Human-AI Interaction, Lecture 12

Communicating Intent

Get exactly this image

Round 1, in pairs, 6 minutes

Words only



FLUX.1 [schnell]

How do people tell an AI **what they want**?

First: what could your pair not get with words?

Lecture 1: the gulf of execution

In most software, you pick an operation.

With a chat model, you **describe the outcome**.

“Natural language programming”

We know in the meantime that the choice of an interface is not just a division of (a fixed amount of) labour, because the work involved in co-operating and communicating across the interface has to be added. We know in the meantime --from sobering experience, I may add-- that a change of interface can easily increase at both sides of the fence the amount of work to be done (even drastically so). Hence the increased preference

E. W. Dijkstra, “On the foolishness of ‘natural language programming’,” EWD 667, 1978

What would this prompt give you?

“Write a toast for Taylor”

Here are a few options for a toast for Taylor, depending on the occasion …

Option 1: Heartfelt & General …
Option 2: Short & Sweet …
Option 3: Funny & Lighthearted …

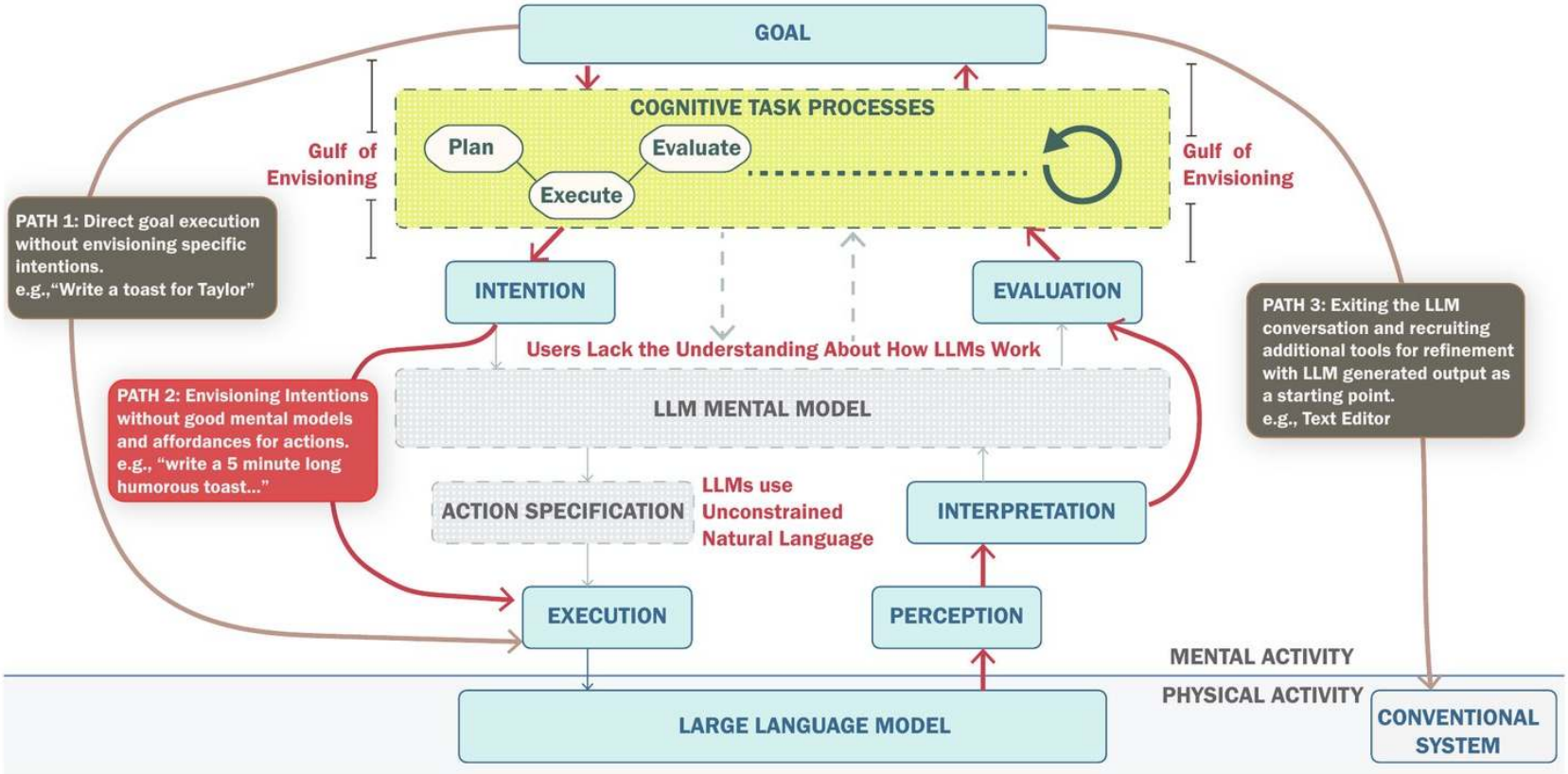
Prompt: Subramonyam et al., CHI 2024. Reply: Qwen3.8-27B

Where did the **second prompt** come from?

