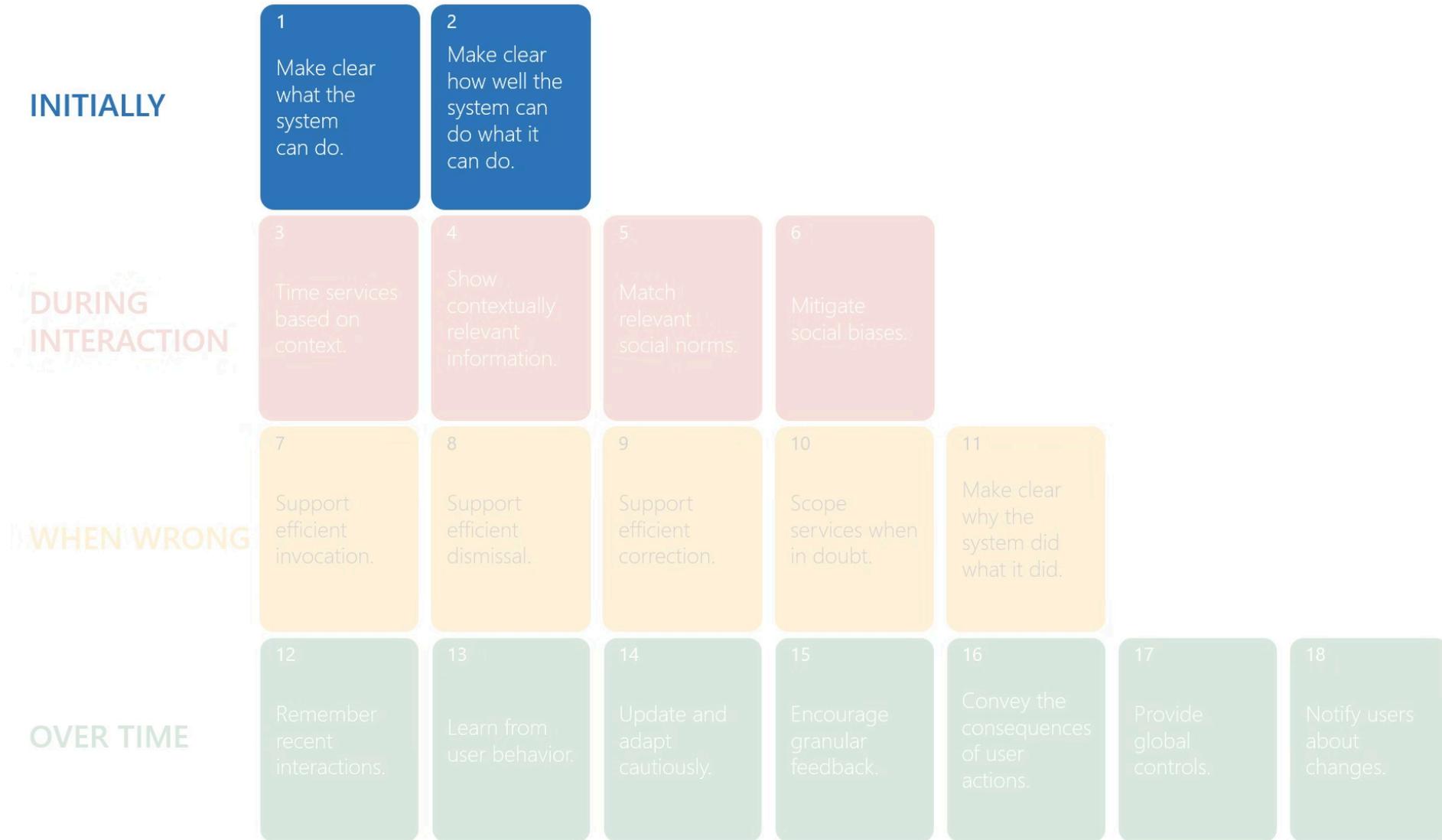
Recognize and recover gracefully from errors or misunderstandings.



Over time

Improve the experience by learning from users while maintaining transparency and user agency.

Guidelines - Initially



Guidelines - During Interaction



Guidelines - When Wrong

INITIALLY

1

Make clear what the system can do.

2

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

DURING INTERACTION

3

Time services based on context.

4

Show contextually relevant information.

5

Match relevant social norms.

6

Mitigate social biases.

WHEN WRONG

7

Support efficient invocation.

8

Support efficient dismissal.

