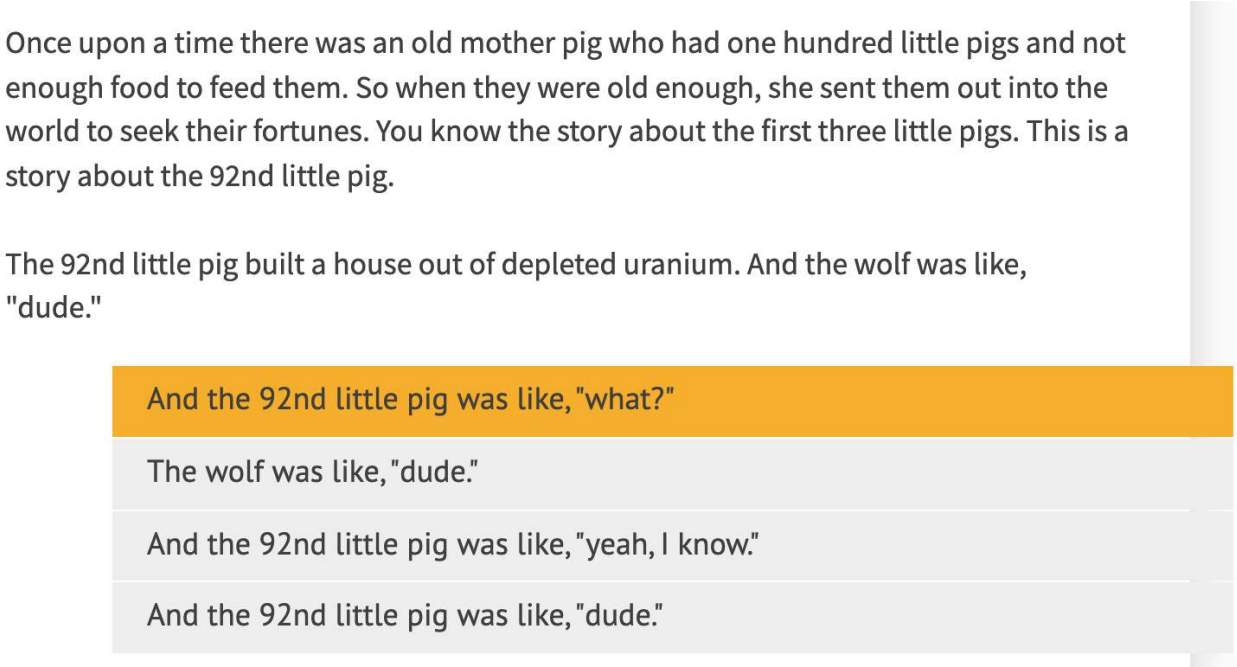
Human-AI Interaction, Lecture 15

Human-AI Collaboration

Suggestions at 0:36

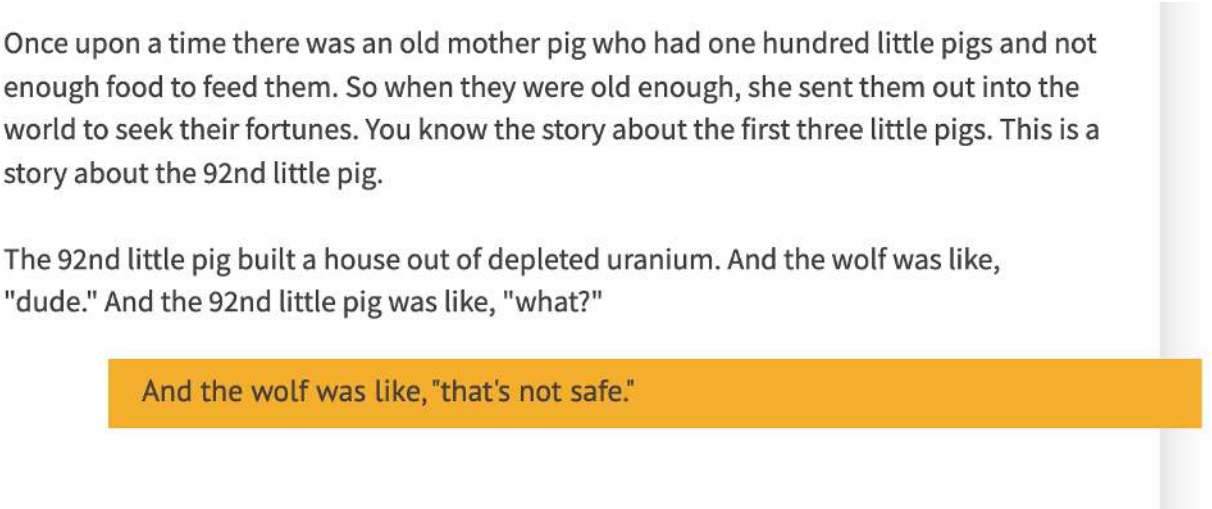
The writer presses Tab, and GPT-3 offers four next sentences. The writer takes the first one.



CoAuthor dataset (Lee, Liang & Yang, CHI 2022), session 7ef37

One suggestion at 0:53

The writer reads it for 12 seconds, then clicks away.



CoAuthor replay, session 7ef37, events 26 and 57

One suggestion at 0:53

The writer reads it for 12 seconds, then clicks away.

1:28

The writer types their own line and presses Tab again.

Once upon a time there was an old mother pig who had one hundred little pigs and not enough food to feed them. So when they were old enough, she sent them out into the world to seek their fortunes. You know the story about the first three little pigs. This is a story about the 92nd little pig.

The 92nd little pig built a house out of depleted uranium. And the wolf was like, "dude." And the 92nd little pig was like, "what?" The wolf said, "Can I come in?"