“Write a heartwarming toast for my best friend Taylor's retirement party, about 5 minutes long, include a humorous twist, and wish them well on the golf course”

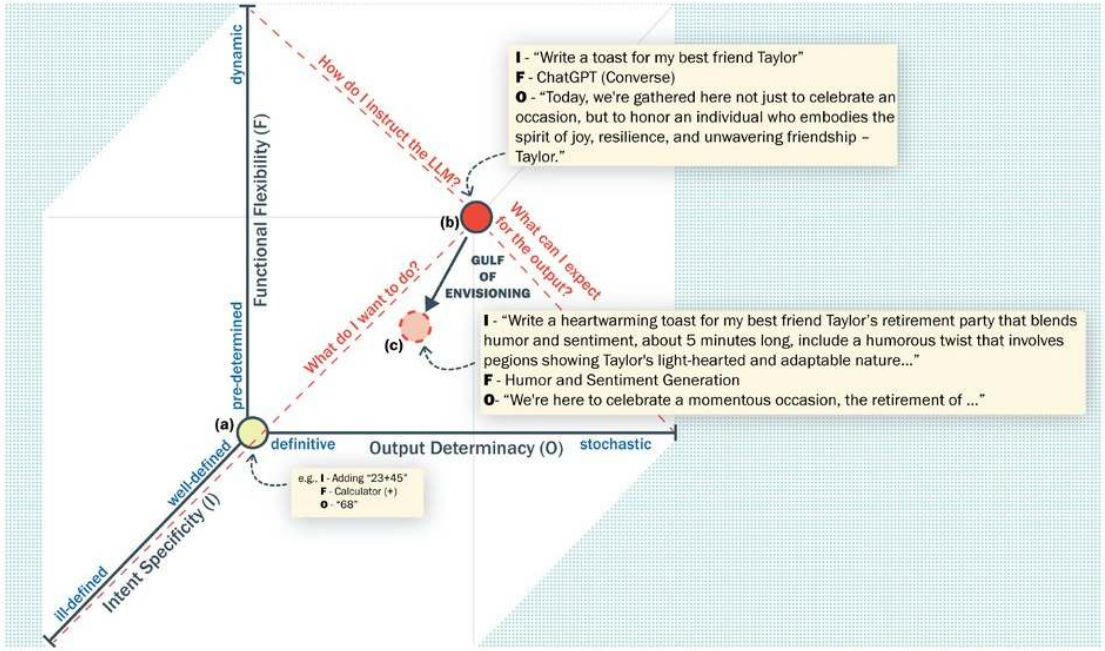
Subramonyam et al., CHI 2024

The gulf of envisioning



Subramonyam, Pea, Pondoc, Agrawala & Seifert, CHI 2024, Fig. 3

Gulf of envisioning: three gaps



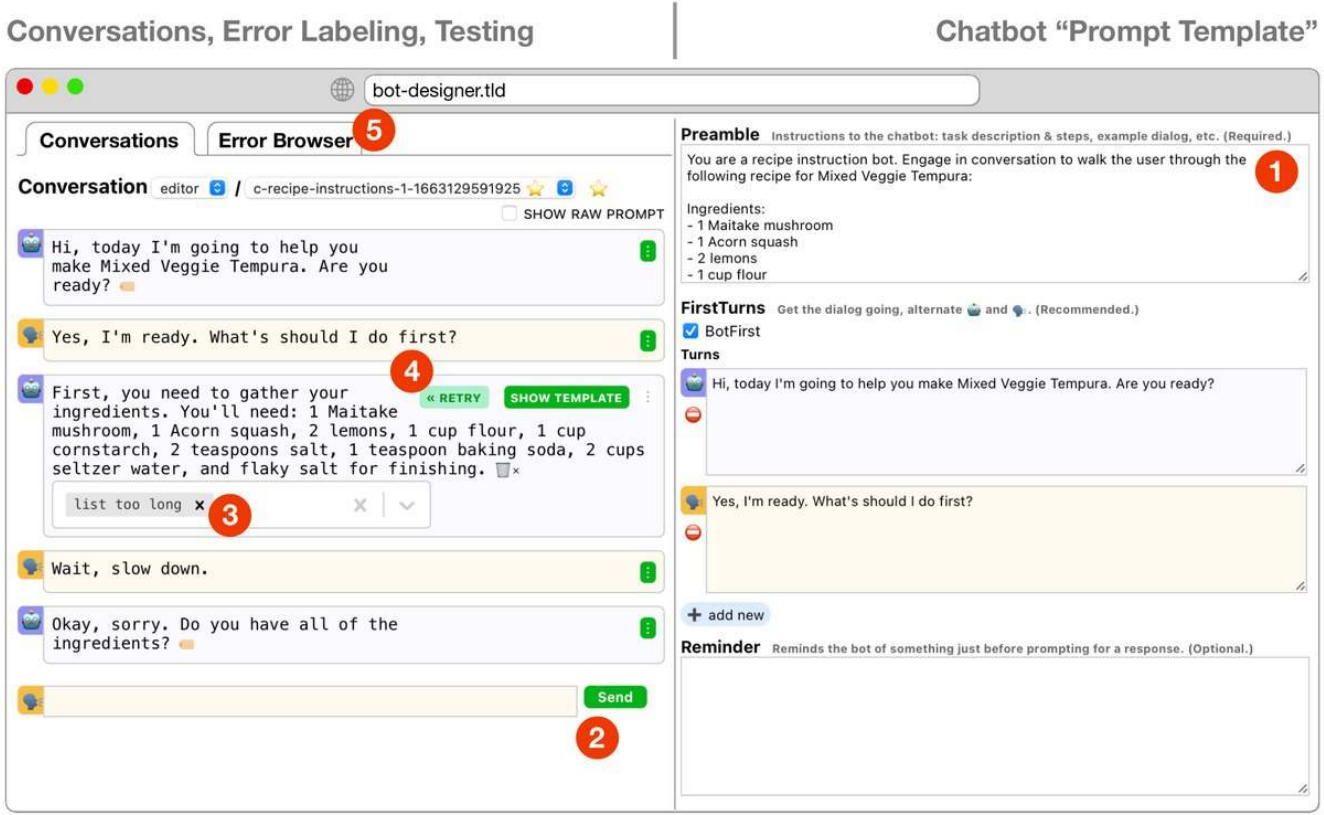
Subramonyam et al., CHI 2024, Fig. 1

Capability gap
What should the task be?

