Human-AI Interaction, Lecture 13

Control and Personalization

Adobe Unveils Future of Creative Cloud With Generative AI as a Creative Co-Pilot in Photoshop

Tuesday, May 23, 2023 07:00 AM

May 2023: Photoshop beta

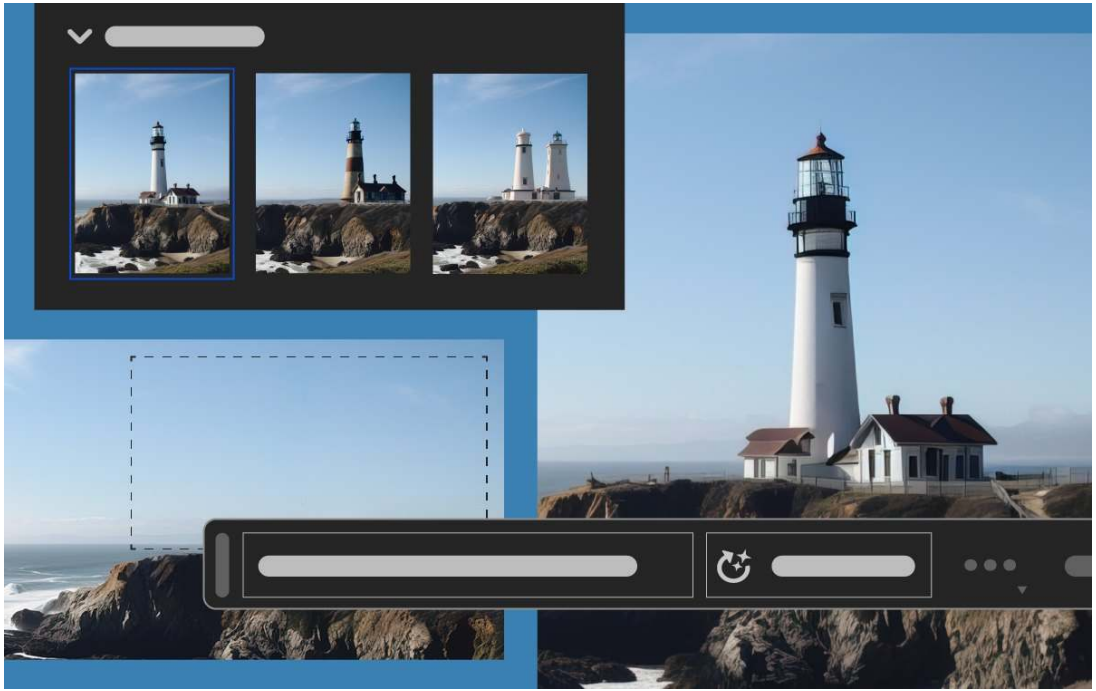
Sept. 2023: regular app

“…allowing users to extend images and add and remove objects”

Adobe Newsroom, May 23 and Sept. 13, 2023

How Generative Fill works

1. **Select** a region
2. **Type** a prompt, or nothing
3. **Pick** one of three

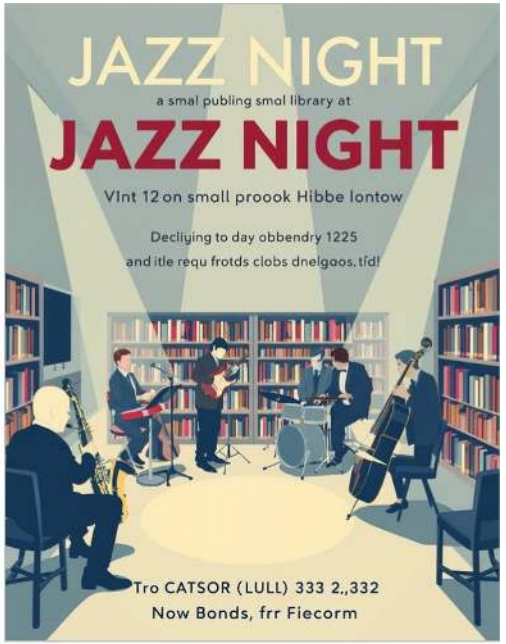


Adobe Help, “Generative Fill,” May 23, 2023; Adobe Blog, May 23, 2023

In Lecture 4, asking for bookshelves also changed **the band, the layout and the lettering**.



"a poster for a jazz night at a small public library"



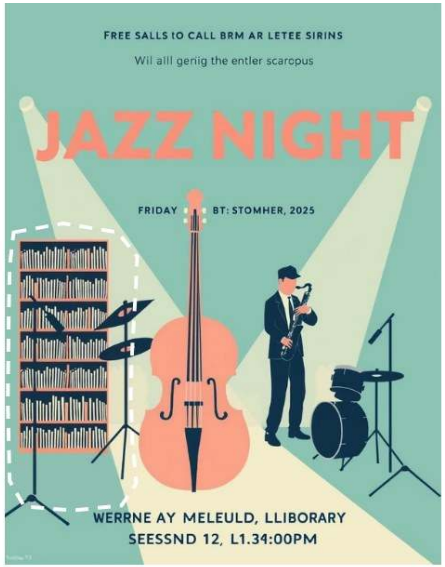
the same prompt + ", with bookshelves behind the band"

FLUX.1 [schnell]

Fill a selected region



selected region



seed 1



seed 2



seed 3

Outside the selection, the poster stays the same, including the misspelled lettering.

FLUX.1 Fill [pro]