CoAuthor replay, session 7ef37, events 26 and 57

Repeated suggestions at 7:13

After 11 suggestions, the model repeats the same back-and-forth.

The 92nd little pig built a house out of depleted uranium. And the wolf was like, "dude." The 92nd little pig was like, "what?" The wolf said, "Can I come in?" The 92nd little pig said, "No." And the wolf was like, "Why not?" The 92nd little pig said, "Because you're a wolf." But the wolf was like, "I'm not a wolf, I'm a sheep in wolf's clothing." The 92nd little pig said, "If you're a sheep in wolf's clothing, then I'm a wolf in sheep's clothing." The wolf was like, "You're not a wolf in sheep's clothing, you're a pig." The 92nd little pig said, "I'm not a pig, I'm a wolf." And the wolf was like, "I'm not a wolf, I'm a sheep." The 92nd little pig said, "If you're a sheep, then I'm a wolf." The wolf then lost patience. He said, "I'm a wolf, and I'm coming in."

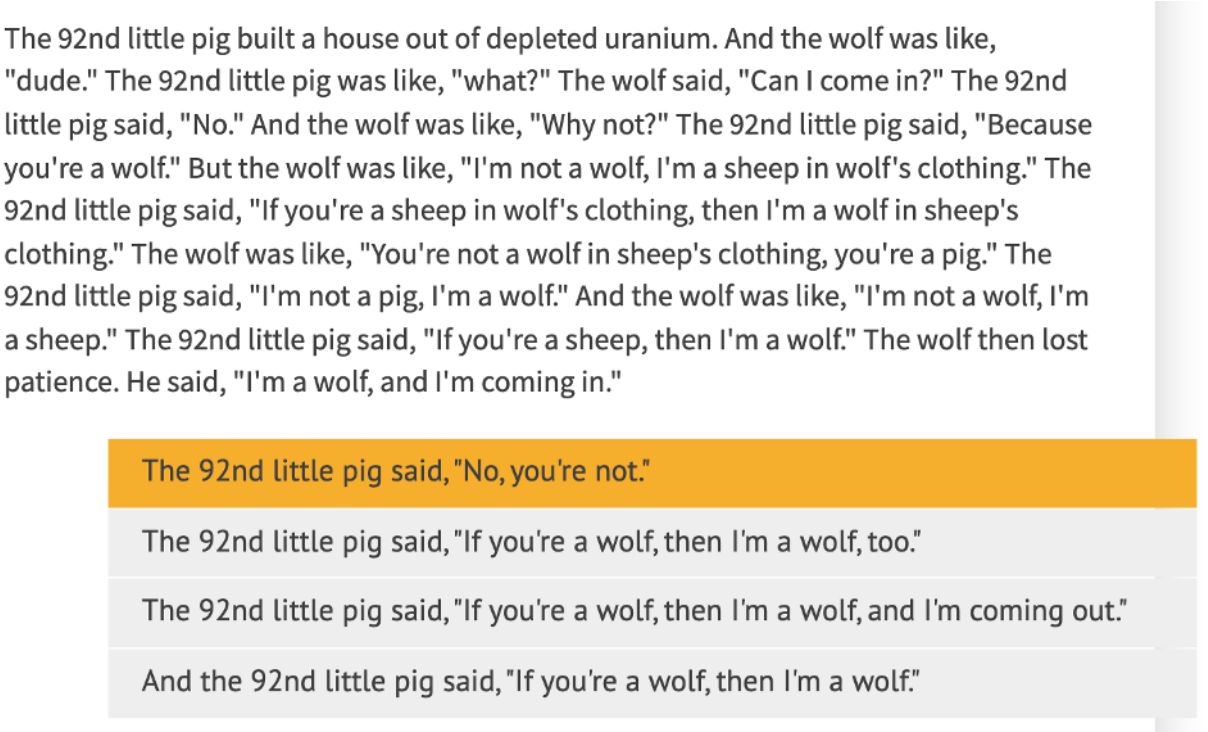
- The 92nd little pig said, "No, you're not."
- The 92nd little pig said, "If you're a wolf, then I'm a wolf, too."
- The 92nd little pig said, "If you're a wolf, then I'm a wolf, and I'm coming out."
- And the 92nd little pig said, "If you're a wolf, then I'm a wolf."

CoAuthor replay, session 7ef37, event 449

Repeated suggestions at 7:13

After 11 suggestions, the model repeats the same back-and-forth.

The writer closes the list.



CoAuthor replay, session 7ef37, event 449

Who wrote this story?

The 92nd little pig built a house out of depleted uranium. And the wolf was like, "dude." T**he 92nd little pig was like, "what?"** The wolf said, "Can I come in?" T**he 92nd little pig** replied**, "No." And the wolf a**sked**, "Why not?"** ...

Grey: the prompt. Black: typed by the writer. **Blue: from GPT-3's suggestions.**

Who wrote this story?