9

Support efficient correction.

10

Scope services when in doubt.

11

Make clear why the system did what it did.

OVER TIME

12

Remember recent interactions.

13

Learn from user behavior.

14

Update and adapt cautiously.

15

Encourage granular feedback.

16

Convey the consequences of user actions.

17

Provide global controls.

18

Notify users about changes.

Guidelines - Over Time

INITIALLY

1

Make clear what the system can do.

2

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

DURING INTERACTION

3

Time services based on context.

4

Show contextually relevant information.

5

Match relevant social norms.

6

Mitigate social biases.

WHEN WRONG

7

Support efficient invocation.

8

Support efficient dismissal.

9

Support efficient correction.

10

Scope services when in doubt.

11

Make clear why the system did what it did.

OVER TIME

12

Remember recent interactions.

13

Learn from user behavior.

14

Update and adapt cautiously.

15

Encourage granular feedback.

16

Convey the consequences of user actions.

17

Provide global controls.

18

Notify users about changes.



HAX Workbook

A tool for structuring early conversations across the multiple roles - **PMs, User Researchers, Designers, Data Scientists, Engineers and Risk & Security stakeholders.** from your team or partner teams

When should we use the workbook?

When you have an idea for a new user-facing AI-based feature and you are starting to define requirements.

OR

When you have an existing user-facing AI feature or prototype that you want to evolve or improve.

How to use the Workbook

1 — Step 1: Select relevant guidelines

Identify all relevant guidelines

2 — Step 2: Imagine the impact of relevant guidelines

For each relevant guideline in step 1, determine impact. Consider how each guideline affects users, outcomes, and potential risks.

3 — Step 3: Define implementation requirements

For each high impact guideline fill the implementation requirements

4 — Step 4: Prioritize

Select guidelines to implement

5 — Step 5: Track

Track the guidelines

Design Library

Refine Results

[Clear filters](#)

Show me:

- Guidelines
- Design Patterns
- Examples

Select Guidelines

- G1: Make clear what the system can do.
- G2: Make clear how well the system can do what it can do.
- G3: Time services based on context.
- G4: Show contextually relevant information.
- G5: Match relevant social norms.
- G6: Mitigate social biases.
- G7: Support efficient invocation.
- G8: Support efficient dismissal.
- G9: Support efficient correction.
- G10: Scope services when in doubt.
- G11: Make clear why the system did what it did.
- G12: Remember recent interactions.
- G13: Learn from user behavior.
- G14: Update and adapt cautiously.
- G15: Encourage granular feedback.
- G16: Convey the consequences of user actions.
- G17: Provide global controls.
- G18: Notify users about changes.

Guideline 1

Guideline 1

Make clear what the system can do >

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

Guideline 1 > Pattern 1C

Pattern 1C: Expose system controls >

G1: Make clear what the syste...

Advertising Chatbot

E-commerce Email

Guideline 1 > Pattern 1A

Pattern 1A: Introductory blurb >

G1: Make clear what the syste...

Advertising Chatbot

E-commerce Email

Guideline 1 > Pattern 1D

Pattern 1D: Demonstrate possible system inputs >

G1: Make clear what the syste...

Advertising Chatbot

E-commerce Email

Guideline 1 > Pattern 1B

Pattern 1B: Use explanation (G11) patterns >

G1: Make clear what the syste...

Advertising Chatbot

E-commerce Email

Guideline 1 > Pattern 1E

Pattern 1E: Show a set of system outputs >

G1: Make clear what the syste...

Advertising Chatbot

E-commerce Email

Design Pattern Example

Copilot in PowerPoint makes clear what the system can do (**Guideline 1**) by exposing system controls (**Pattern 1C**)



Phishing Awareness
Helping teams recognize and respond to phishing threats effectively

Create a phishing awareness training deck with 5 slides, include a quiz.

 Copilot 

Here's a suggested outline for a 5-slide phishing awareness training deck, including a quiz. You can copy this structure into your PowerPoint and customize as needed:

Slide 1: Title Slide
Phishing Awareness
Helping teams recognize and respond to

Add more quiz questions

Include speaker notes for slides

Suggest images for each slide



Design Pattern Example

Copilot supports explainable alerts (**Guideline 11**) by allowing users to access an explanation of why the AI system behaved as it did (**Pattern 11A**).