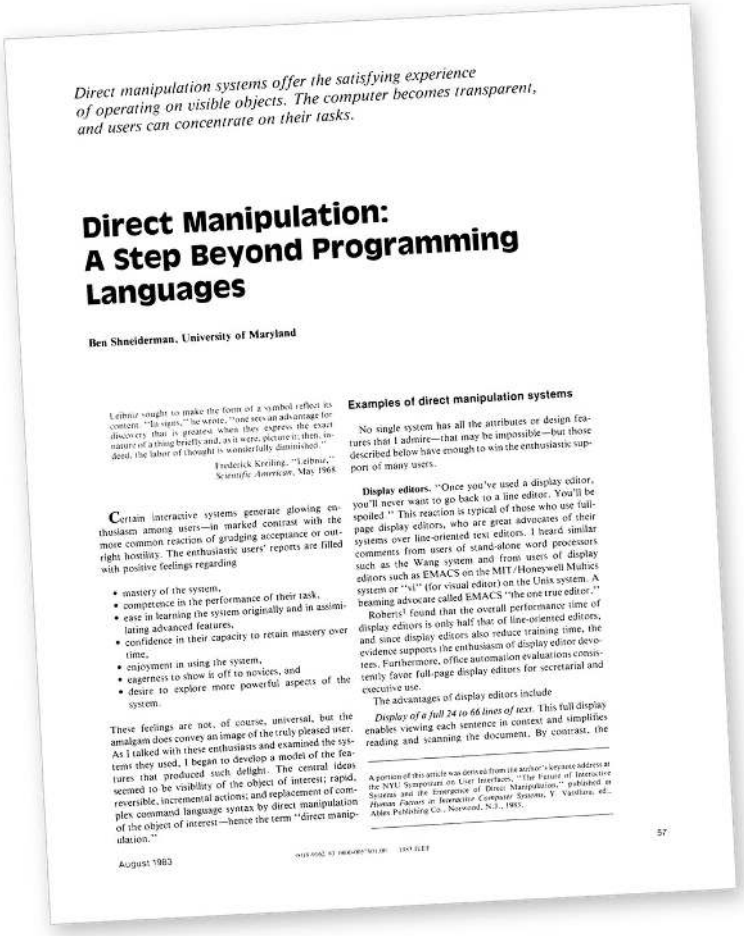
How can people **control** what a model makes, and **teach** it what they want?

Today

1. Direct manipulation
2. Painting with prompts
3. Sketches and sliders
4. Discussion
5. Activity: Sketch a control
6. Personalizing a model

1. Direct manipulation

Today's paper: Shneiderman, “Direct Manipulation: A Step Beyond Programming Languages,” *IEEE Computer* 1983



Direct manipulation systems offer the satisfying experience of operating on visible objects. The computer becomes transparent, and users can concentrate on their tasks.

Direct Manipulation: A Step Beyond Programming Languages

Ben Shneiderman, University of Maryland

Leibniz sought to make the form of a symbol reflect its content. "In signs," he wrote, "one sees an advantage for discovery that is greatest when they express the exact nature of a thing briefly and, as it were, picture it; then, indeed, the labor of thought is wonderfully diminished."

Frederick Kreiling, "Leibniz," *Scientific American*, May 1968

Certain interactive systems generate glowing enthusiasm among users—in marked contrast with the more common reaction of grudging acceptance or outright hostility. The enthusiastic users' reports are filled with positive feelings regarding

- mastery of the system,
- competence in the performance of their task,
- ease in learning the system originally and in assimilating advanced features,
- confidence in their capacity to retain mastery over time,
- enjoyment in using the system,
- eagerness to show it off to novices, and
- desire to explore more powerful aspects of the system.

These feelings are not, of course, universal, but the amalgam does convey an image of the truly pleased user. As I talked with these enthusiasts and examined the systems they used, I began to develop a model of the features that produced such delight. The central ideas seemed to be visibility of the object of interest; rapid, reversible, incremental actions; and replacement of complex command language syntax by direct manipulation of the object of interest—hence the term "direct manipulation."

Examples of direct manipulation systems

No single system has all the attributes or design features that I admire—that may be impossible—but those described below have enough to win the enthusiastic support of many users.

Display editors. "Once you've used a display editor, you'll never want to go back to a line editor. You'll be spoiled." This reaction is typical of those who use full-page display editors, who are great advocates of their systems over line-oriented text editors. I heard similar comments from users of stand-alone word processors such as the Wang system and from users of display editors such as EMACS on the MIT/Honeywell Multics system or "vi" (for visual editor) on the Unix system. A beaming advocate called EMACS "the one true editor."

Roberts[1] found that the overall performance time of display editors is only half that of line-oriented editors, and since display editors also reduce training time, the evidence supports the enthusiasm of display editor devotees. Furthermore, office automation evaluations consistently favor full-page display editors for secretarial and executive use.

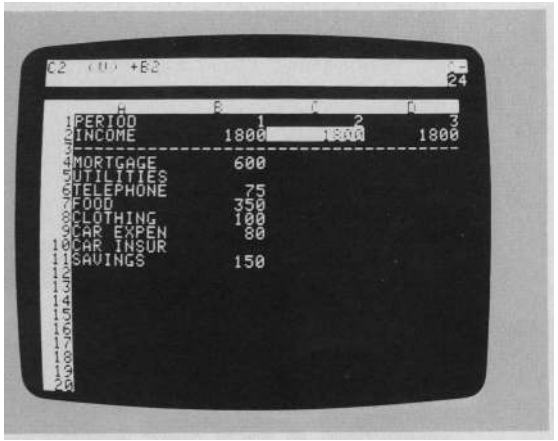
The advantages of display editors include

Display of a full 24 to 66 lines of text. This full display enables viewing each sentence in context and simplifies reading and scanning the document. By contrast, the

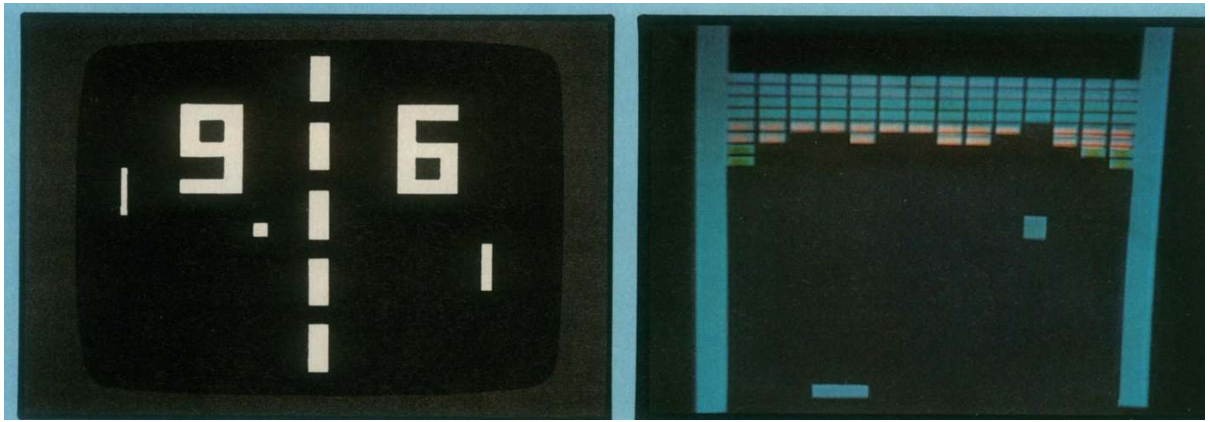
A portion of this article was derived from the author's keynote address at the NYU Symposium on User Interfaces, "The Future of Interactive Systems and the Emergence of Direct Manipulation," published as *Human Factors in Interactive Computer Systems*, Y. Vassiliou, ed., Ablex Publishing Co., Norwood, N.J., 1983.