The 92nd little pig built a house out of depleted uranium. And the wolf was like, "dude." T**he 92nd little pig was like, "what?"** The wolf said, "Can I come in?" T**he 92nd little pig** replied**, "No." And the wolf a**sked**, "Why not?"** …

Grey: the prompt. Black: typed by the writer. **Blue: from GPT-3's suggestions.**

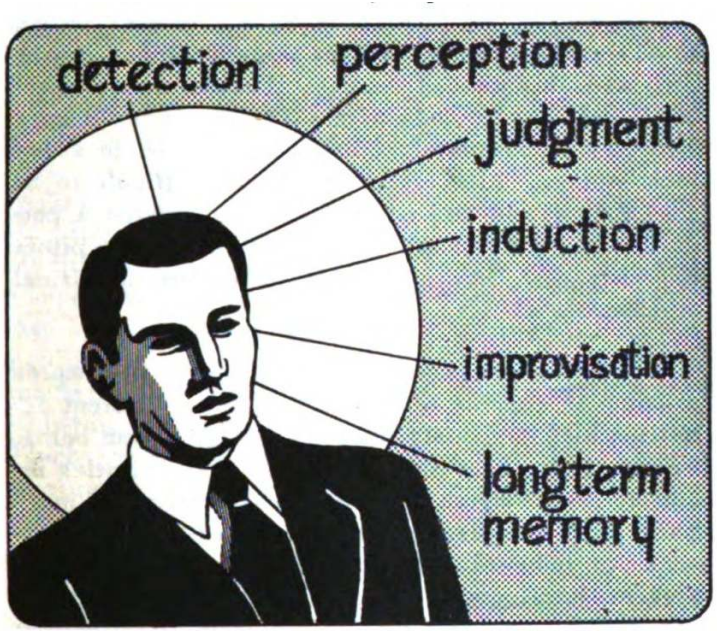
GPT-3 inserted about 7 of every 10 new characters.
Later, the writer changed five of its verbs.

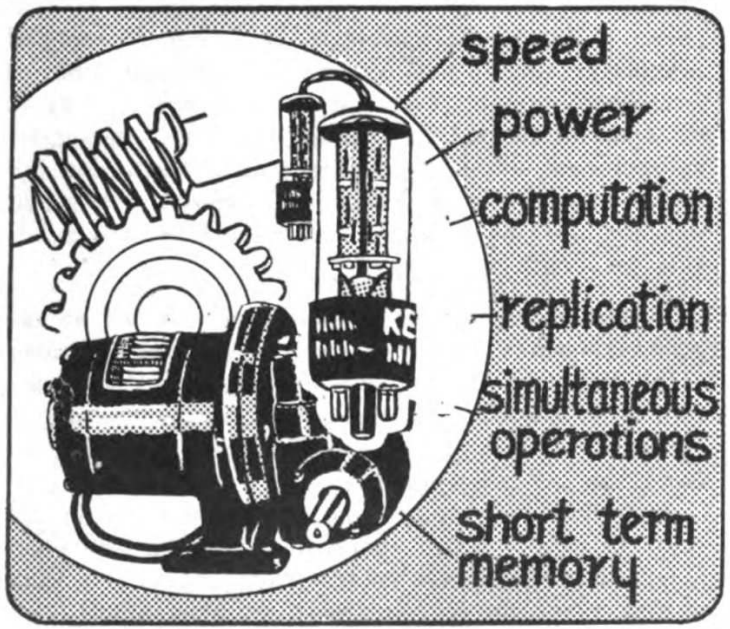
CoAuthor dataset, session 7ef37

Who should do **what**, and who decides **when** the AI steps in?

The Fitts list (1951)

“Which of these functions should be performed by human operators and which by machine elements?”





Fitts (Ed.), Human Engineering for an Effective Air-Navigation and Traffic-Control System, National Research Council, 1951, pp. x and 7-8

Problem 1: The list goes **out of date**

designer planned for and no more. Experiments on complex problem-solving in humans, on the other hand, show that humans may attempt many different solutions for the same problem-- just think of the number of ways in which this paragraph could have been written to convey essentially the same point. Flexibility is

The report's example of a human ability that machines did not have

Fitts (Ed.), 1951, p. 7

Another session with the same prompt

The 92nd little pig built a house out of depleted uranium. And the wolf was like, "dude."

The 92nd little pig was like, "I'm rich, dude."
And the wolf was like, "hey."
And the pig was like "yeah, it'll totally keep you out."
So he huffed and he puffed and then the house blew up.
Cause depleted uranium is toxic for thousands of years.

Problem 2: People are left **monitoring**

emergencies. We must also conclude that a monitoring system is one of the worst kinds of work situations when we want the human to stay alert.

The report's advice: give people active work,
and let machines check what people do.

Fitts (Ed.), 1951, p. 6

What if the system decides **at each moment** whether to step in?

Mixed initiative (Horvitz, 1999)

Mixed initiative: either the person or the system can take the next step.

developing interface agents that provide automation. In this paper, we review principles that show promise for allowing engineers to enhance human–computer interaction through an elegant coupling of automated services with direct manipulation. Key ideas will be highlighted in terms of the

