



# 人工智能前沿专题 人机交互

彭振辉

中山大学人工智能学院

2023.11.13



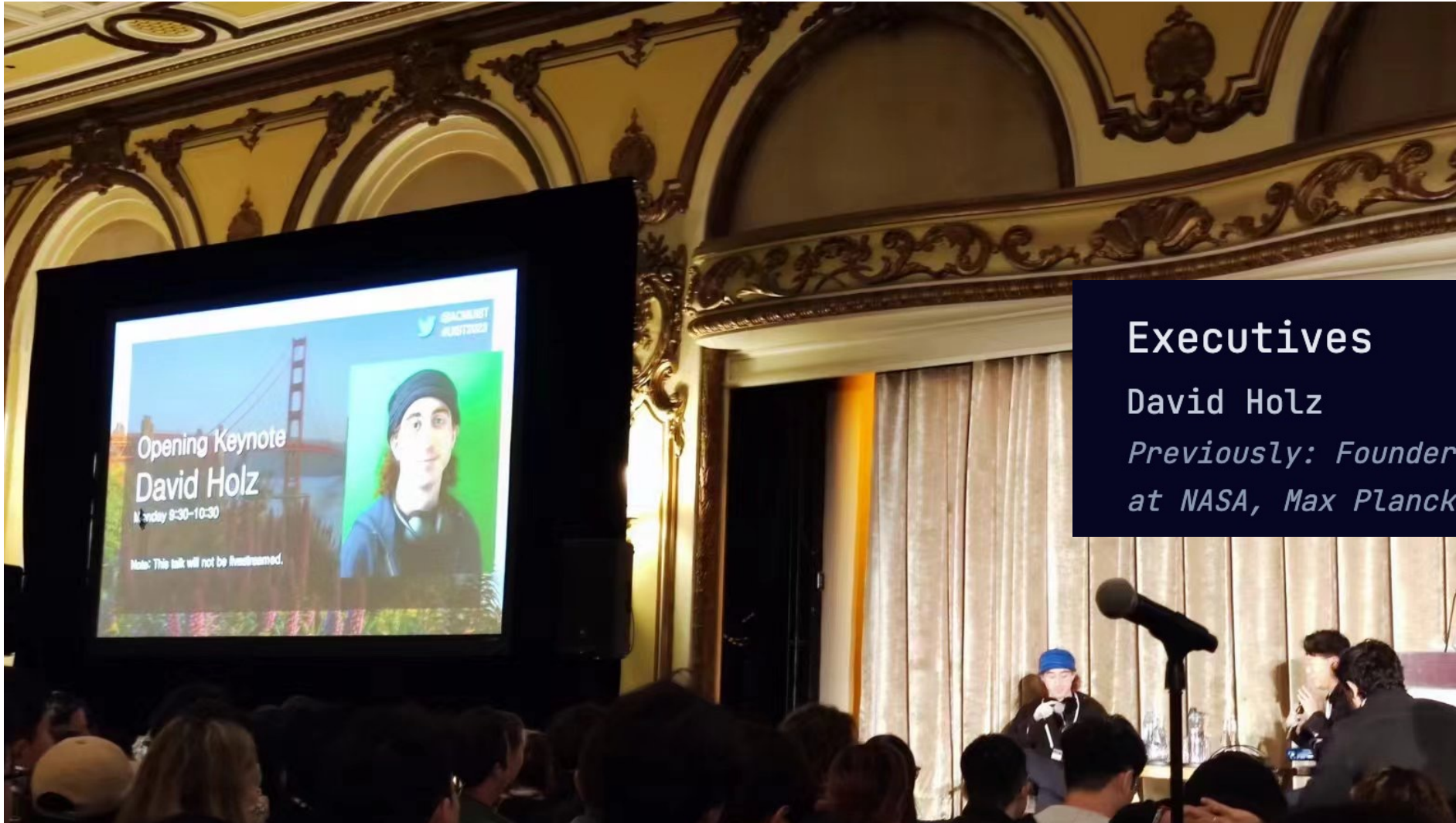
---

### **The Fairmont, San Francisco, California USA | Oct. 29 – Nov. 1, 2023**

The ACM Symposium on User Interface Software and Technology (UIST) is the premier forum for innovations in human-computer interfaces. Sponsored by ACM special interest groups on computer-human interaction (SIGCHI) and computer graphics (SIGGRAPH), UIST brings together people from diverse areas including graphical & web user interfaces, tangible & ubiquitous computing, virtual & augmented reality, multimedia, new input & output devices, Human-Centered AI, and CSCW. The intimate size and intensive program make UIST an ideal opportunity to exchange research results and ideas.



# Opening Keynotes



MidJourney CEO

## Executives


David Holz

*Previously: Founder Leap Motion, Researcher at NASA, Max Planck*

# MidJourney



`/v5_upscale` 14 hrs ago  
Asian mom and daughter, chinese half english, funny, shopping, walking, shopping, happiness, fashion, studio lighting, daylight,...

 ductstore01

# 访谈简记

---



- Don't start a startup because you want to be in charge of something
- 用视觉元素交互进行图片生成，效果很好。
- Do you publish? —> Bring something out that people can use. 每周都跟别人分享自己的想法，这可以看做我的publish... Am I helping the world? I feel good.

# Opening Keynotes



## GenAssist: Making Image Generation Accessible

Mina Huh  
The University of Texas at Austin  
Austin, TX, USA  
minahuh@cs.utexas.edu

Yi-Hao Peng  
Carnegie Mellon University  
Pittsburgh, PA, USA  
yihao@cs.cmu.edu

Amy Pavel  
The University of Texas at Austin  
Austin, TX, USA  
apavel@cs.utexas.edu

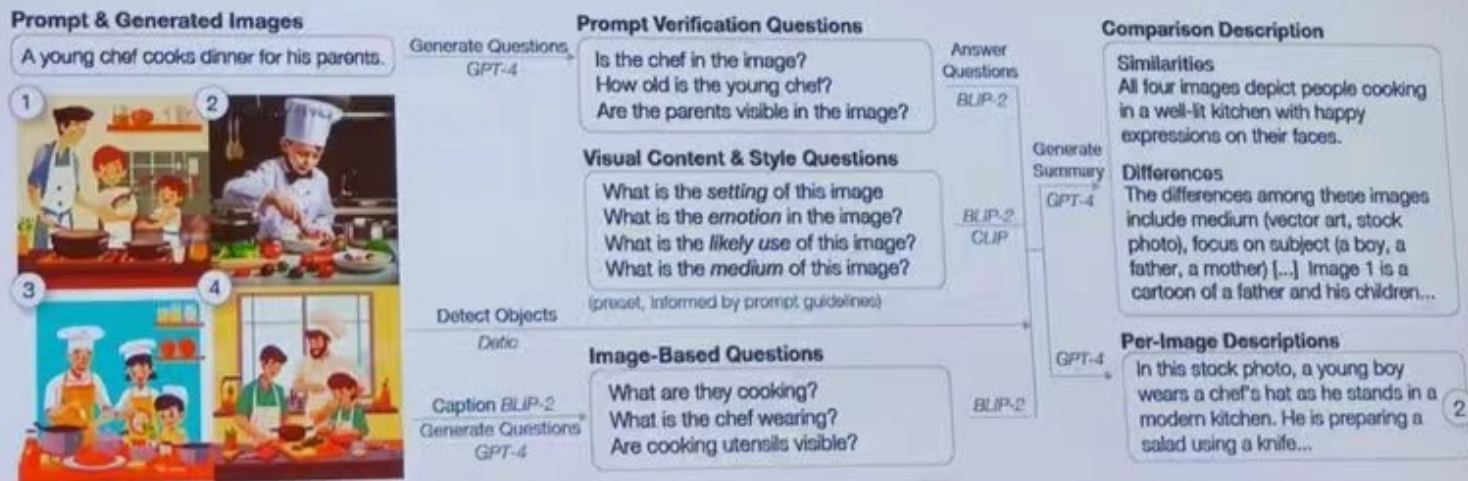


Figure 1: GenAssist makes image generation accessible by providing a comprehensive analysis of image generation results. Given

看标题，它是干啥的？

## Prompt & Generated Images

A young chef cooks dinner for his parents.



## Prompt Verification Questions

Generate Questions  
GPT-4

Is the chef in the image?  
How old is the young chef?  
Are the parents visible in the image?

Answer Questions  
BLIP-2

## Visual Content & Style Questions

What is the *setting* of this image?  
What is the *emotion* in the image?  
What is the *likely use* of this image?  
What is the *medium* of this image?

(preset, informed by prompt guidelines)

BLIP-2  
CLIP

Detect Objects  
Detic

## Image-Based Questions

What are they cooking?  
What is the chef wearing?  
Are cooking utensils visible?

BLIP-2

Caption BLIP-2  
Generate Questions GPT-4

## Comparison Description

### Similarities

All four images depict people cooking in a well-lit kitchen with happy expressions on their faces.

### Differences

The differences among these images include medium (vector art, stock photo), focus on subject (a boy, a father, a mother) [...] Image 1 is a cartoon of a father and his children...

### Per-Image Descriptions

In this stock photo, a young boy wears a chef's hat as he stands in a modern kitchen. He is preparing a salad using a knife...

Generate Summary

GPT-4

GPT-4

GenAssist is a system that enables blind or low vision creators to generate images by providing rich visual descriptions of the generation results. With the given text prompt and a set of generated images, GenAssist uses a large language model (GPT4) to generate prompt verification questions based on the text prompt, and image-based questions based on individual image captions (BLIP-2). GenAssist also extracts the visual content and style of the images using the vision-language model (CLIP, BLIP2), and object detection model (Detic). All of the information is then summarized using the GPT-4 to generate the comparison descriptions and per-image descriptions .

<https://minahuh.com/GenAssist/> 用的都是现有技术, why best paper? 框架, for social good



# Opening Keynotes



## Generative Agents: Interactive Simulacra of Human Behavior

Joon Sung Park  
Stanford University  
Stanford, USA  
joonspk@stanford.edu

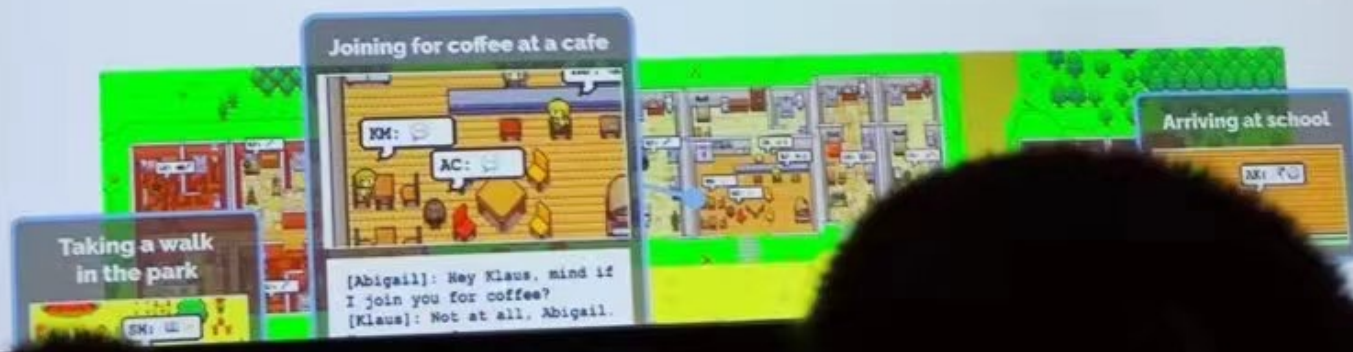
Joseph C. O'Brien  
Stanford University  
Stanford, USA  
jobrien3@stanford.edu

Carrie J. Cai  
Google Research  
Mountain View, CA, USA  
cjcai@google.com

Meredith Ringel Morris  
Google DeepMind  
Seattle, WA, USA  
merrie@google.com

Percy Liang  
Stanford University  
Stanford, USA  
плиang@cs.stanford.edu

Michael S. Bernstein  
Stanford University  
Stanford, USA  
msb@cs.stanford.edu



建模人类行为很难，chatgpt提供了一种解决思路，prompt+perceive+memory; generate plan, reflection and action by the language model。通过一个simulation，分析这些生成式agents的交互行为。通过interviews with the agents来做分析



# Generative Agents

## Interactive Simulacra of Human Behavior



**Joon Sung Park**  
Stanford University



**Joseph C. O'Brien**  
Stanford University



**Carrie J. Cal**  
Google Research



**Meredith Ringel Morris**  
Google DeepMind



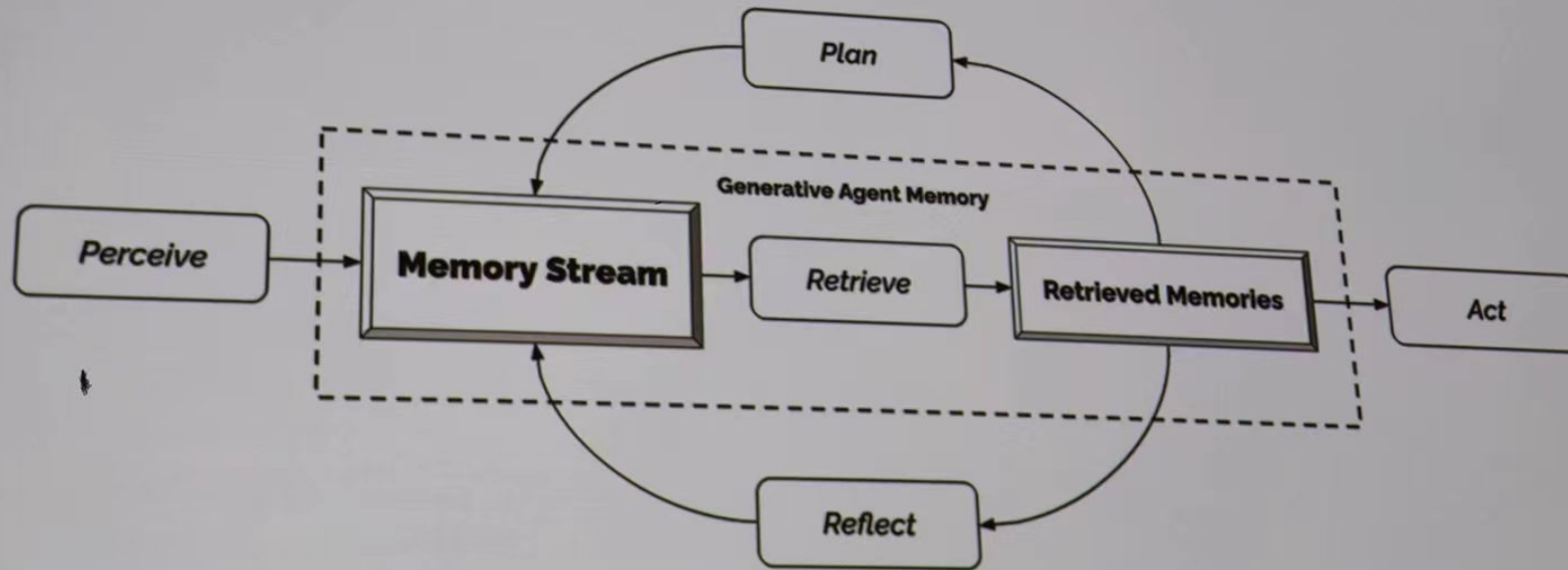
**Percy Liang**  
Stanford University



**Michael S. Bernstein**  
Stanford University

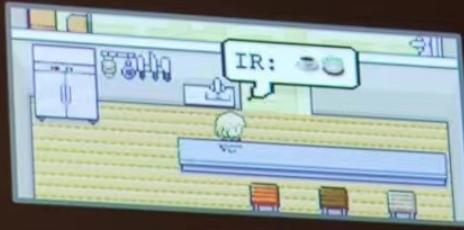


# Generative Agents

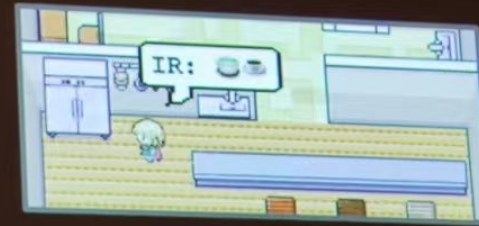


# Agents' actions impact the game environment

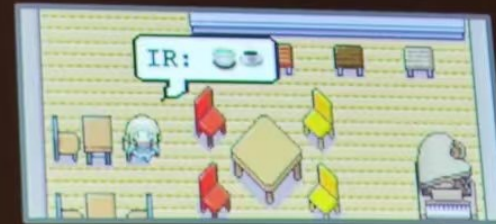
"Make and drink coffee"



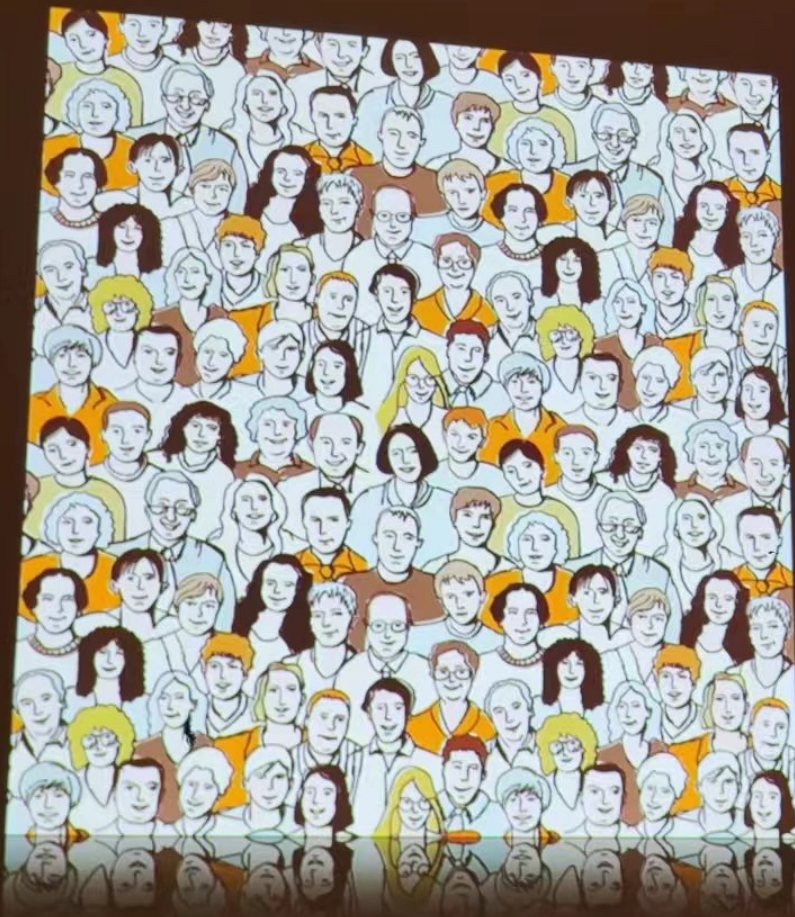
Cup -> **cleaned**



Coffee machine  
-> **turned on**



Chair -> **occupied**



Large language models can be prompted to generate human behavior conditioned on a variety of experiences.

GPT

"[name] is a [description]"

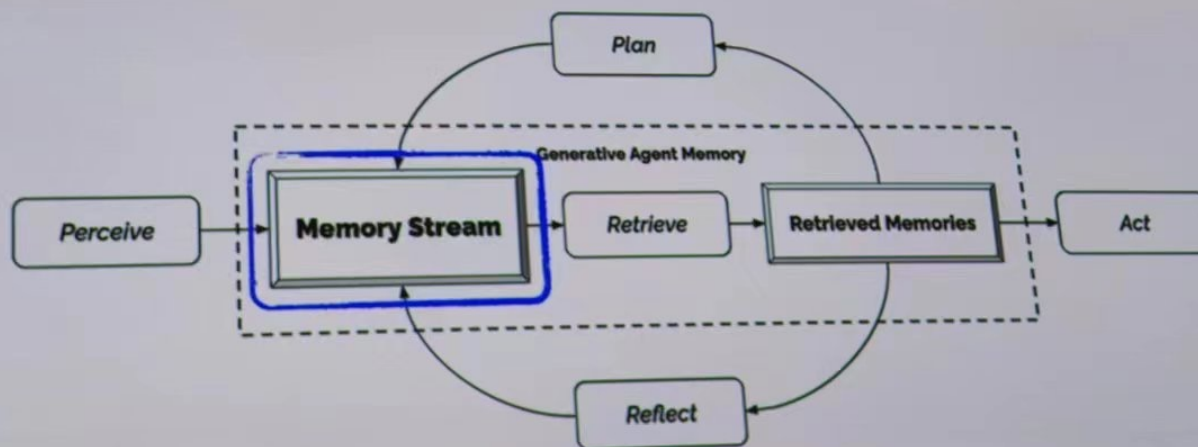
Social Simulacra (UIST '22)

Joon Sung Park, Lindsay Popowski, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein.  
Social Simulacra: Creating Populated Prototypes for Social Computing Systems. UIST 2022.

- Maria is chatting with Klaus
- The chair is empty
- Giorgio is playing the piano

### Isabella's Memory Stream

2023-02-13 22:48:20: Maria is chatting with Klaus  
2023-02-13 22:48:20: The chair is empty  
2023-02-13 22:48:20: Giorgio is playing the piano  
2023-02-13 22:48:20: Giorgio is playing the piano  
2023-02-13 22:48:20: Giorgio is playing the piano  
...



# Paper Presentation



## Generative Agents: Interactive Simulacra of Human Behavior

**Fastest growing repo on GitHub**

# of stars

Days	# of stars
0	0
1	~1000
2	~2000
3	~4000
4	~6000
5	~8000
6	~10000
7	~12000

7 days

**Half a million visited the demo**

We thank Stanford HAI, Google Research, Hasso Plattner Design Thinking Research Program, the Siegel Family Endowment, OpenAI, and the MSR PhD Fellowship for their funding support.

**Joon Sung Park**

Demo: [reverie.herokuapp.com/arXiv\\_Demo](https://reverie.herokuapp.com/arXiv_Demo/)  
Contact: [joonspk@stanford.edu](mailto:joonspk@stanford.edu) / [joonsungpark.com](https://joonsungpark.com)  
Twitter: [joon\\_s\\_pk](https://twitter.com/joon_s_pk)



# Paper Presentation



## Sensecape: Enabling Multilevel Exploration and Sensemaking with Large Language Models



**Sangho Suh**  
University of California  
San Diego  
sanghosuh@ucsd.edu



**Bryan Min**  
University of California  
San Diego  
bdmin@ucsd.edu



**Srishti Palani**  
University of California  
San Diego  
srpalani@ucsd.edu



**Haijun Xia**  
University of California  
San Diego  
haijunxia@ucsd.edu

**UC San Diego**  
The Design Lab

2

帮助更好地使用  
大模型的交互工  
具研究



SA

I plan to move to San Francisco as a recent graduate. What are some nice neighborhoods in SF?



San Francisco has several neighborhoods that are popular among recent graduates and young professionals due to their vibrant atmospheres, proximity to job opportunities, and entertainment options. Here are some nice neighborhoods in San Francisco for recent graduates:

1. **The Mission District:** The Mission is known for its artsy vibe, vibrant street art, and a wide range of dining options, from trendy restaurants to taquerias. It has a young and diverse population, making it a popular choice for recent graduates.
2. **SoMa (South of Market):** SoMa is a bustling neighborhood with a mix of tech companies, startups, and cultural venues. It's great for professionals looking to be close to job opportunities and enjoy a lively nightlife scene.
3. **Lower Haight:** This neighborhood offers a hip and eclectic atmosphere with unique boutiques, bars, and a strong sense of community. It's a great option for those who want a slightly quieter neighborhood but still want access to nightlife.

...

Remember that San Francisco's cost of living is high, so consider your budget when choosing a neighborhood. It's also a good idea to visit these neighborhoods in person to get a feel for their unique atmospheres and decide which one aligns best with your preferences and lifestyle. Additionally, factors like proximity to public transportation, work, and amenities should be considered in your decision-making process.



## How can we support information exploration & sensemaking with LLMs?

把LLM的搜索进行层级化，应对用GPT生成时答案过长过乱的问题

7



## Simple

SA

How many people live in San Francisco?



As of my last knowledge update in September 2021, the population of San Francisco was estimated to be around **883,305** people.

## Complex

SA

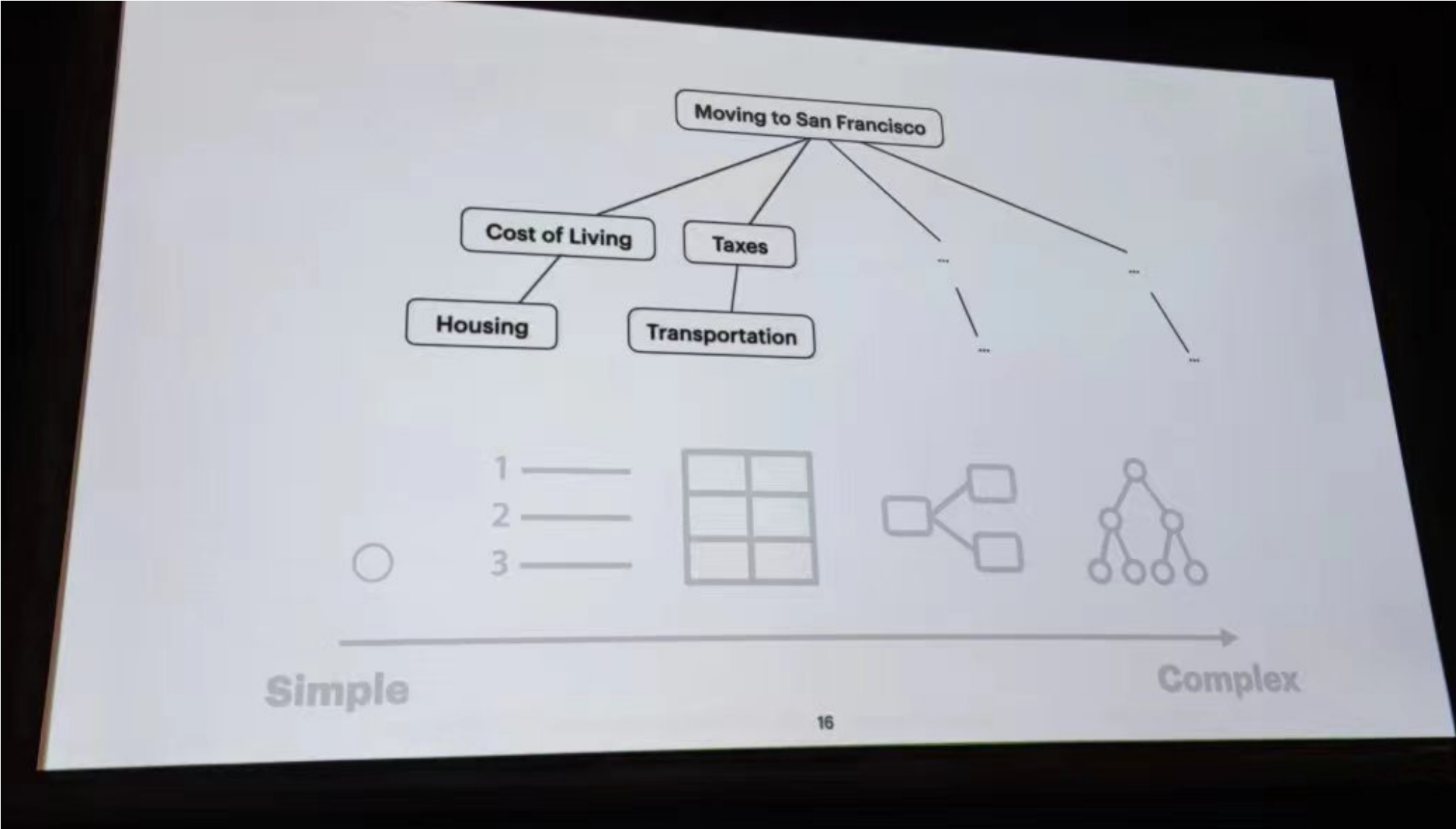
I plan to move to SF as a new graduate. What factors should I consider?



Moving to San Francisco as a new graduate is an exciting prospect, but there are several factors to consider to ensure a successful transition. Here are some key factors to keep in mind:

1. **Cost of Living:** San Francisco is known for its high cost of living. Be prepared for expensive housing, groceries, transportation, and entertainment. Make sure to create a budget and have a clear understanding of your expenses.
2. **Housing:** Finding affordable and suitable housing can be challenging. Research neighborhoods, consider roommates to share rent costs, and be prepared to act quickly when you find a place that suits your needs.
3. **Transportation:** San Francisco has a well-developed public transportation system, including buses, trams (Muni), and BART (Bay Area Rapid Transit). Consider the proximity of your home to work or school when choosing a neighborhood.
4. **Taxes:** Understand California's tax system, which

# Paper Presentation



Keyword(s): AI's impact on future of work

Why is Artificial Intelligence becoming increasingly important?  
 AI: Cost, Security, Efficiency, Innovation

Efficiency, Accuracy, Personalization, Innovation, Cost savings

Why is it important that automation and AI not stand for significant effects for people's work?  
 When will we begin to see a meaningful gap between using artificial intelligence for tasks traditionally done by humans?  
 When might ethics become a hurdle after companies in certain sectors of industries?  
 When should organizations begin to consider introducing AI solutions into their business operations?  
 Where are businesses turning to get help with understanding and implementing artificial intelligence technology in their companies?  
 Where do most firms see the biggest changes due to automation and artificial intelligence will occur? And in terms of job roles or sectors of jobs?  
 How can organizations ensure they are taking advantage of all available opportunities presented through advancements in artificial intelligence technology?  
 How might artificial intelligence revolutionize traditional working practices?