Why did people love some systems?

“Certain interactive systems generate glowing **enthusiasm** among users…”



VisiCalc



Pong and Breakout

Shneiderman, IEEE Computer 1983, p. 57, Fig. 2 and p. 60

Principles of direct manipulation

- Continuous representation of the object of interest.

1. **See** the object

Shneiderman, IEEE Computer 1983, p. 64

Principles of direct manipulation

- Continuous representation of the object of interest.

- Physical actions (movement and selection by mouse, joystick, touch screen, etc.) or labeled button presses instead of complex syntax.

1. See the object

2. **Act on it**

Shneiderman, IEEE Computer 1983, p. 64

Principles of direct manipulation

- Continuous representation of the object of interest.

1. See the object

- Physical actions (movement and selection by mouse, joystick, touch screen, etc.) or labeled button presses instead of complex syntax.

2. Act on it

- Rapid, incremental, reversible operations whose impact on the object of interest is immediately visible.

3. **Undo small steps**

Shneiderman, IEEE Computer 1983, p. 64

Benefits and problems

- Users experience less anxiety because the system is comprehensible and because actions are so easily reversible.
- Users gain confidence and mastery because they initiate an action, feel in control, and can predict system responses.

Problems he names:

Users “must learn the meaning of the components”

“The graphic representation may be misleading”

Shneiderman, IEEE Computer 1983, pp. 64-65

Text prompts and the three principles

1. See the object	The image does not **exist** yet.
2. Act on it	You **describe** the change.
3. Undo small steps	Five words changed **the band and layout**.

What does the selection **add**?

