Machine Learning for Design, Week 1

AI as a Design Material

Material and form



Turned maple
American, 1650–1700



Bent beech and cane
Michael Thonet, c. 1851



Steer horn
Texas, 1870–80

Three of the 56 chairs we use all term; The Metropolitan Museum of Art and the Art Institute of Chicago, open access (CC0)

AI as a new design material

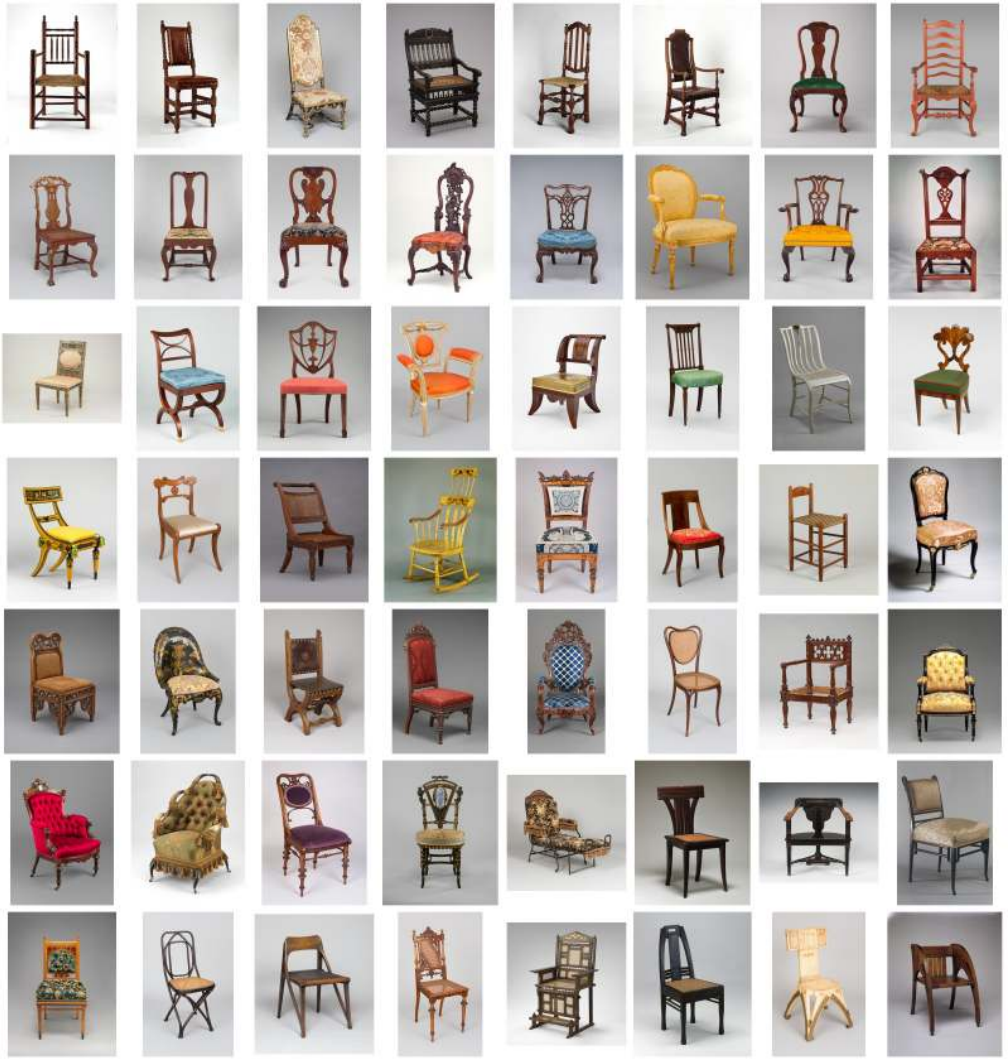
“As AI becomes a more and more vital part of everyday products, designers will have to figure out how to work with intelligence as a new material, with its own specific quirks and opportunities.”

Lars Erik Holmquist, 2017



Holmquist, “Intelligence on Tap: Artificial Intelligence as a New Design Material,” Interactions 2017, p. 31, optional reading; cover composition by Andrij Borys Associates

A new side chair for a museum café, drawing on 56 historical chairs



Chairs from The Metropolitan Museum of Art and the Art Institute of Chicago, 1650 to 1904, open access (CC0)