Instruction gap
How do I say it?

Intentionality gap
What should I expect?

The BotDesigner study



Ten people with little prompting experience.

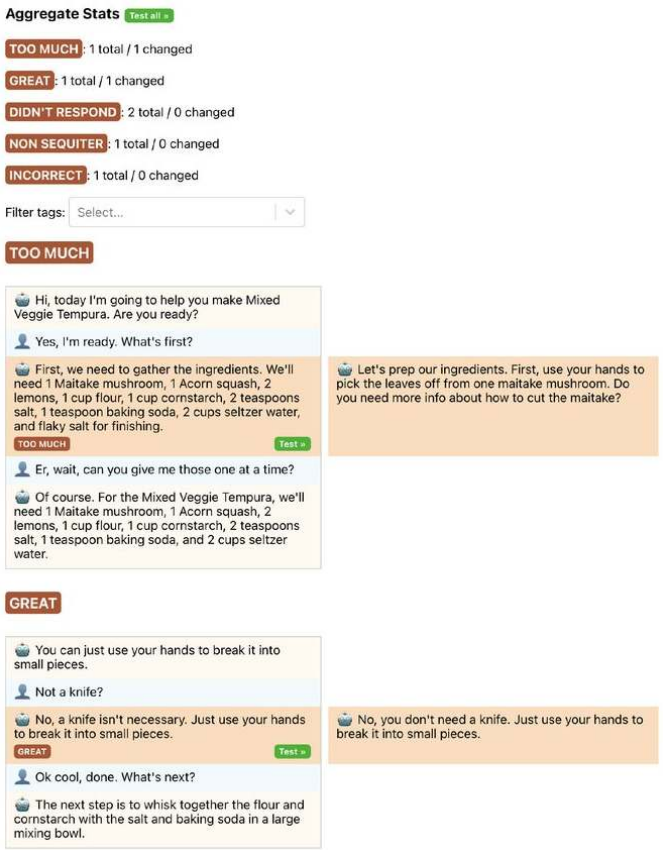
Zamfirescu-Pereira, Wong, Hartmann & Yang, “Why Johnny Can't Prompt,” CHI 2023, Fig. 1

Testing a prompt

It failed once:

“ok, I guess it doesn't like to tell jokes.” (P2)

None of them used the panel on the right.



Zamfirescu-Pereira et al., CHI 2023, Sections 4.1-4.3 and Fig. 2

Treating the bot like a person

P8 began every instruction with “please”.

P9 and P10 expected it to remember earlier chats.

P4 mixed steps for the cook and for the bot.

Zamfirescu-Pereira et al., CHI 2023, Section 4.3.1

“Don't say …”

“Don't say that when you use your hands you get a little bit more of that natural flavor in there.”

The line had **no discernible effect**.

Zamfirescu-Pereira et al., CHI 2023, Section 4.3.3

Instructions or examples?

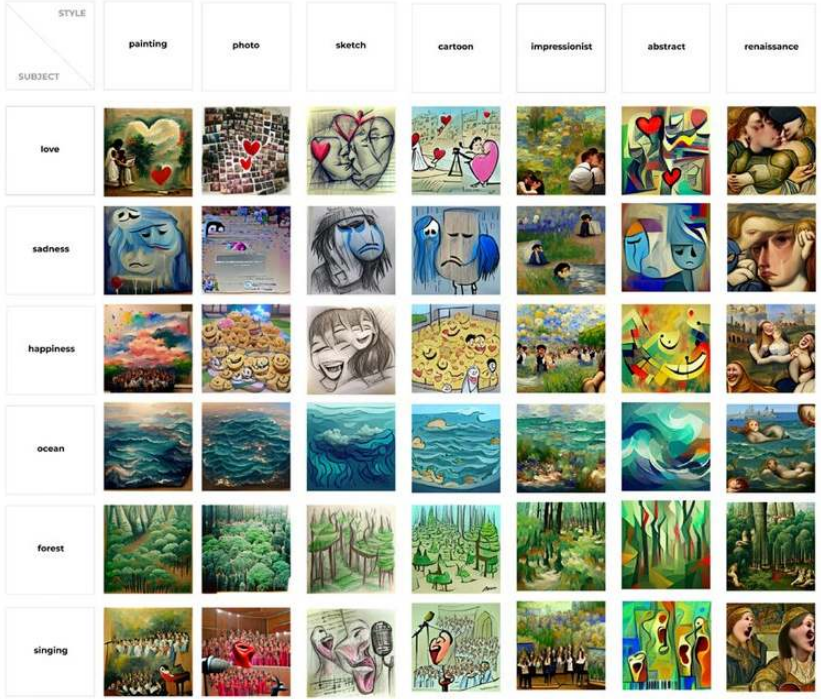
Instruction: “Tell some jokes”

Example: say “Isn't that a line from a Woody Allen movie?”