prompt only

input → **words**

output → **a whole new image**

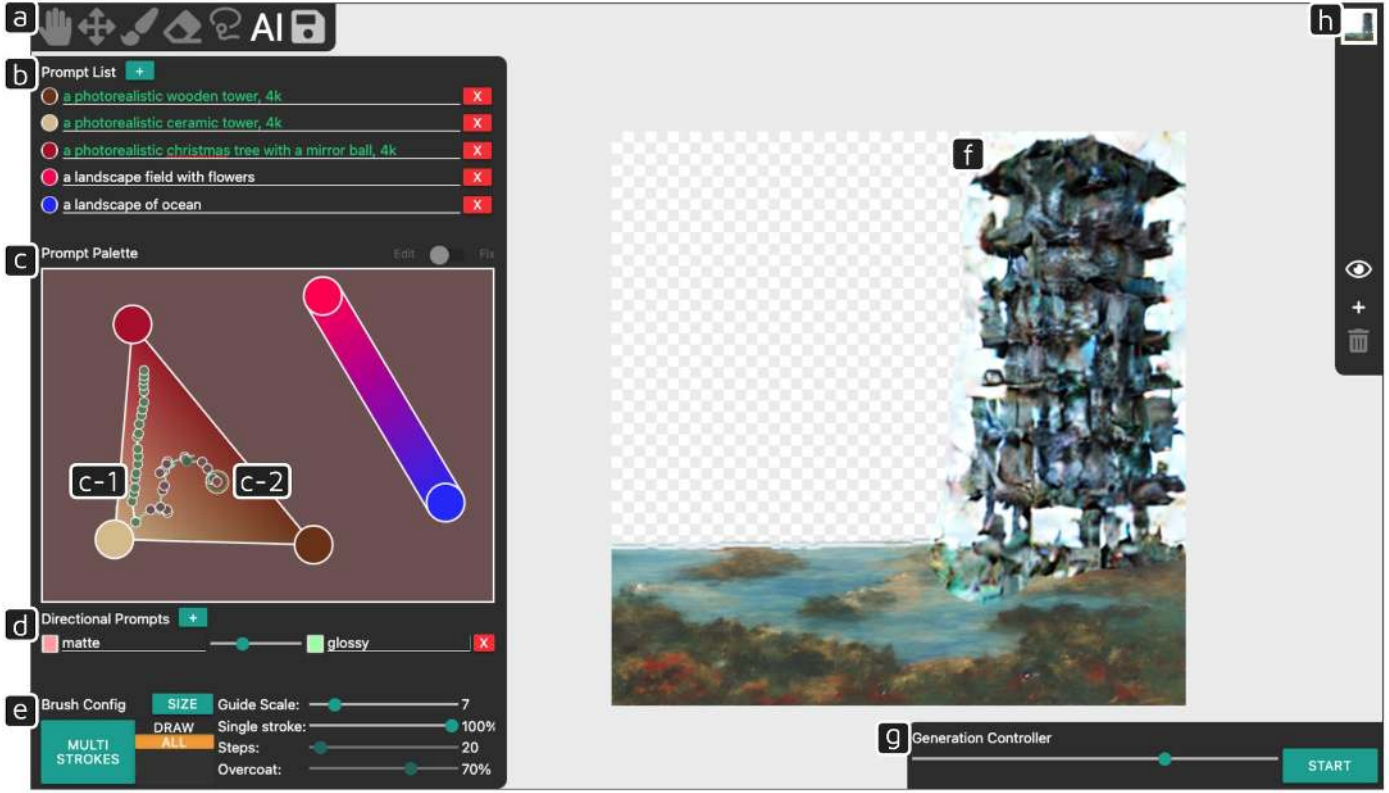
select, then prompt

input → **region + words**

output → **that region**

then → **pick 1 of 3**

PromptPaint



b the prompts

c a palette to mix them

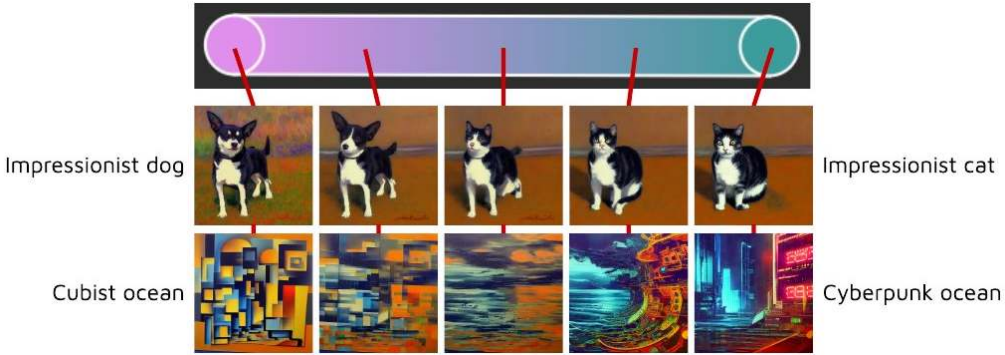
d a slider between two

f where a prompt goes

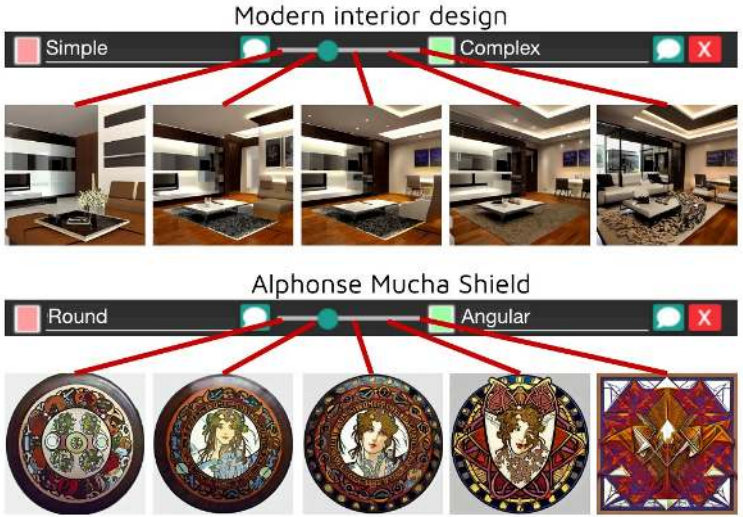
g when it changes

Chung & Adar, UIST 2023, Fig. 1

Prompt mixing and directional prompts



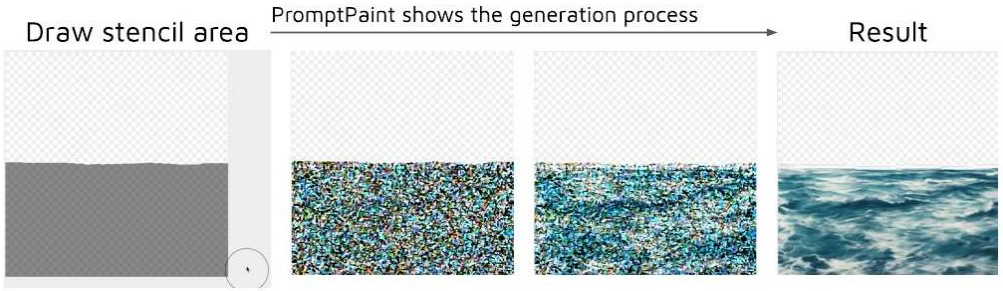
Prompt mixing: a point between two prompts



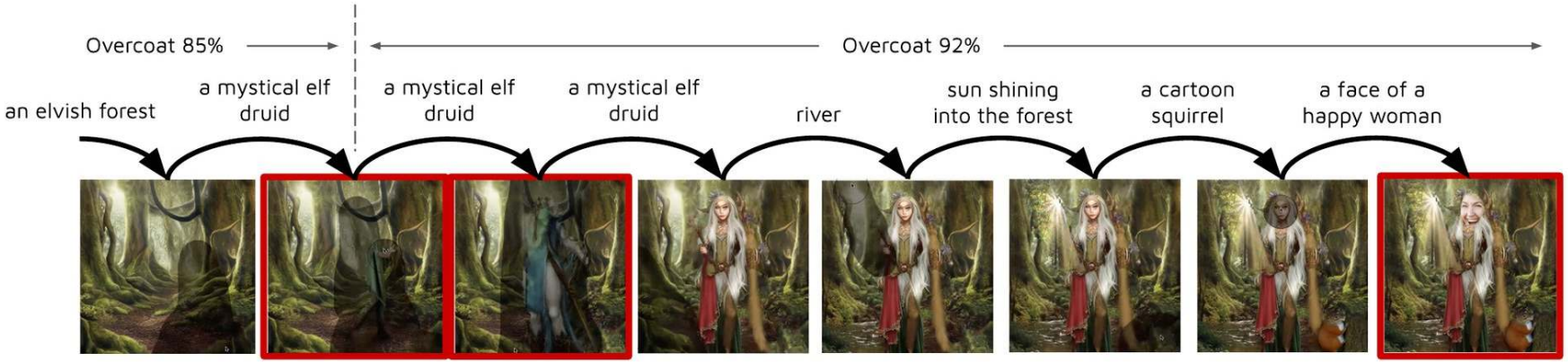
Directional prompt: a slider between two end prompts