	A	B	C	D	E	F	G	H	I	J	K	L
1	User ID	Login Time	Location	Failed Attemp	Result	AI Risk Flag						
2	u101	2025-09-0	Chicago, USA	0	Success	Normal login						
3	u102	2025-09-0	ABC city, XYZ country	5	Success	High Risk – unusual login from ABC at 2:45 AM with 5 failed attempts						
4	u103	2025-09-0	Chicago, USA	1	Success	Normal login						
5	u104	2025-09-0	New York, USA	0	Success	Normal login						
6	u105	2025-09-0	ABC city, XYZ country	4	Failed	Suspicious – failed login attempt from ABC at 3:30 AM with 4 retries						
7	u106	2025-09-0	Chicago, USA	0	Success	Normal login						
8	u107	2025-09-0	ABC city, XYZ country	6	Success	High Risk – login from ABC at 3:05 AM after 6 failed attempts						
9	u108	2025-09-0	New York, USA	0	Success	Normal login						
10	u109	2025-09-0	Chicago, USA	2	Failed	Suspicious – failed login attempt at 3:15 AM with 2 retries						
11	u110	2025-09-0	Chicago, USA	0	Success	Normal login						
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												

Copilot

u107	2025-09-03 03:05	ABC city, XYZ cou
u109	2025-09-04 03:15	Chicago, USA

These entries were flagged due to:

- Multiple failed login attempts
- Unusual login times (early morning hours)
- Logins from less common locations

Would you like a summary, visualization, or further analysis of these suspicious logins?

Design Pattern Example

Copilot in Excel encourages granular feedback (**Guideline 15**) on individual system outputs (**Pattern 15A**) when it enables the user to provide feedback for each response it generates.

County	State	State FIPS Code	County FIPS Code	FIPS	Total Population	Male Population	Female Population	Total Race Responses	White Alone	Black or African
Osceola County	Michigan	26	133	26133	23122	11803	11319	23122	21537	
Oscoda County	Michigan	26	135	26135	8338	4270	4068	8338	7856	
Otsego County	Michigan	26	137	26137	25352	12868	12484	25352	23760	
Ottawa County	Michigan	26	139	26139	298614	148156	150458	298614	255201	
Presque Isle County	Michigan	26	141	26141	13142	6578	6564	13142	12575	
Roscommon County	Michigan	26	143	26143	23639	11919	11720	23639	22144	
Saginaw County	Michigan	26	145	26145	189210	92309	96901	189210	135723	
St. Clair County	Michigan	26	147	26147	160089	79999	80090	160089	146571	
St. Joseph County	Michigan	26	149	26149	60870	30677	30193	60870	53093	
Sanilac County	Michigan	26	151	26151	40574	20463	20111	40574	38046	
Schoolcraft County	Michigan	26	153	26153	8093	4042	4051	8093	6820	
Shiawassee County	Michigan	26	155	26155	68061	33855	34206	68061	64514	
Tuscola County	Michigan	26	157	26157	53071	26977	26094	53071	49654	
Van Buren County	Michigan	26	159	26159	75681	37702	37979	75681	62002	
Washtenaw County	Michigan	26	161	26161	368394	183761	184633	368394	257686	
Wayne County	Michigan	26	163	26163	1773767	858831	914936	1773767	886114	
Wexford County	Michigan	26	165	26165	33868	17041	16827	33868	31309	
Aitkin County	Minnesota	27	1	27001	15930	8036	7894	15930	14887	
Anoka County	Minnesota	27	3	27003	367095	184597	182498	367095	283489	
Becker County	Minnesota	27	5	27005	35260	17667	17593	35260	30380	
Beltrami County	Minnesota	27	7	27007	46421	23392	23029	46421	32925	
Benton County	Minnesota	27	9	27009	41458	20690	20768	41458	36188	
Big Stone County	Minnesota	27	11	27011	5152	2612	2540	5152	4889	
Blue Earth County	Minnesota	27	13	27013	69386	35092	34294	69386	60199	
Brown County	Minnesota	27	15	27015	25826	12928	12898	25826	24319	
Carlton County	Minnesota	27	17	27017	36457	19307	17150	36457	32077	
Carver County	Minnesota	27	19	27019	108622	54415	54207	108622	95421	
Cass County	Minnesota	27	21	27021	30688	15808	14880	30688	25087	
Chippewa County	Minnesota	27	23	27023	12388	6160	6228	12388	10553	
Chisago County	Minnesota	27	25	27025	57434	29865	27569	57434	52665	
Clay County	Minnesota	27	27	27027	65628	32675	32953	65628	56650	

Copilot

MISSOURI 0100101
Montana 1105072
Nebraska 1020750

...