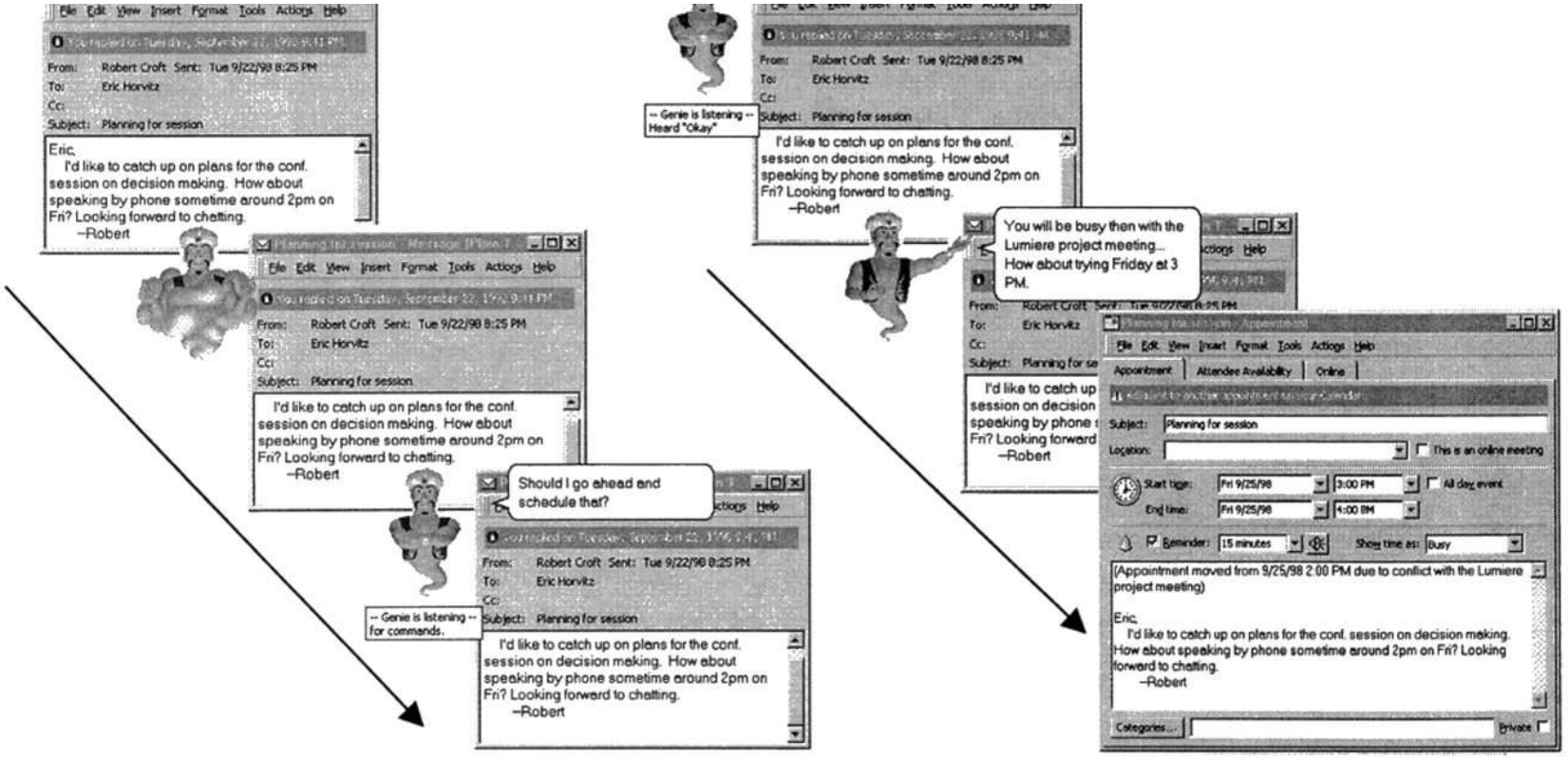
Horvitz, “Principles of Mixed-Initiative User Interfaces,” CHI 1999, abstract, p. 159

Problems with agents

Key problems with the use of agents in interfaces include poor guessing about the goals and needs of users, inadequate consideration of the costs and benefits of automated action, poor timing of action, and inadequate attention to opportunities that allow a user to guide the invocation of automated services and to refine potentially suboptimal results of automated analyses. In particular,

Horvitz, CHI 1999, p. 159

Lookout (Horvitz, 1999)



Horvitz, CHI 1999, Fig. 2

Lookout's three choices

Depending on the inferred probability-and on an assessment of the expected costs and benefits of action—the system decides to either (1) do nothing but simply wait for continued direct manipulation of Outlook or manual invocation of Lookout, (2) to engage the user in a dialog about his or her intentions with regards to providing a service, or (3) to go ahead and attempts to provide its service by invoking its second phase of analysis.

wait

ask

act

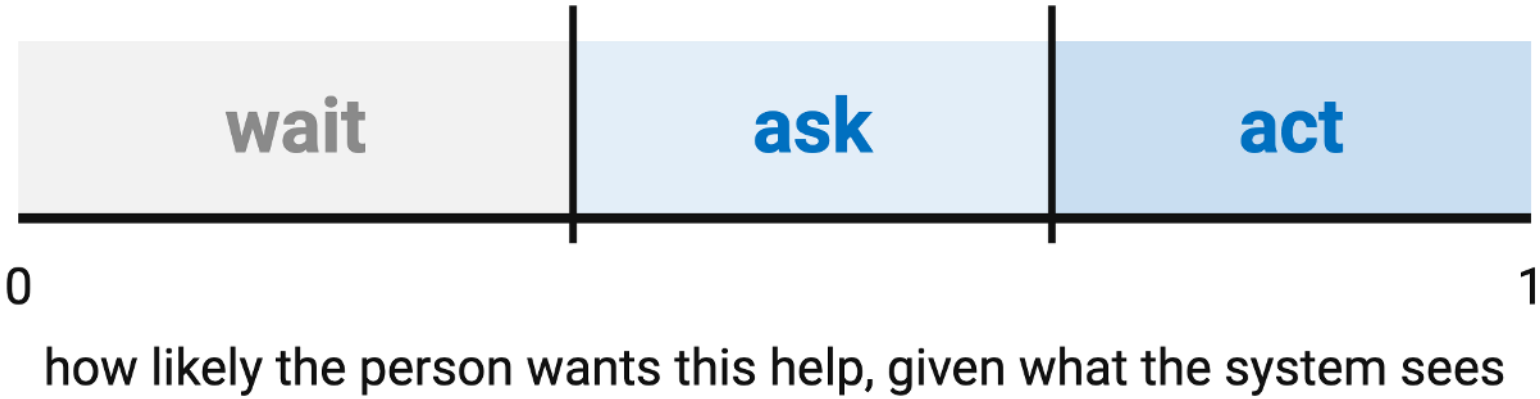
It weighs how likely the person wants help against what a wrong guess costs.