Chung & Adar, UIST 2023, Figs. 6 and 7

Prompt stencil

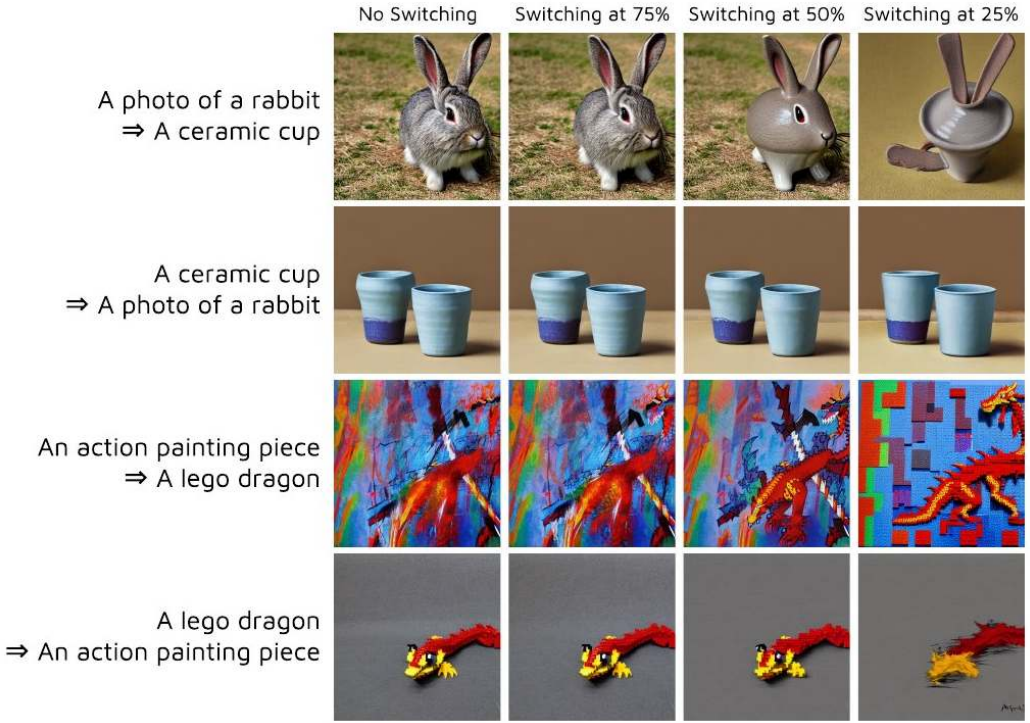


The model fills only the brushed area, and people watch it happen.



Chung & Adar, UIST 2023, Figs. 8 and 21

Prompt intervention



People switch the prompt partway through generation.

Chung & Adar, UIST 2023, Fig. 11

User study with 8 hobbyists and novices

+ mixing: images “difficult to describe verbally” (7 of 8)

+ stencils: **local** control (7 of 8)

− switching partway: **hardest to use** (5 of 8)

Chung & Adar, UIST 2023

ControlNet



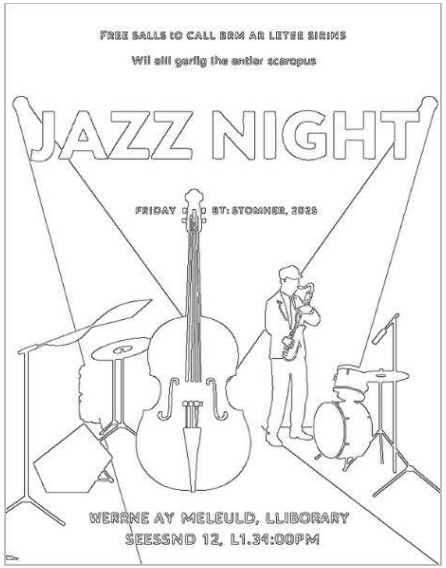
The original model stays locked. A trainable copy learns the edges or the pose.

Zhang, Rao & Agrawala, ICCV 2023, Fig. 1

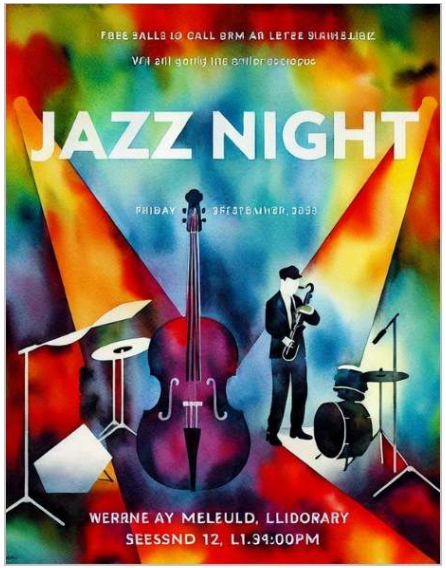
ControlNet with the edges of our poster



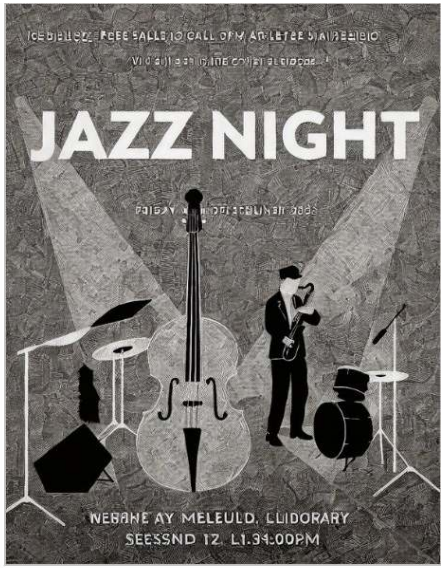
the poster



its edges



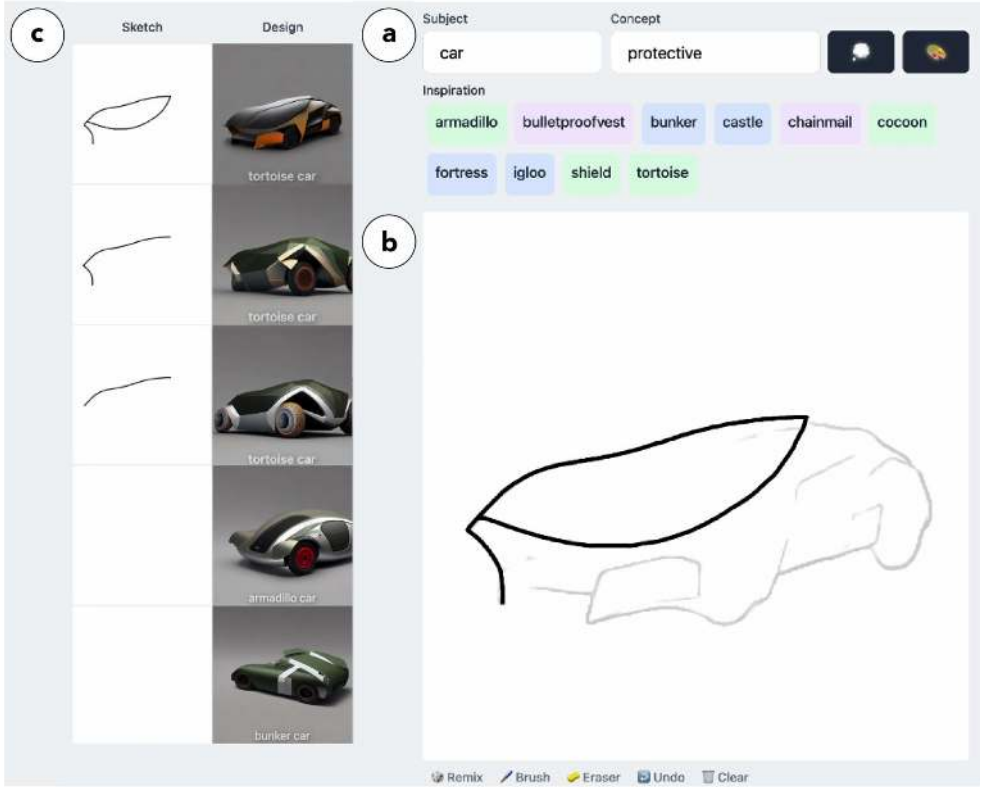
"a jazz night poster, watercolor painting"



"a jazz night poster, black and white linocut print"

With the same edges, both prompts keep the band, spotlights and lettering in place.