What challenges come with introducing AI into the workplace?  
 AI: Cost, Security, Efficiency, Innovation

Resistance to change, Bias and discrimination, Security risks, Ethics and transparency, Cost and complexity

What strategies can employers use to manage changes brought about by advances in AI technology?  
 AI: Cost, Security, Efficiency, Innovation

workforce, AI, collaboration, communication, strategic goals

Defining a Metrics Framework for AI

Hierarchy View

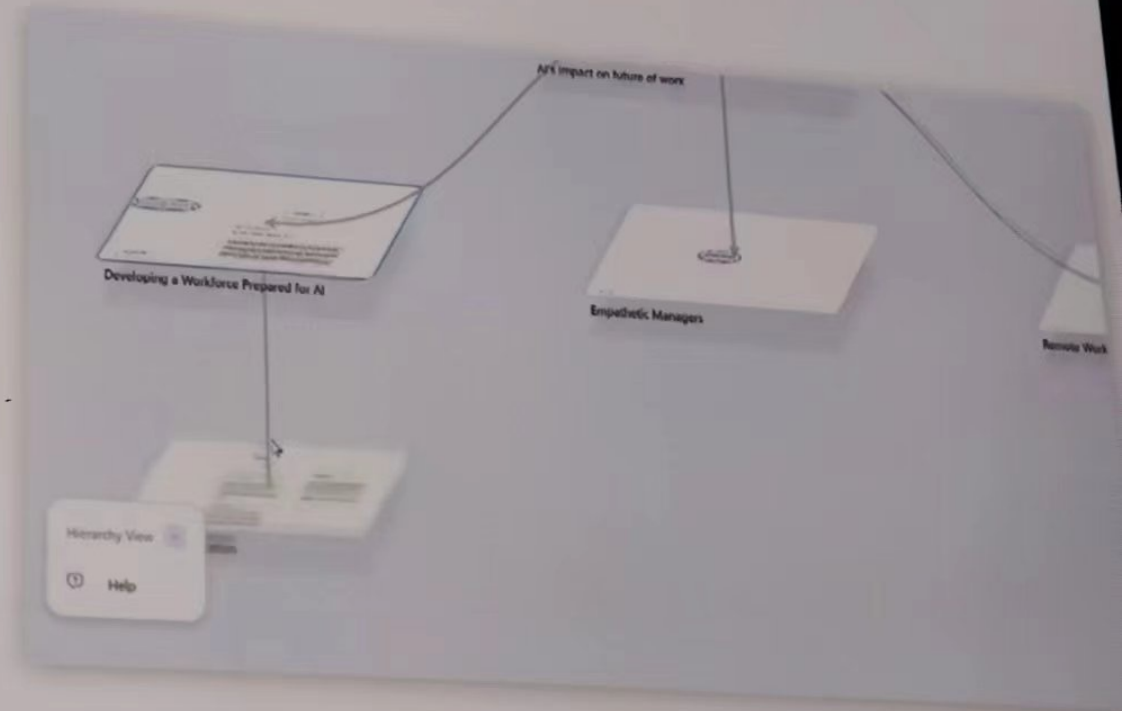
Help

18

# Canvas <-> Hierarchy

**Canvas**  
Space for exploring  
each *topic*

**Hierarchy**  
Space for visualizing  
relationships between  
topics





...

bicarbonate (HCO<sub>3</sub><sup>-</sup>)

...

Tell me about bicarbonate (HCO<sub>3</sub><sup>-</sup>)

Bicarbonate (HCO<sub>3</sub><sup>-</sup>) is an important molecule in the regulation of-base balance in the body.

which enables users to navigate across different

0:24 / 0:30

## Sensecape: Enabling Multilevel Exploration and Sensemaking with Large Language Models



ACM SIGCHI ✓  
2.03万位订阅者

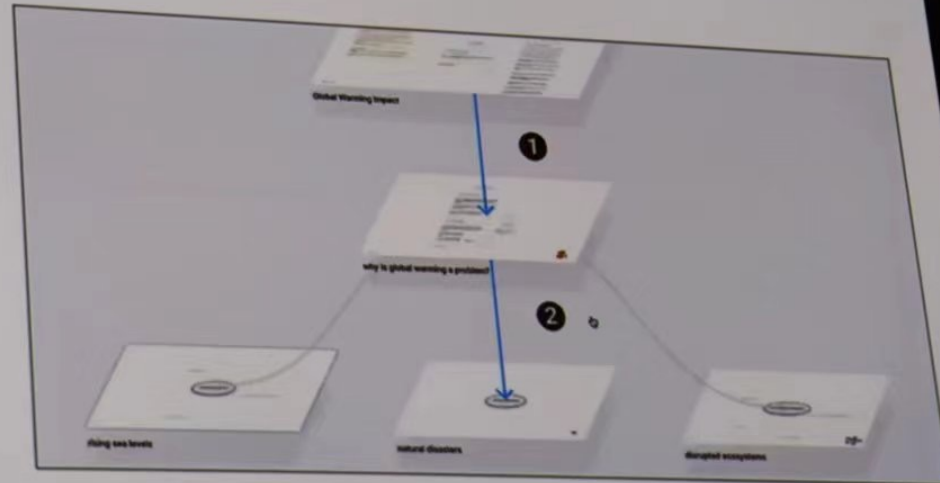




# Sensecape motivates exploration

"... **hierarchical interface** gave me clues on how to search, what to search.

It **motivated** me to **search** and **explore more** complicated information."



# Paper Presentation



## **G**raphologue

Exploring Large Language Model Responses  
with Interactive Diagrams



Peiling Jiang\*



Jude Rayan\*



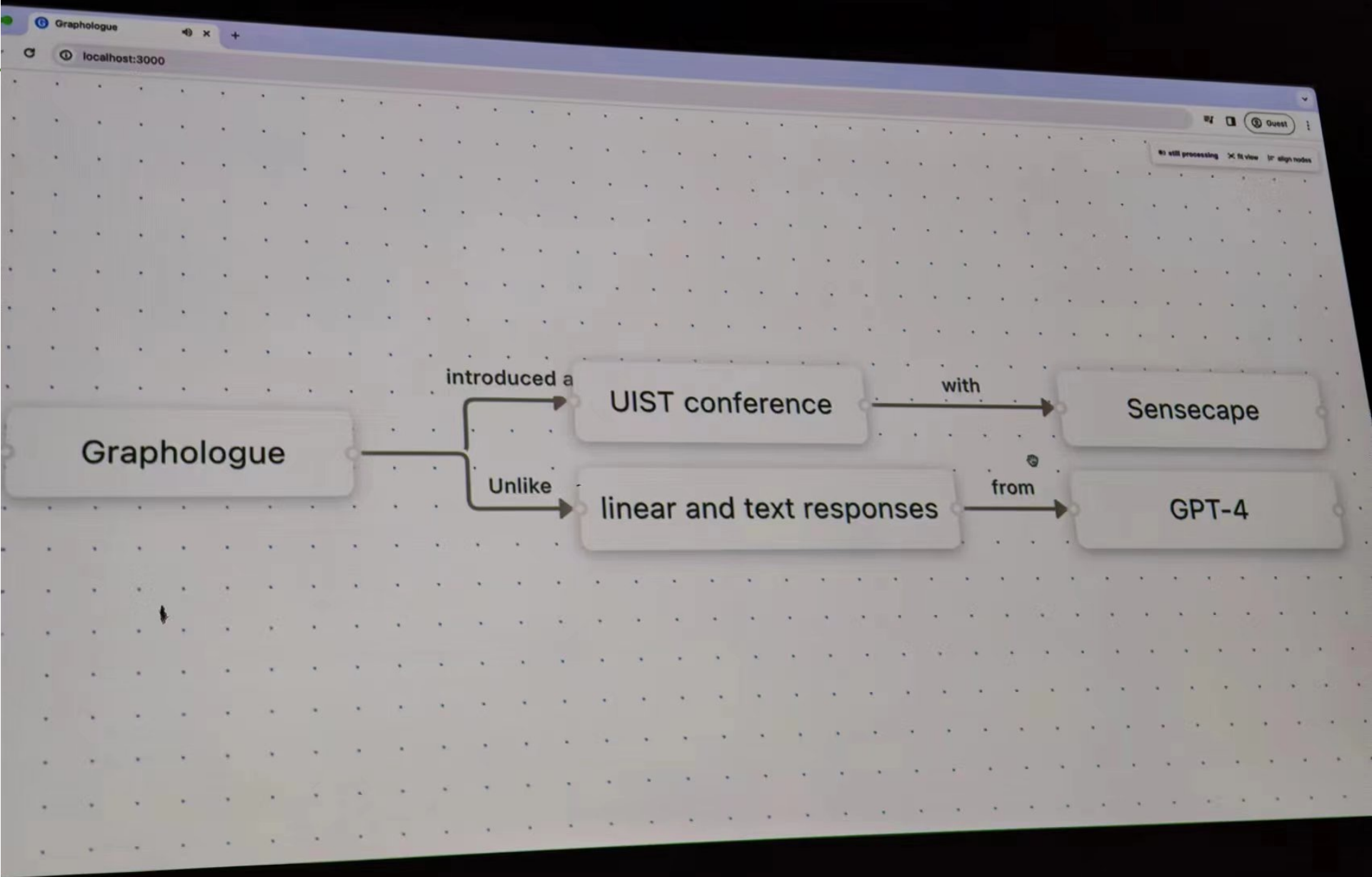
Steven Dow

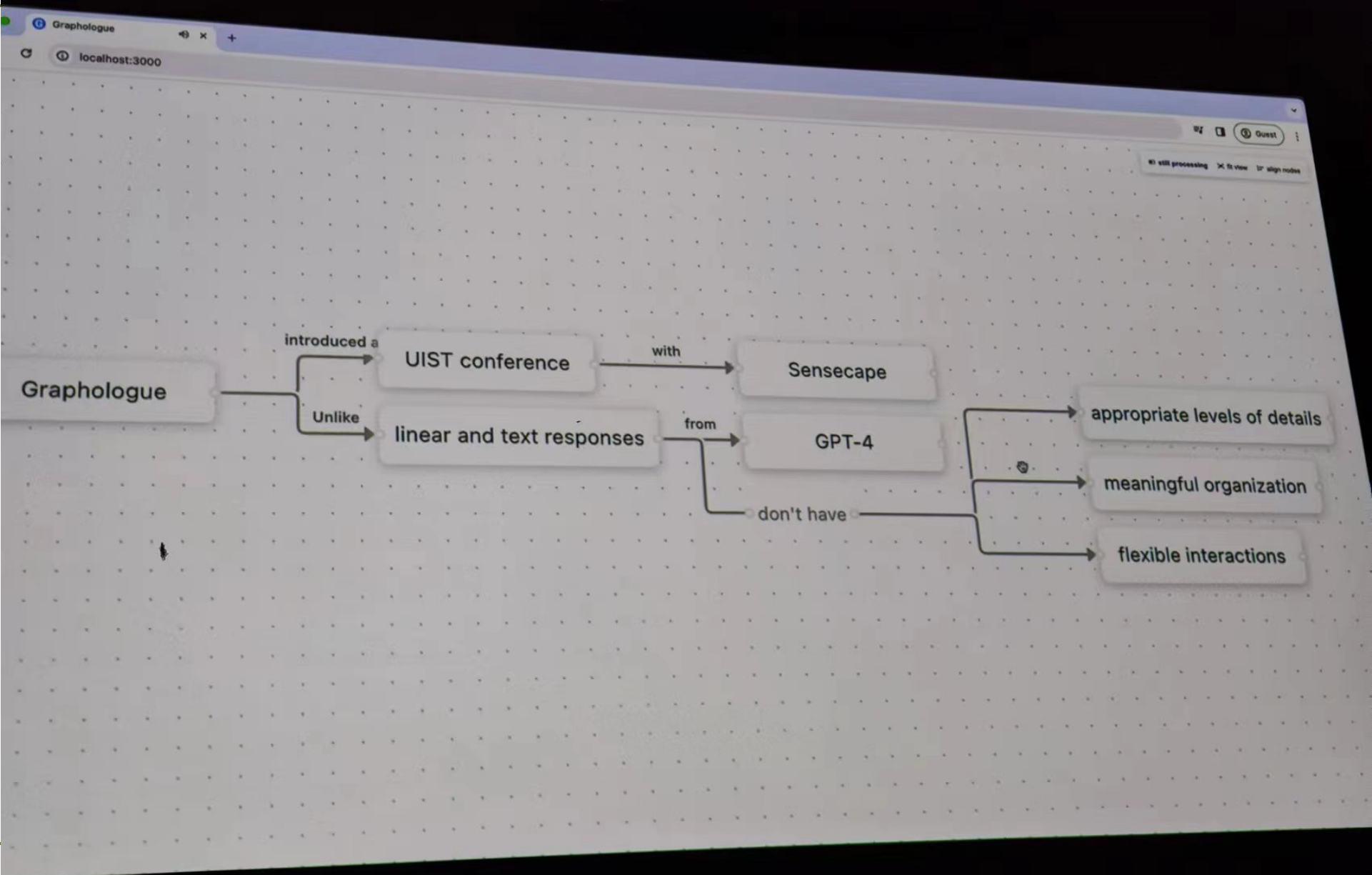


Haijun Xia

**Creativity Lab** University of California San Diego  
[creativity.ucsd.edu](http://creativity.ucsd.edu)

\* equal contribution







Text, then diagram

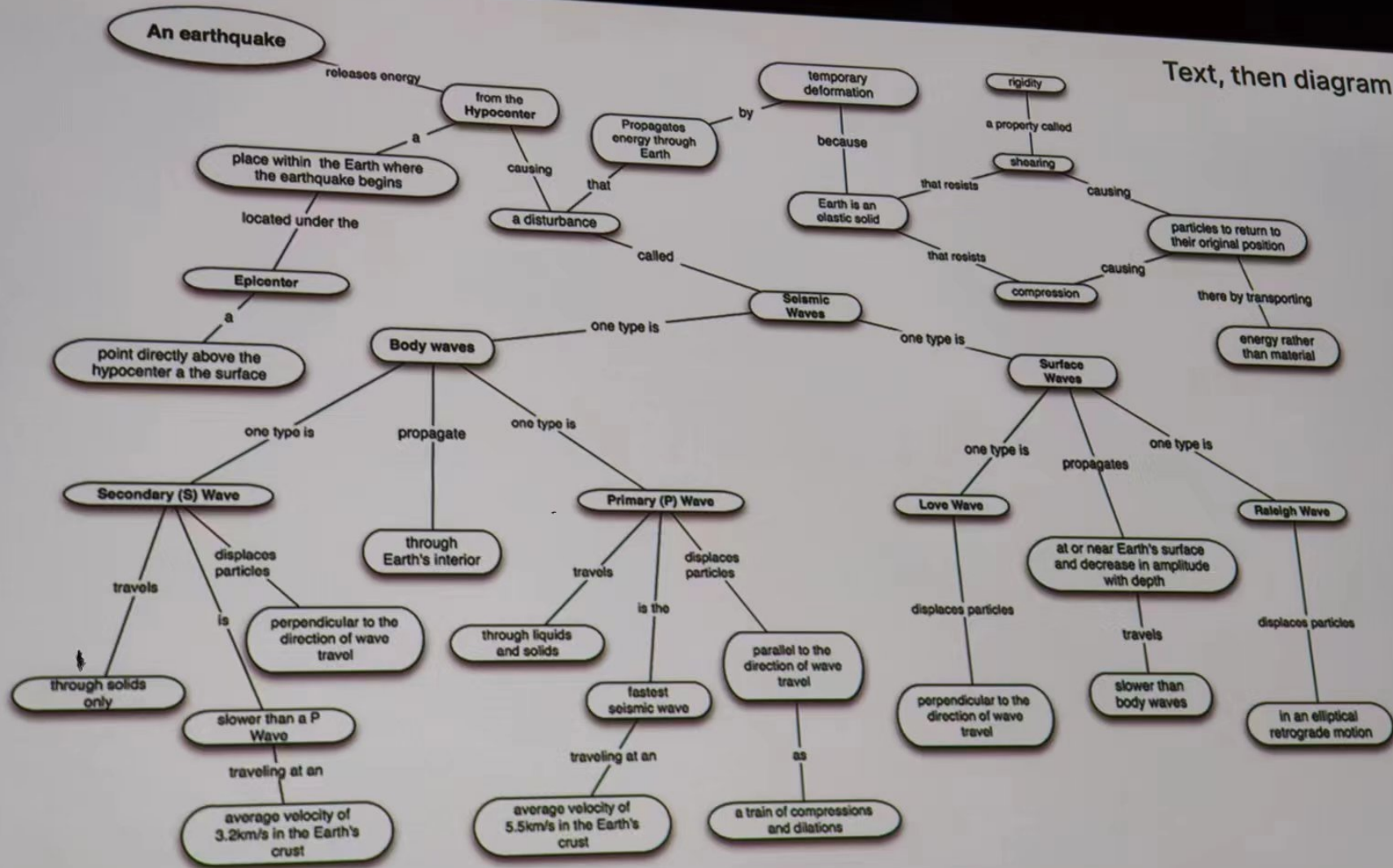


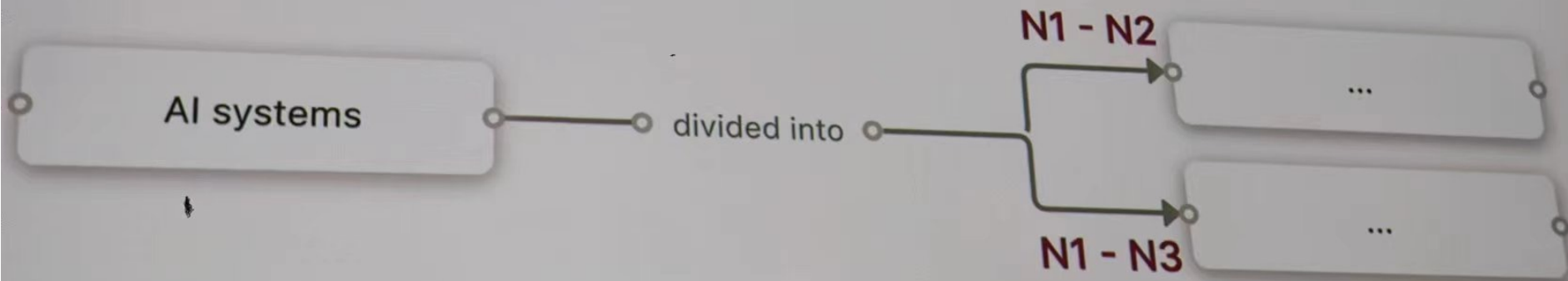


Diagram only

[Earthquakes] - (due to) - [complex processes]  
[complex processes] - (involving) - [the Earth's crust]  
[complex processes] - (involving) - [mantle]

GPT-4 response to parse and render as diagrams

Relationship annotation  
[AI systems (\$N1)] can be [divided into (\$H, \$N1, \$N2; \$H, \$N1, \$N3)]





original | outline | summary | highlight | hide

Earthquakes occur due to complex processes involving the Earth's crust, mantle, tectonic plates, and associated forces. One major process involved in earthquake formation is plate tectonics, which refers to the movement and interaction of large sections of the Earth's lithosphere.

original | outline | summary | highlight | hide

Plate boundaries play a crucial role in earthquake formation. Three main types of Plate Boundaries: Divergent Boundaries, convergent Boundaries, and Transform Boundaries. At these boundaries, tectonic plates interact in different ways, leading to the formation of earthquakes.

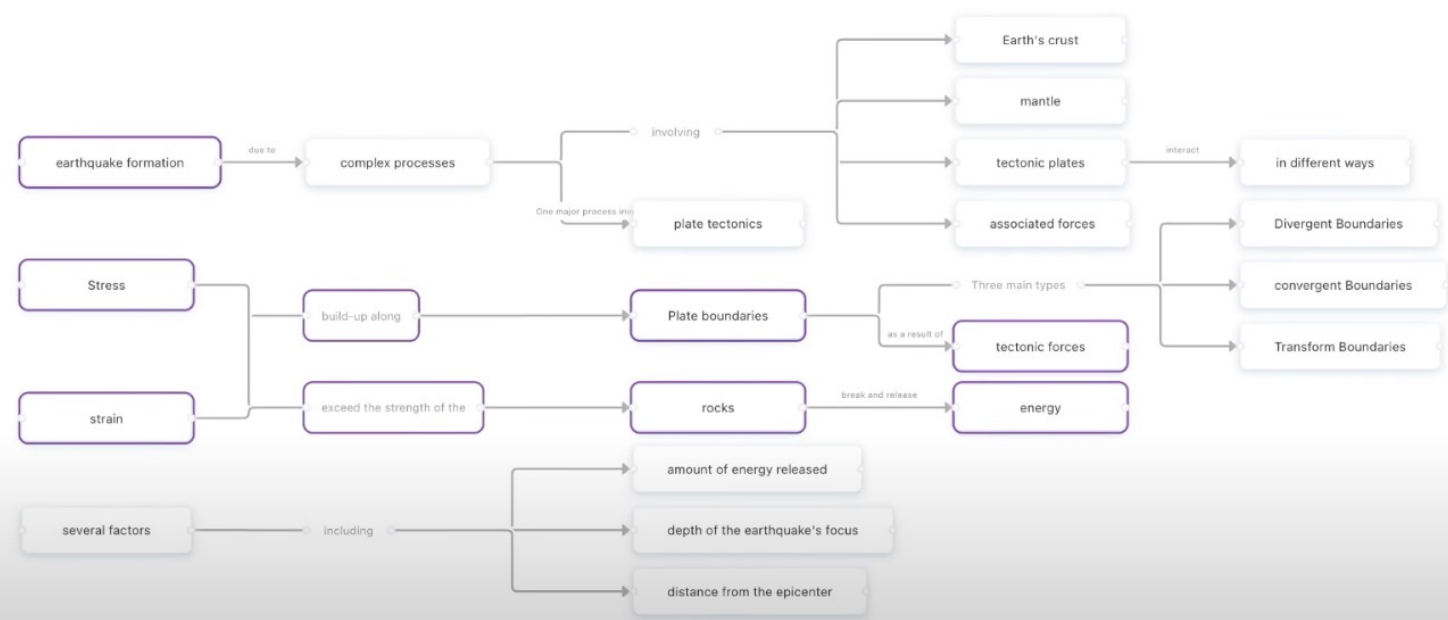
original | outline | summary | highlight | hide

Stress and strain build-up along plate boundaries as a result of tectonic forces. When Stress and strain exceed the strength of the rocks, they break and release energy in the form of seismic waves, which is felt as an earthquake.

original | outline | summary | highlight | hide

Finally, the magnitude and intensity of an earthquake depend on several factors, including the amount of energy released, the depth of the earthquake's focus, and the distance from the epicenter.

split diagram | merged diagram | saliency | annotation



0:12 / 0:26

⏪ ⏩ 🔊 ⏸ 📄 ⚙️ HD 🖥️ 🖱️

## Graphologue: Exploring Large Language Model Responses with Interactive Diagrams (Preview)

**C Creativity Lab**  
389位订阅者

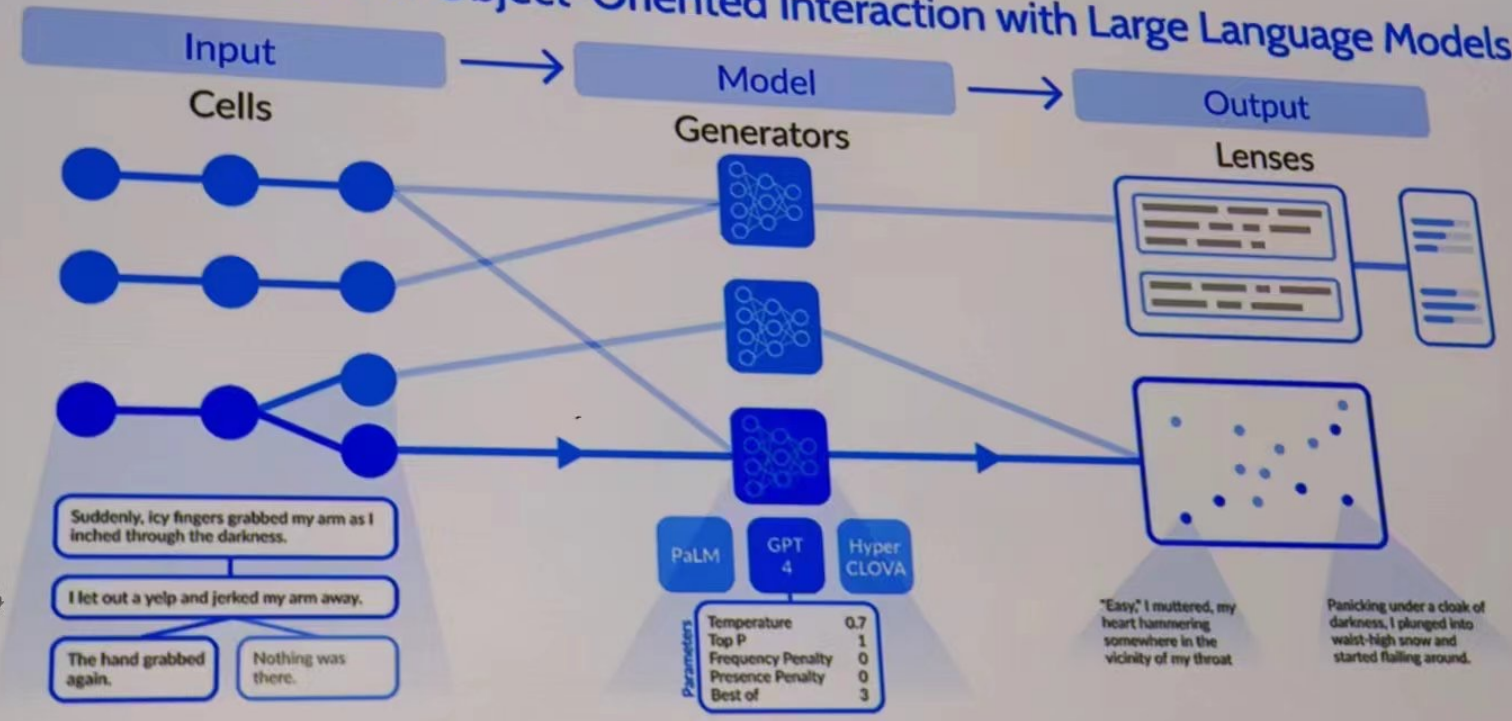
订阅

👍 8 | 🗨️ | ➦ 分享 | ≡+ 保存 | ⋮



# Cells, Generators, and Lenses

## Design Framework for Object-Oriented Interaction with Large Language Models



Tae Soo Kim, Yoonjoo Lee, Minsuk Chang, Juho Kim

[llm-objects.kixlab.org](http://llm-objects.kixlab.org)



Keynote File Edit Insert Slide Format Arrange View Play Share Window Help

Mon Oct 30 12:00 PM

## Motivation

# Traditional Text Editor-Like Design

Once upon a time in a quaint coastal town, there lived a sailor named Captain James O'Malley. He was known far and wide for his tales of the sea and the adventures he had encountered on his trusty ship, the "Siren's Call."

Captain O'Malley had been sailing the open waters for as long as he could remember. His weathered face bore the marks of countless sunrises and sunsets over the vast expanse of the ocean. His eyes sparkled with the wisdom of a man who had seen both the beauty and the fury of the sea.

One brisk morning, as the sun painted the sky in hues of pink and orange, Captain O'Malley stood on the shore, gazing out at the horizon. The salty breeze tousled his salt-and-pepper hair, when he noticed something in the horizon.

Model

Temperature

Top-P

8

# Design Framework Object-Oriented Interaction

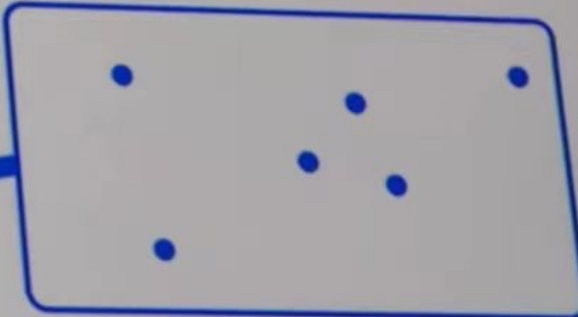
## Cells

- He was known far and wide for his tales of the sea and the adventures he had encountered on his trusty ship, the "Siren's Call."
- Captain O'Malley had been sailing the open waters for as long as he could remember.
- His weathered face bore the marks of countless sunrises and sunsets over the vast expanse of the ocean.
- His eyes sparkled with the wisdom of a man who had seen both the beauty and the fury of the sea.
- One brisk morning, as the sun painted the sky in hues of pink and orange, Captain O'Malley stood on the shore, gazing out at the horizon.
- The salty breeze tousled his salt-and-pepper hair, and he felt a longing in his heart for one last great adventure.