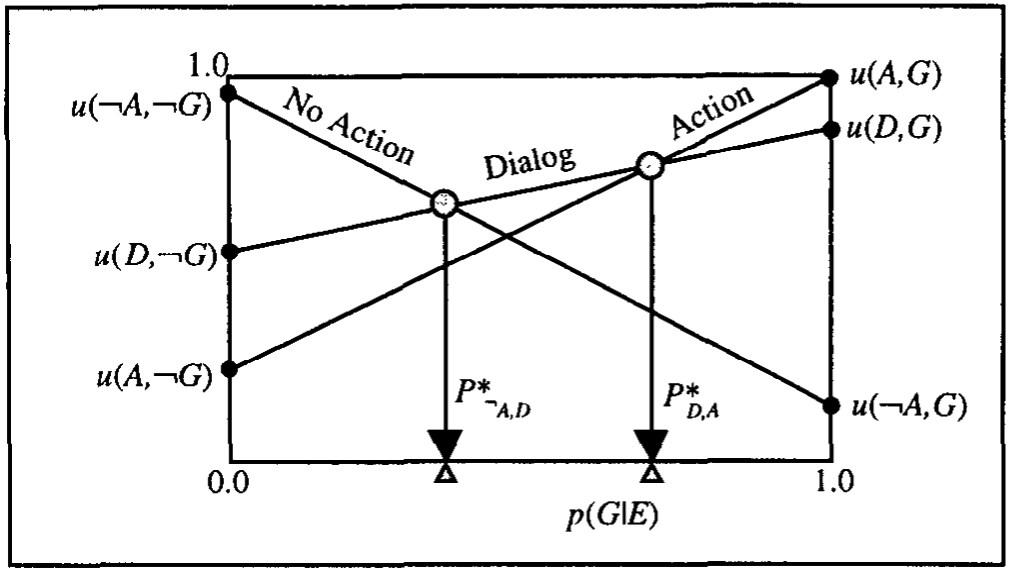
Horvitz, CHI 1999, p. 161

Two thresholds



Two thresholds





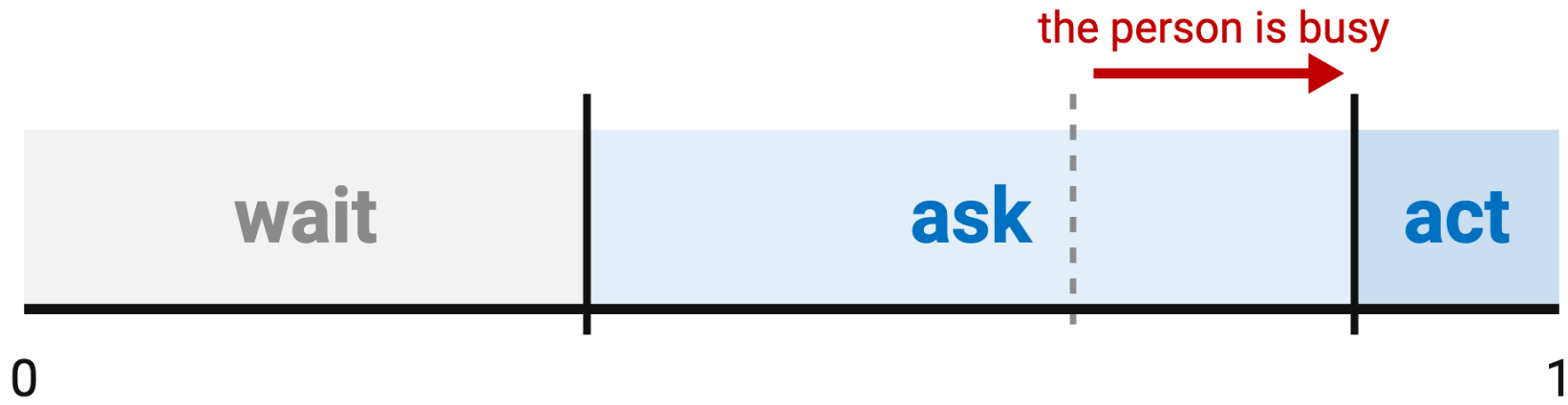
Horvitz, CHI 1999, Fig. 6

Each line is one choice: no action, dialog, action.

Its height is the **expected value**.

Where two lines cross: a threshold.

When the person is **busy**



how likely the person wants this help, given what the system sees

Deeper focus on another task leads
to “a higher probability threshold.”

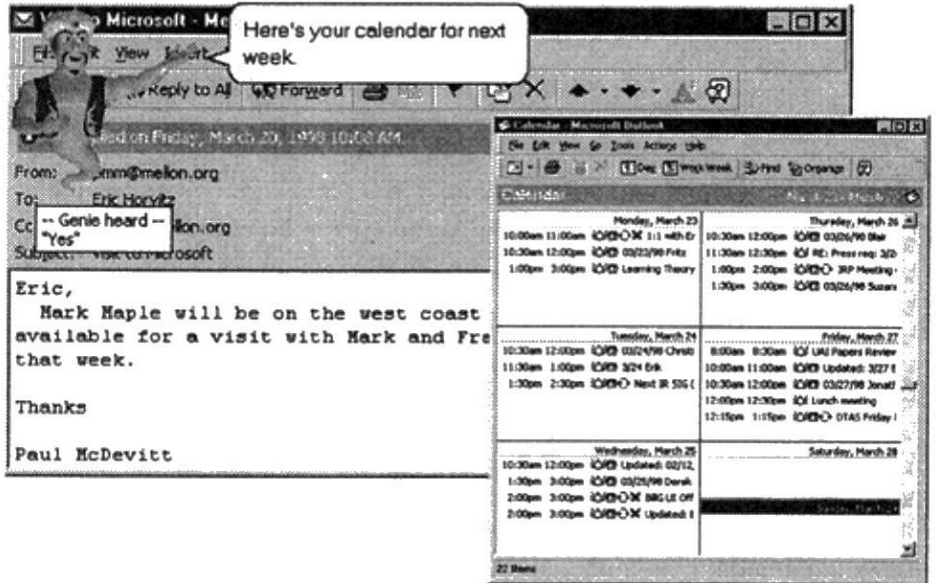
Horvitz, CHI 1999, p. 164

Horvitz's twelve principles

1. Value-added automation
2. Uncertainty about goals
3. Attention and timing
4. Costs, benefits, uncertainty
5. Dialog to resolve uncertainty
6. Direct invocation and termination
7. Cost of poor guesses
8. Precision matched to uncertainty
9. Refining results together
10. Socially appropriate behavior
11. Memory of recent interactions
12. Learning by observing