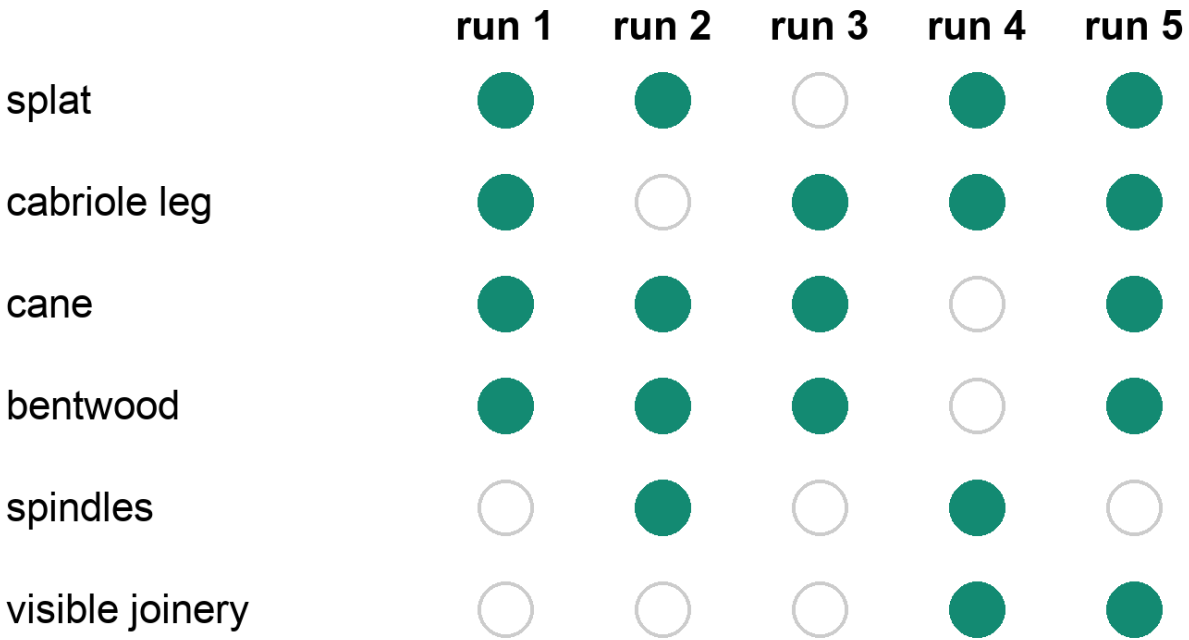
Asking Gemini for chair concepts

“Give me 5 concepts for a new side chair for a museum café. The museum's collection has 56 historical chairs from 1650 to 1904, and the new chair should draw on them. One short line per concept.”

1. The Ghost Splat
2. The Bent Essence
3. The Structural Ornament
4. The Neoclassical Silhouette
5. The Victorian Texture

gemini-2.5-flash through the Gemini API, run 1 of 5; concept names only, each came with one line of description

Five runs of the same prompt



Same name in two runs:
“The Ghost Splat” (1, 2)
“Cane Weave Abstraction” (2, 5)

A dot means the run's five concepts mention that feature. gemini-2.5-flash, same prompt, five separate runs.

Anderson et al.'s study

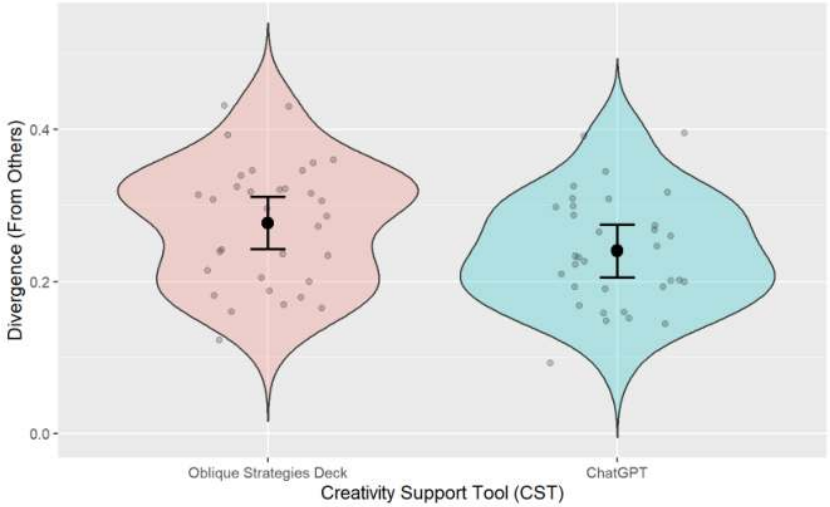


Figure 2: Participant responses were more homogenous at the group level (i.e., more semantically similar to the average embedding of *all* participant ideas) when using ChatGPT.[1]

33 participants came up with ideas with ChatGPT and with the Oblique Strategies deck

With ChatGPT, their ideas were closer to everyone else's ideas

Anderson et al., C&C 2024, Fig. 2 (CC BY-NC-SA 4.0), optional reading

What kind of design material is AI?

Today

1. How this course runs
2. Machine learning as a design material
3. Why AI is uniquely difficult to design
4. Inside a design assistant
5. Activity: ideas with and without AI
6. Discussion
7. Your project

When does AI help designers, and how can they direct what it makes?