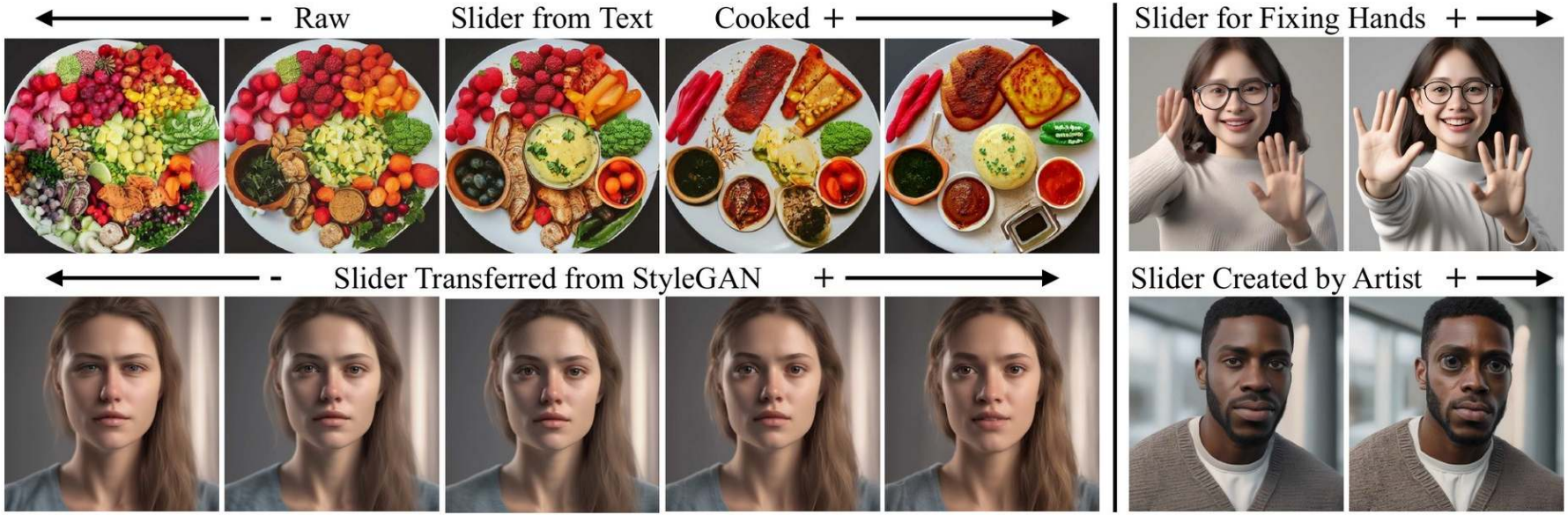
Inkspire



a analogies for an abstract word
b a new design after each stroke
c the designs so far

Lin et al., CHI 2025, Fig. 1 and Section 7.1

Concept Sliders



Each slider changes one concept.

Gandikota et al., ECCV 2024, Fig. 1

Discussion: Shneiderman and Chung & Adar

1. Which of his principles can a generative tool never fully meet?
2. Why was PromptPaint's most controllable function the hardest to use?
3. Does P5 own the image they made?

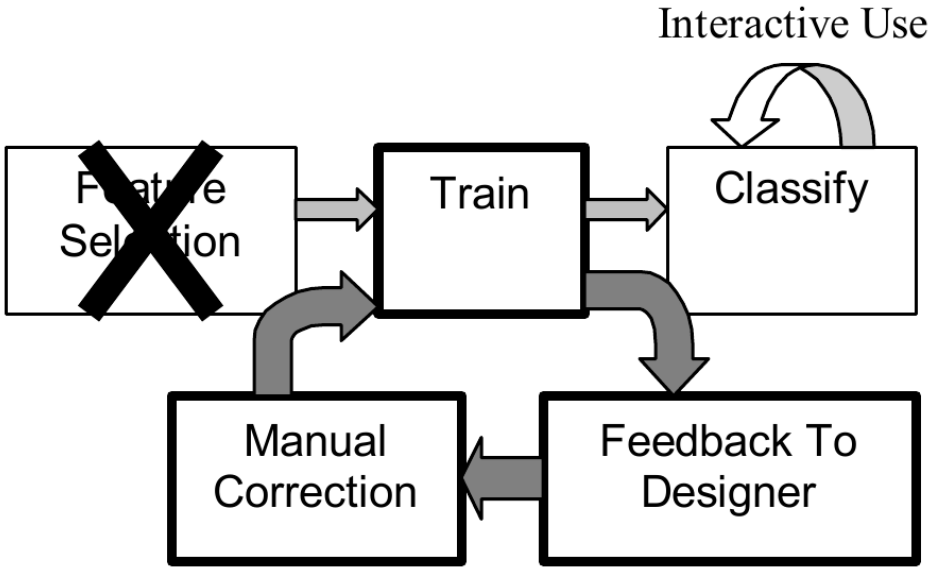
Chung & Adar, UIST 2023, p. 13

Activity: Sketch a control

In teams, 15 minutes

1. Pick a change your users often want
2. Sketch a control for it. No text box.
3. Write one change for the other team
4. Swap, make that change, mark where it fails
5. Share: where did it fail?