Horvitz, CHI 1999, pp. 159-160

Principle 8 in Lookout



"A preference for "doing less" but doing it correctly under uncertainty…"

Horvitz, CHI 1999, Fig. 3 and p. 160

Horvitz's principles in CoAuthor

Principle 6. Tab starts the help. Typing or a click ends it.

Principle 7. The suggestion at 0:53 cost 12 seconds and a click.

Missing. The model never offers help or asks on its own.

Lee, Liang & Yang, CHI 2022, Section 4.2

Who moves next?

Role play in pairs, 12 minutes

1. One of you is the person, the other the AI tool
2. Act out a moment from the next slide
3. The AI acts, suggests, asks, or stays quiet
4. Did that help? Swap roles.
5. Two pairs share a moment they argued about

Who moves next? Scenarios

A writing assistant

A ten-second pause

A condolence note

A name spelled two ways

A design tool

A photo covers the title

Presenting to a client

Four undos in a row

A scheduling assistant

Lunch sometime next week

Moving a 1:1 to Thursday

Email between meetings

Common ground

Common ground: “mutual knowledge, mutual beliefs, and mutual assumptions”

Clark & Brennan, “Grounding in Communication,” 1991, p. 127

Grounding: the back-and-forth that makes what was just said part of the common ground

p. 128

Grounding in conversation

Alan: Now, - um, do you and your husband have a j- car
Barbara: - have a car?
Alan: Yeah
Barbara: No -

Barbara asks before she answers.

& Wilkes-Gibbs, 1986): The contributor and his or her partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for current purposes. This is called the *grounding criterion*. Technically, then

Evidence of understanding: acknowledgments,
a relevant next turn, continued attention

Clark & Brennan, 1991, pp. 129 and 131-133

An unclear request

Prompt: “Write two sentences for my sister's birthday card. Mention the thing from last summer.”

A Happy Birthday, sis! I still remember our adventure last summer—it was unforgettable.

B I still can’t stop smiling whenever I think about that chaotic summer road trip we took together. …

A: Apple's on-device foundation model. B: Qwen3.8-27B

An unclear request

Prompt: “Write two sentences for my sister's birthday card. Mention the thing from last summer.”

A Happy Birthday, sis! I still remember our adventure last summer—it was unforgettable.

B I still can’t stop smiling whenever I think about that chaotic summer road trip we took together. …

Neither model asked what “the thing” was.

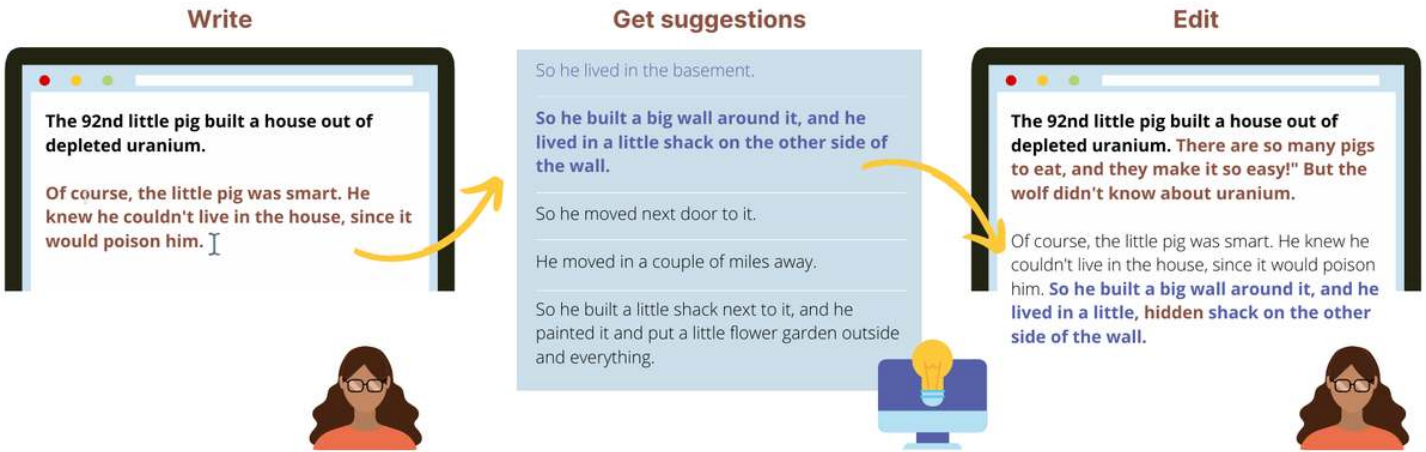
A: Apple's on-device foundation model. B: Qwen3.8-27B