Three parts

Part 1: **Designing with AI**, weeks 1-3

Part 2: **Appropriating AI**, weeks 4-6

Part 3: **Agentic Design**, weeks 7-9

Week 10: **Final Bake-Off**

Grading

Reading posts: **15%**

Hand-ins, weeks 1-6: **15%**

Project: **50%**

Position essay: **10%**

Participation: **10%**

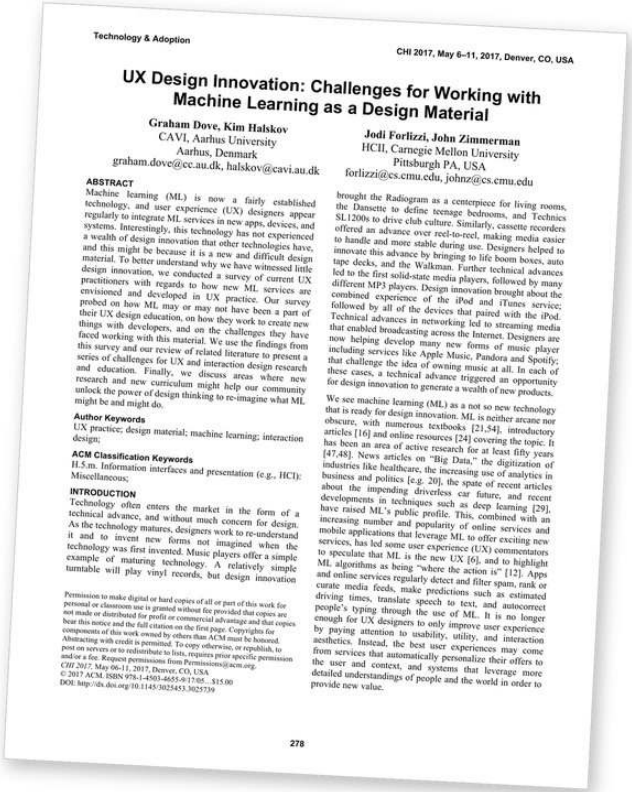
The project

Your team of 3 or 4 works as a small design consultancy for an organization you can reach.

Week 2: Team and partner
Week 3: Partner interviews
Week 5: Midpoint presentation
Week 7: Working prototype
Weeks 8-9: Partner tests
Week 10: Bake-off and final paper

2. Machine learning as a design material

Today's reading: Dove et al., “UX Design Innovation: Challenges for Working with Machine Learning as a Design Material,” CHI 2017



Technology & Adoption CHI 2017, May 6–11, 2017, Denver, CO, USA

UX Design Innovation: Challenges for Working with Machine Learning as a Design Material

Graham Dove, Kim Halskov
CAVI, Aarhus University
Aarhus, Denmark
graham.dove@cc.au.dk, halskov@cavi.au.dk

Jodi Forlizzi, John Zimmerman
HCII, Carnegie Mellon University
Pittsburgh PA, USA
forlizzi@cs.cmu.edu, johnz@cs.cmu.edu

ABSTRACT
Machine learning (ML) is now a fairly established technology, and user experience (UX) designers appear regularly to integrate ML services in new apps, devices, and systems. Interestingly, this technology has not experienced a wealth of design innovation that other technologies have, and this might be because it is a new and difficult design material. To better understand why we have witnessed little design innovation, we conducted a survey of current UX practitioners with regards to how new ML services are envisioned and developed in UX practice. Our survey probed on how ML may or may not have been a part of their UX design education, on how they work to create new things with developers, and on the challenges they have faced working with this material. We use the findings from this survey and our review of related literature to present a series of challenges for UX and interaction design research and education. Finally, we discuss areas where new research and new curriculum might help our community unlock the power of design thinking to re-imagine what ML might be and might do.

Author Keywords
UX practice; design material; machine learning; interaction design;

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

INTRODUCTION
Technology often enters the market in the form of a technical advance, and without much concern for design. As the technology matures, designers work to re-understand it and to invent new forms not imagined when the technology was first invented. Music players offer a simple example of maturing technology. A relatively simple turntable will play vinyl records, but design innovation brought the Radiogram as a centerpiece for living rooms, the Dansette to define teenage bedrooms, and Technics SL1200s to drive club culture. Similarly, cassette recorders offered an advance over reel-to-reel, making media easier to handle and more stable during use. Designers helped to innovate this advance by bringing to life boom boxes, auto tape decks, and the Walkman. Further technical advances led to the first solid-state media players, followed by many different MP3 players. Design innovation brought about the combined experience of the iPod and iTunes service; followed by all of the devices that paired with the iPod. Technical advances in networking led to streaming media that enabled broadcasting across the Internet. Designers are now helping develop many new forms of music player including services like Apple Music, Pandora and Spotify; that challenge the idea of owning music at all. In each of these cases, a technical advance triggered an opportunity for design innovation to generate a wealth of new products.

We see machine learning (ML) as a not so new technology that is ready for design innovation. ML is neither arcane nor obscure, with numerous textbooks [21,54], introductory articles [16] and online resources [24] covering the topic. It has been an area of active research for at least fifty years [47,48]. News articles on “Big Data,” the digitization of industries like healthcare, the increasing use of analytics in business and politics [e.g. 20], the spate of recent articles about the impending driverless car future, and recent developments in techniques such as deep learning [29], have raised ML’s public profile. This, combined with an increasing number and popularity of online services and mobile applications that leverage ML to offer exciting new services, has led some user experience (UX) commentators to speculate that ML is the new UX [6], and to highlight ML algorithms as being “where the action is” [12]. Apps and online services regularly detect and filter spam, rank or curate media feeds, make predictions such as estimated driving times, translate speech to text, and autocorrect people’s typing through the use of ML. It is no longer enough for UX designers to only improve user experience by paying attention to usability, utility, and interaction aesthetics. Instead, the best user experiences may come from services that automatically personalize their offers to the user and context, and systems that leverage more detailed understandings of people and the world in order to provide new value.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
CHI 2017, May 06-11, 2017, Denver, CO, USA
© 2017 ACM. ISBN 978-1-4503-4655-9/17/05…$15.00
DOI: http://dx.doi.org/10.1145/3025453.3025739