Examples worked, but **none kept using them**.

Zamfirescu-Pereira et al., CHI 2023, Section 4.3.1

Image prompts: subject and style



Liu & Chilton, CHI 2022, Fig. 1

Rewording a prompt with the same subject and style words made no significant difference.

Round 2: anything but the target

6 minutes. Try a reference picture, several images, or an edit.

Rate each round's best image, 1 to 5.



The target's prompt

“a flat illustration of a red bicycle leaning against a pale yellow wall, a **round window** above it with an orange cat asleep on the sill, three small cacti in pots on the ground, long morning shadows, soft pastel colors”



FLUX.1 [schnell]

Seeds 8, 9 and 10









The shutters and the basket came from **seed 7**.

FLUX.1 [schnell]

Definition of a prompt

A prompt is an input to a Generative AI model, that is used to guide its output (Meskó, 2023; White et al., 2023; Heston and Khun, 2023; Hadi et al., 2023; Brown et al., 2020). Prompts may consist of text, image, sound, or other media. Some ex-

Schulhoff et al., "The Prompt Report," arXiv 2024, Section 1.1

By this definition, a picture or a sound can also be a prompt.

Few-shot examples

[English] Hello! [French] Bonjour!
[English] Do you like the weather? [French]

Model: “Vous aimez le temps?”

A reference picture and a short instruction



reference: the target

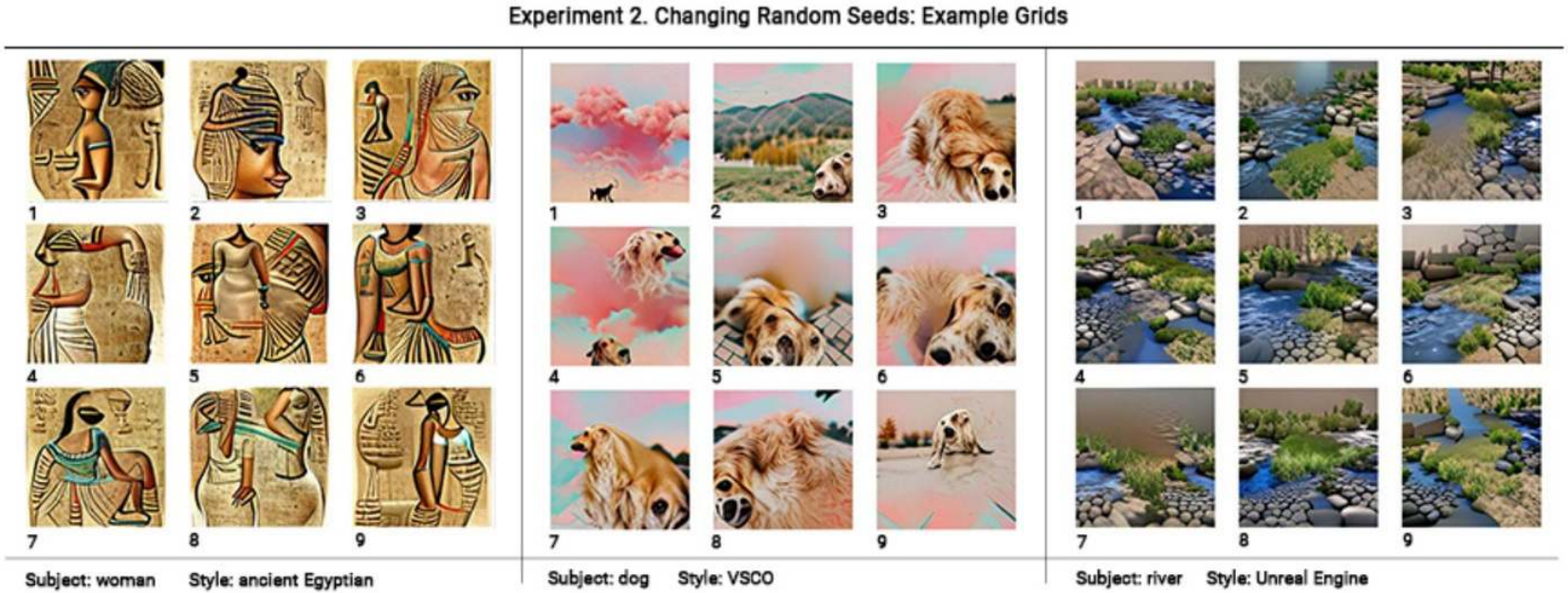
the same picture, but at night



result

FLUX.2 [klein] 9B

Several outputs from one prompt



Look at 3 to 9 seeds before judging a prompt.