Copilot AI-generated content may be incorrect

I created a PivotTable summarizing the total population by state. For example, Michigan has a total population of 3,223,885 and Minnesota has 5,713,716. See the PivotTable above for more details.

Get deeper analysis results using Python

Are there any outliers in my data?

Total 'White Alone' for 'State' excluding 'Montana'

Search for content, ask a question, or tell me what you'd like to do with A1:L400

HAX Playbook

The HAX Playbook is a tool for proactively and systematically exploring common human-AI interaction failures.

The screenshot displays the HAX Playbook interface. The left sidebar, titled "Describe your AI feature", contains three sections: "Which of these systems is closest to what you are designing?" with radio buttons for Search, Recommendation (selected), Conversational AI, Text prediction and assistance, and Classification; "What is the primary input modality for your Recommendation system?" with radio buttons for Text (selected), Speech, and Action sequence; and "Will your Recommendation system have a clear way of knowing when it should trigger?" with radio buttons for Yes (selected) and No. The main area, titled "Test these scenarios", shows a total of 6 scenarios. A filter bar includes "Correct operation" (1), "Input errors" (1), "Response generation errors" (4), "Trigger errors" (0), and "Delimiter errors" (0). The list of scenarios includes: "Correct operation" (expanded to show "Normal operation" with a description), "Input errors" (expanded to show "Action sequence includes spurious events" with a list of simulation methods), "Trigger errors", "Delimiter errors", and "Response generation errors".

HAX Playbook

Export

Describe your AI feature

Restart Undo

Which of these systems is closest to what you are designing?

- Search
A search system locates items that correspond to user-specified keywords, phrases, or characters.
- Recommendation
A recommendation system makes preferred-content suggestions, sometimes by predicting the user's rating of the content.
- Conversational AI
Conversational agents interact with people through one or more exchanges of natural-language dialogue, similar to conversing with a person.
- Text prediction and assistance
Text prediction and writing assistants help with the writing process, including research, grammar, tone, and localization.
- Classification
Text or document classification categorizes text in groups organized by predefined labels.

What is the primary input modality for your Recommendation system?

- Text
The system operates on text that a person types, or on text-bearing documents.
- Speech
The system operates on speech that a person speaks into a microphone, or on audio recordings.
- Action sequence
An action sequence is a sequence of one or more events, or choices, that triggers a system response.

Will your Recommendation system have a clear way of knowing when it should trigger?

- Yes
The system is triggered by a button, or has some other clear signal that is unlikely to be misinterpreted.
- No

Test these scenarios

Total scenarios: 6

Correct operation 1 Input errors 1 Trigger errors 0 Delimiter errors 0

Response generation errors 4

- ▼ Correct operation
 - ▼ Normal operation
As usual, include scenarios where the system demonstrates normal operation.
- ▼ Input errors
Your system will take an action sequence as its primary input modality. Here are some scenarios that you can consider including in your user-testing protocol.
 - ▼ Action sequence includes spurious events
People may misclick or trigger accidental actions which they may then try to undo. The resultant action sequence will contain spurious events that can confuse the system. Here are some ways to simulate spurious events:
 - Introduce accidental clicks on nearby, or neighboring buttons or links.
 - Click buttons that have similar iconography to the intended buttons.
 - Include all actions needed to recover from, or undo, the misclicks.
- ▼ Trigger errors
Your system will have a clear way of knowing when to **trigger**.
- ▼ Delimiter errors
Your system will have a clear way of **delimiting** the user's input.
- ▼ Response generation errors
Your system will **select** list items from an existing set. However, failure to understand user's intent or context can result in

Microsoft HAX Toolkit

1 Guidelines

Best practices for how AI systems should behave during interaction helping ensure trust, usability, and resilience.

3 Design Library

Learn the Guidelines for Human-AI Interaction and how to apply them, using patterns and examples.

2 HAX Workbook

Work together with your team to prioritize which Guidelines to implement in your product.

4 HAX Playbook

For applications using natural language processing, identify common failures so you can plan for mitigating them. Use it like a red-team tool for AI: anticipate failure modes, adversarial inputs, and design mitigations before deployment.



**You have more power at your fingertips than
ever before.**



With Great Power Comes.....

Thank You – Let's Continue the Conversation

Scan this QR code



"We're not just building systems; we're building futures."