Dove et al.'s survey

Type of Involvement	#
The **UX design team collaborated with engineers**, product managers or others, and jointly developed an idea for a new product or service that utilizes machine learning	12
The **UX design team gave an interactive form** to a machine learning idea that came from others (e.g. software developers or engineers)	8
The **UX design team generated a novel design concept** utilizing machine learning, which was presented and then selected for integration into a new product or service	7

Table 2: Involvement of UX team in generating ideas for novel ML products

Of 37 UX practitioners with a design degree, **3** had learned about machine learning in school

In **7** of 27 projects, the UX team came up with the idea

Dove et al., CHI 2017, Table 2; an online survey of 51 UX practitioners

What the practitioners wrote

Another stated that, *"It is black magic to designers. ...Designers don't understand the constraints of the technology and how to employ it appropriately."* This

for designers to collaborate with skilled technologists. One noted, *"Machine learning is hard to prototype. Machine learning requires highly skilled collaborators, when a lot of companies are only able to hire 'warm bodies'."* Additionally, another respondent stated, *"...making interactive prototypes that incorporates machine learning is hard (haven't found a way to do that yet in an easy fashion)*," and another indicated that UX designers make statements such as *" ...inputs come in ...some magic happens ...and all your business needs are met!"* Clearly, in

30 respondents listed their three biggest challenges with machine learning

Dove et al., CHI 2017, p. 283; highlights mine

Dove et al.'s four challenges

1. Consider the interplay between ML statistical intelligence and human common sense intelligence
2. Envision opportunities to apply ML in less obvious ways
3. Represent ML's dependency on data in early prototypes
4. Foreground ethical considerations of ML

Dove et al., CHI 2017, p. 284: challenges for design research and education

3. Why AI is uniquely difficult to design

Today's reading: Yang et al., “Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design,” CHI 2020



CHI 2020 Paper

CHI 2020, April 25–30, 2020, Honolulu, HI, USA

Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design

Qian Yang [yangqian,steinfeld]@cmu.edu
Aaron Steinfeld Carolyn Rosé Carnegie Mellon University, Pittsburgh PA, USA cp3a@andrew.cmu.edu
John Zimmerman johnz@cs.cmu.edu

ABSTRACT
Artificial Intelligence (AI) plays an increasingly important role in improving HCI and user experience. Yet many challenges persist in designing and innovating valuable human-AI interactions. For example, AI systems can make unpredictable errors, and these errors damage UX and even lead to undesired societal impact. However, HCI routinely grapples with complex technologies and mitigates their unintended consequences. What makes AI different? What makes human-AI interaction appear particularly difficult to design? This paper investigates these questions. We synthesize prior research, our own design and research experience, and our observations when teaching human-AI interaction. We identify two sources of AI's distinctive design challenges: 1) uncertainty surrounding AI's capabilities, 2) AI's output complexity, spanning from simple to adaptive complex. We identify four levels of AI systems. On each level, designers encounter a different subset of the design challenges. We demonstrate how these findings reveal new insights for designers, researchers, and design tool makers in productively addressing the challenges of human-AI interaction going forward.

Author Keywords
User experience, artificial intelligence, sketching, prototyping.

CCS Concepts
•**Human-centered computing** → **Human computer interaction (HCI)**; *Interaction design process and methods;*

INTRODUCTION
Advances in Artificial Intelligence (AI) have produced exciting opportunities for human-computer interaction (HCI). From mundane spam filters to autonomous driving, AI holds many promises for improved user experiences (UX), and it enables otherwise impossible forms of interaction. This trend has led to the idea of *AI as a design material* in the research community, with the hope that HCI researchers and designers can effectively envision and refine new uses for AI that have yet to be imagined [13, 26, 58].

This work is licensed under a Creative Commons Attribution International 4.0 License.

CHI '20, April 25–30, 2020, Honolulu, HI, USA.
© 2020 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-6708-0/20/04.
http://dx.doi.org/10.1145/3313831.3376301

The growing interest in how AI can improve UX implies that HCI designers have become skilled at integrating AI's capabilities into their practices. Interestingly, the research shows something else, that HCI designers largely struggle to envision and prototype AI systems [2, 3, 13, 18, 22, 26, 49]. For example, even simple AI applications can make inference errors that are difficult to anticipate. These errors impact the intended user experience, and can sometimes raise serious ethical concerns or result in societal-level consequences. However, current HCI design methods meant to mitigate unintended consequences (i.e. sketching and prototyping) can seem ill-fitted for AI. HCI professionals cannot easily *sketch* the numerous ways an AI system might adapt to different users in different contexts [13, 58]. Nor can they easily *prototype* the types of inference errors a not-yet developed AI system might make [29, 42, 49].