## Generators

- |          |       |
|----------|-------|
| Model    | Temp  |
| Chat GPT | 0.7   |
| Presence | Top-P |
| 0        | 1     |
- |          |       |
|----------|-------|
| Model    | Temp  |
| Chat GPT | 0.7   |
| Presence | Top-P |
| 0        | 1     |
- |          |       |
|----------|-------|
| Model    | Temp  |
| Chat GPT | 0.7   |
| Presence | Top-P |
| 0        | 1     |

## Lenses

- With a determined gleam in his eye, he ordered his crew to prepare the ship and set sail.
  - As he squinted his eyes to get a better look, he realized it was a ship in distress.
  - As he squinted his eyes and adjusted his telescope, Captain O'Malley let out a gasp.
- 



# User Study (N=18) Are the Objects Effective?

## Between-Subjects

Participants  
Novice writers  
(N=18)

Tasks  
Writing advertisements  
for 2 products

## Conditions

### Treatment

### Control



It all started when I accidentally picked up the wrong suitcase at the airport. It all happened because I forgot to check my yellow ribbon.

GPT-3	
Engine	Temp
D	0.7
Presence	Best Of
0	1

I'd just arrived in London and was on my way to my hotel.

I was really tired from my trip and just wanted to go to bed, so I decided to unpack the suitcase first.

Inside, I found someone else's clothes, passport, and other personal belongings.

Inside, I found someone else's clothes, passport, and other belongings.

GPT-3	
Engine	Temp
D2	1
Presence	Best Of
0	1

GPT-3	
Engine	Temp
D	1
Presence	Best Of
0	3

I had a lot of luggage with me, so I was in a hurry to get to my hotel.

0:25 / 0:29

## Cells, Generators, and Lenses: Design Framework for Object-Oriented Interaction with Large Langua...



ACM SIGCHI ✓  
2.03万位订阅者

订阅

0 | 分享 | 下载 | ...



# VISAR

## A Human-AI Argumentative Writing Assistant with Visual Programming and Rapid Draft Prototyping



Zheng Zhang



Jie Gao



Ranjodh Singh  
Dhaliwal



Toby Jia-Jun Li



应用大语言模型的文章



In The Dark Knight, Batman has a villainous character.

Please select the keywords to explore:

Actions and behavior of Batman

Motivations of Batman

Contrast with traditional heroic characteristics

Batman's use of fear and intimidation tactics

Batman's willingness to bend or break the law

Batman's disregard for personal and public safety

Batman's ambiguous moral code

GENERATE DISCUSSION POINTS



# VISAR

1

Hierarchical writing goal recommendation

2

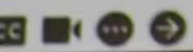
Argumentative sparks

Varying prototype editing strategies

4

Synchronized text and visual planning

3







Universities should require every student to take a variety of courses outside the student's field of study. From a career preparation standpoint, it is essential for universities to require students to take a variety of courses outside their field of study. Taking a variety of courses outside of one's field of study is crucial for building a well-rounded resume and skillset that can set students apart in the job market. Employers are looking for candidates who possess a diverse range of skills and knowledge, and taking courses outside of one's field of study demonstrates a willingness to learn and adapt. For example, a business major who takes courses in environmental science or sustainability will have a unique perspective and skillset that can be attractive to companies with a focus on corporate social responsibility. Additionally, taking courses outside of one's field of study can lead to unexpected career opportunities. A computer science student who takes a course in technical writing or user experience design, which can open up new career paths. Therefore, requiring students to take a variety of courses outside their field of study is a valuable way to provide them with a breadth of knowledge and skills that will serve them well in their personal and professional lives.

One of the main reasons why universities should require every student to take courses outside their field of study is to provide them with a breadth of knowledge. Your argument that universities should require students to take courses outside of their field of study is to provide them with a breadth of knowledge. Exposure to diverse subjects and perspectives provides numerous benefits. By taking courses in areas outside of their field of study, students can develop a well-rounded skillset that can be applied to any subject or profession. Additionally, taking courses in different fields can help students discover new passions and interests, which can lead to future career opportunities. Overall, requiring students to take a variety of courses outside their field of study is a valuable way to provide them with a breadth of knowledge and skills that will serve them well in their personal and professional lives.

Counter Arguments:

1. Limited Time and Resources: For many students, taking courses outside their field of study can be a significant time and financial commitment. This can be particularly challenging for students who are working part-time or supporting themselves financially while in school. In some cases, taking additional courses outside of their major may not be feasible or practical, and students may need to focus on completing their degree requirements in a timely manner.
2. Irrelevant Skills: Some courses outside a student's field of study may not be directly applicable to their major or future career. This can lead to a waste of time and resources, as students may not gain the skills or knowledge they need to succeed in their chosen field.
3. Lack of Focus: Taking a variety of courses can lead to a lack of focus on a student's primary field of study. This can result in lower grades and a less thorough understanding of their major, which can be detrimental to their long-term success.
4. Limited Practical Experience: While taking courses outside of their field of study can provide students with a breadth of knowledge, it may not provide them with the practical experience they need to succeed in their chosen profession. Hands-on learning through internships, co-op programs, and other experiential learning opportunities is often more valuable than classroom instruction.
5. Lack of Practical Skills: While taking courses outside of their field of study can provide students with a breadth of knowledge, it may not provide them with the practical skills they need to succeed in their chosen profession. Hands-on learning through internships, co-op programs, and other experiential learning opportunities is often more valuable than classroom instruction.

REVIEW AND SKETCH

student's field of study is to provide them with a breadth of knowledge. Your argument that universities should require students to take courses outside of their field of study is to provide them with a breadth of knowledge. Exposure to diverse subjects and perspectives provides numerous benefits. By taking courses in areas outside of their field of study, students can develop a well-rounded skillset that can be applied to any subject or profession. Additionally, taking courses in different fields can help students discover new passions and interests, which can lead to future career opportunities. Overall, requiring students to take a variety of courses outside their field of study is a valuable way to provide them with a breadth of knowledge and skills that will serve them well in their personal and professional lives.

思考：但这整个流程，会不会跟通常写作的流程有冲突？值得探讨。



## VISAR: A Human-AI Argumentative Writing Assistant



Toby Li  
34位订阅者



1 | 分享 | 下载 | ...

# PromptPaint: Steering Text-to-Image Generation Through Paint Medium-like Interactions

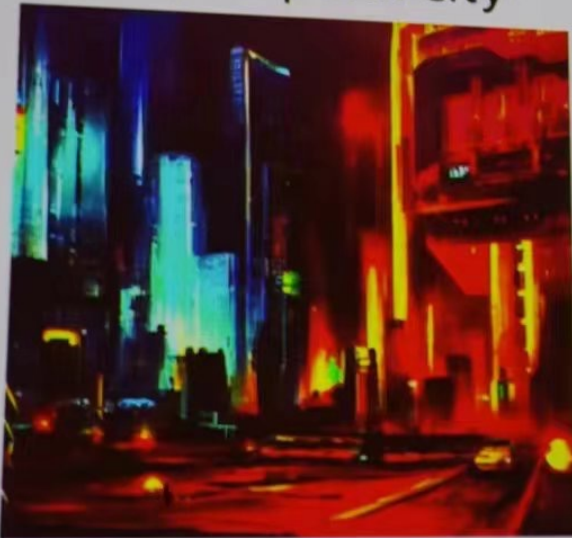


John Joon Young Chung\*  
SpaceCraft Inc.  
Los Angeles, USA  
jjyc@spacecraft.inc

Eytan Adar  
University of Michigan  
Ann Arbor, USA  
eadar@umich.edu

Sometimes, it is not steerable enough

"rough oil painting  
of cyberpunk city"



Flatter image...

How much flat though?



Impressionist dog

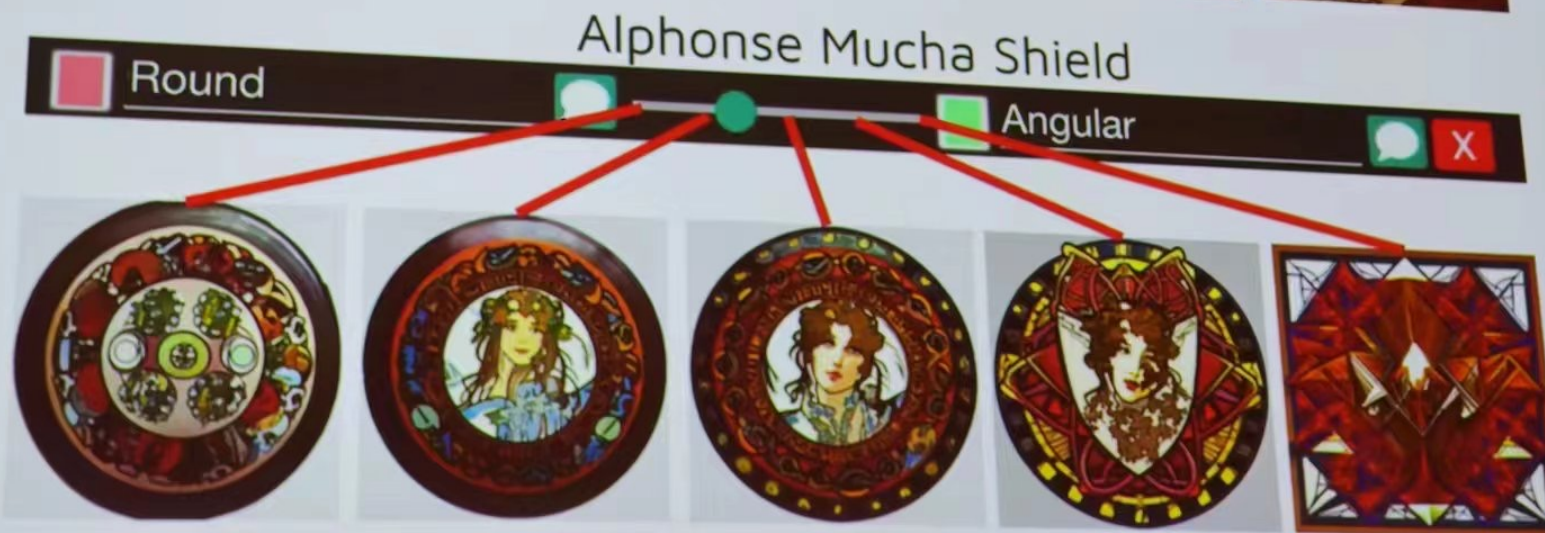
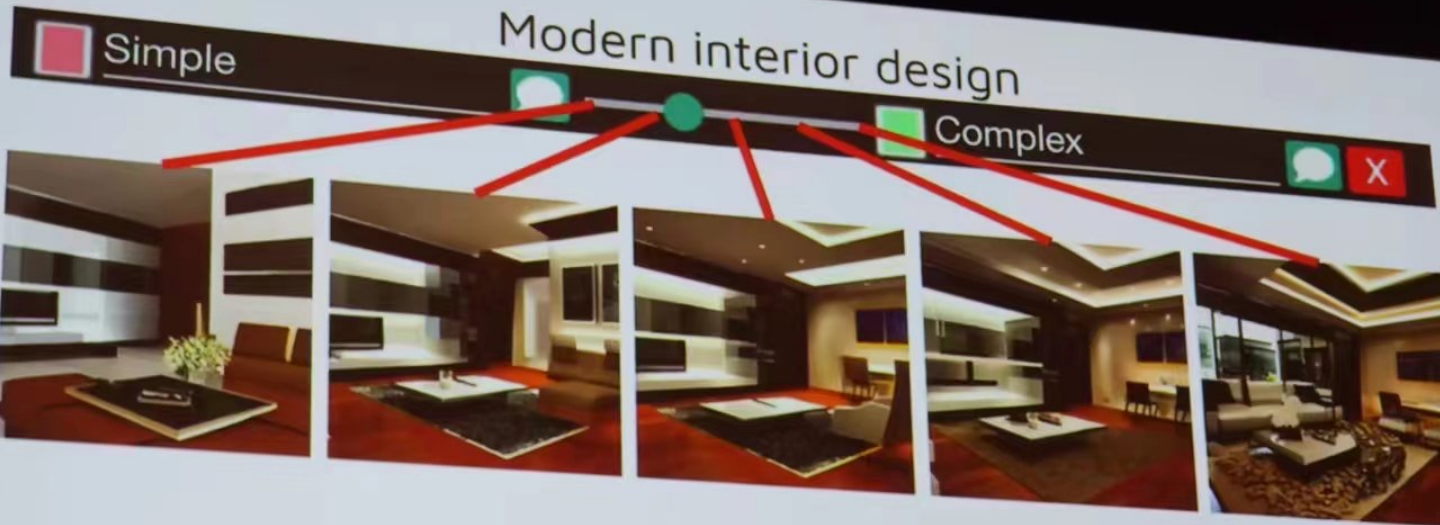


Impressionist cat

Cubist ocean



Cyberpunk ocean



A photo of a rabbit  
⇒ A ceramic cup

No Switching



Switching at 75%



Switching at 50%

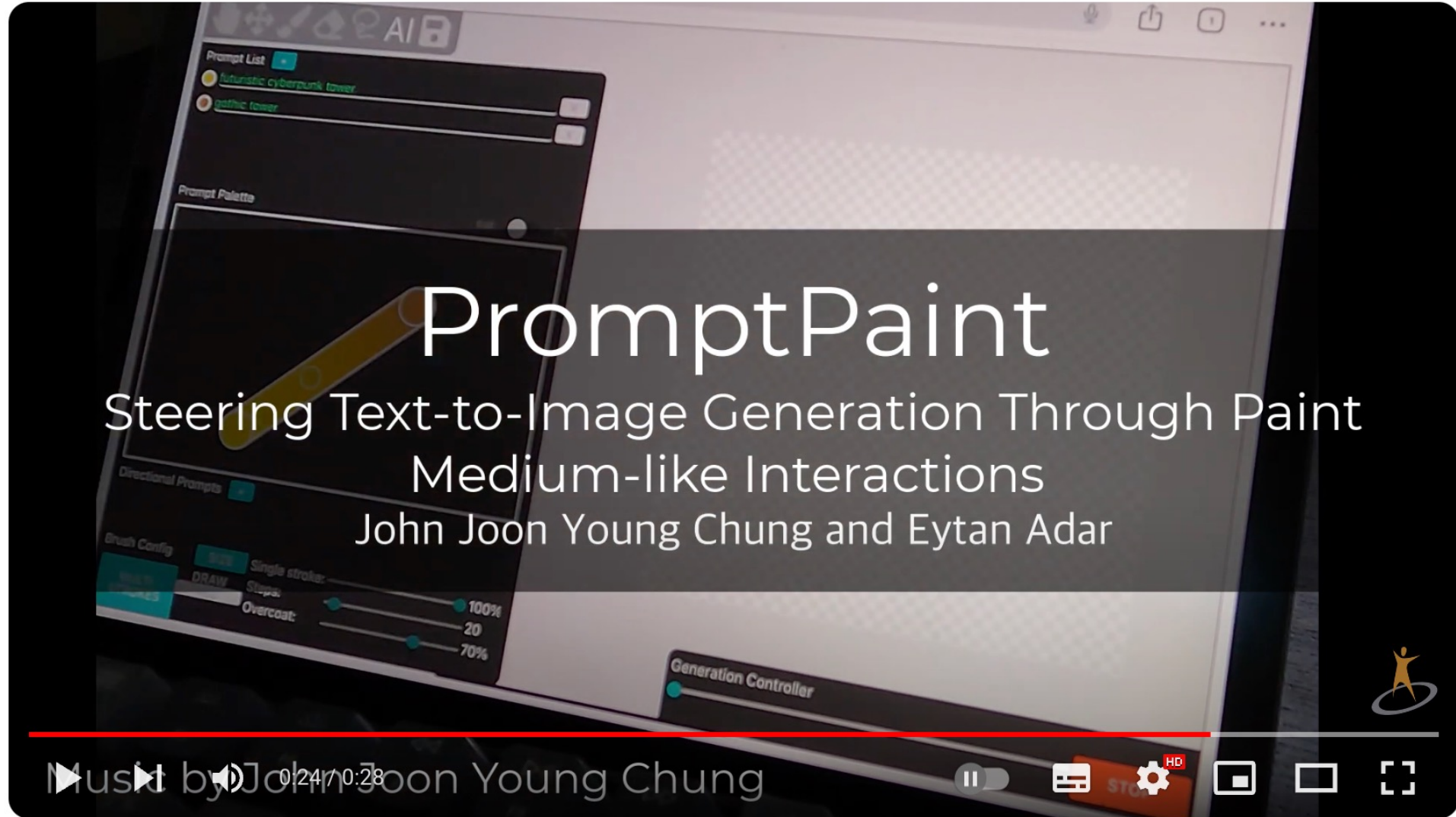


Switching at 25%



An action painting piece  
⇒ A lego dragon





## PromptPaint: Steering Text-to-Image Generation Through Paint Medium-like Interactions



ACM SIGCHI ✓  
2.03万位订阅者

订阅

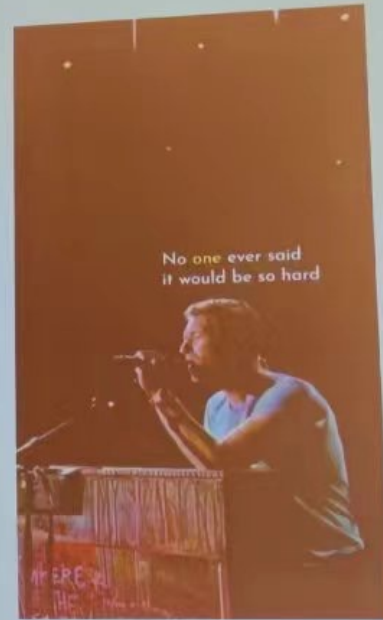
0 | 分享 | 下载 | ...

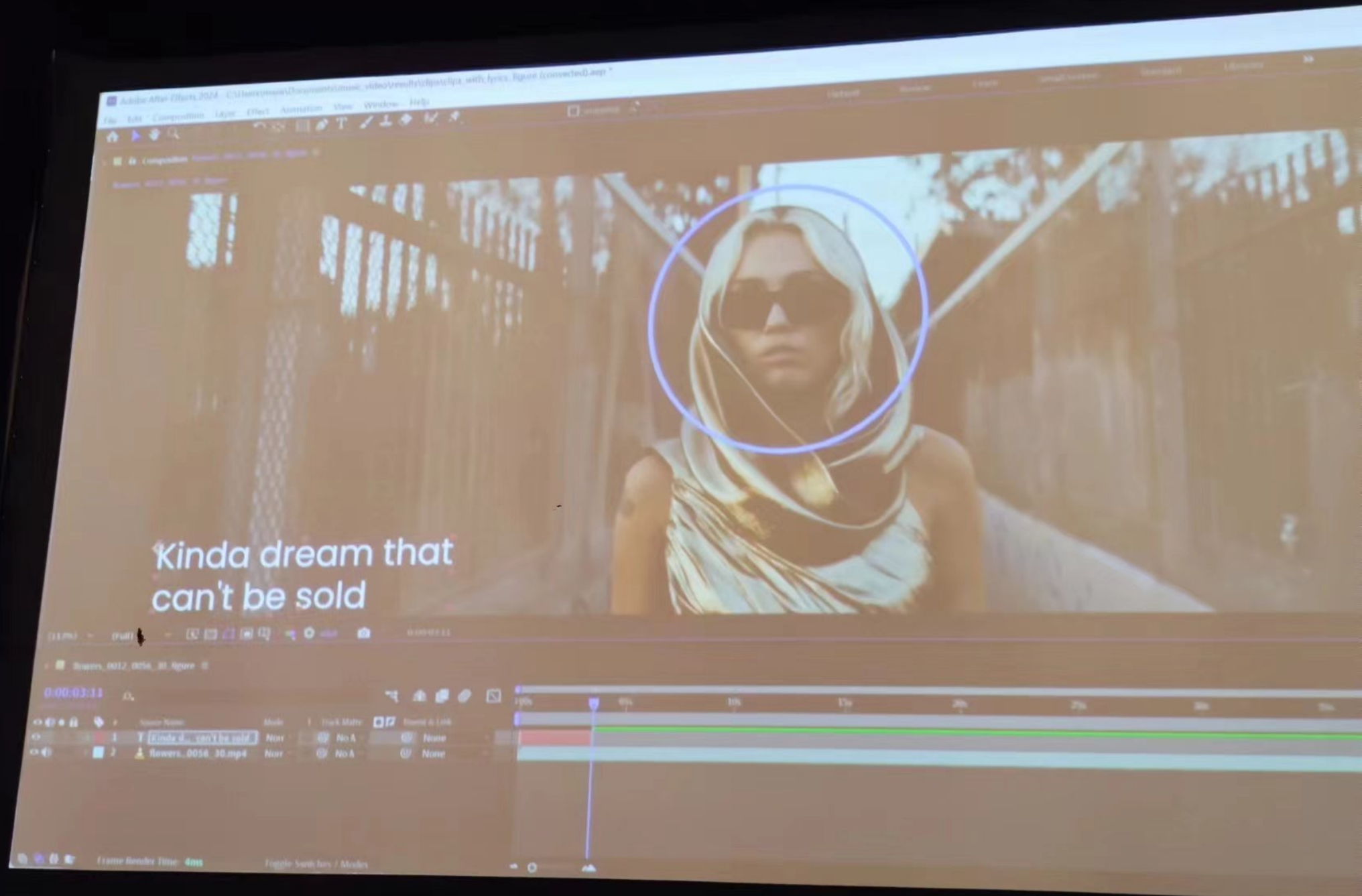
# Automated Conversion of Music Videos into Lyric Videos

Jiaju Ma<sup>\*†</sup>, Anyi Rao<sup>\*</sup>, Li-Yi Wei<sup>†</sup>, Rubaiat Habib Kazi<sup>†</sup>,  
Hijung Valentina Shin<sup>†</sup>, and Maneesh Agrawala<sup>\*\*‡</sup>



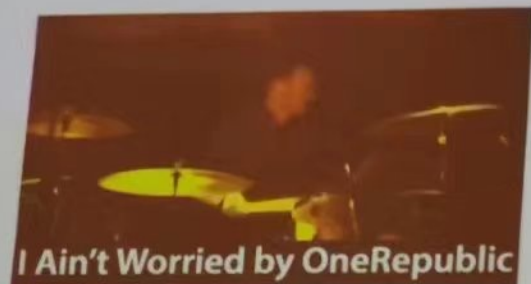
Stanford University<sup>\*</sup>, Adobe Research<sup>†</sup>, Roblox<sup>‡</sup>  
UIST 2023







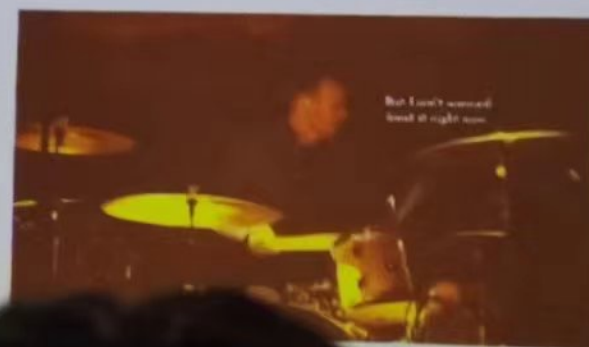




Stage 1  
Lyrics Phrasing  
and Line Breaking

Stage 2  
Lyrics Placement  
via Optimization

Stage 3  
Lyrics Rendering  
with Animation



~~DG1: Follow song phrasing~~

~~DG2: Break long lines~~

DG3: Highlight word being sung

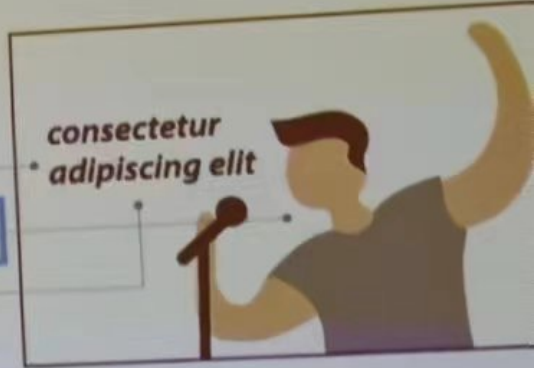
DG4: Maintain color contrast



~~DG5: Synchronize to music~~

DG6: Place near focus of attention

DG7: Place consistently



DG6: Place near focus of attention



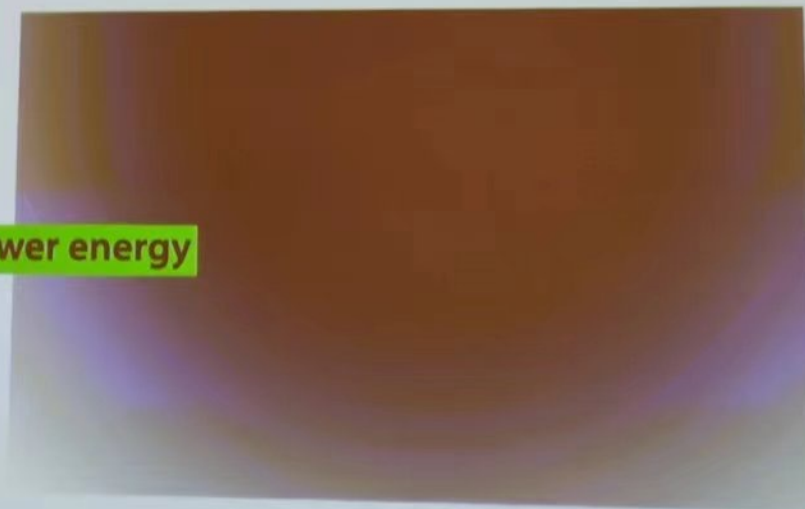
↓ Convolution



Foreground Cost Map



↓ Distance



Focus of Attention Cost Map

Darker means lower energy



## Automated Conversion of Music Videos into Lyric Videos



ACM SIGCHI ✓  
2.03万位订阅者





# Mirrorverse

## LIVE TAILORING OF VIDEO CONFERENCING INTERFACES

Jens Emil Grønbæk,<sup>1</sup> Marcel Borowski,<sup>1</sup> Eve Hoggan,<sup>1</sup> Wendy E. Mackay,<sup>2</sup>  
Michel Beaudouin-Lafon,<sup>2</sup> Clemens N. Klokrose<sup>1</sup>

<sup>1</sup> Aarhus University | <sup>2</sup> Université Paris-Saclay, CNRS, Inria

UIST 2023



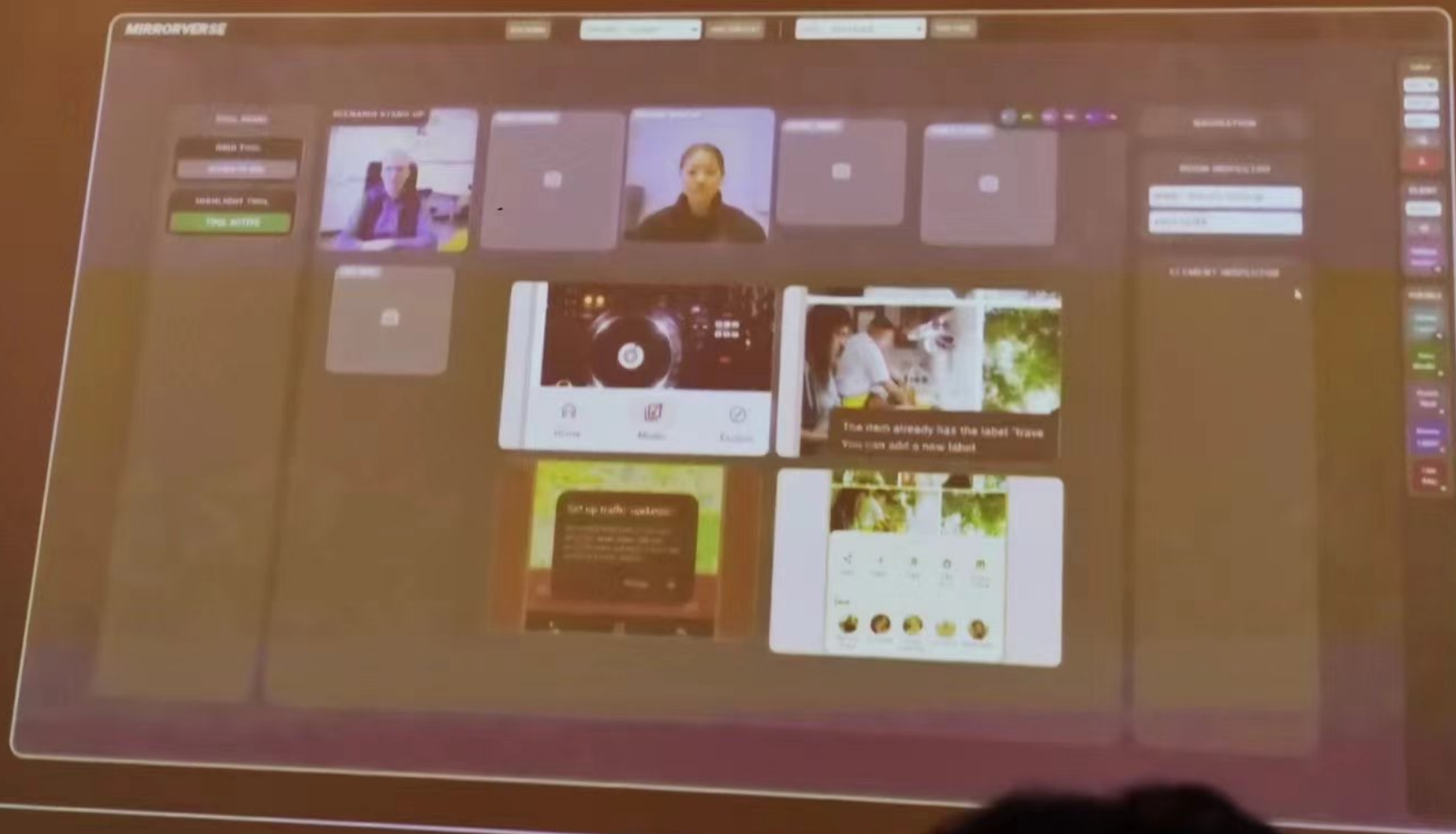
AARHUS  
UNIVERSITY

Inria

# REPLICATING MEETING EXPERIENCES



# IMPROMPTU REPROGRAMMING OF TOOLS







## Mirrorverse: Live Tailoring of Video Conferencing Interfaces



ACM SIGCHI ✓  
2.03万位订阅者

订阅

1 1

分享

保存

134次观看 2周前

<https://www.youtube.com/watch?v=jBiG48rW6Ck>



# Papeos

Augmenting Research **Papers** with Talk Videos



Tae Soo Kim  
KAIST



Matt Latzke  
AI2



Jonathan Bragg  
AI2



Amy X. Zhang  
University of  
Washington



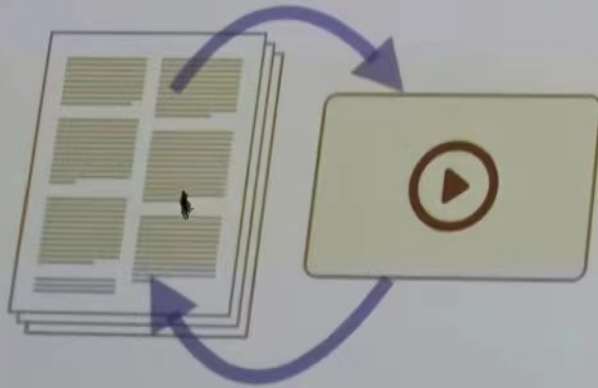
Joseph Chee Chang  
AI2



## Design Goals

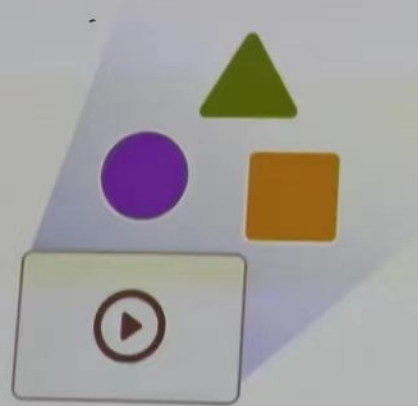
**DG1**

Fluid Switching



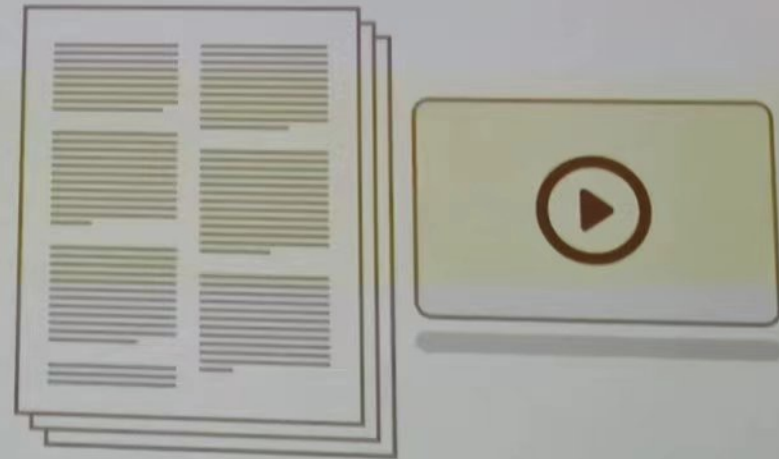
**DG2**

Surface Visuals



**DG3**

Coarse & Fine



### 3 FORMATIVE AND CO-DESIGN STUDY

To explore the design space for combining research papers and talk videos, we conducted a formative study where participants explored the opportunities and challenges in combining the two formats from the perspectives of both readers and authors.