Liu & Chilton, CHI 2022, Fig. 3

Editing the closest image



seed 8, the closest?

add open mauve shutters on both sides of the window, put a basket on the front of the bicycle, and keep only three cactus pots

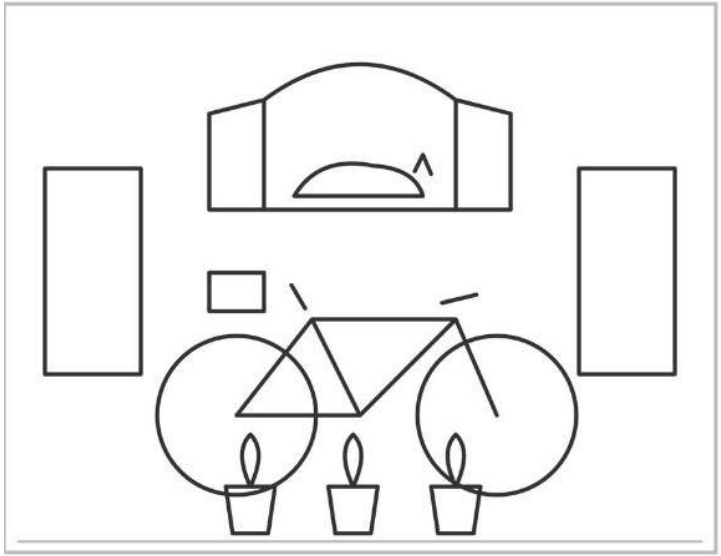


result

still four pots

FLUX.2 [klein] 9B

A sketch for the layout



the sketch

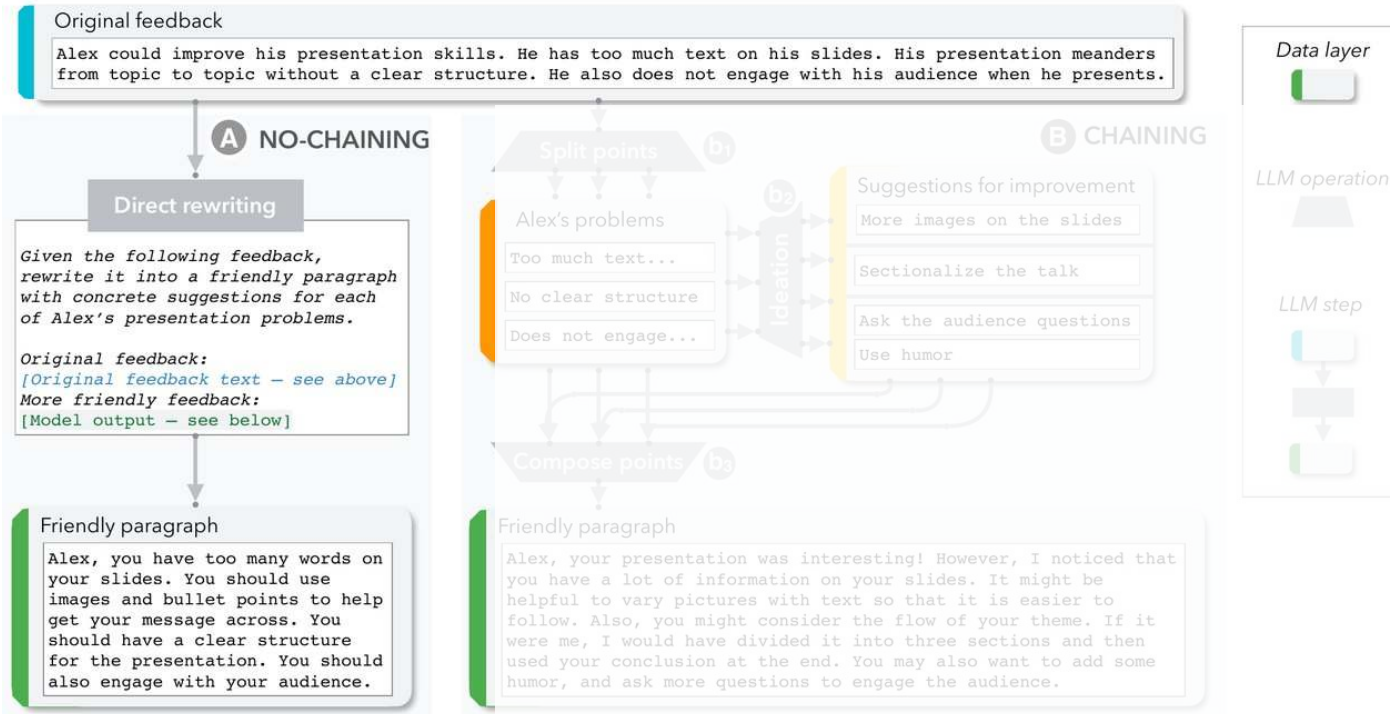


result

The layout followed the sketch.