Existing research frequently attributes these challenges to AI's technical complexity, demand for data, and unpredictable interactions [49, 26, 13, 42]. Less discussed is that HCI routinely grapples with complex, resource-intensive technologies using simple, inexpensive artifacts, e.g., paper prototypes and Wizard of Oz systems. What makes AI different and distinctly difficult to prototype? Equally important, designers routinely choreograph complex, dynamic, sometimes unpredictable interactions, with a focus on mitigating technologies' unintended consequences (e.g., [61]). What makes AI interactions particularly difficult to sketch? A critical first step in designing valuable human-AI interactions is to articulate the unique qualities of AI that made it so difficult to design.

The goal of this paper is to delineate *whether*, *why*, and *how* human-AI interaction is distinctly difficult to design and innovate. The paper has four parts:

1. We set the stage by cataloging the many human-AI interaction design challenges that literature has reported as well as solutions proposed.
2. We ask three provocative questions as a critique of current related work. These questions serve as a springboard for rethinking how to approach the question of why human-AI interaction appears so difficult to design.
3. We synthesize our own research, including studies on the challenges HCI practitioners faced when working with AI, our insights from making AI things via research through design, and our insights from teaching students in human-AI interaction design and innovation. This synthesis identified

Challenges at each design stage

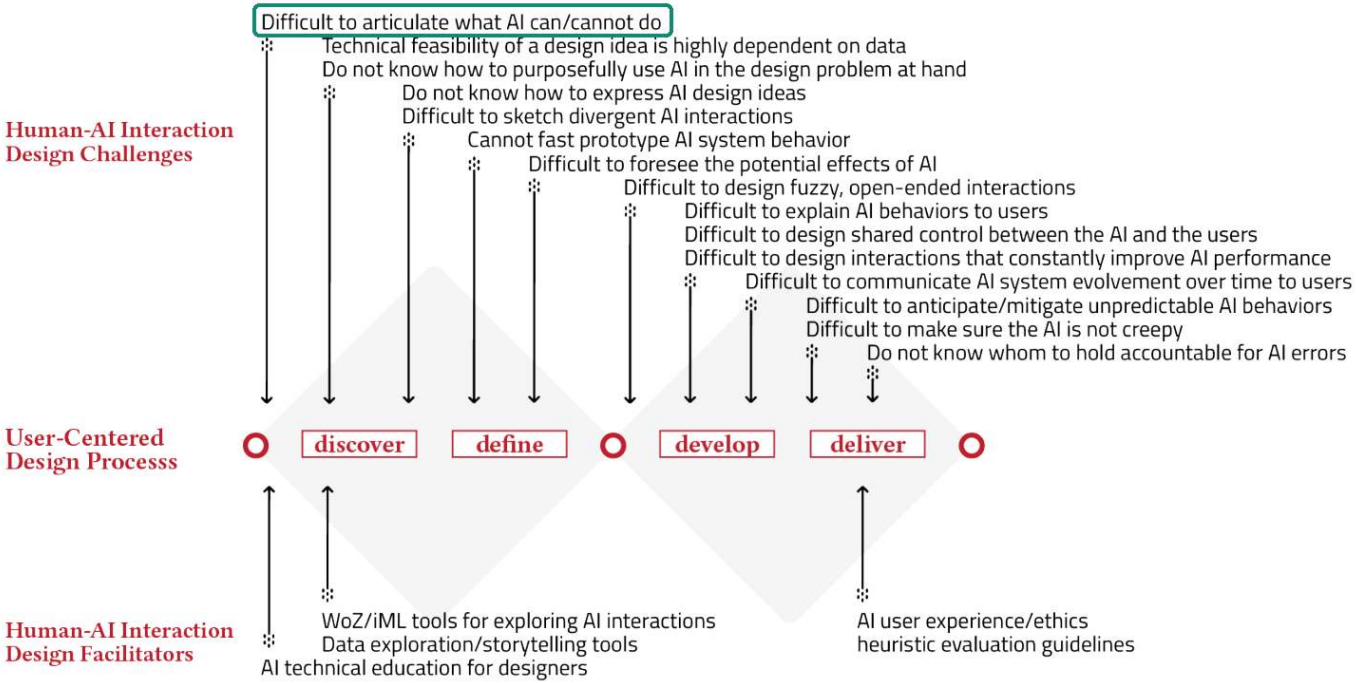


Figure 1: Mapping the human-AI interaction design challenges in the literature [58, 13, 26, 53] onto a user-centered design process (Double Diamond [10])