#### 3.1 Participants

We invited 14 researchers who had previously published at least one paper and created accompanying talk videos. 10 were doctoral students, 2 were Master's students, and the remaining 2 were a postdoc and an undergraduate student. 10 of the 14 participants identified their discipline as human-computer interaction (HCI) or related sub-fields (e.g., visualizations, AI fairness), 3 as natural language processing (NLP), 2 as machine learning (ML), and 1 as computer vision (CV).<sup>1</sup>

#### 3.2 Apparatus

Consuming scholarly papers and talk videos at the same time is a new experience that may be hard for participants to imagine. In a preliminary version of this formative study, we gave participants (n=4) a paper and talk video pair side-by-side and instructed them to "understand the content of the paper based on your real-life habits". Although participants could freely choose how they wished to consume the paper and video, they all watched the whole video first and then delved into the paper. Participants expressed how this was not due to a lack of desire to jump to the paper while watching the video, but due to the prohibitively high cost of cross-referencing between formats. This preliminary study revealed that unaugmented papers and videos were inadequate to explore how readers wanted to leverage both formats together.

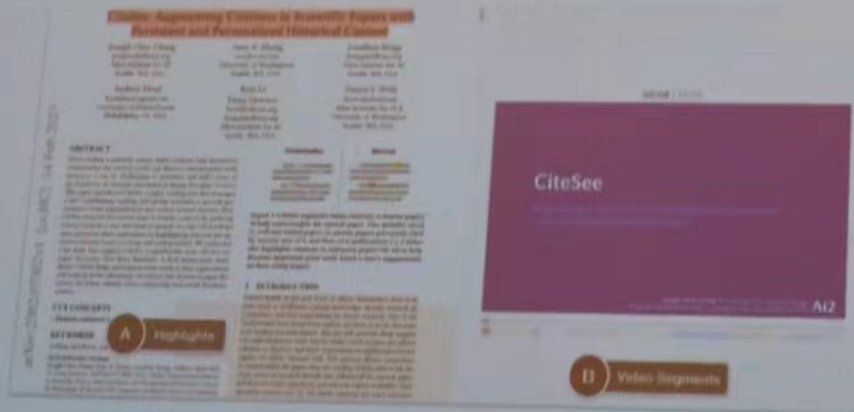
<sup>1</sup>Several participants identified with multiple disciplines.

Thus, we developed a technology probe [35] (Fig. 2) where we could pre-link segments of a talk video to relevant passages in the paper (e.g., paragraphs, figures) and color-code them so that participants could switch between the two formats with lower cost. Before the study, one of the authors manually created the links between the papers and videos for three papers in each of the recruited participants' research fields (e.g., empirical HCI, systems HCI, NLP, CV). To create these links, the author followed criteria that were based on insights from the preliminary study: segment the video on slide transitions, and link segments to paragraphs based on content similarity (e.g., phrases, figures) while following the paper's reading order.

#### 3.3 Study Procedure

The study consisted of two consecutive sessions. First, there was a formative session where participants took the perspective of paper readers and used the technology probe (Fig. 2) to read a paper where several passages were pre-linked to relevant segments of the talk video. Then, in a co-design session, participants took the perspective of paper authors and considered designs for combining their own research papers and talk videos.

For the formative session, participants chose their preferred paper from the set of pre-linked paper-video pairs and, while thinking aloud, read the paper using the technology probe for 20 minutes. In the probe, linked passages in the paper were highlighted, and participants could click on a linked passage to automatically navigate to the corresponding segment in the video. The video segments were also displayed under the video timeline, and participants could click on a video segment to scroll to the corresponding passage in the paper. After the reading period, participants were asked about the benefits and drawbacks of using the probe and the talk video during paper reading.



00:06

### Academic Papers and Talk Videos



**Transcript** To explore the design space for combining these two formats,

## Video Notes

00:20

### Formative and Co-Design Study As Readers and As Authors

Explore the Design Space for Combining Papers and Videos (N=14)

Formative Study



Talk: Use a research probe to consume a paper and video together

**Transcript** In the formative part of the study, participants took the perspective of

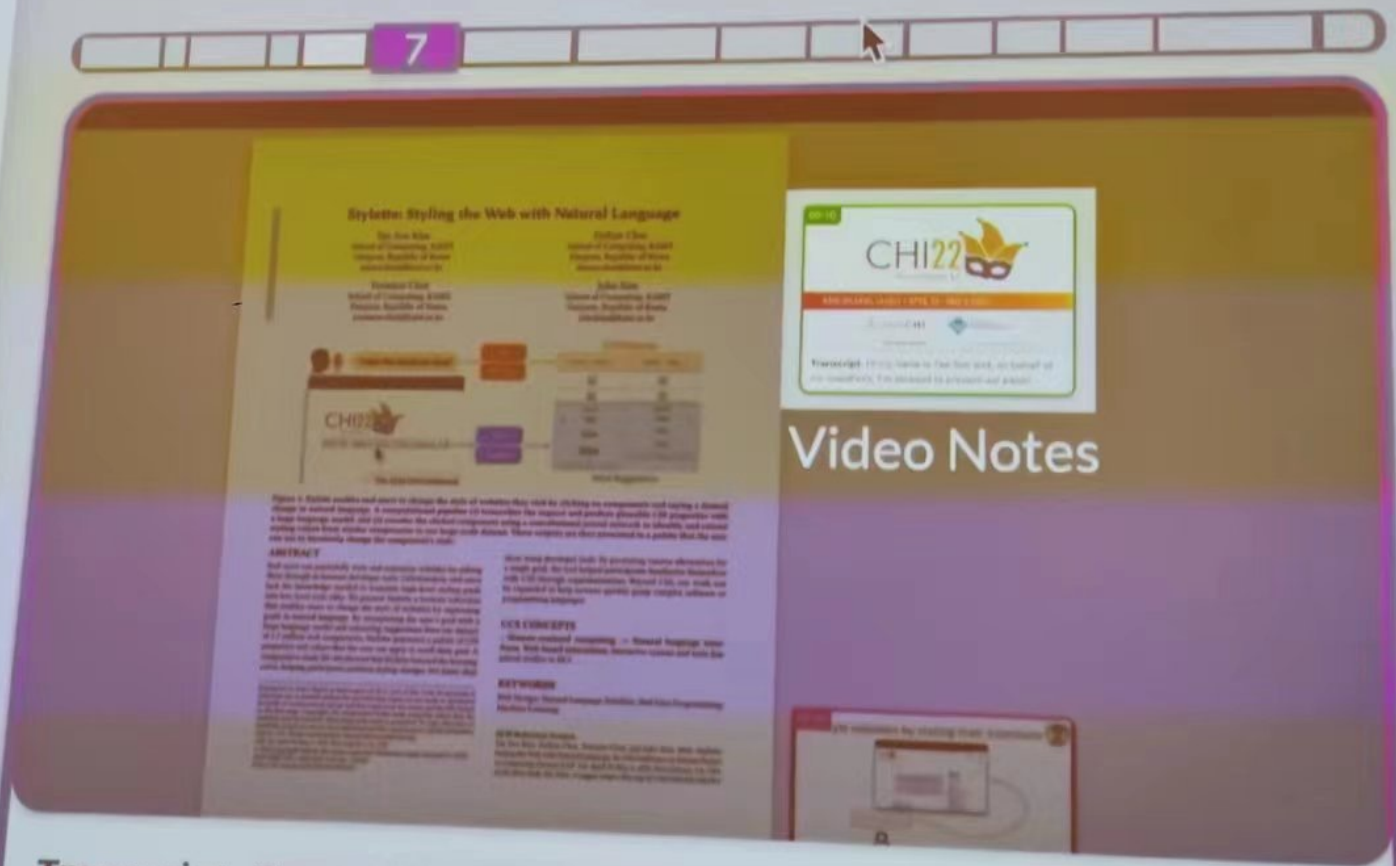
Figure 2: Technology probe used during the formative studies. On the left, a PDF reader for the paper where passages linked to



ber 01, 2023, San Francisco, CA, USA

respectively, visually represent importance. By skimming based reader could prioritize reading interface and a few important are included in the conference implementation details and remembering the thumbnails per, the user can also develop paper to help them return to If the thumbnail or transcript about the video segment, the the highlight bar to peek into (fig. 4a).

and Video. As the user is read- ruggle to understand certain particular sections. For exam- rn the implementation details already familiar with the back- ases, if a video note is linked, d/or summarized explanation highlight bar or video note it- ote "activates" the video note to a video player that starts



**Transcript** To combine papers and videos, we propose Papeos, a novel reading experience that augments papers with segments from talk videos. The Papeo read... segments as



## Stylette: Styling the Web with Natural Language

Tae Soo Kim  
School of Computing, KAIST  
Daejeon, Republic of Korea  
taeso.kim@kaist.ac.kr

Yoonsu Choi  
School of Computing, KAIST  
Daejeon, Republic of Korea  
yoonsu.choi@kaist.ac.kr

DaEun Choi  
School of Computing, KAIST  
Daejeon, Republic of Korea  
daeun.choi@kaist.ac.kr

Juho Kim  
School of Computing, KAIST  
Daejeon, Republic of Korea  
juhokim@kaist.ac.kr

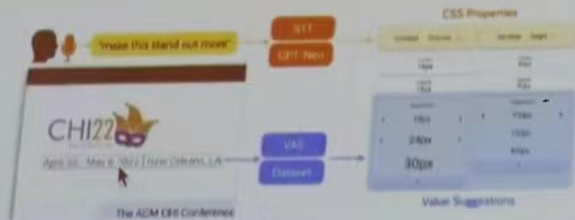


Figure 1: Stylette enables end users to change the style of websites they visit by clicking on components and saying a desired change in natural language. A computational pipeline (1) transcribes the request and predicts plausible CSS properties with a large language model, and (2) encodes the clicked component using a convolutional neural network to identify and extract styling values from similar components in our large-scale dataset. These outputs are then presented in a palette that the user can use to iteratively change the component's style.

### ABSTRACT

End users can potentially style and customize websites by adding them through in-browser developer tools. Unfortunately, end users lack the knowledge needed to translate high-level styling goals into low-level code edits. We present *Stylette*, a browser extension that enables users to change the style of websites by expressing goals in natural language. By interpreting the user's goal with a large language model and extracting suggestions from our dataset of 1.7 million web components, *Stylette* generates a palette of CSS properties and values that the user can apply to reach their goal. A comparative study (N=60) showed that *Stylette* lowered the learning curve, helping participants perform styling changes 35% faster than

using developer tools. By presenting various alternatives for a single goal, the tool helped participants familiarize themselves with CSS through experimentation. Beyond CSS, our work can be expanded to help novices quickly grasp complex software or programming languages.

### CCS CONCEPTS

Human-centered computing → Natural language interfaces; Web-based interaction: Interactive systems and tools; Empirical studies in HCI

### KEYWORDS

Web Design; Natural Language Interface; End-User Programming; Machine Learning

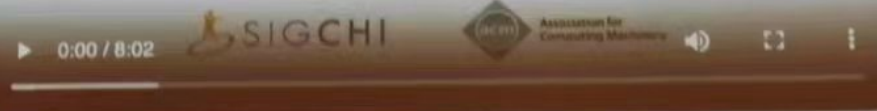
### ACM Reference Format

Tae Soo Kim, DaEun Choi, Yoonsu Choi, and Juho Kim. 2022. Stylette: Styling the Web with Natural Language. In *CHI Conference on Human Factors in Computing Systems* (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 17 pages. <https://doi.org/10.1145/3541161.3541171>

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 the author(s) must be acknowledged. Copying to reproduce copies is permitted. In case of republication, the prior consent of the publisher is required. For all other uses, permission should be sought from the publisher or the copyright owner. Copyright © 2022, ACM. <https://doi.org/10.1145/3541161.3541171>



NEW ORLEANS, LA (US) / APRIL 30 - MAY 5, 2022



- 100% +

Current Time: 00:00:00

Current Clip: 00:10:30 - 00:17:08

00:07:07 Stylette: styling the web with natural language.

00:10:30 Websites are inherently malleable as the documents used to render a website are accessible to

00:15:32 end-users through their browsers.

00:17:08 This means that users can edit these documents to in turn customize the website that they frequently browse.

00:23:59 However, editing these documents requires knowledge and expertise of the languages used

00:28:04 in these documents.

00:29:51 Knowledge and expertise that most end-users might not have.

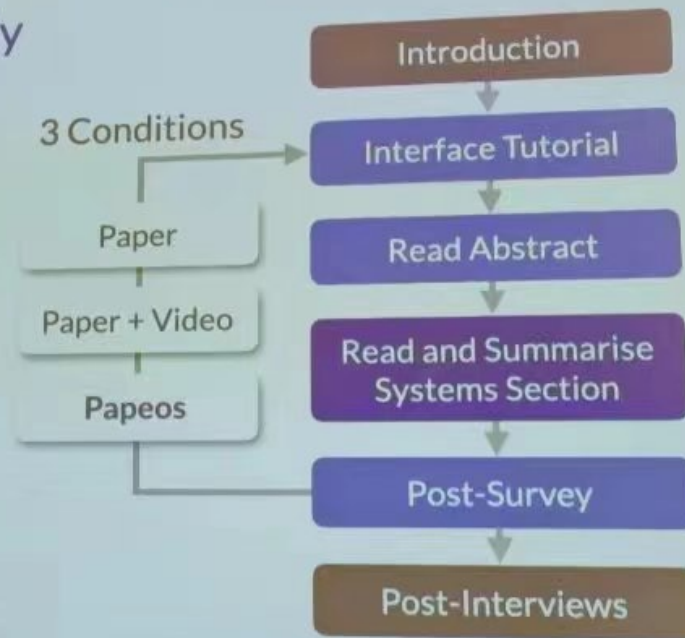
00:33:08 So, even if users have an intention to change a website, they cannot translate their

## Within-Subjects User Study

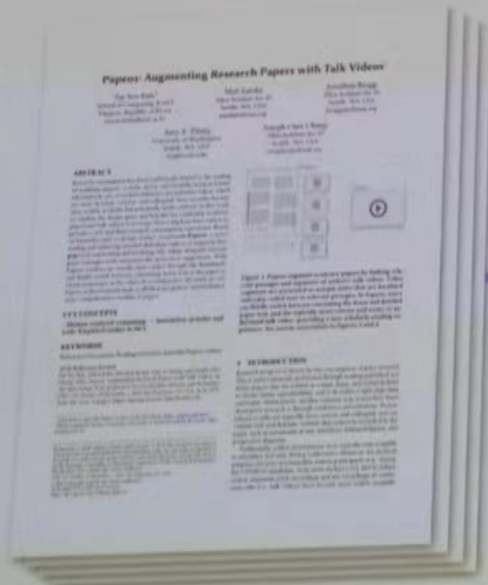
**Participants** Early-stage HCI researchers (N=16)

**Materials** 3 HCI systems papers

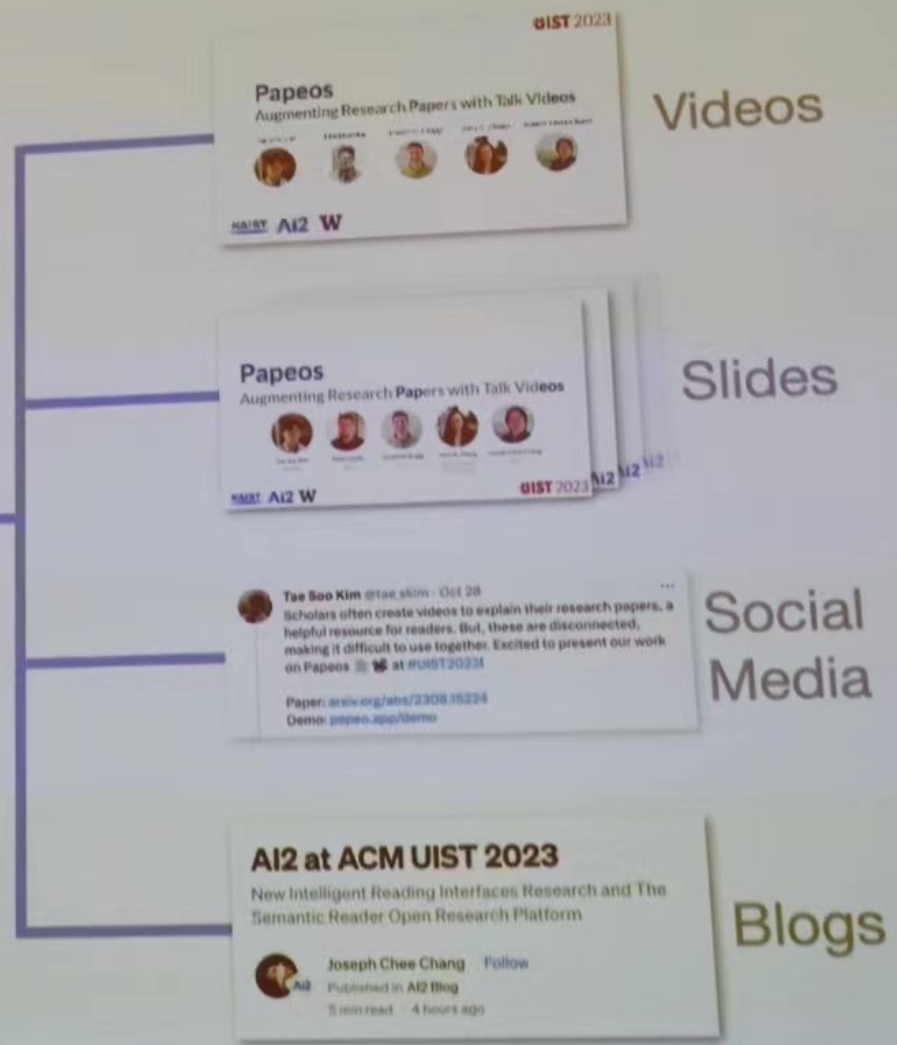
**Task** Summarize the systems section



# Augment Static Papers...



# ... with Existing Content



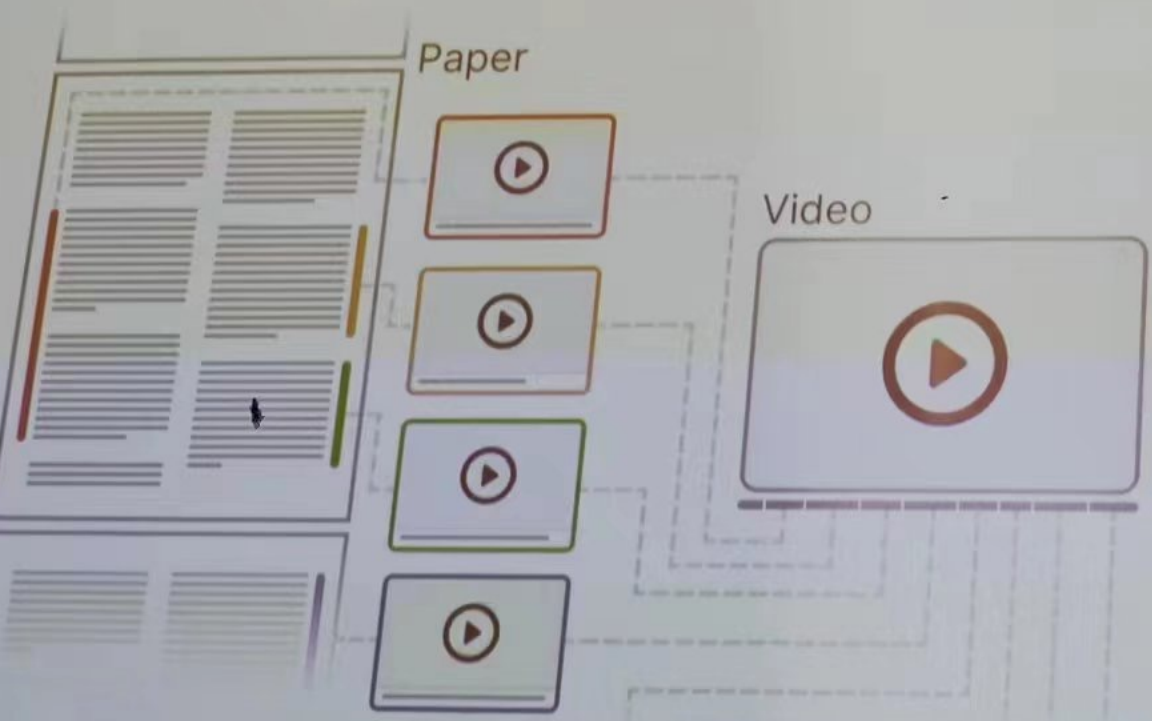




# Papeos

Augmenting Research  
Papers with Talk Videos

taesoo.kim@kaist.ac.kr  
@tae\_skim



Play with the Papeo of Papeo  
[papeo.app/demo](https://papeo.app/demo)

<https://papeo.app/demo>



MONASH  
INFORMATION  
TECHNOLOGY

## **Video2Action: Reducing Human Interactions in Action Annotation of App Tutorial Videos**

**Sidong Feng<sup>1</sup>, Chunyang Chen<sup>1</sup>, Zhenchang Xing<sup>2</sup>**

<sup>1</sup> Monash University

<sup>2</sup> CSIRO



## However

App Tutorial Videos may lack of **accessibility**.

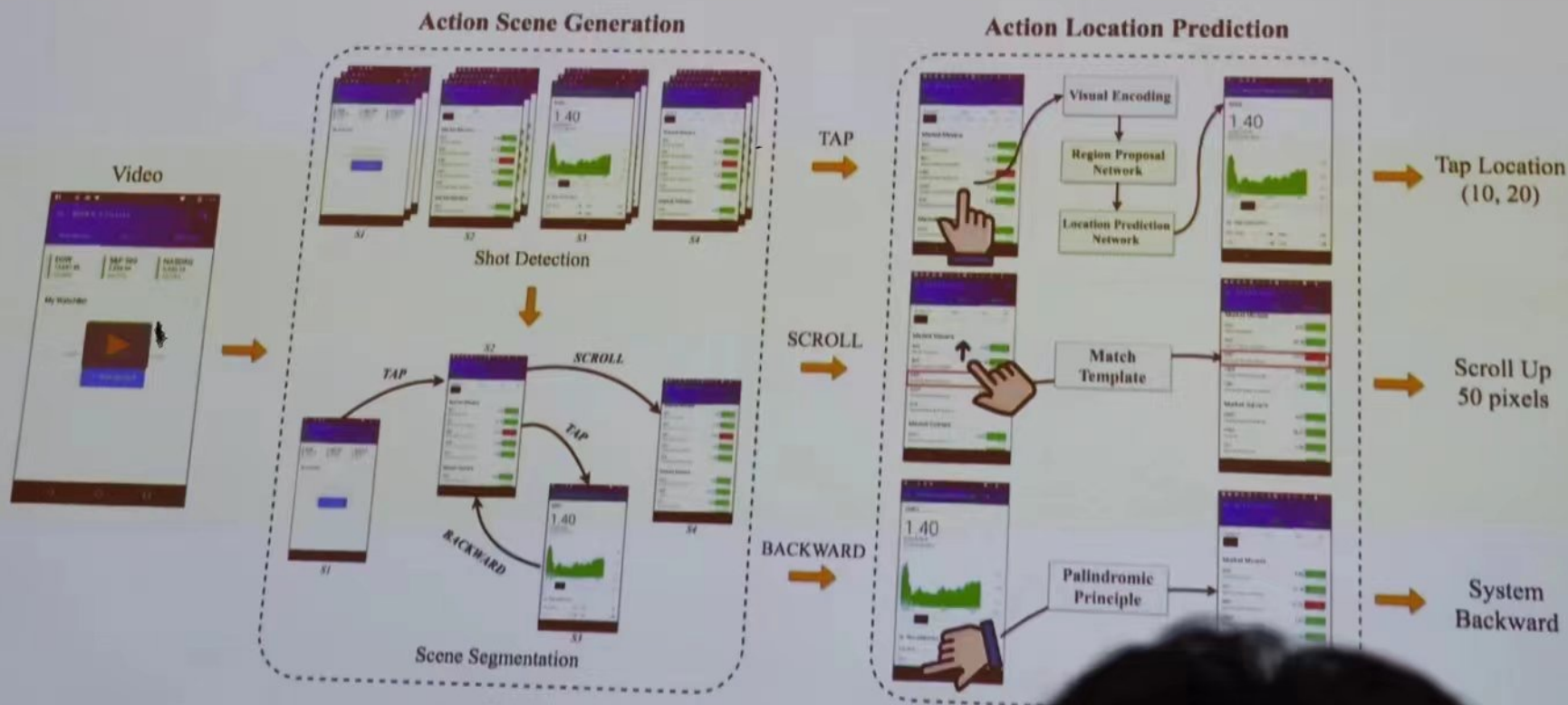


- Video may play too **fast** to watch, and the actions performed sometimes are **not obvious** to be realized
- Auditory information may not accessible to **non-native-language** users or **hearing-impaired** users



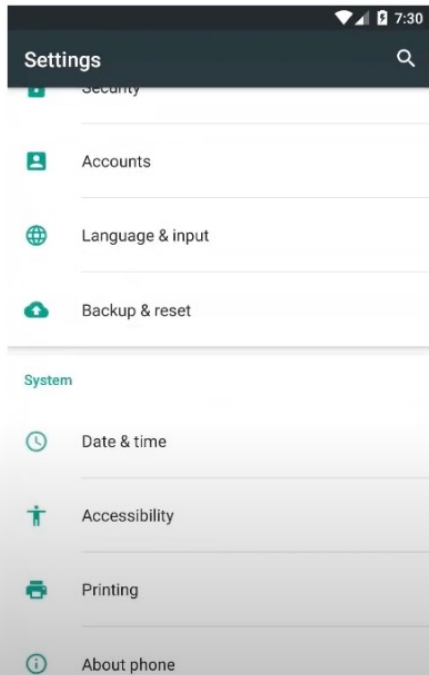


**Video2Action**, a lightweight non-intrusive approach to automatically acquire the actions from the video, reducing the burdens of video creators in action annotation.