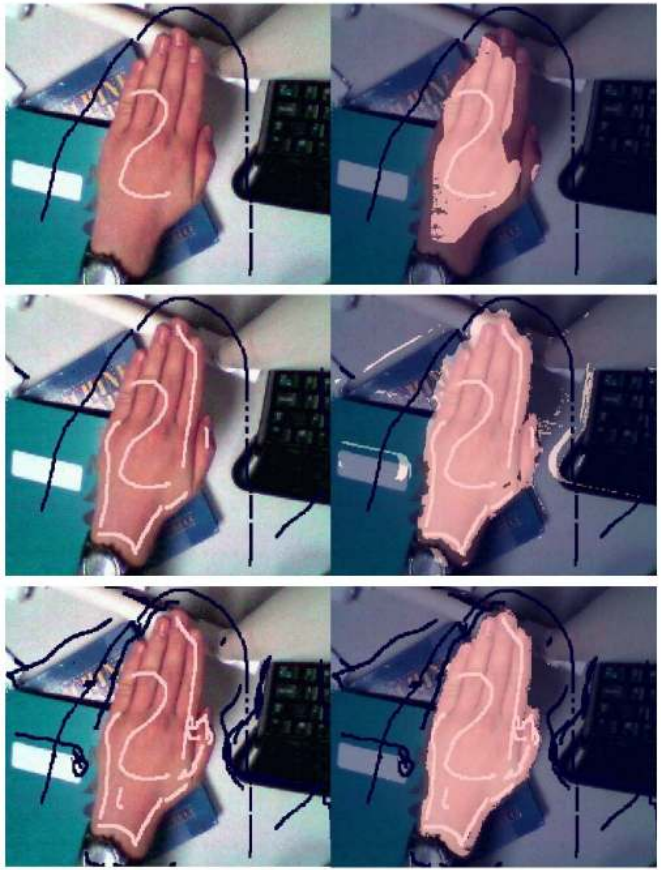
Interactive machine learning



Fails & Olsen, IUI 2003, Fig. 2

Each round of training had to take **under five seconds**.

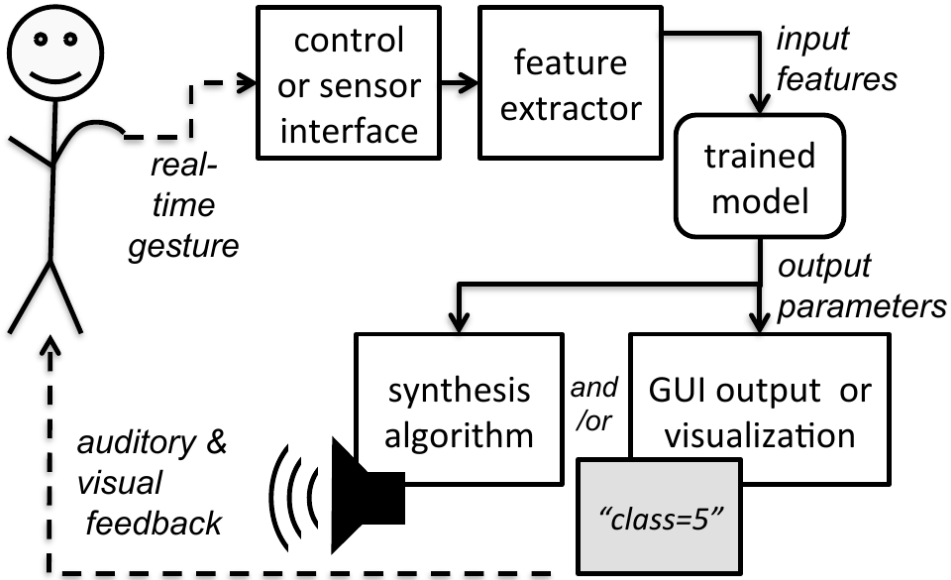
Crayons: training a skin classifier



1. Paint a little skin and background
2. Paint skin where it did poorly
3. A skin classifier **in a few rounds**

Fails & Olsen, IUI 2003, Fig. 5

Wekinator



People show it gestures and sounds, train it, then play it.

Fiebrink, Cook & Trueman, CHI 2011, Fig. 3

+ they tested it by **playing** more often than by checking accuracy

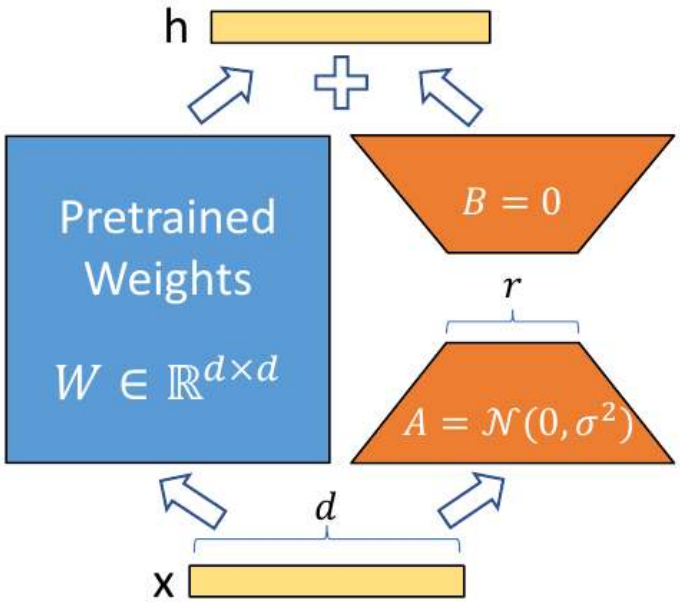
Fine-tuning: DreamBooth



It learns one subject from “just a few images.”