Yang et al., CHI 2020, Fig. 1 (CC BY 4.0); highlight mine

Two views on why AI is hard

The view they question

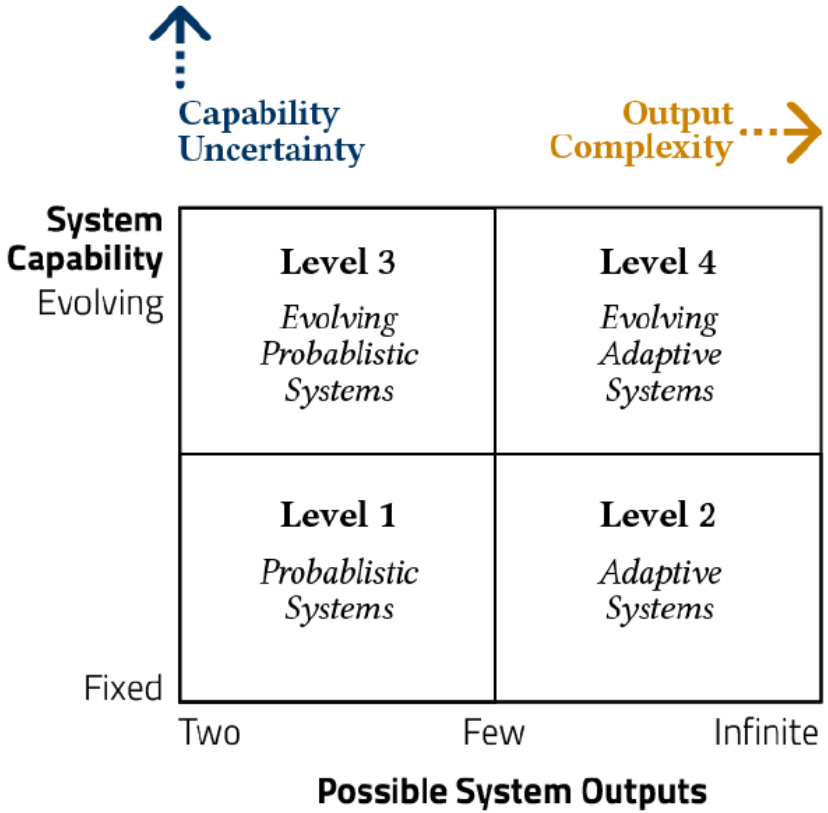
“AI appeared to be difficult to design because AI is just ‘a new and difficult design material,’ suggesting that over time, known HCI methods will likely address these challenges”

Their answer

“the root challenges are that AI’s capabilities are adaptive and its outputs can autonomously diverge at a massive scale”

Yang et al., CHI 2020, p. 3, describing Dove et al., and p. 10

The AI design complexity map



Yang et al., CHI 2020, Fig. 4

Capability uncertainty: what it can do

Output complexity: how many different outputs it can give

Level 1 and Level 4 systems

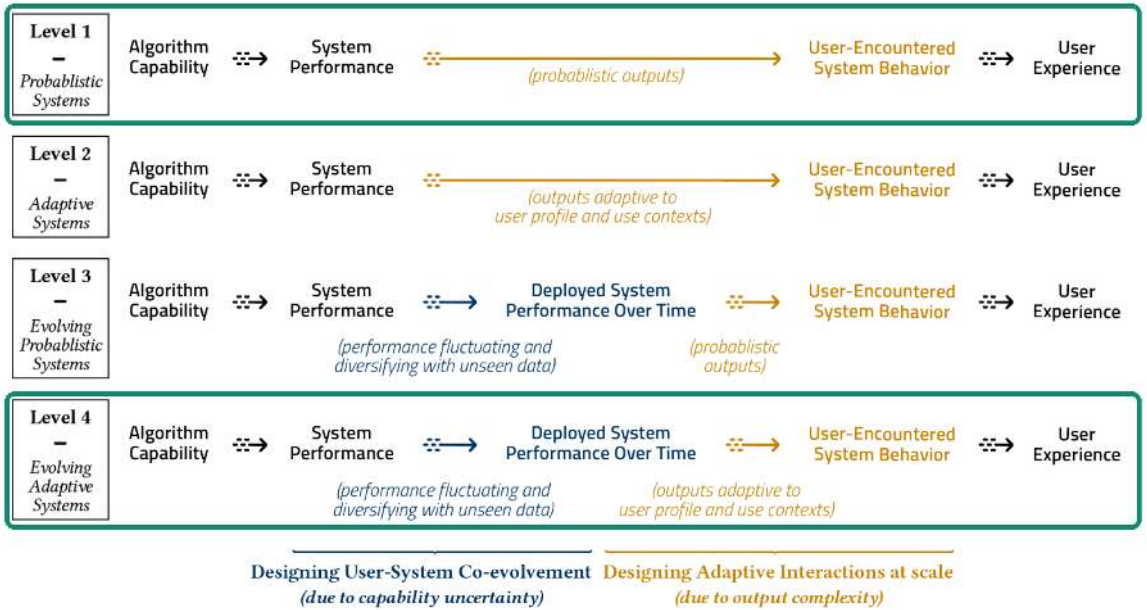


Figure 5: Four levels of AI systems according to design complexity.