## App tutorial video

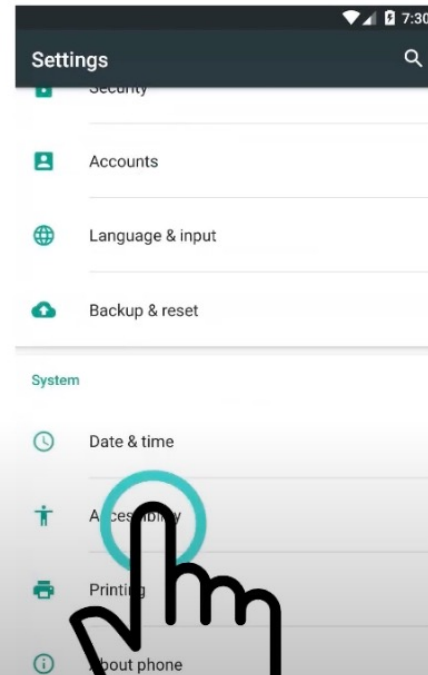


Video2Action



Tutorial Creators

## Annotated video



### Video2Action: Reducing Human Interactions in Action Annotation of App Tutorial Videos



ACM SIGCHI ✓  
2.03万位订阅者

订阅

👍 0



🔗 分享

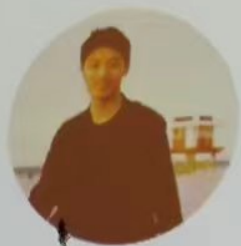
📄 下载





UIST 2023

# PEANUT: A Human-AI Collaborative Tool for Annotating Audio-visual Data



Zheng Zhang \*



Zheng Ning \*



Chenliang Xu



Yapeng Tian



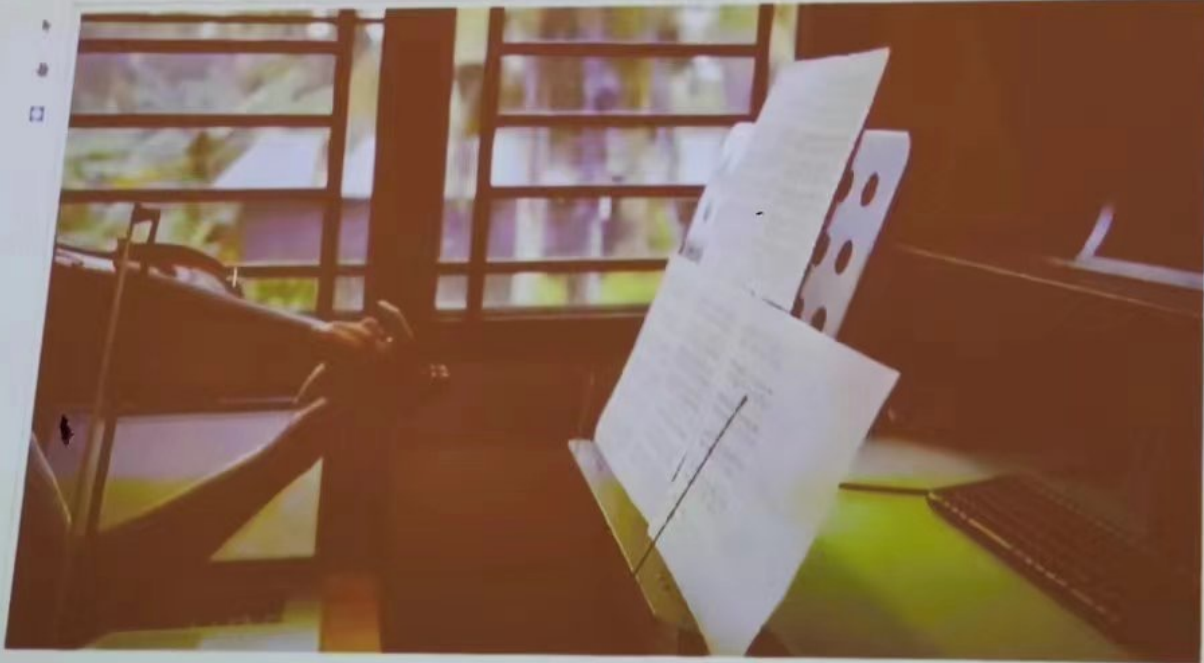
Toby Jia-Jun Li



# Baseline Annotation Method

02BAKBRW\_b4.mp4 (1/82) Time: 0:125s

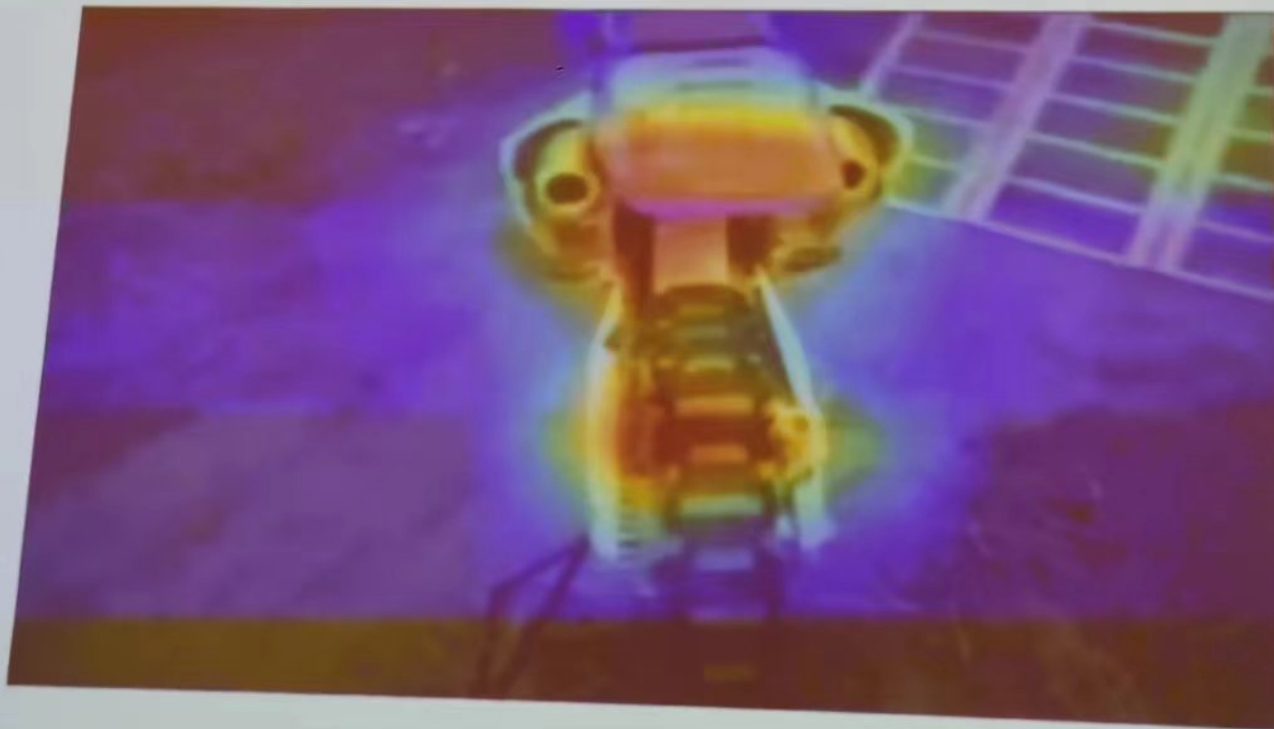
Video Audio Text Media Tools Annotations History



- Task Description
- Classifications
- Regions
- History

# Challenge of Fully Automation

- Existing multi-modal models are not robust for achieving accurate annotation
- Single-modality models are hard to automatically correspond annotations in visual and auditory modality









1fioenvrmUQ.mp4 (1/82) Time: 0.125s

NextLabel Video Audio Prev Next MoveTo Thumbnail Review Fullscreen Save

Task Description  
Classifications  
Regions  
History

100%

## PEANUT: A Human-AI Collaborative Tool for Annotating Audio-Visual Data



ACM SIGCHI ✓  
2.03万位订阅者

订阅

👍 0



🔗 分享

↓ 下载



<https://www.youtube.com/watch?v=PpPaxbxV5BI>

# Study Result

**Avg. Seconds of Completion/Frame**

**Number of Frames**

**Accuracy (IoU)**

Baseline  
(Human only)

7.73

149.45

0.72

PEANUT  
(Human+AI)

5.12

488.85

0.93

↓ 33.8%

↑ 227%

↑ 29.2%



# Soundify: Matching Sound Effects to Video

David Chuan-En Lin<sup>1</sup>

Anastasis Germanidis<sup>2</sup>

Cristóbal Valenzuela<sup>2</sup>

Yining Shi<sup>2</sup>

Nikolas Martelaro<sup>1</sup>

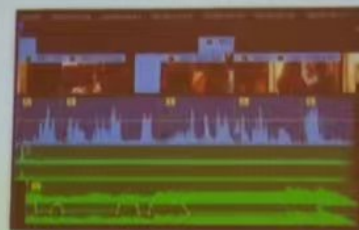
<sup>1</sup>Carnegie Mellon University, <sup>2</sup>Runway

# Design Principles

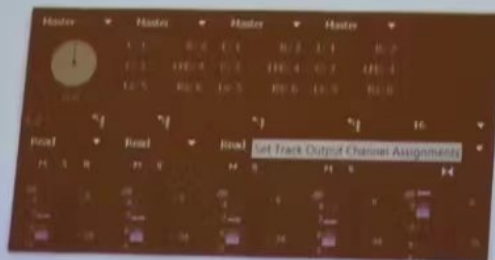
1. Surface suitable sounds



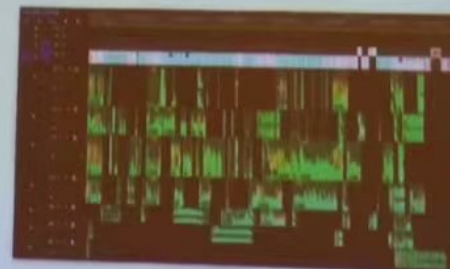
2. Synchronize sounds to video



3. Tune spatial parameters (pan and volume)



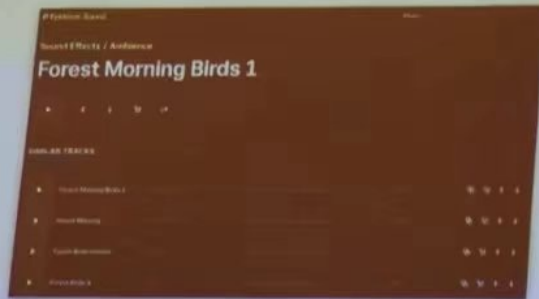
4. Stack multiple soundtracks





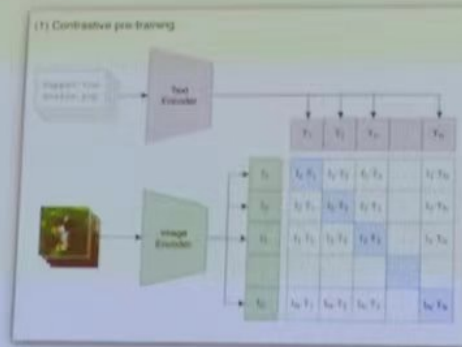
# This Work

Studio quality  
sound effects libraries



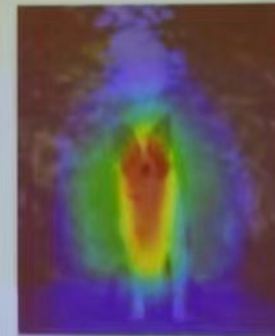
+

CLIP (open vocab  
classification)



+

Activation map



1. 2022. Epidemic Sound. Retrieved August 15, 2022 from <https://www.epidemicsound.com/>
2. Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. 2021. Learning transferable visual models from natural language supervision. In International Conference on Machine Learning. PMLR, 8748–8763.
3. Ramprasaath R Selvaraju, Abhishek Das, Ramakrishna Vedantam, Michael Cogswell, Devi Parikh, and Dhruv Batra. 2016. Grad-CAM: Why did you say that? arXiv preprint arXiv:1611.07450 (2016).



Motivation

System Overview

Evaluation

Future Work



## Soundify Demo



David Chuan-En Lin's Research  
53位订阅者

订阅



9



分享



下载



保存



<https://www.youtube.com/watch?v=VhsD699osIM>

# User Study

## Task

Create a layered sound environment for a video

## Baseline

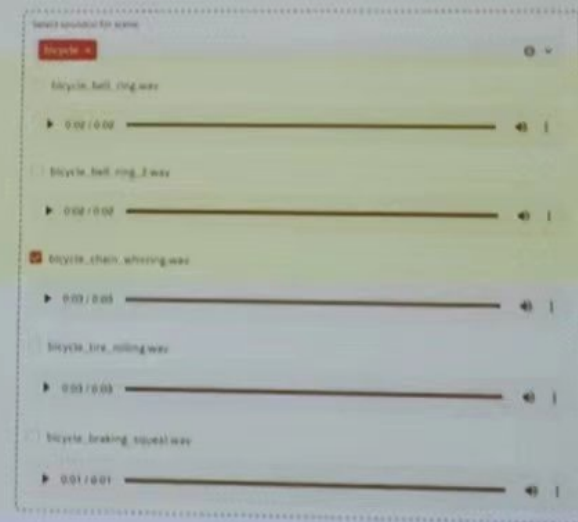
Manual editing with Adobe Premiere Pro

## Participants

12 professional video editors

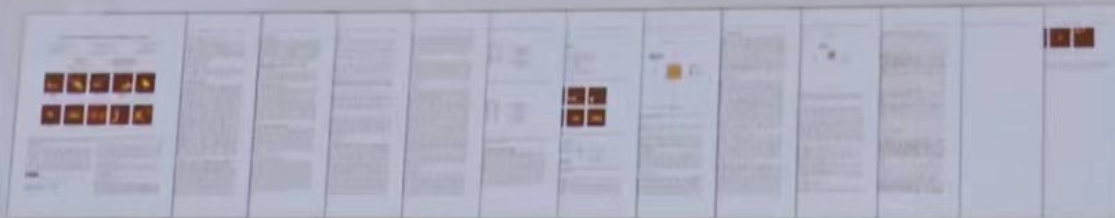
## Measures

- Workload (NASA Task Load Index, 5-point Likert)
- Task Completion Time (seconds)
- Usability (System Usability Scale, 5-point Likert)





## Paper



Lin, D., Germanidis, A., Valenzuela, C., Shi, Y., Martelaro, N. (2023). Soundify: Matching Sound Effects to Video. In Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology

Project Page <https://chuanenlin.com/soundify>



Live Demo  
(lightweight version  
using image input) <https://soundify.cc>



<https://soundify.cc/>



Ucla

vibint



UNIVERSITY  
of VIRGINIA



# Robust Finger Interactions with COTS Smartwatches via Unsupervised Siamese Adaptation

Wenqiang Chen, Ziqi Wang, Pengrui Quan, Zhencan Peng, Shupeil Lin,  
Mani Srivastava, Wojciech Matusik, John Stankovic.

Massachusetts Institute of Technology  
University of California, Los Angeles  
VibInt AI  
University of Virginia

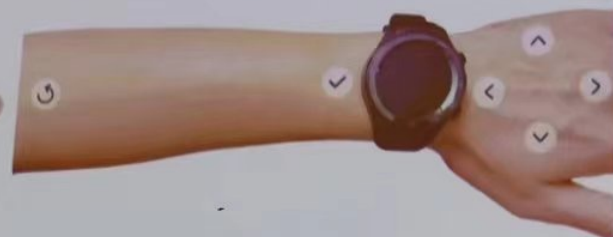


Play games

## Finger/hand gesture recognition using IMU



(a)



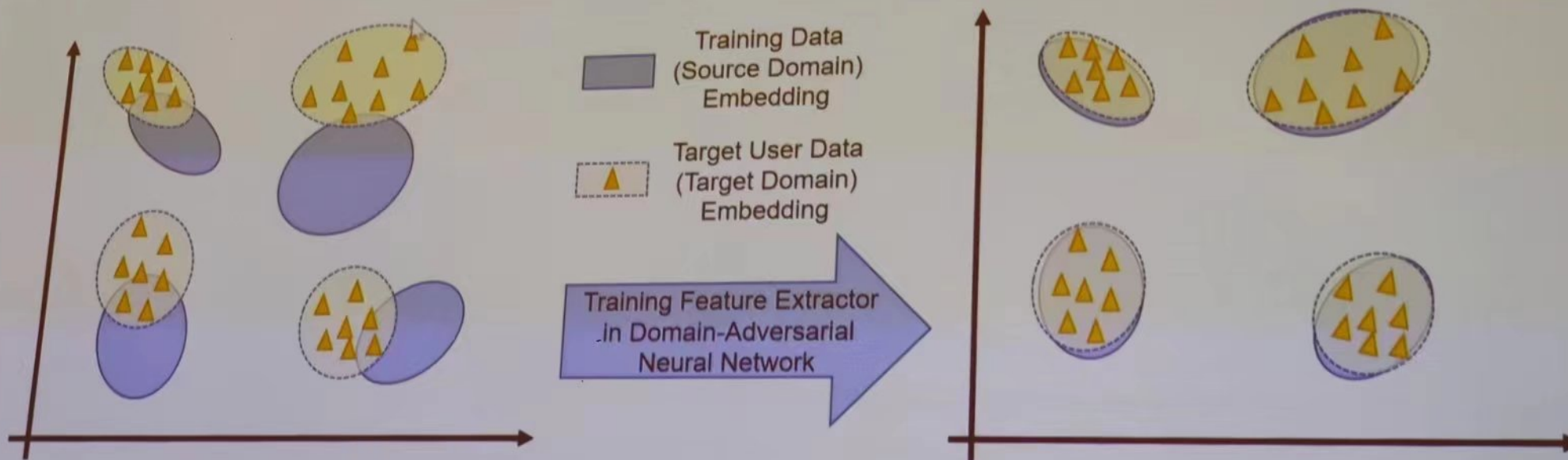
(b)



(c)

1. Wenqiang Chen, Lin Chen, Yandao Huang, Xinyu Zhang, Lu Wang, Rukhsana Ruby, Kaishun Wu.  
Taprint: Secure Text Input for Commodity Smart Wearables. ACM Mobicom 2019
2. Wenqiang Chen, Maoning Guan, Yandao Huang, Lu Wang, Rukhsana Ruby, Wen Hu, Kaishun Wu.  
ViType: A Cost Efficient On-body Typing System through Vibration. IEEE SECON 2018. **(Best Paper Award)**
3. Wenqiang Chen, Lin Chen, Kenneth Wan, and John Stankovic.  
Demo Abstract: A Smartwatch Product Provides On-body Tapping Gestures Recognition. **(Best Demo Award)**

# Domain Adaptation



- ✓ Target new users produce tapping samples in daily usage.
- ✗ There is no label in the target new users' data.



# Different wearing positions



## Robust Finger Interactions with COTS Smartwatches via Unsupervised Siamese Adaptation



ACM SIGCHI ✓  
2.03万位订阅者

订阅

👍 1



🔗 分享

📄 下载



# Color Field



Matthew  
Beaudouin-Lafon  
mattbl@ucsd.edu



Jane L. E



Haijun Xia

- Color Field is a novel visualization of color filters
- Professional Vision as a framework for expertise
- Design CSTs to help users understand their domain

Prototype:  
<https://creativity.ucsd.edu/colorfield>

<https://creativity.ucsd.edu/colorfield>

Raise your hand if you have used ...  
Color Filters on your phone?



Apple Clips.





The screenshot shows a software interface for 'Color Field'. At the top, there is an 'Intensity' slider. Below it are two image panels: 'Sample Color' and 'Result'. The 'Sample Color' panel shows a woman in a white shirt against a colorful, abstract background. The 'Result' panel shows the same woman, but the background colors are shifted towards pink and purple, and the overall saturation is lower. Below the images are two color field visualizations. The left one is labeled 'Sample Color' and shows a color field with a gradient from blue to purple. The right one is labeled 'Result' and shows a color field with a gradient from blue to pink. A red line is drawn across the bottom of the color field visualizations. The video player interface includes a play button, a progress bar showing 0:21 / 0:30, and various control icons like volume, settings, and full screen.

## Color Field: Developing Professional Vision by Visualizing the Effects of Color Filters



ACM SIGCHI ✓  
2.03万位订阅者

订阅

👍 2



🔗 分享

↓ 下载



<https://www.youtube.com/watch?v=p08FIV1utFQ>



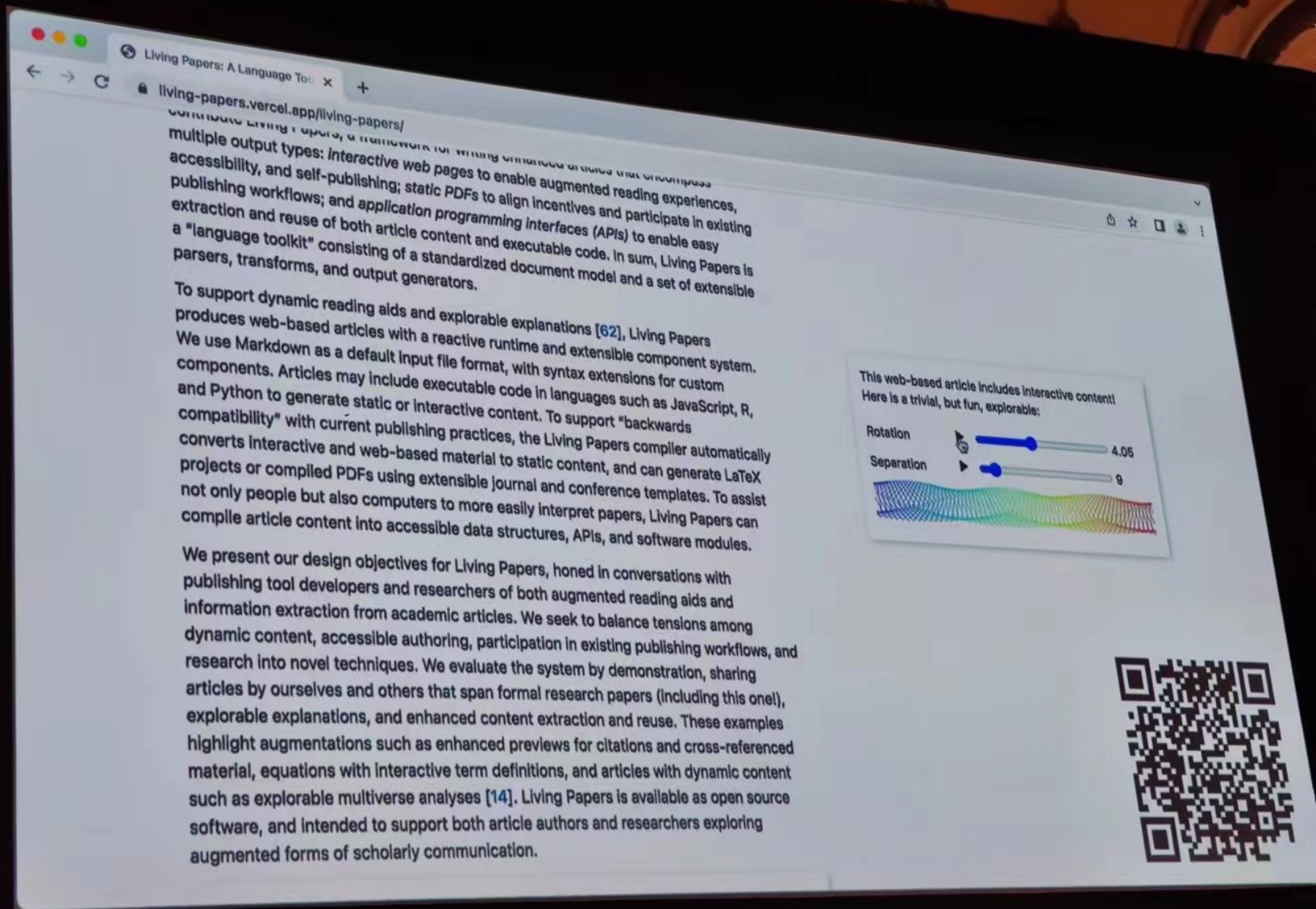
# Living Papers

A Language Toolkit for Augmented  
Scholarly Communication

Jeffrey Heer, Matthew Conlen, Vishal Devireddy,  
Tu Nguyen, and Joshua Horowitz

University of Washington







# R Knitr Integration

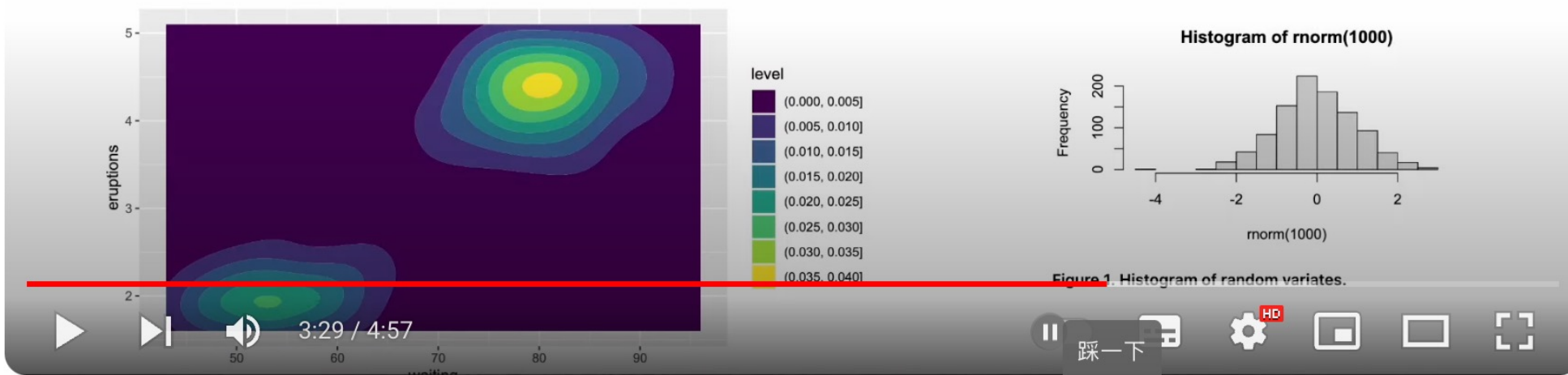
Living Papers Team

This article demonstrates compile-time evaluation of R code, with results stitched back into the Living Papers article. JSON-serialized values can also be bound to variables in the reactive runtime.

R version: **R version 4.1.3 (2022-03-10)**

## 1 Plots

Let's make some plots.



## Living Papers



Jeffrey Heer

7位订阅者



12 | 分享 | 下载 | ...

## The Future of Scholarly Communication

### A "Standard Toolkit" of Reading Augmentations

Definitions, equations, figures, and more...

Helpful to readers, easy for authors to apply

Further design space exploration & evaluation

### Collaborative Editing & Reviewing

Novel reviewing tools, pre- and post-publication

### Improved Conversion & Accessibility

How might we effectively convert rich, dynamic content to similarly effective static content?



Course Readings for The Future of Scholarly Communication, Fall '22



# Living Papers

A Language Toolkit for Augmented  
Scholarly Communication

[github.com/uwdata/living-papers](https://github.com/uwdata/living-papers)

<https://uwdata.github.io/living-papers-template/>



# Artificial General Intelligence

## AGI is Coming... Is HCI Ready?

Meredith Ringel Morris  
Google DeepMind

## AGI = Capabilities, not Processes

- ✓ Breadth & Depth of Non-Physical Capabilities
- ✗ Learning, Thinking, or Understanding in a Human-Like Way
- ✗ Consciousness, Sentience, etc.