FLUX.2 [klein] 9B

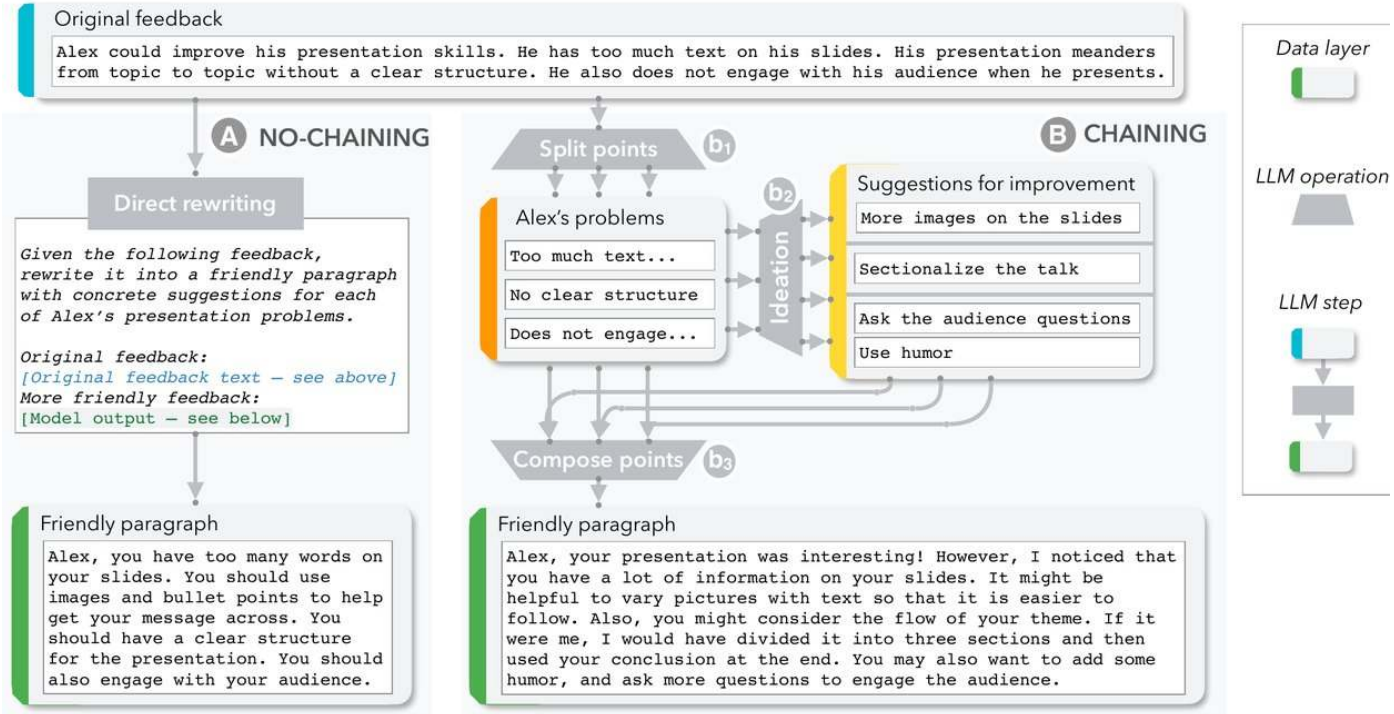
One prompt for the whole task



Wu, Terry & Cai, "AI Chains," CHI 2022, Fig. 1

With one prompt (A), the paragraph did not give concrete suggestions for all three problems.

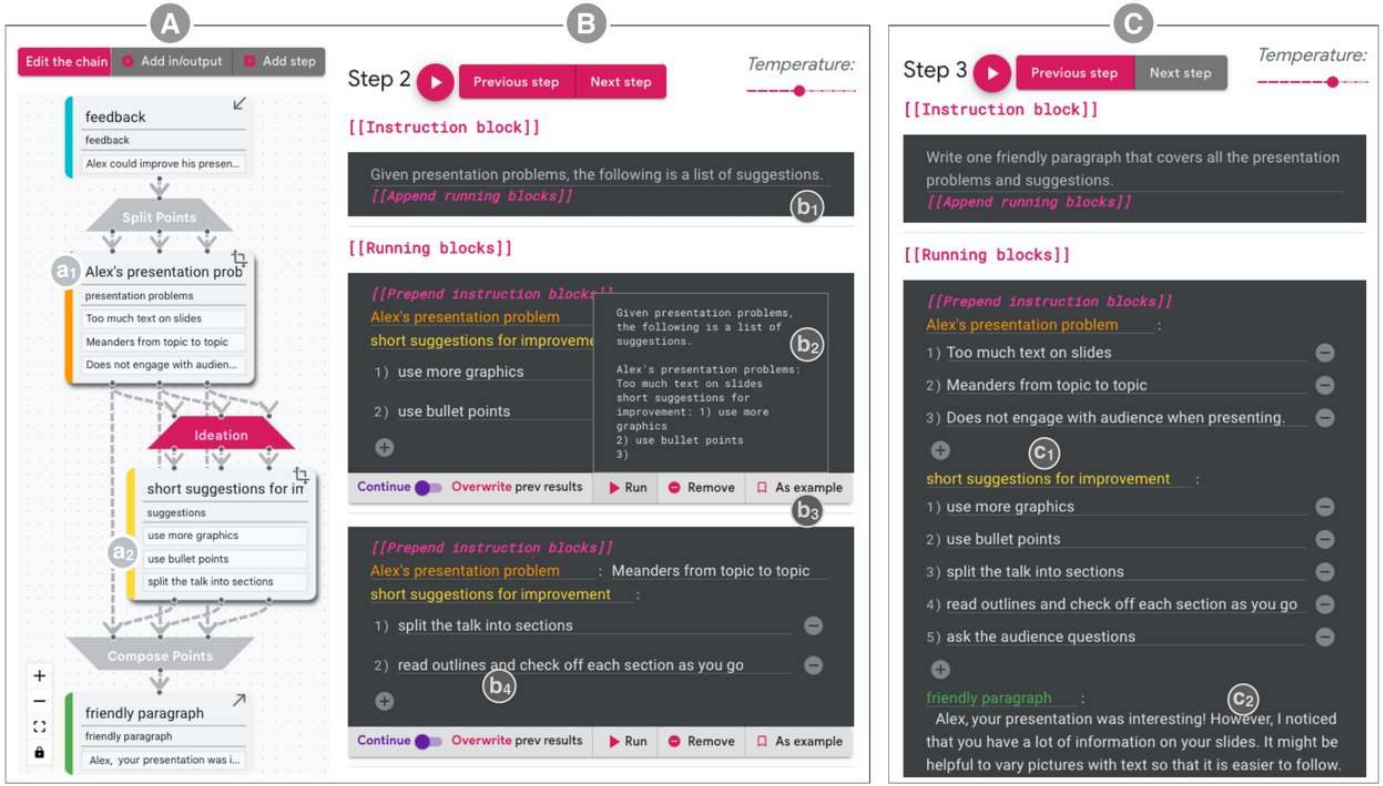
A chain of three prompts



Wu, Terry & Cai, "AI Chains," CHI 2022, Fig. 1

Split points, then ideation, then compose points
(B): the output of each step is the input of the next.