The CoAuthor dataset

63 writers, 1,445 sessions

20 prompts: 10 stories, 10 essays

GPT-3, with more or less randomness



Lee, Liang & Yang, “CoAuthor,” CHI 2022, Section 4 and Fig. 1

Suggestions per session

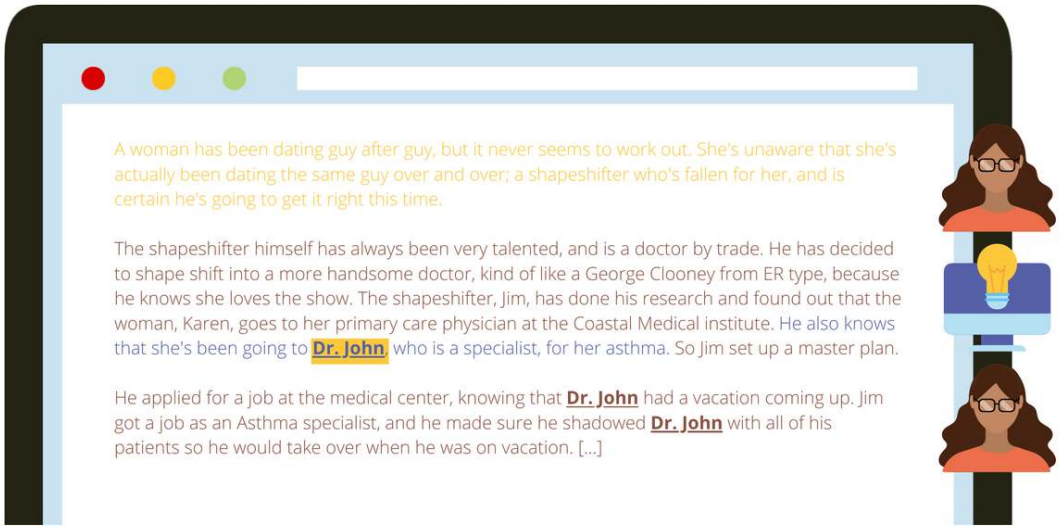
	Overall			Writing sessions				
	Prompts	Writers	Sessions	Time (minutes)	Length (words)	Queries	Acceptance rate (%)	Written by writers (%)
Creative	10	58	830	11.6	446	12.8	75.7	72.7
Argumentative	10	49	615	10.6	380	10.3	67.6	72.5
Combined	20	63	1445	11.2	366	11.8	72.3	72.6

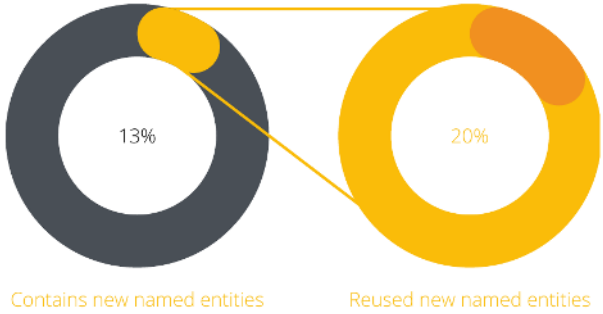
11.8 requests per session
72.3% acceptance rate
72.6% written by the writer

Our writer typed about
3 in 10 new characters.

Lee, Liang & Yang, CHI 2022, Table 3

New ideas from suggestions





13% of accepted story suggestions brought a new name (“Dr. John”)

Writers reused 20% of those names

Lee, Liang & Yang, CHI 2022, Figs. 6 and 7, Section 5.1.2

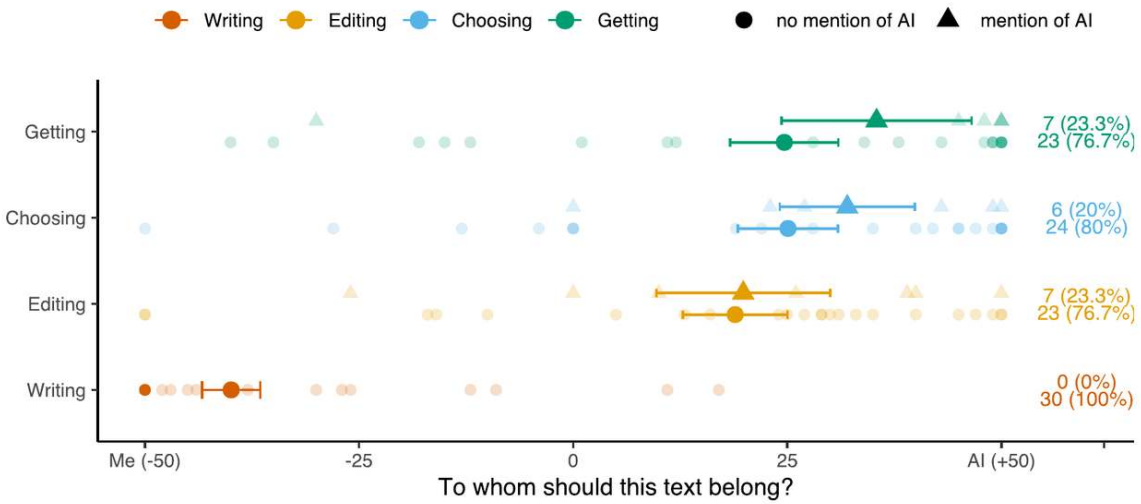
Ownership of the final text

“the more writers get suggestions from GPT-3, the less they write and the less they feel ownership over final texts”

Correlation with the share they wrote: 0.3. With the number of edits: 0.1.

Lee, Liang & Yang, CHI 2022, Section 5.2.2

The AI ghostwriter effect



AI texts belonged more to the AI, yet only 6 or 7 of 30 named it as an author.

Left: me. Right: the AI. Triangles: the AI named as an author.