Typically, HCI and other sociotechnical domains are not considered as core to the ML research community as areas like model building. **However, I argue that research on Human-AI Interaction and the societal impacts of AI is vital and central to this moment in computing history. HCI must not become a “second class citizen” to AI, but rather be recognized as fundamental to ensuring the path to AGI and beyond is a beneficial one.**





## Performance x Generality: AGI Level Estimates

	<b>Narrow</b>	<b>General</b>
<b>Emerging</b> <i>≅ unskilled person</i>	<b>"Narrow Emerging AI"</b> GOF AI (e.g., SHRDLU)	<b>"Emerging AGI"</b> SOTA LLMs (ChatGPT, Bard)
<b>Competent</b> <i>≥50th % skilled</i>	<b>"Narrow Competent AI"</b> Siri, Watson, PaLI, Jigsaw	<b>"Competent AGI"</b> <nothing yet>
<b>Expert</b> <i>≥90th % skilled</i>	<b>"Narrow Expert AI"</b> Grammarly, Dall-E 2	<b>"Expert AGI"</b> <nothing yet>
<b>Virtuoso</b> <i>≥99th % skilled</i>	<b>"Narrow Virtuoso AI"</b> AlphaGo, Deep Blue	<b>"Virtuoso AGI"</b> <nothing yet>
<b>Superhuman</b> <i>additional skills</i>	<b>"Narrow Superhuman AI"</b> AlphaFold	<b>"ASI"</b> <nothing yet>



## Model Capabilities Unlock Interaction Paradigms

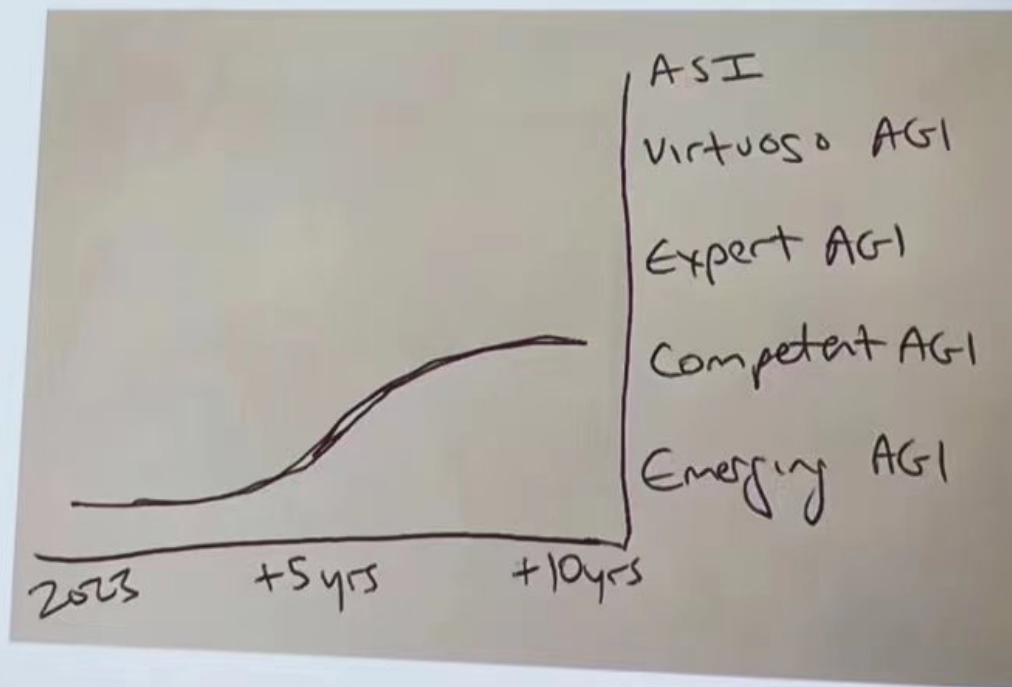
Autonomy Level	Unlocked At...	Examples
AI as a Tool	Emerging Narrow AI	Running a grammar-checker over text
AI as a Consultant	Competent Narrow AI; Emerging AGI	Using an LLM to summarize a set of documents Programming with a code-generating model
AI as a Collaborator	Expert Narrow AI; Emerging or Competent AGI	Training as a chess player via a chess AI Socializing with AI bots
AI as an Expert	Narrow AI Virtuoso; Expert AGI	Using an AI system to advance scientific discovery (e.g., protein-folding)
AI as an Agent	Virtuoso AGI; ASI	Autonomous AI-powered personal assistants

# My AGI Timelines

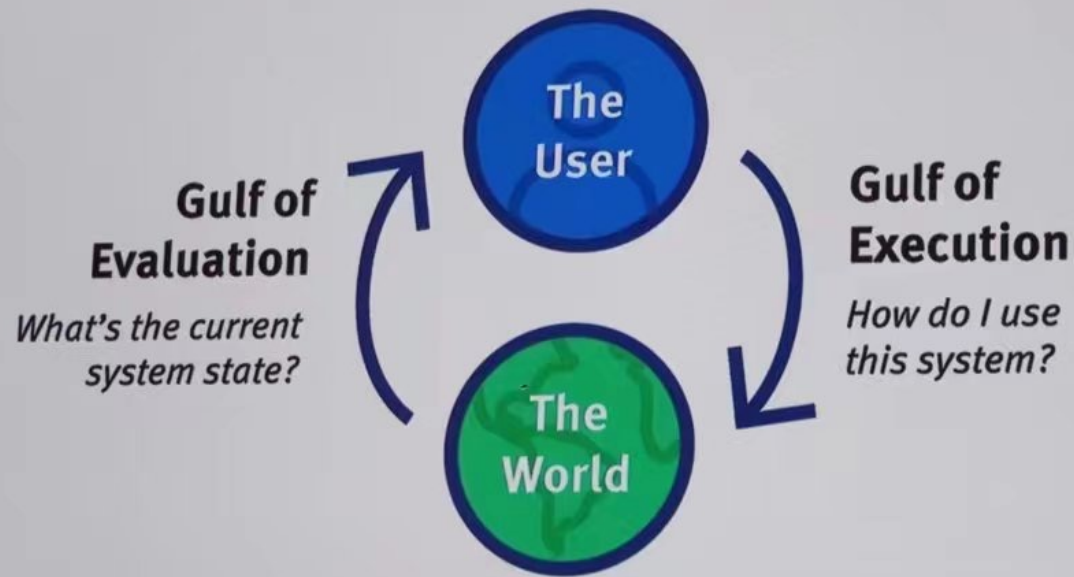
Timeline 1:  $\geq$  Expert AGI in 5-10 yrs



Timeline 2: Competent AGI Plateau



# Back to Basics: Norman's Gulfs 鸿沟



nngroup.com **NN/g**

Image credit: Nielsen Norman Group



A PERSON MIGHT  
THINK THAT  
BECAUSE

HE HAS PRESSED A  
BUTTON ON AS  
REMOTE CONTROL  
THAT THE TV WILL  
TURN OFF

IF THE REMOTE  
CONTROL DOES NOT  
WORK BECAUSE OF  
TECHNICAL  
PROBLEMS

THERE IS STILL A

**GULF OF EXECUTION**



Gulf of execution VS Gulf of evaluation | Project Management Professional | PMP | CT Academy



CT Academy

409位订阅者

订阅

14



分享

下载



<https://www.youtube.com/watch?v=isAAu8s3OtM>

## Updating Norman's Gulfs for the AGI Era

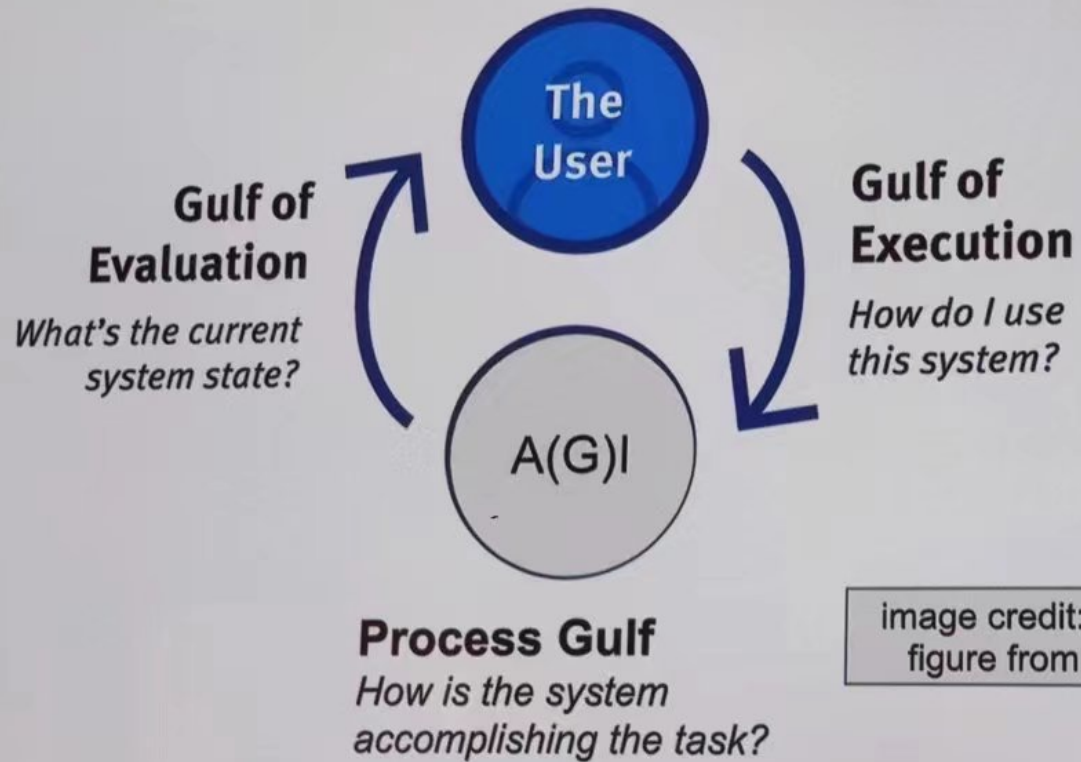
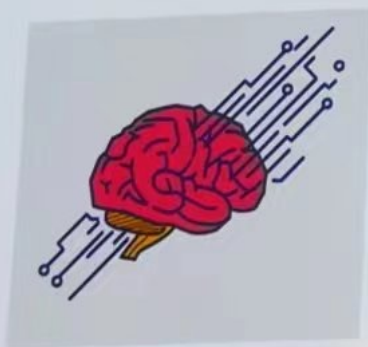


image credit: modification of figure from nngroup.com

## Human-AGI Interaction: Novel Interaction Techniques

BCI\* for “ESP” Interactions?



\*Tang, et al. “[Semantic reconstruction of continuous language from non-invasive brain recordings](#)” *Nature Neuroscience*, 2023

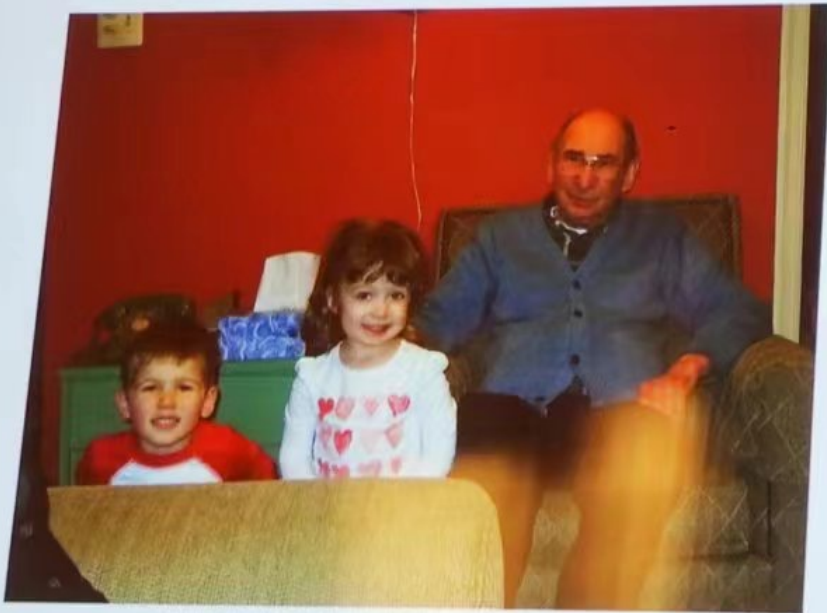
Cross-Species Interactions?



What are Merrie’s chickens saying?

## AGI Design Challenge: Anticipating Societal Impacts

Merrie's kids interact with her grandfather?



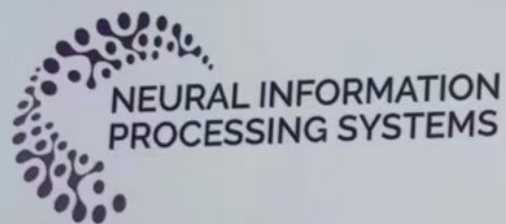
Hire Merrie's generative ghost to edit your UIST 2100 paper?



\*joint "wild idea" w/ Jed Brubaker from CU Boulder... possible CHI 2024 Workshop 👍

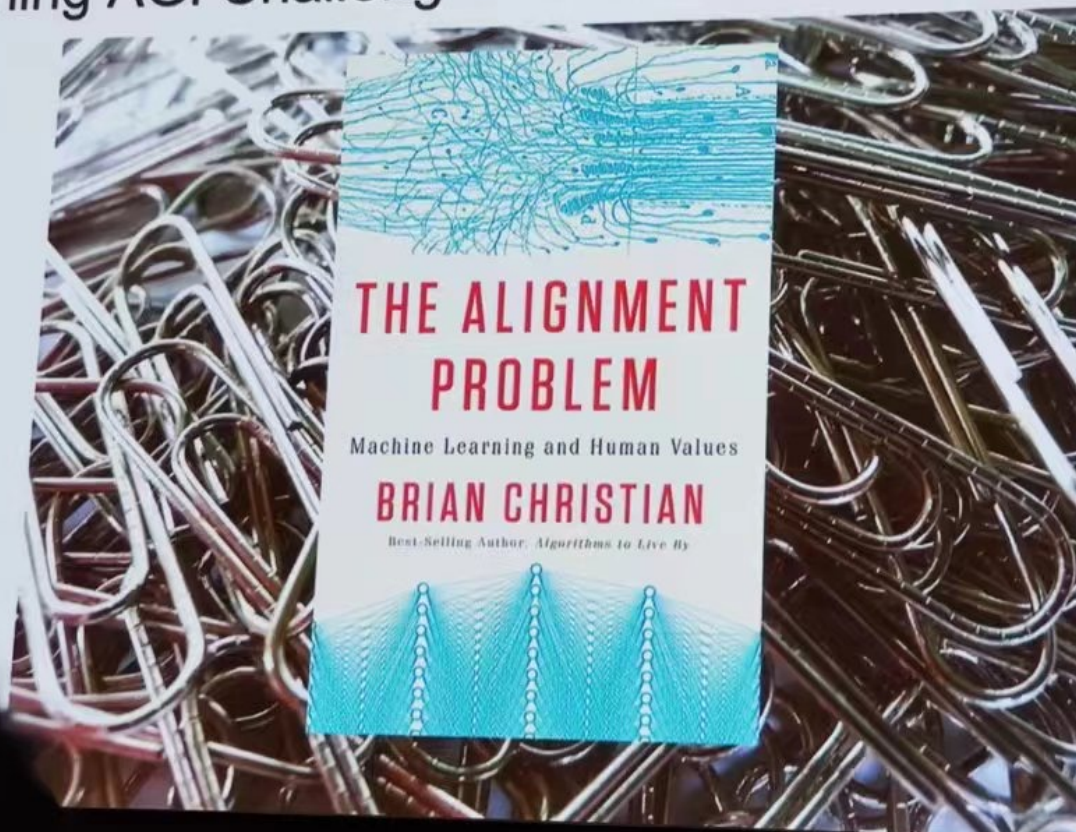


$A(G)I = ML + HCI$



**GPT3 → GPT3.5 也是HCI的一个大事件：1) 它有一个简单好用的interface；2) 它的训练改进过程是采用了human的feedback去加强的。**

## Reframing AGI Challenges as HCI Challenges

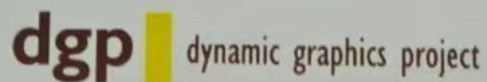


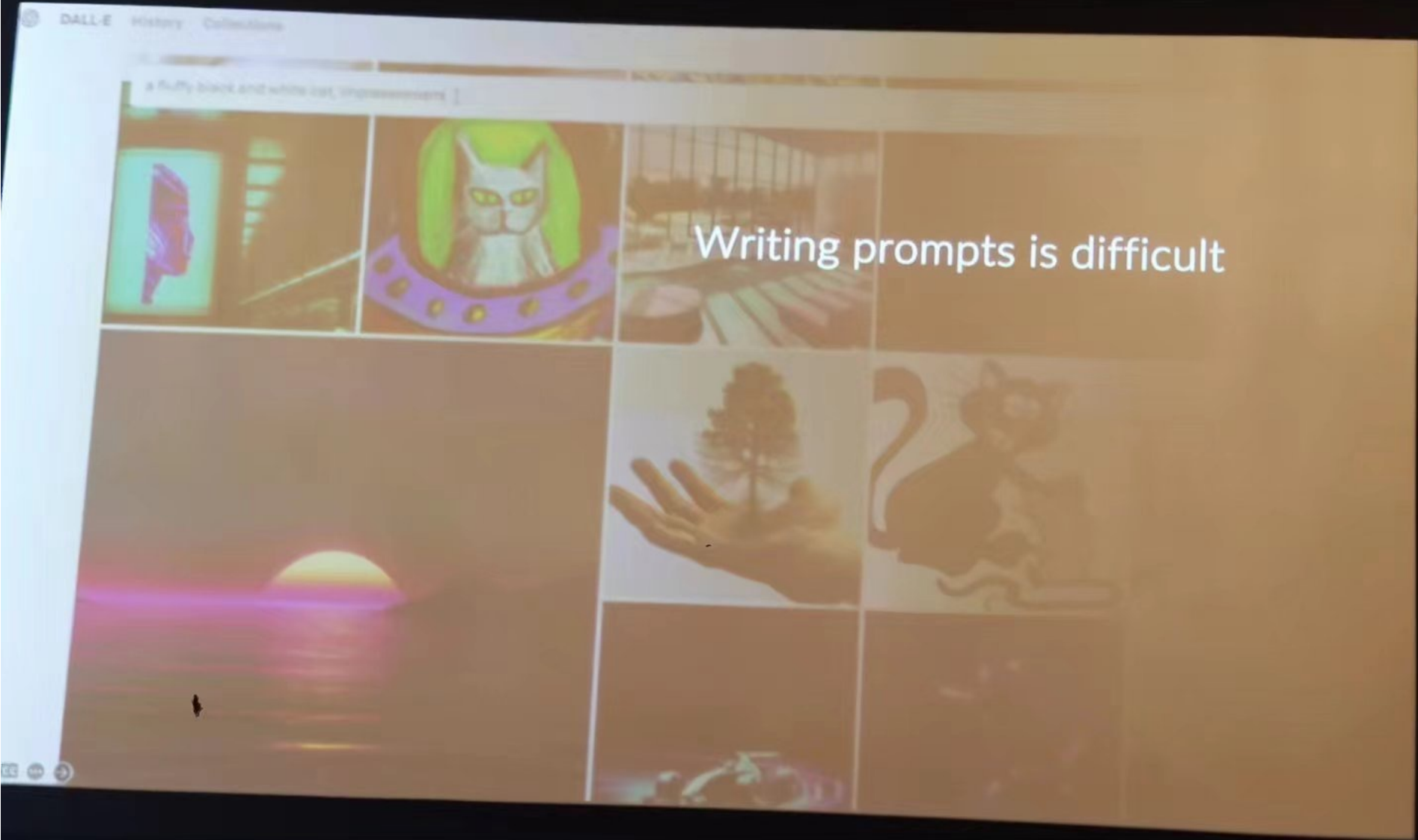


# PROMPTIFY

Text-to-Image Generation through Interactive Prompt  
Exploration with Large Language Models

Stephen Brade\*, Bryan Wang\*, Máuricio Sousa\*, Sageev Oore†, Tovi Grossman\*  
University of Toronto\*, Dalhousie University†





# Promptify

## Stable Diffusion Controls

### Prompt Writing

An ancient castle perched atop a mountain peak in the breathtaking Dolomites, impressionism style, with soft and dreamy brushstrokes, capturing the beautiful landscape and the vast open sky above, by John Singer Sargent, oil painting on canvas

Enter negative prompt

frame

### Prompt Extension

Steer the subject matter

Keep similar subject matter but make the setting Italy

Steer the style

impressionism

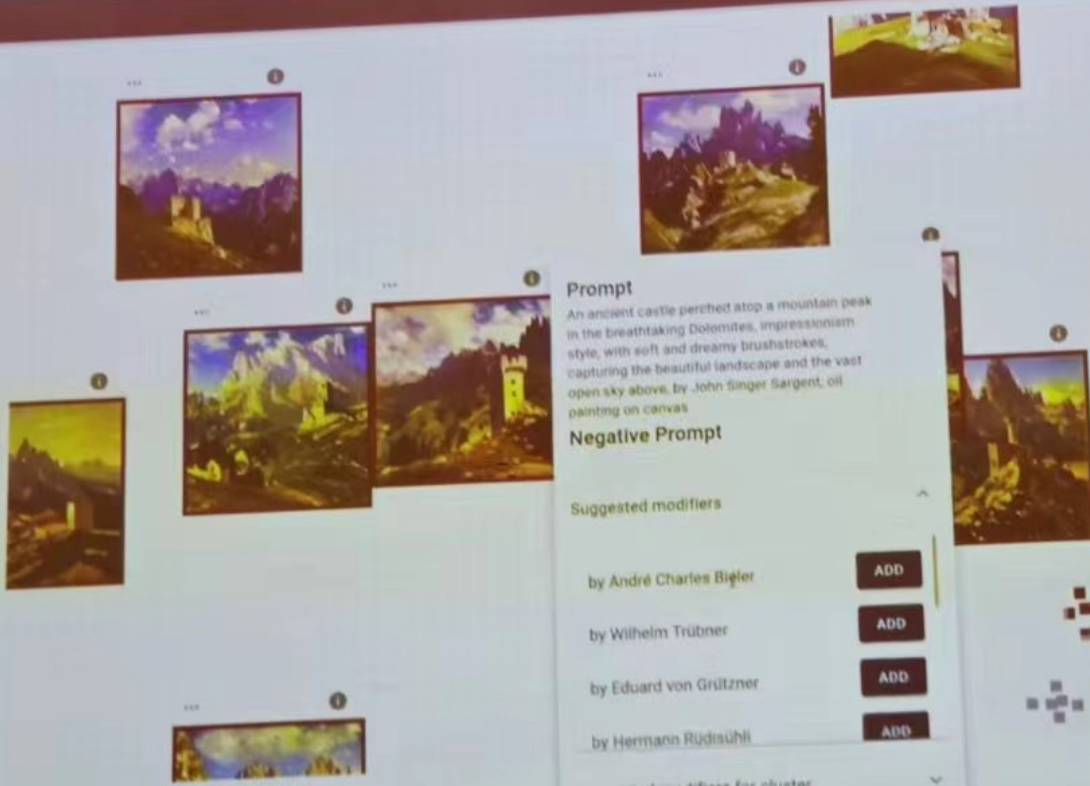
EXTEND STYLE

CLEAR SUBJECT

CLEAR STYLES

Batch Size

GENERATE



### Prompt

An ancient castle perched atop a mountain peak in the breathtaking Dolomites, impressionism style, with soft and dreamy brushstrokes, capturing the beautiful landscape and the vast open sky above, by John Singer Sargent, oil painting on canvas

### Negative Prompt

### Suggested modifiers

by André Charles Bièler

ADD

by Wilhelm Trübner

ADD

by Eduard von Grützner

ADD

by Hermann Rüdigsühli

ADD

Suggested modifiers for cluster

Suggested modifiers unique to cluster

## Clustering Controls

CLUSTER

POSITION BY SIMILARITY

Scale

Select prompts to show

**Prompt:** An ancient castle perched atop a mountain peak in the breathtaking Dolomites, impressionism style, with soft and dreamy brushstrokes, capturing the beautiful landscape and the vast open sky above, by John Singer Sargent, oil painting on canvas

SAVE CURRENT STATE

LOAD STATE



# Style Extension





**Promptify**

**Stable Diffusion Controls**

Prompt Writing  
Enter prompt  
a mountain range

Enter negative prompt

Prompt Extension  
**EXTEND SUBJECT**

Steer the subject matter  
**STEER SUBJECT**

Steer the style  
**EXTEND STYLE**

Batch Size  
**GENERATE**

**Clustering Controls**

**CLUSTER**

**POSITION BY SIMILARITY**

Scale

**SAVE CURRENT STATE**

**LOAD STATE**

0:57 / 3:22

HD

React Flow

## UIST 2023 Promptify: Text-to-Image Generation through Interactive Prompt Exploration with LLMs



**Bryan Wang**  
147位订阅者



# Comparative Evaluation

## 14 – Participant Study

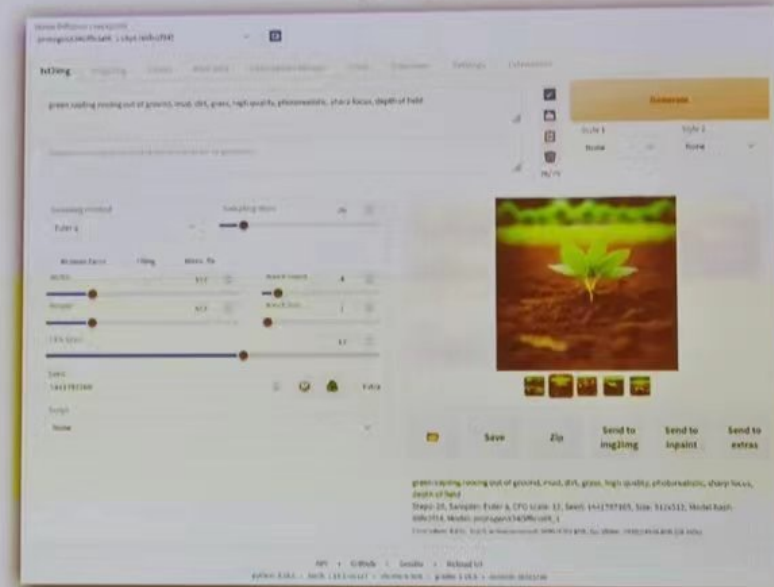
Within subjects comparison with Automatic1111

Solicited Subjective Ratings and NASA-TLX

Concluded with free usage observation + interview

**Next step: 对话式生成图片， or  
direct manipulation**

## Automatic1111





## PoseVEC: Authoring Adaptive Pose-aware Effects using Visual Programming and Demonstrations



well done!



Yongqi Zhang  
Cuong Nguyen  
Rubaiat Habib Kazi  
Lap-Fai Yu

George Mason University  
Adobe Research  
Adobe Research  
George Mason University

# Pose-aware Visual Effects



Entertainment ©JP Pirie

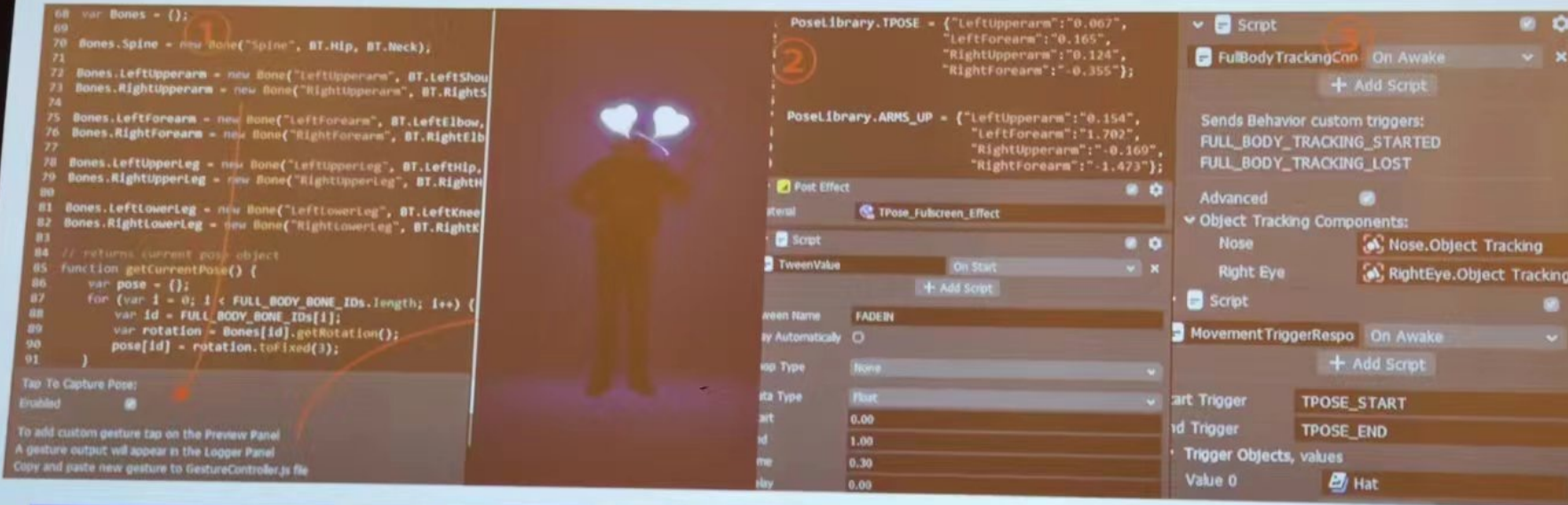


At-home workout ©Jeremy Ethier



Education ©Grant VanderHayden

# Programming-based approach is expressive but too technical



The screenshot displays a Unity development environment with three main panels:

- Code Editor (Left):** Shows JavaScript code for bone management and pose retrieval.
 

```

68 var Bones = {};
69
70 Bones.Spine = new Bone("Spine", BT.Hip, BT.Neck);
71
72 Bones.LeftUpperarm = new Bone("LeftUpperarm", BT.LeftShoulder, BT.RightShoulder);
73 Bones.RightUpperarm = new Bone("RightUpperarm", BT.RightShoulder, BT.LeftShoulder);
74
75 Bones.LeftForearm = new Bone("LeftForearm", BT.LeftElbow, BT.LeftWrist);
76 Bones.RightForearm = new Bone("RightForearm", BT.RightElbow, BT.RightWrist);
77
78 Bones.LeftUpperleg = new Bone("LeftUpperleg", BT.LeftHip, BT.LeftKnee);
79 Bones.RightUpperleg = new Bone("RightUpperleg", BT.RightHip, BT.RightKnee);
80
81 Bones.Leftlowerleg = new Bone("Leftlowerleg", BT.LeftKnee, BT.LeftAnkle);
82 Bones.Rightlowerleg = new Bone("Rightlowerleg", BT.RightKnee, BT.RightAnkle);
83
84 // returns current pose object
85 function getCurrentPose() {
86   var pose = {};
87   for (var i = 0; i < FULL_BODY_BONE_IDS.length; i++) {
88     var id = FULL_BODY_BONE_IDS[i];
89     var rotation = Bones[id].getRotation();
90     pose[id] = rotation.toFixed(3);
91   }
92 }

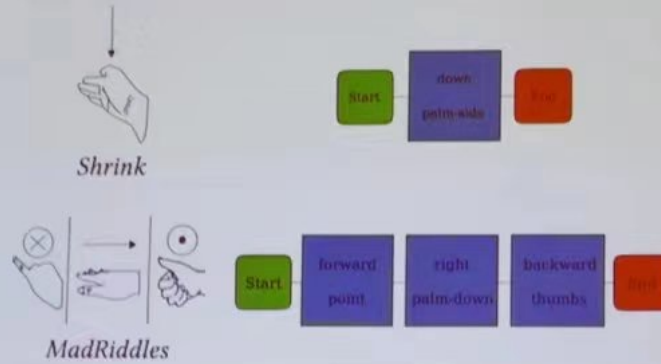
```
- 3D Viewport (Center):** Shows a stylized 3D character model with glowing eyes and a purple aura.
- Hierarchy and Inspector (Right):** Shows a 'Script' component for 'FullBodyTrackingConn' with various settings:
  - Object Tracking Components:** Includes 'Nose.Object Tracking' and 'RightEye.Object Tracking'.
  - MovementTriggerRespo:** Set to 'On Awake'.
  - Start Trigger:** 'TPOSE\_START'.
  - End Trigger:** 'TPOSE\_END'.
  - Trigger Objects, values:** Value 0 is set to 'Hat'.