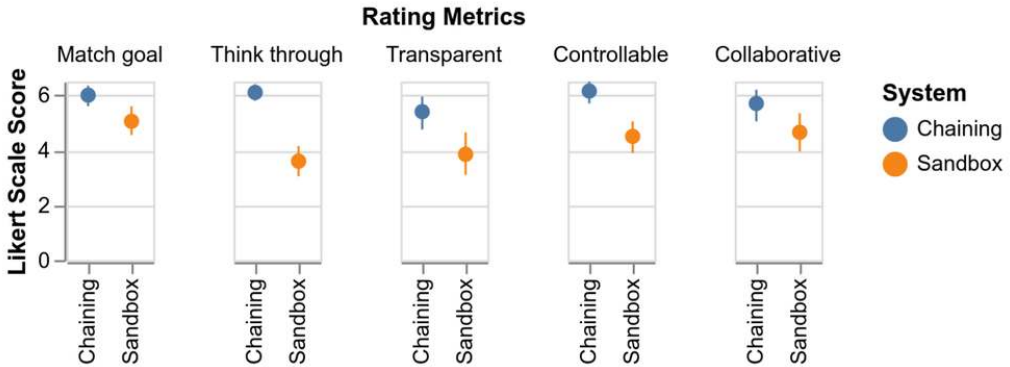
The AI Chains interface

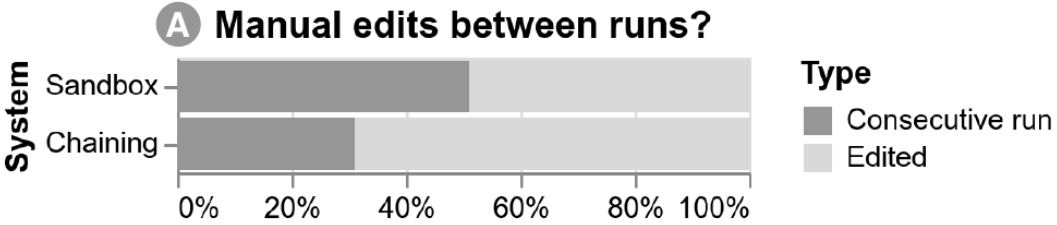


People can edit each step's outputs before the next step runs.