Ruiz et al., CVPR 2023, Fig. 1

LoRA



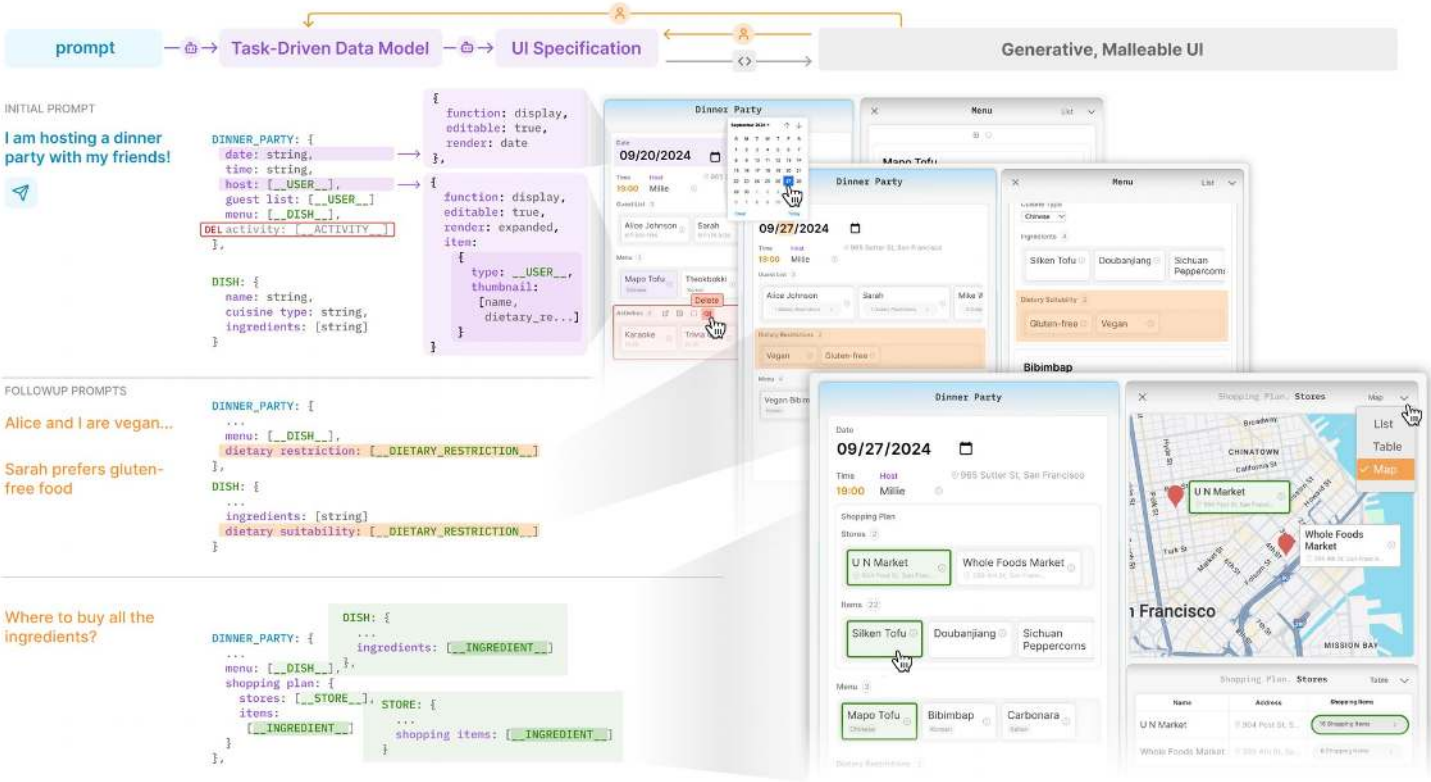
Hu et al., ICLR 2022, Fig. 1

W stays frozen. Only A and B are trained.

10,000 times fewer trainable parameters

Concept Sliders are LoRA add-ons too.

Customizing the interface: Jelly



People ask in words or edit the interface directly. Both change the data model underneath.

Cao, Jiang & Xia, CHI 2025, Fig. 1

Design questions for a personalized model

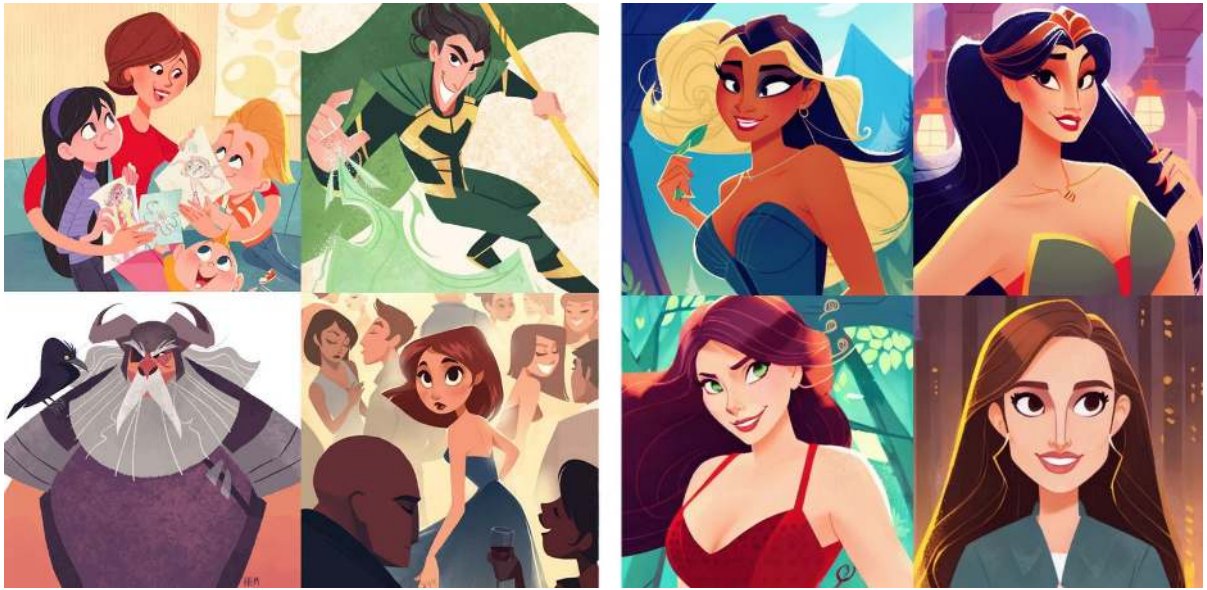
What did it learn, and how do people **see** that?

How do they **undo** it?

Who **owns** it?

Invasive Diffusion: How one unwilling illustrator found herself turned into an AI model

POSTED NOVEMBER 1, 2022 BY ANDY BAIO



A. Baio, waxy.org, Nov 1, 2022

A model fine-tuned on
32 of her illustrations

“If I had been asked if they could do this, I wouldn't have said yes.”

Hollie Mengert

Milestone 3: Make it steerable

Due Thursday of week 8

Next: Lecture 14, AI and Creativity

Read before class:

- Shneiderman, “Creativity Support Tools: Accelerating Discovery and Innovation,” Communications of the ACM 2007
- Doshi & Hauser, “Generative AI Enhances Individual Creativity but Reduces the Collective Diversity of Novel Content,” Science Advances 2024

Reading post due at 9 pm the night before.