Authoring Workflow	Example	Cons
Keyframe-based approach	Adobe Premiere, After Effects	Output effect isn't reusable
Programming-based approach	Lens studio, Spark AR	Require programming and configuration

# Previous research on Programming by Demonstration



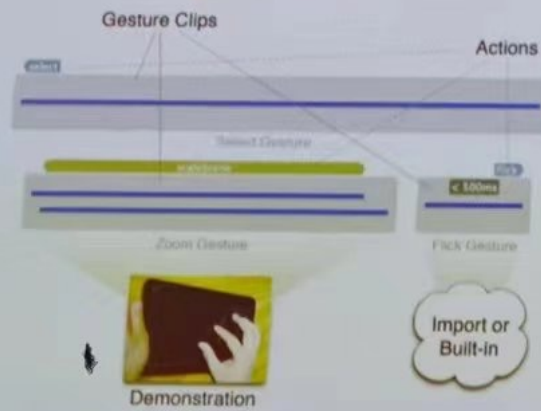
GesturAR



Gesture Knitter

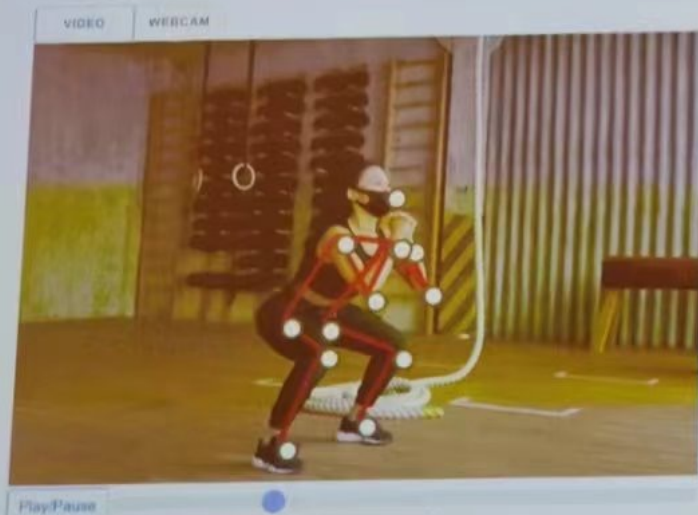


Rapido



GestureStudio

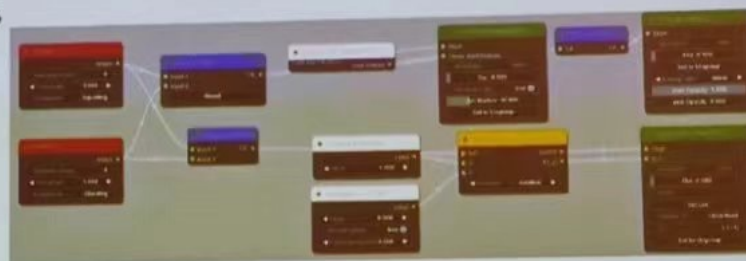
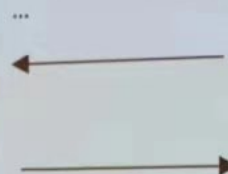
# PoseVEC(Pose-aware Visual Effect Creator)



Video Canvas

Direct manipulation on video canvas to streamline authoring process.

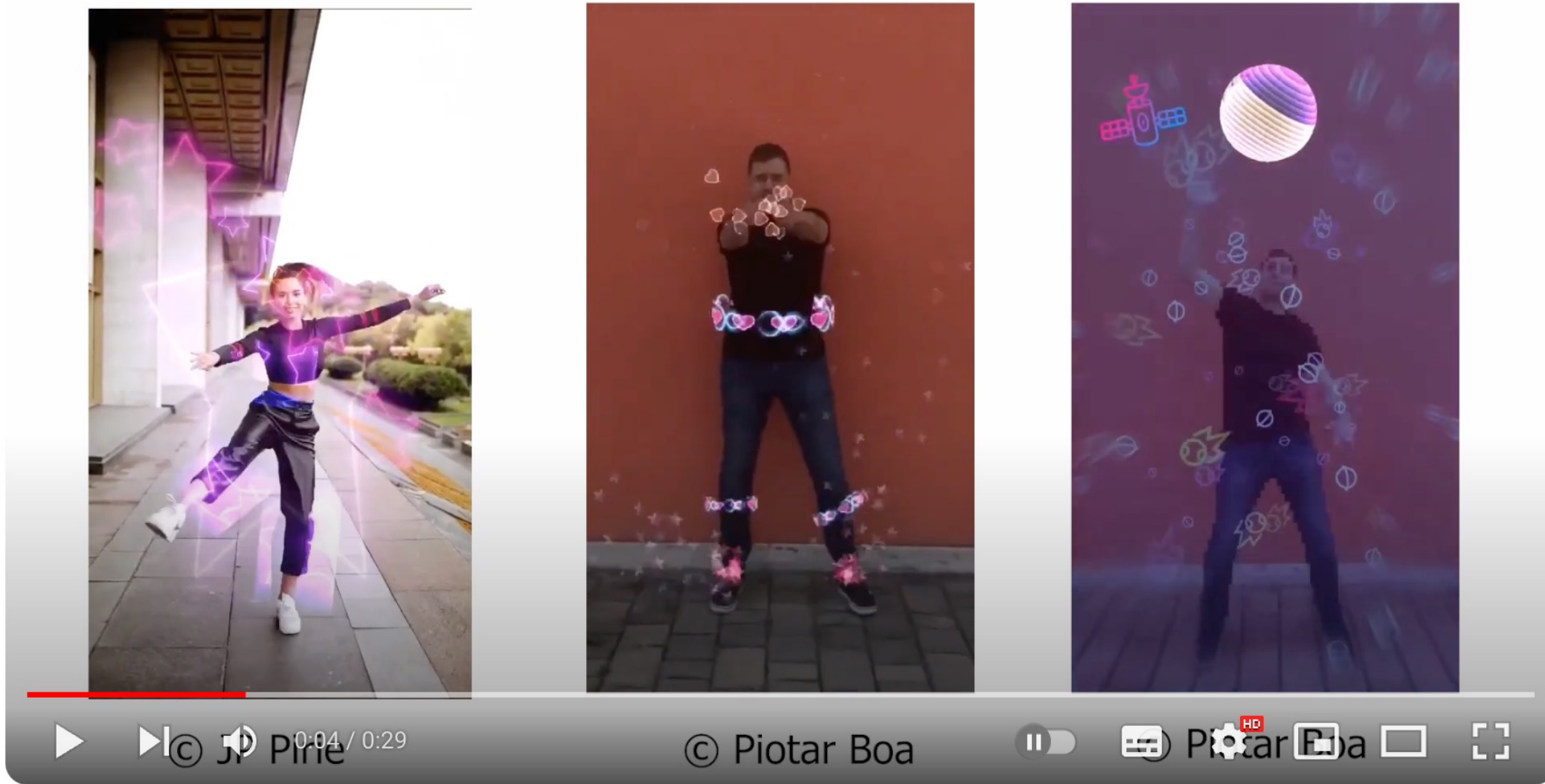
Load videos,  
Manipulate graphics,  
...



Create pose recognizer,  
Create nodes,  
...

Node Canvas

Reduce programming complexity using node UI.



## PoseVEC: Authoring Adaptive Pose-aware Effects Using Visual Programming and Demonstrations



ACM SIGCHI ✓

2.03万位订阅者



# Summary

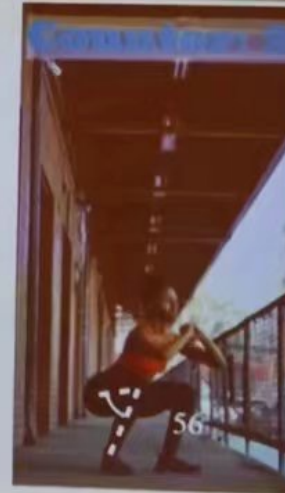
## PoseVEC

- programming by demonstration and visual programming
- direct manipulation with video canvas
- node-based programming UI

Project website



FOLLOW ME



GOOD FORM





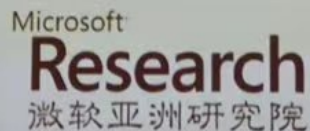
# UIST 2023

## Wakey-Wakey: Animate Text by Mimicking Characters in a GIF

Liwenhan Xie\*, Zhaoyu Zhou\*, Kerun Yū, Yun Wang, Huamin Qu, and Siming Chen

\*: equal contribution

Presenter: Qingyu Guo

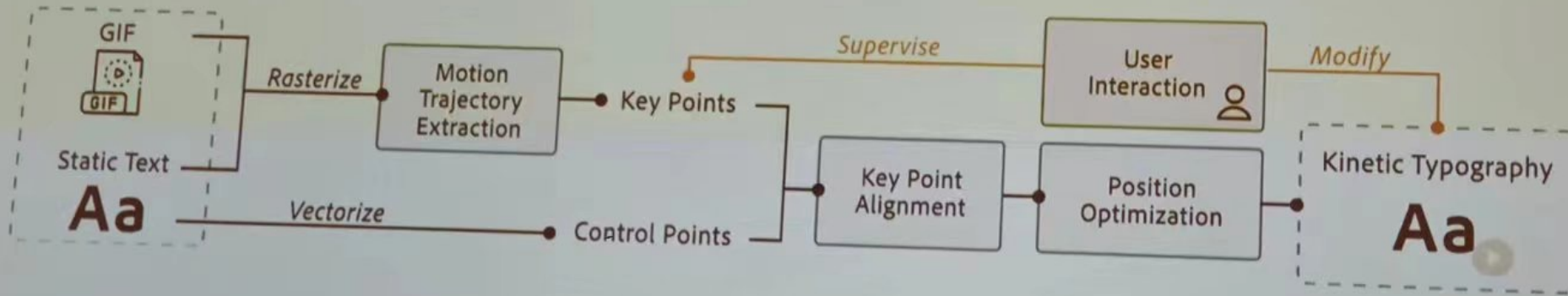




# Workflow

UIST 2023

## Input



## Output

讨论了generalizability, motion transfer model的局限



The screenshot displays the AnimaText software interface, which is used for creating animated GIFs from text. The interface is divided into three main panels:

- Left Panel:** Shows the text "wakey" in a large, bold font. Below it, there are options to select the type of animation (text, point, line) and a color picker. A list of GIF assets is visible, including a bear character.
- Middle Panel:** Shows a preview of the GIF. The word "wakey" is displayed in a stylized font, and the bear character is shown in a dynamic, jumping pose. A timeline at the bottom of this panel allows for keypoint correction.
- Right Panel:** Shows the final output of the GIF, which is a sequence of frames where the word "wakey" is animated with the bear character's movements.

The video player interface at the bottom shows the video is at 0:17 / 0:29. The video title is "[UIST'23 Preview] Wakey-Wakey: Animate Text by Mimicking Characters in a GIF". The creator is Liwenhan Xie, with 2 subscribers. The video has 5 likes and a share button.

### [UIST'23 Preview] Wakey-Wakey: Animate Text by Mimicking Characters in a GIF

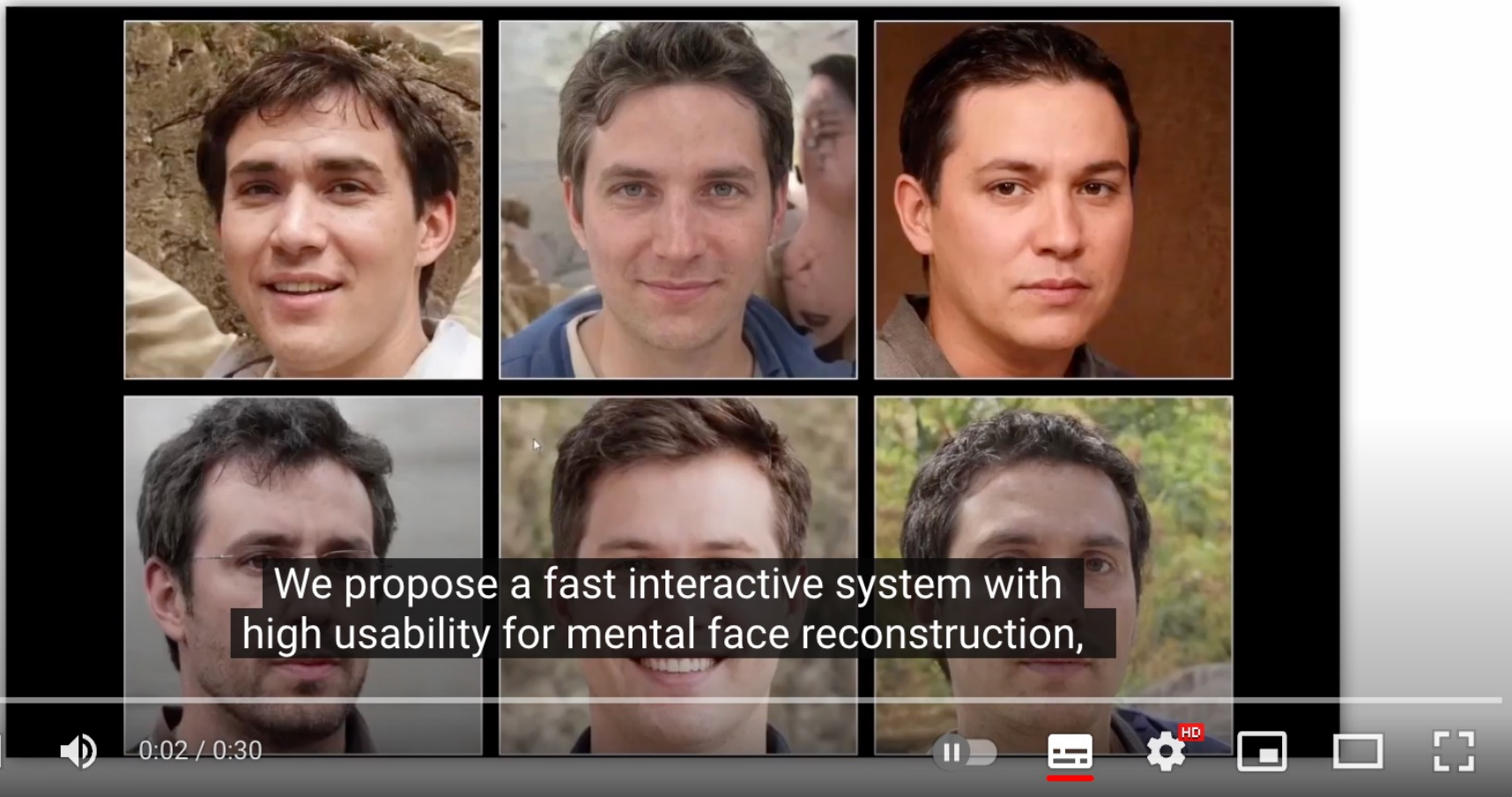


Liwenhan Xie  
2位订阅者



5 | | 分享 | 下载 | ...

# Our Mental Face Reconstruction System (MFRS)



Usable and Fast Interactive Mental Face Reconstruction

记忆中的人脸图片重建



ACM SIGCHI ✓  
2.03万位订阅者

订阅

👍 0



🔗 分享

↓ 下载



# Usable and Fast Interactive Mental Face Reconstruction



## MFRS – Quantitative Evaluation

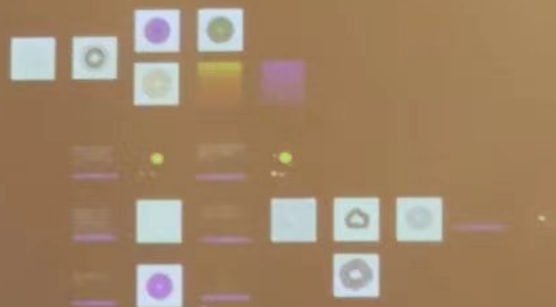
Method	Mental Rating $\uparrow$	Visual Rating $\uparrow$	SUS $\uparrow$	NASA-TLX $\downarrow$	Time (mins) $\downarrow$
Ours	4.0* $\pm$ .82	4.1 $\pm$ .95	84.6* $\pm$ 12.9	27.2* $\pm$ 18.4	10.0* $\pm$ 4.1
CG-GAN	4.8* $\pm$ .80	3.9 $\pm$ .95	59.0* $\pm$ 12.7	43.4* $\pm$ 10.9	17.0* $\pm$ 5.6

mental and visual quality的打分分析，有意思



# Spellburst

A Node-based Interface for  
Exploratory Creative Coding with  
Natural Language Prompts



[spellburst-paper.replit.app](https://spellburst-paper.replit.app)



Tyler Angert  
Replit, Inc.  
tyler@replit.com



Miroslav Suzara  
Stanford University  
msuzara@stanford.edu



Jenny Han  
Stanford University  
jennyhan@cs.stanford.edu



Christopher Pondoc  
Stanford University  
clpondoc@stanford.edu

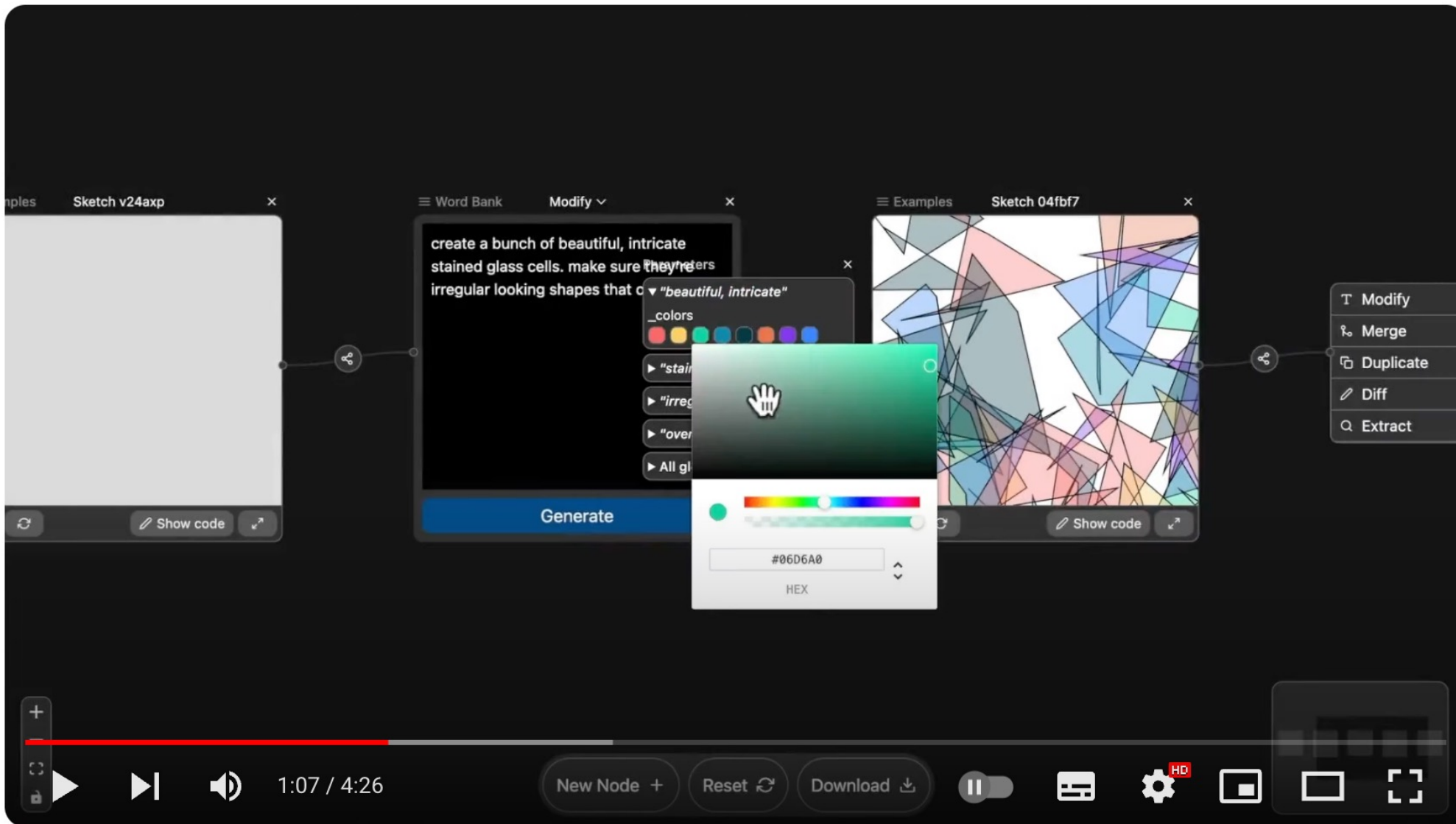


Hariharan Subramonyam  
Stanford University  
harihars@stanford.edu

# Semantic sliders

- 1. Idea → code
- 2. Variations
- 3. Combinations





## Spellburst: A Node-based Interface for Exploratory Creative Coding with Natural Language Prompts



Hariharan Subramonyam

订阅

59



分享

剪辑

<https://www.youtube.com/watch?v=5ww0jsS9hTI>

# Other Scenarios in UIST



Banquet  
at Holloween





# Poster session



# Poster session



# Demo session

## Storyfier: Exploring Vocabulary Learning Support with Text Generation Models

Zhenhui Peng\*, Xingbo Wang\*, Qushi Han, Junkai Zhu, Xiaojian Ma, Haimin Qi

**Want to have meaningful contexts for learning any target word?**

**User need**

- When learning a foreign language, it is beneficial to have vocabulary in meaningful contexts, such as short stories or dialogues, to learning vocabulary effectively.
- Maximizing the advantage of target words.
- Being able to use target words in the expression.

**Traditional approach**

- Readers reading materials or manually create contexts.
- It could not provide a meaningful context for any set of target words that users wish to learn.

**Proposed work and contributions**

- A vocabulary learning system Storyfier that facilitates users to explore the meanings and usage of any target English words on its generated stories and writing activities.
- Design process and user studies involving students, teachers, and NLP researchers that provide first-hand findings on the feasibility, effectiveness, and user experiences of applying generative models to vocabulary learning.
- Insights and design considerations of leveraging generative models to support learning tasks.

**We take an iterative approach to design Storyfier.**

**Phase 1**

**Phase 2**

**Phase 3**

**Our study offers reliable insights.**

**Key findings**

**Design considerations**

**Users can learn any interested target word set with meaningful story contexts in Storyfier.**

**Four types of interfaces for evaluating the value of generated stories and adaptive support**

**Read and rate** provides dictionary features such as example sentences, grammar, and target word (Part A, B, C).

**Read and rate** additionally provides a question every time users target words (Part A, B, C).

**Read and rate** also provides a question every time users target words (Part A, B, C).

**Read and rate** also provides a question every time users target words (Part A, B, C).

## CriTrainer: An Adaptive Training Tool for Critical Paper Reading

Kangyu Yuan\*, Hehai Lin\*, Shilei Cao\*, Zhenhui Peng†, Qingyu Guo, Xiaojian Ma

**Here are the training process and AI support!**

**Want to learn how to read a paper critically?**

**User need**

- Read critically paper contents is important for building the better comprehension. It requires:
  - Understanding the paper's main idea.
  - Finding critical thinking questions on the paper.

**Traditional approach**

- Teachers for critical paper reading or critical thinking.
- Control, not active in the paper's content.
- Discussion with a feedback loop given and between students, maybe qualified papers.

**Proposed work and contributions**

- A novel tool CriTrainer to actively user challenges and work for the support in the critical reading training process.
- A critical paper reading training tool CriTrainer which offers user-specific guidance and personalized information elements in a contextual reading process.
- Knowledge user enter that demonstrates CriTrainer's effectiveness in supporting participants' critical paper reading skills.
- Insights and design considerations for better critical paper reading training tools.

**QRAC training process**

**QR Question Read**

- Users read each paper section with questions about corresponding content.
- CriTrainer generates comprehension questions about the selected section to guide users in the QR stage.

**AC-1 (Answer): Check in a summarization task**

- Users summarize the reading section and check if it was a good summary.
- Users summarize the reading section and check if it was a good summary.
- Users summarize the reading section and check if it was a good summary.

**AC-2 (Answer): Check in a question-asking task**

- Users raise critical thinking questions in the reading section and check if they were relevant to the paper content and reflect critical thoughts.
- CriTrainer generates complete based user-specific critical thinking questions about the reading section.

**Example text-specific generated critical thinking questions**

- (Introduction) Is the method of this paper, e.g., infrared depth sensor (Method keyword), novel?
- (Introduction) Is the research about complexity non-made environment (Background keyword) important?
- (Discussion) Do the authors point out any potential concerns of their findings about directly attacking human (Result keyword)? Can I solve them?

**We develop an adaptive training tool CriTrainer.**

**Our study proves its effectiveness.**

**Study design**

- A mixed-design (test as between-subjects, time as within-subjects factor) with 24 university students.
- Baseline test: without text-specific and interaction features, with the general guidance in the QRAC process.

**Key findings**

- Compared to those with the baseline tool, participants with CriTrainer had significantly more improvement regarding their ability to raise understandable, relevant, and critical questions after the training session.
- After training with either tool, participants can also better express their understanding of the paper contents in their drafted summaries.

**Design considerations**

- Offer adaptive and interpretable suggestions on how to improve the drafted summary.
- Increase the diversity of generated critical thinking questions.
- Encourage necessary effort spent in the learning tasks.
- Balance the tradeoff between flexibility and sufficient support.

# Human-Cat Interaction



Coffee  
break





# Coffee break



# People I Met



# Yukang Yan

I am an incoming Assistant Professor of Computer Science at University of Rochester, also as a member of ROC-HCI group. I am currently a postdoctoral researcher in the Augmented Perception Lab, Human-Computer Interaction Institute at Carnegie Mellon University. Before that, I earned my Ph.D. degree and Bachelor's degree from Tsinghua University. My research focus lies in Human-Computer Interaction and Mixed Reality.

My research is focused on understanding, predicting and enhancing user behavior in Mixed Reality, which follows three connected threads: 1) capture the user's behavioral and perceptual patterns with computational methods, 2) develop input techniques and adaptive user interfaces to facilitate human-computer two-way communication, 3) explore unique behavioral enhancements enabled by Mixed Reality. I publish at ACM CHI, UIST, IMWUT and IEEE VR.

I will join UR in Januray 2024, I am looking for students (including Ph.D. students in 2024 Fall) to join my lab, please reach out if you are interested.