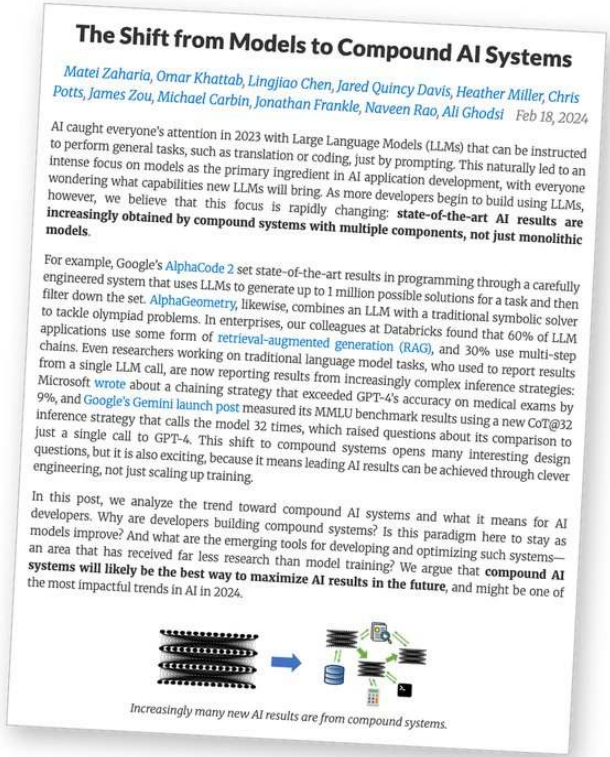
Level 1: a toxic comment classifier

Level 4: face tagging in a photos app, which learns from the photos and tags people add

Yang et al., CHI 2020, Fig. 5; highlights mine

4. Inside a design assistant

Optional reading: Zaharia et al., “The Shift from Models to Compound AI Systems,” BAIR Blog, 2024



The Shift from Models to Compound AI Systems

Matei Zaharia, Omar Khattab, Lingjiao Chen, Jared Quincy Davis, Heather Miller, Chris Potts, James Zou, Michael Carbin, Jonathan Frankle, Naveen Rao, Ali Ghodsi Feb 18, 2024

AI caught everyone's attention in 2023 with Large Language Models (LLMs) that can be instructed to perform general tasks, such as translation or coding, just by prompting. This naturally led to an intense focus on models as the primary ingredient in AI application development, with everyone wondering what capabilities new LLMs will bring. As more developers begin to build using LLMs, however, we believe that this focus is rapidly changing: **state-of-the-art AI results are increasingly obtained by compound systems with multiple components, not just monolithic models**.

For example, Google's AlphaCode 2 set state-of-the-art results in programming through a carefully engineered system that uses LLMs to generate up to 1 million possible solutions for a task and then filter down the set. AlphaGeometry, likewise, combines an LLM with a traditional symbolic solver to tackle olympiad problems. In enterprises, our colleagues at Databricks found that 60% of LLM applications use some form of retrieval-augmented generation (RAG), and 30% use multi-step chains. Even researchers working on traditional language model tasks, who used to report results from a single LLM call, are now reporting results from increasingly complex inference strategies: Microsoft wrote about a chaining strategy that exceeded GPT-4's accuracy on medical exams by 9%, and Google's Gemini launch post measured its MMLU benchmark results using a new CoT@32 inference strategy that calls the model 32 times, which raised questions about its comparison to just a single call to GPT-4. This shift to compound systems opens many interesting design questions, but it is also exciting, because it means leading AI results can be achieved through clever engineering, not just scaling up training.

In this post, we analyze the trend toward compound AI systems and what it means for AI developers. Why are developers building compound systems? Is this paradigm here to stay as models improve? And what are the emerging tools for developing and optimizing such systems—an area that has received far less research than model training? We argue that **compound AI systems will likely be the best way to maximize AI results in the future**, and might be one of the most impactful trends in AI in 2024.

Increasingly many new AI results are from compound systems.

One request to Gemini

```
POST models/gemini-2.5-flash:generateContent

{
 "contents": [
  {
   "role": "user",
   "parts": [
    {
     "text": "Give me 5 concepts for a new side chair for a museum café.
   The museum's collection has 56 historical chairs from 1650 to 1904, and
   the new chair should draw on them. One short line per concept."
    }
   ]
  }
 ]
}

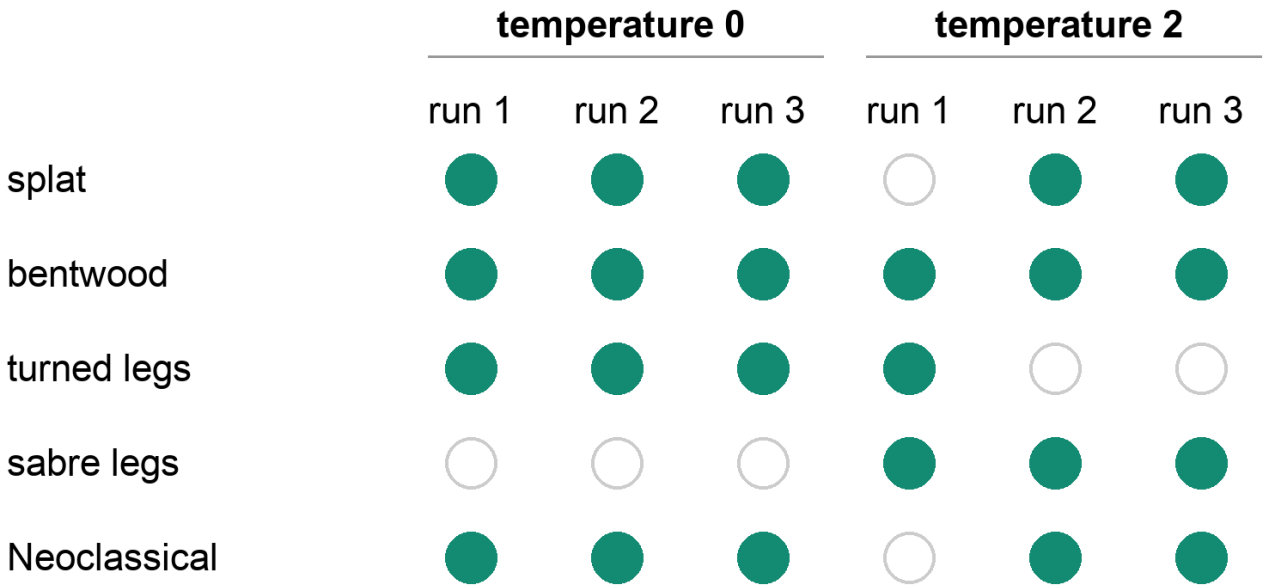
response:
 "text": "Here are 5 concepts for a new side chair, drawing on histori…"
 "modelVersion": "gemini-2.5-flash"
 "promptTokenCount": 55
 "candidatesTokenCount": 209
 "thoughtsTokenCount": 1894
```