Wu, Terry & Cai, CHI 2022, Fig. 3

AI Chains: study results

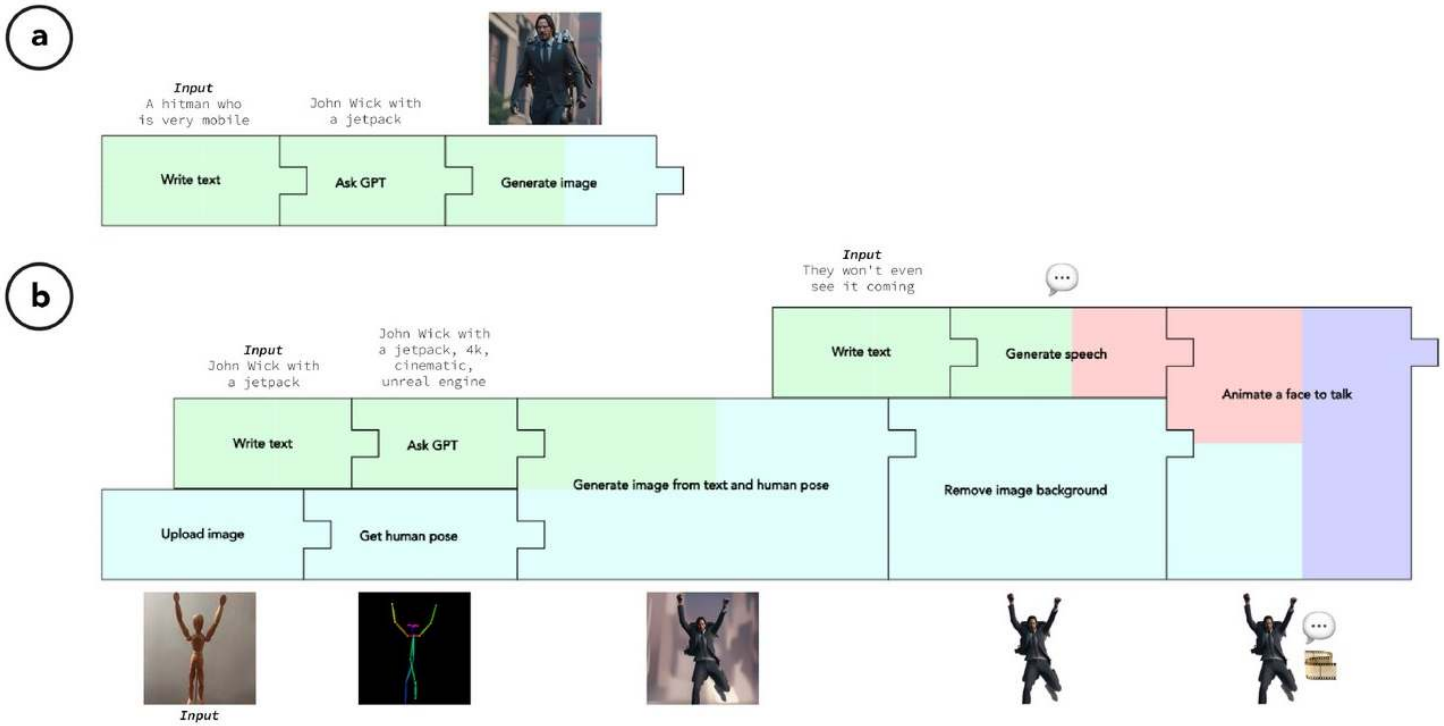




Two raters preferred the chained results **85%** and **80%** of the time.

Wu, Terry & Cai, CHI 2022, Fig. 5 and Fig. 6A

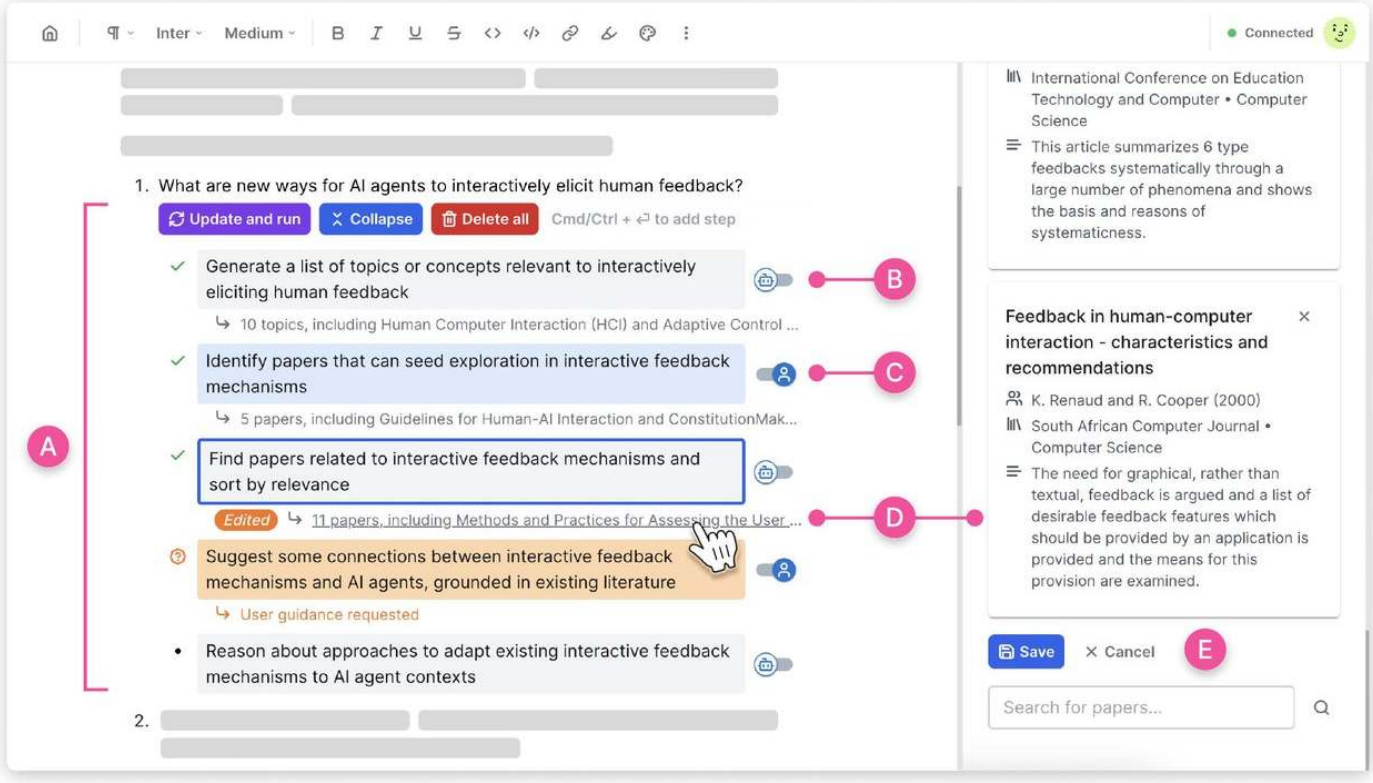
Jigsaw: chaining different models



Lin & Martelaro, CHI 2024, Fig. 9

A game designer posed a wooden mannequin, and a pose model passed the pose to an image model.

Cocoa: editing an agent's plan



Cocoa puts an agent's plan in the document.

Feng et al., CHI 2026, Fig. 2

Discussion: Zamfirescu-Pereira and Wu

1. Why did people drop examples that worked?
2. Who should design a chain: the user, the designer, or the model?
3. When is typing a prompt the best way to say what you want?

15 minutes

Milestone 3: Make it steerable

Due Thursday of week 8

Takeaways

1. A text box makes people **describe** what they want
2. People have to **envision** the result before they ask
3. Non-experts prompt **by trial and error**
4. Prompts can also be **examples, pictures or plans**

Next: Lecture 13

Control and Personalization

Read before class:

Shneiderman, IEEE Computer 1983
Direct Manipulation

Chung & Adar, UIST 2023
PromptPaint