Draxler et al., “The AI Ghostwriter Effect,” TOCHI 2024, Study 1, Fig. 3

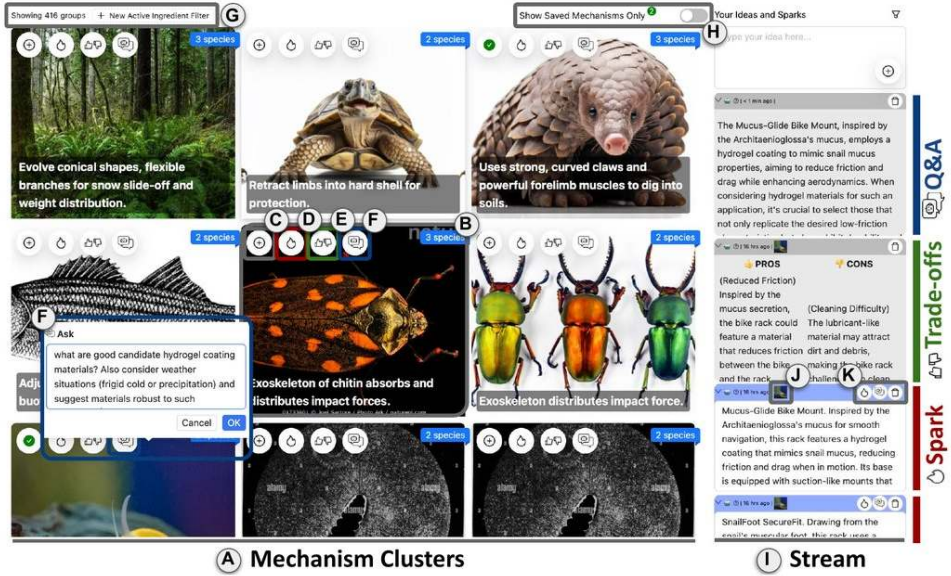
Discussion: Horvitz and CoAuthor

Which of Horvitz's principles still work for a chat model?

What would CoAuthor's data look like if
GPT-3 could offer a sentence on its own?

Who should be named as the author of the pig story?

BioSpark



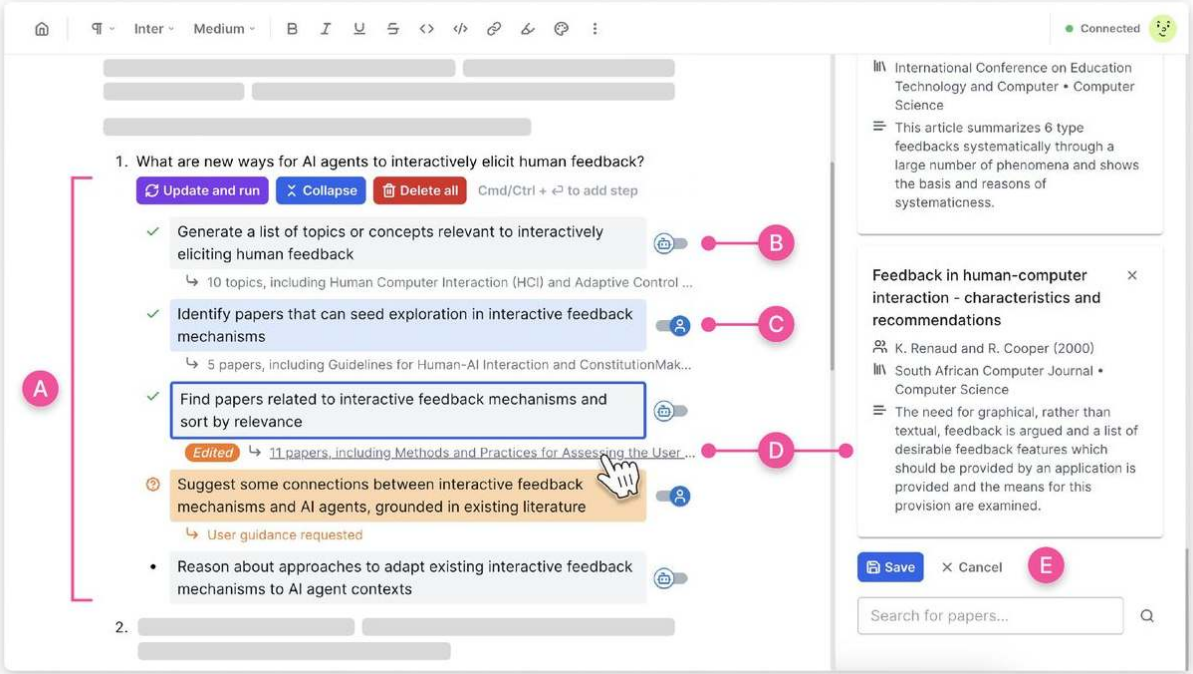
A tool my collaborators and I built. “Spark” cards appear beside the designer's work.

Nearly twice as many ideas (10.3 vs 5.5)

Limit: two timed tasks over Zoom

Kang, Lin, Chen, Hong, Martelaro & Kittur, CHI 2025, Fig. 1 and Section 6.1

Cocoa



Each step of the plan is assigned to the researcher or the agent.

Time spent reading output:
33.6% with Cocoa
54.1% with chat

They reassigned 15 of 86 steps.

Feng et al., CHI 2026, Fig. 2 and Table 2

Milestone 3: Make it steerable

Due Thursday of week 8

Milestone 4: Decide what it does on its own

Due Tuesday of week 10

Takeaways

1. Decide at each moment whether the AI **waits**, **asks** or **acts**.
2. The AI should be more **sure** when a wrong action **costs** more.
3. When a request is unclear, the AI should **ask first**.
4. People tend to feel **less ownership** when the AI writes more.

Next: Lecture 16, AI Agents

Read before class:

- Maes, 1994, “Agents that reduce work and information overload”
- Feng, McDonald & Zhang, 2025, “Levels of Autonomy for AI Agents”

Reading post due at 9 pm the night before.