55 tokens in

1,894 thinking tokens

209 tokens out

The same prompt at two temperatures

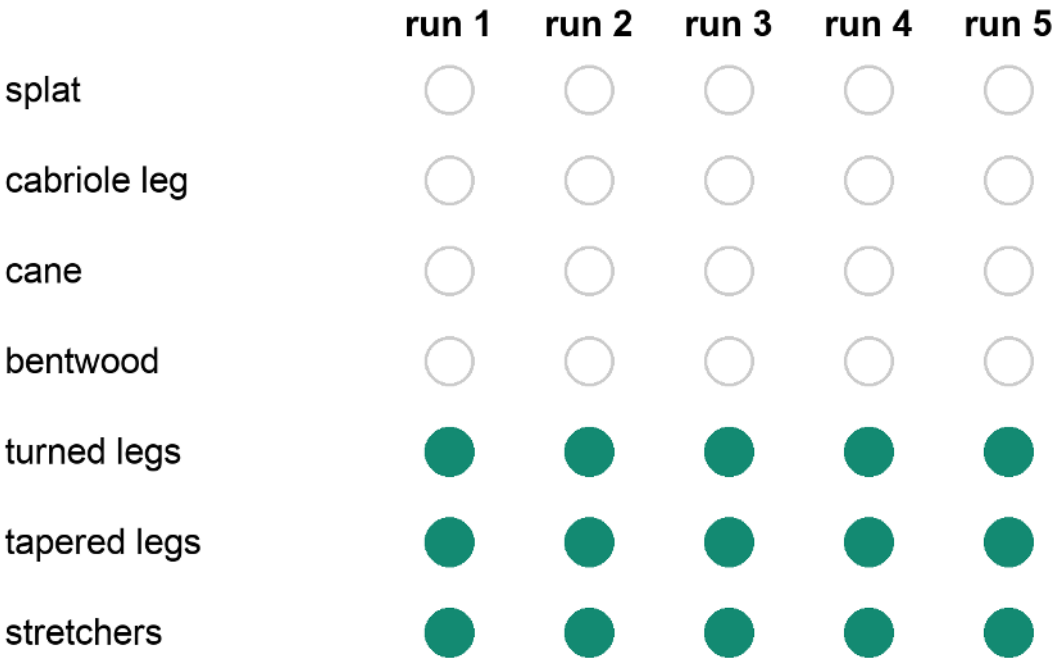


At temperature 0, the three answers are the same, word for word

gemini-2.5-flash, the prompt from run 1, three runs at each temperature; the default is 1

Asking it to leave out four features

Added to the prompt: “Do not use splats, cabriole legs, cane or bentwood.”



None of the four came back

Turned legs, tapered legs and stretchers: all five runs

gemini-2.5-flash, five separate runs of the new prompt

Three properties from our runs

1. At the default temperature, the same prompt gives a different answer each run
2. The answers pull toward the same few features
3. Told to leave those out, it moves to the next most common ones

gemini-2.5-flash, all on the café chair brief

Activity: ideas with and without AI

1. 10 minutes: sketch 5 ideas for the café chair on your own
2. 10 minutes: ask an assistant for 5 more, then change one thing in your prompt and ask again
3. Mark the ideas you would not have had alone

In pairs, then a few of you share

Discussion

1. What did the assistant's ideas have in common?
2. Which of Yang's two sources did you run into?
3. Where in your own process would you use an assistant first?

Hand-in: ideas with and without AI

Your 5 sketches and the assistant's ideas side by side, with both prompts, and a few sentences on which ones you would develop and why

Due before week 2

Team and partner

Due week 2: form a team of 3 or 4
and find an organization to work with

Next: Week 2

Mood Boards and Inspiration

Read before the studio:

- Koch et al., “May AI? Design Ideation with Cooperative Contextual Bandits,” CHI 2019
- Koch et al., “SemanticCollage: Enriching Digital Mood Board Design with Semantic Labels,” DIS 2020