## Selected Publications

### HandAvatar: Embodying Non-Humanoid Virtual Avatars through Hands

Yu Jiang, Zhipeng Li, Mufei He, David Lindlbauer, **Yukang Yan**



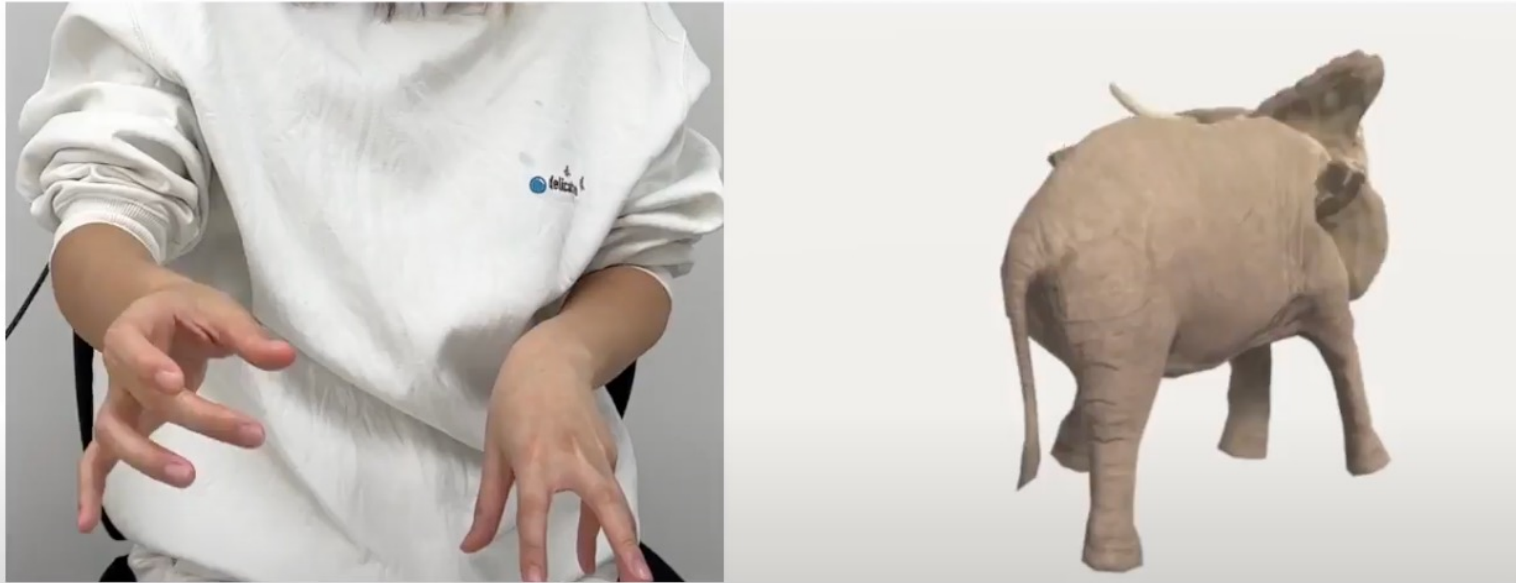
[\[Google Scholar\]](#)[\[CV/Resume\]](#)

## Latest News

October 28 - November 1: Attending UIST. I will chair the session "Digital Dexterity: Touching and Typing Techniques".



# HandAvatar: Embodying Non-Humanoid Virtual Avatars through Hands



1.5X

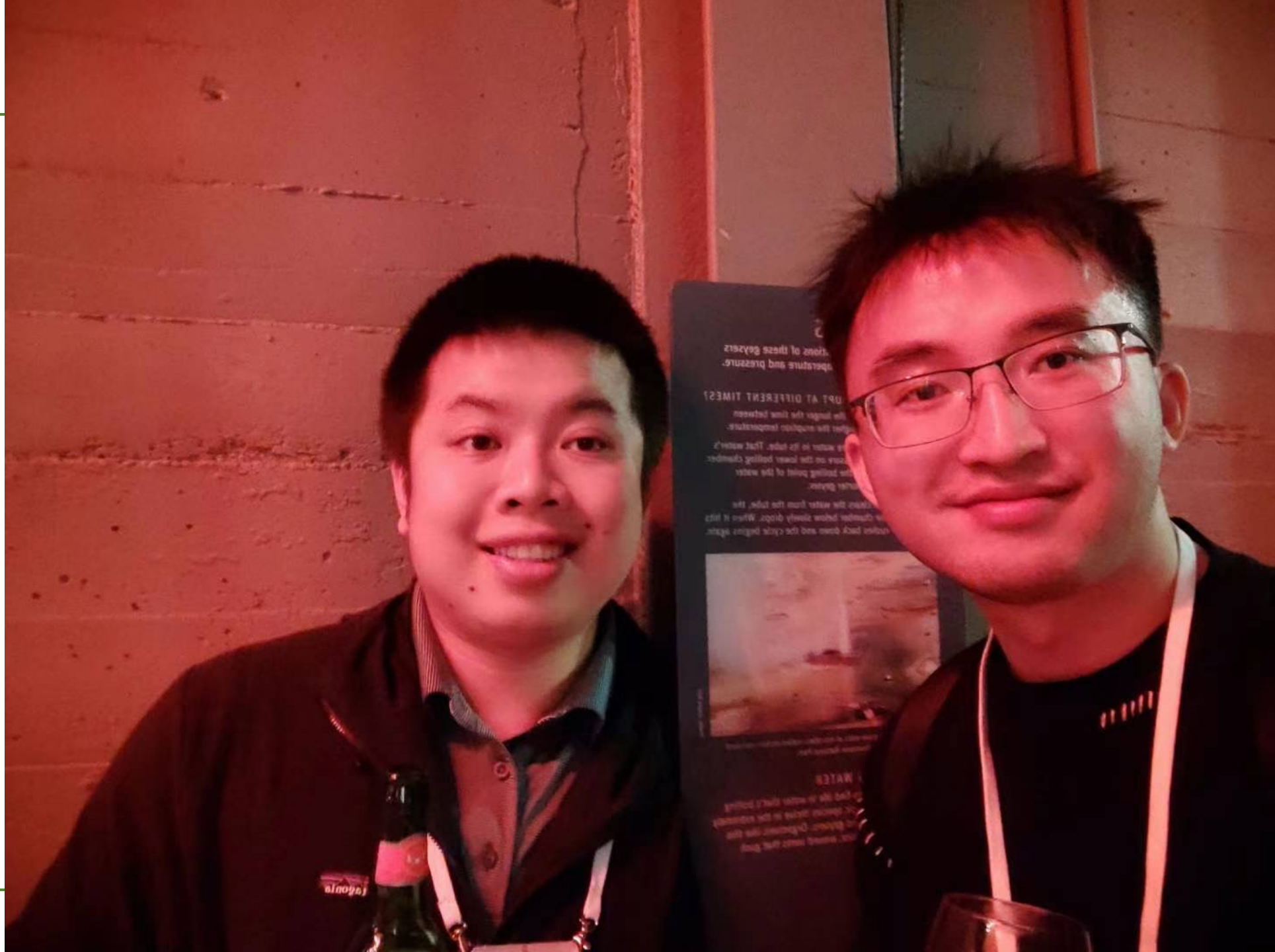


0:12 / 5:18



HD





**Toby Jia-Jun Li**

Pronoun: he/him

Assistant Professor

[Department of Computer Science and Engineering](#)[University of Notre Dame](#)

📍 214B Cushing Hall

✉ Email

🐦 Twitter

🏠 Github

🎓 Google Scholar

📢 I am recruiting Ph.D. students, undergraduate researchers, and visiting researchers. See the **Students** page for details.

Toby Jia-Jun Li is an Assistant Professor in the **Department of Computer Science and Engineering** at the **University of Notre Dame** where he directs the SaNDwich Lab. Toby received a Ph.D. degree in **Human-Computer Interaction** at **Carnegie Mellon University**, where he was advised by **Brad A. Myers**. Toby also worked closely with **Tom M. Mitchell**. Prior to Ph.D., Toby received his Bachelor's degree in **Computer Science** from the **University of Minnesota** where he worked with **Brent Hecht** as a member of **GroupLens Research**.

Toby works at the intersection of **Human-Computer Interaction (HCI)**, **End-User Software Engineering**, **Machine Learning (ML)**, and **Natural Language Processing (NLP)** applications, where he uses human-centered methods to design, build, and study interactive systems to empower individuals to create, configure, and extend AI-powered computing systems. His recent work seeks to address the **societal challenges in the future of work** through a bottom-up **human-AI collaborative** approach that helps individual workers automate and augment their tasks with AI systems.



< Story Buddy read with parent

- Select Storybooks
- Configuration
- Dashboard
- Switch Mode

Select a story for your kid!

### Featured Stories

Three Bears in a Boat

→ READ THE STORY

Chris P. Bacon

→ READ THE STORY

1:05 / 4:03

⏪ ⏩ 🔊 ⏸ ⏴ ⚙️ HD 📺 🖥️ 🗑️

## StoryBuddy: A Human-AI Collaborative Agent for Parent-Child Interactive Storytelling with ...



ACM SIGCHI ✓  
2.03万位订阅者

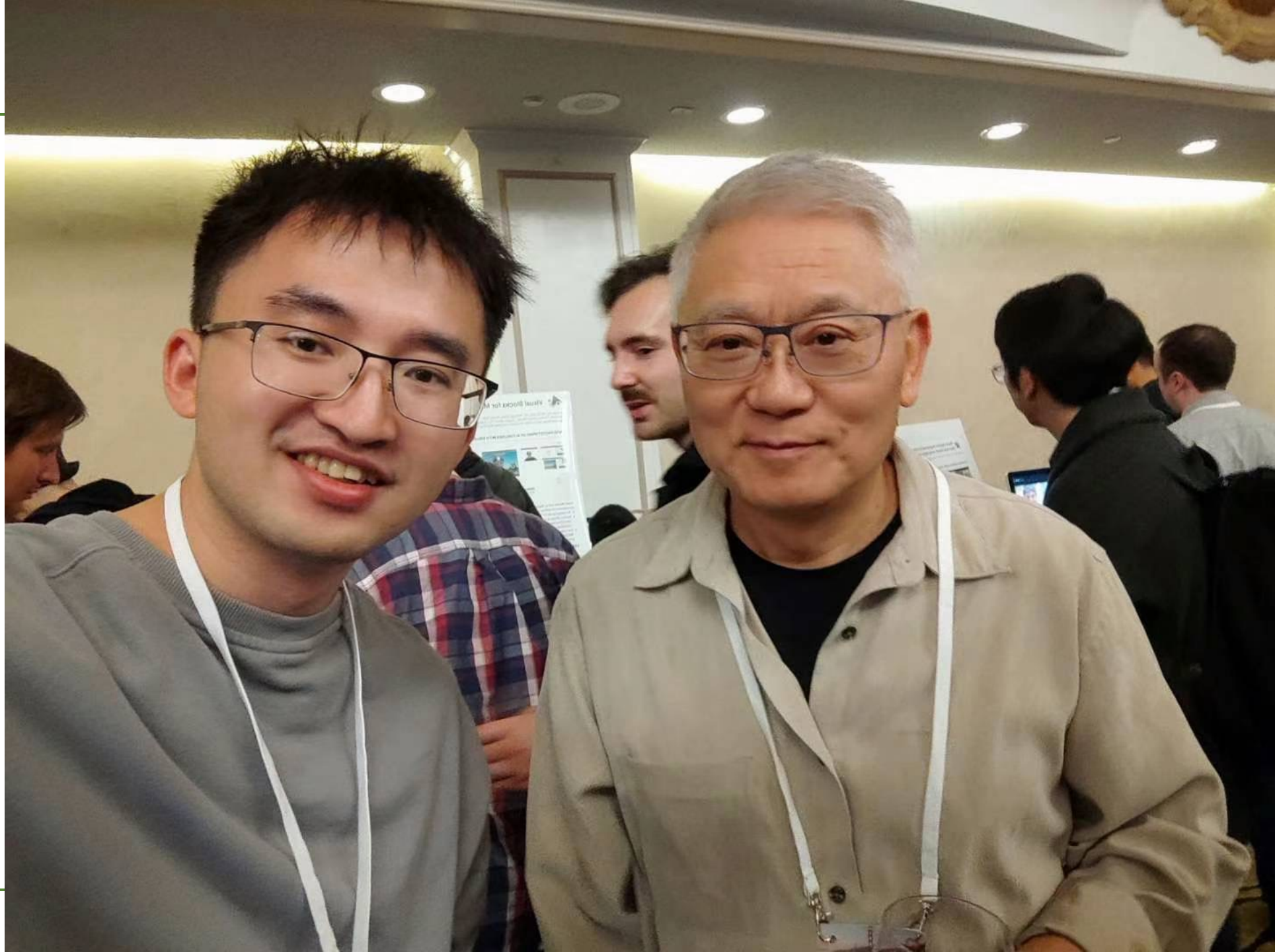


👍 6 | 🗨️

🔗 分享

⬇️ 下载







## Shumin Zhai

---



Shumin Zhai is a Principal Scientist and Director of Gboard and Interaction Methods at Google. His research has advanced foundational models and understandings of human-computer interaction (HCI), as well as practical user interface designs, inventions, and flagship products. His publications have won numerous awards, including the ACM UIST Lasting Impact Award, the ACM Mobile HCI Impact Award, and the IEEE Computer Society Best Paper Award. He was the Editor-in-Chief of the *ACM Transactions on Computer-Human Interaction* from 2009 to 2015. He received his Ph.D. degree from the University of Toronto in 1995. In 2006, he was named to ACM's inaugural class of Distinguished Scientists. In 2010, he was named a Member of the ACM CHI Academy and a Fellow of the ACM.



Gboard: now available for Android

Sure. Umami burger?

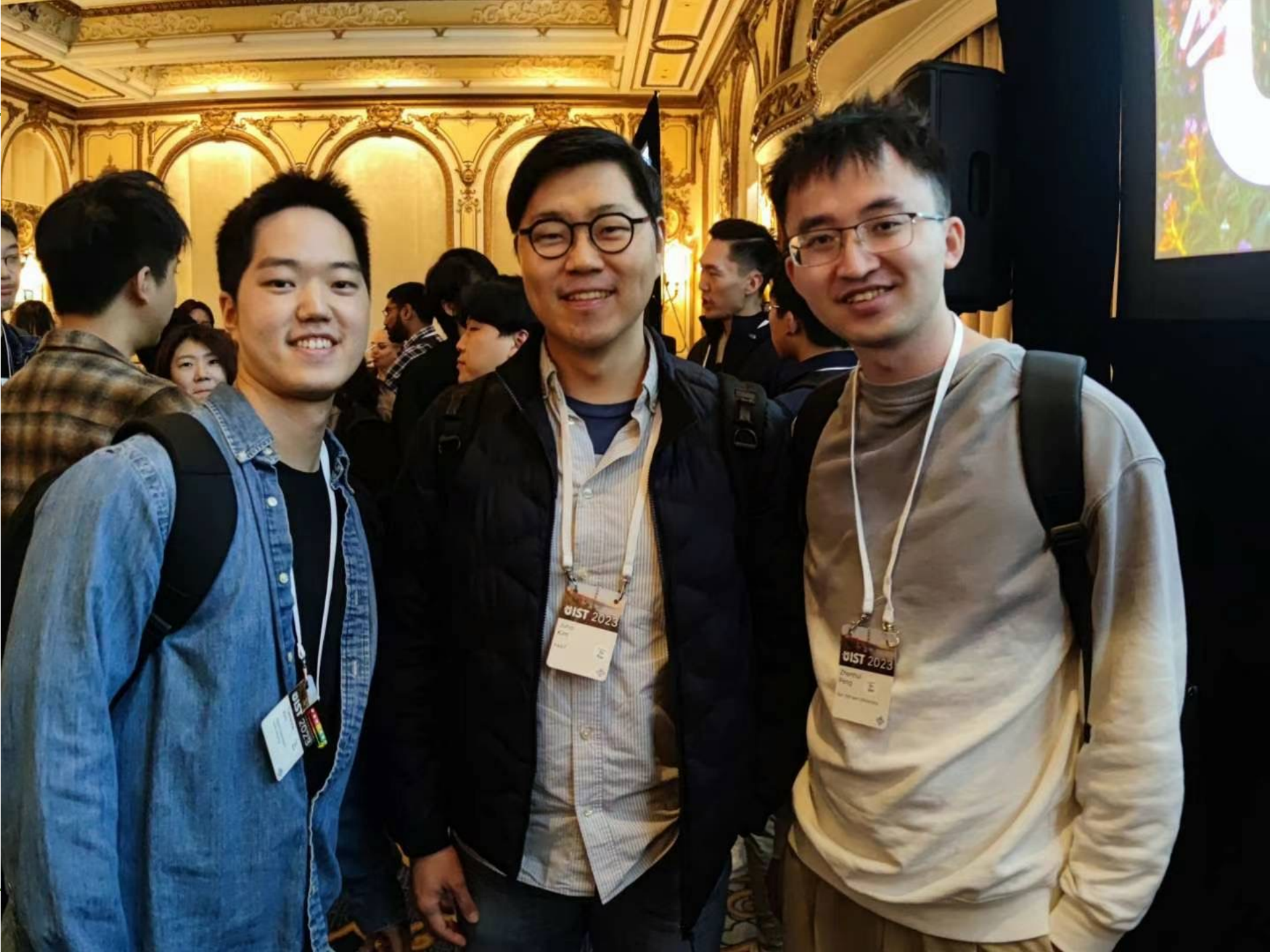
Yeah. Know the address?

738 E. 3rd St.

More videos

0:29 / 1:09

YouTube





# Juho Kim

Home

Publications

Students

Research

Fun



Associate Professor  
School of Computing  
KAIST

<https://juhokim.com/>

Juho Kim is an Associate Professor in the **School of Computing** at **KAIST**, and directs **KIXLAB** (the KAIST Interaction Lab). He is also affiliated faculty in the **Kim Jaechul Graduate School of AI** at KAIST. His research in human-computer interaction and human-AI interaction focuses on building interactive and intelligent systems that support interaction at scale, aiming to improve the ways people learn, collaborate, discuss, make decisions, and take action online.

He earned his Ph.D. from **MIT**, M.S. from **Stanford University**, and B.S. from **Seoul National University**. In 2015–2016, he was a Visiting Assistant Professor and a **Brown Fellow** at **Stanford University**. He is a recipient of a KIISE/IEEE-CS Young Computer Researcher Award, KAIST’s Songam Distinguished Research Award, Grand Prize in Creative Teaching, Q-Day Creative Education Award, and Excellence in Teaching Award, as well as 14 paper awards from ACM CHI, ACM CSCW, ACM Learning at Scale, ACM IUI, ACM DIS, and AAI HCOMP. He recently spent his sabbatical year as a chief scientist at Ringle Inc. to transfer his research on AI-powered analysis and diagnosis of English learners’ proficiency into a real product. He gave a keynote at NeurIPS 2022 titled “Interaction-Centric AI”.

If you're interested in working with me at KAIST, please **read this page**. Also, in **this interview video for HCI Korea**, I summarize KIXLAB's representative projects and share my thoughts on research and mentoring. English transcription is available.

Video of my recent talks on "Interaction-Centric AI": (1) **NeurIPS 2022 keynote** (targeted at AI audience), (2) **Stanford HCI Seminar** (targeted at HCI audience)



# People I Met



SlidesLive LIBRARY Categories Search Log in Sign

NeurIPS > NeurIPS 2022 > Live Streams > Interaction-Centric AI

NEURAL INFORMATION PROCESSING SYSTEMS  
INTERACTION-CENTRIC AI

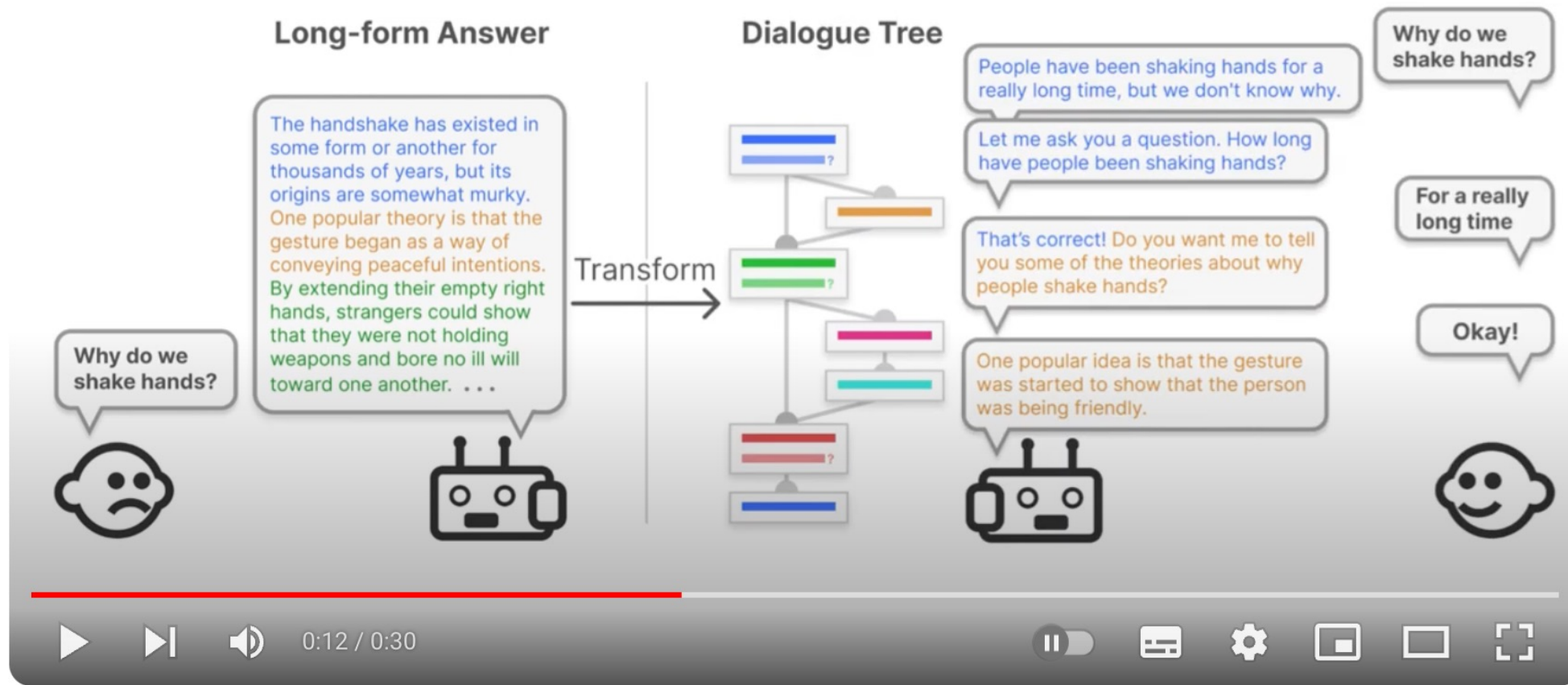
## Interaction-Centric AI

Juho Kim  
juhokim.com | kixlab.org | @imjuhokim | juhokim@kaist.ac.kr

Slide 1 / 118

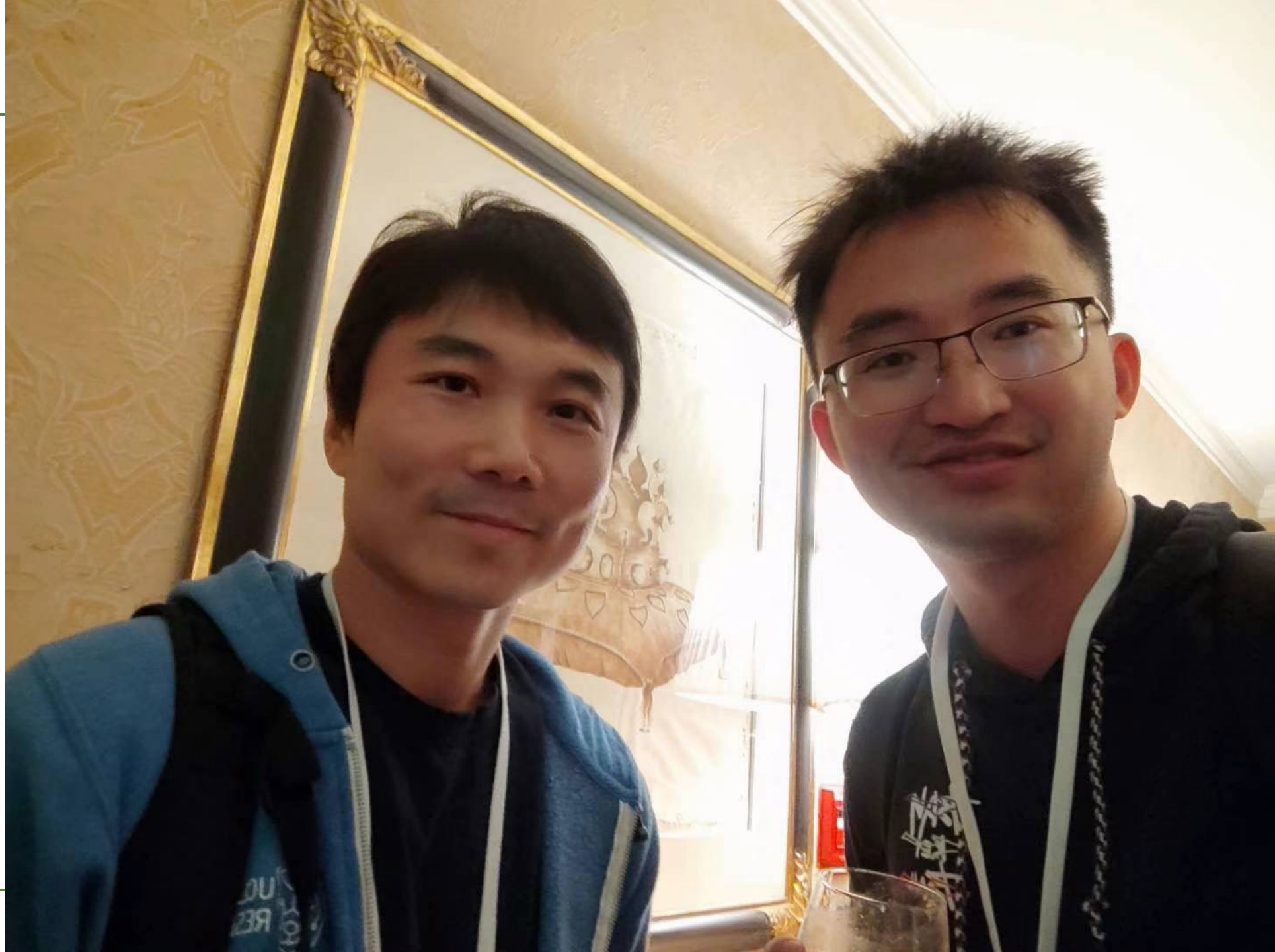
KAIST School of Computing

Should this presentation be stored for 1000 years?  
[How do we store presentations](#)



### DAPIE: Interactive Step-by-Step Explanatory Dialogues to Answer Children's Why and How Questions

ACM SIGCHI  2.03万位订阅者  0 |  |  分享 |  下载 | 





## Xiang 'Anthony' Chen

xac @ ucla.edu

Assistant Professor

ECE Department, UCLA

UCLA HCI Research

 Fall 2023 Office Hours

My research employs human-centered design methods to build systems that catalyze advances in AI to augment human activities, supported by NSF CAREER Award, ONR YIP Award, Google Research Scholar Award, Intel Rising Star Award, Hellman Fellowship, and NSF CRII Award.

Vitæ

Papers

Tweets

Reads



I was also awarded a Google PhD Fellowship in Human Computer Interaction in 2022!

Northwestern wrote a nice [article featuring me as a fellowship recipient](#).



I am a Postdoctoral Researcher at Stanford University CS, where I collaborate with [Diyi Yang](#) and [Emma Brunskill](#) to advance AIs potential to assist in Mental Health and Psychotherapy. Mental health providers--from trained therapists to volunteer peer supporters--remain crucial, and we hypothesize that AI-assisted training and practice can improve their efficacy, confidence, and growth. Human-AI collaboration and interfaces for effectively imbuing AIs with domain-specific and personal-knowledge of this area will be important.

I received my PhD from Northwestern's [Technology and Social Behavior](#) program, where I worked closely with Prof. [Haoqi Zhang](#) and [Darren Gergle](#). My research aimed to enhance people's communication of intent and goals when using AI systems to assist in personally-meaningful pursuits, such as creating music compositions and sharing meaningful social experiences at distance. To do this, I advanced Human-AI Interface Layers, a technology sitting on top of existing AIs that help people in communicating their intent and expectations to their AI assistants and collaborators.



## Anyi Rao

Postdoctoral Scholar

Computer Science Department  
Stanford University

Email: anyirao [at] stanford.edu



### Bio

Anyi Rao is a Postdoctoral Scholar at **Stanford** with **Maneesh Agrawala**. He has research experiences at **Meta Reality Lab**, Vector Institute, University of Toronto, Hong Kong University. He received the Ph.D. at **MMLab**, Chinese University of Hong Kong in 2022, advised by **Dahua Lin** and **Bolei Zhou**. He studies human-centered AI for creativity and multimodality, with focuses on content generation, intelligent video editing and creation, video semantic and cinematic analysis, aiming to build connections between AI and humans for collaborative intelligence. His works include **ControlNet**, **AnimateDiff**, **MovieNet**, **CityNeRF**, **Shoot360**, and **Virtual Dynamic Storyboard**, with a Marr Prize (ICCV best paper award).

*If you also have some exciting ideas and insights on the aforementioned research, please drop me an email. Let's push it forward together.*



# Adding Conditional Control to Text-to-Image Diffusion Models



Lvmin Zhang, Anyi Rao, and Maneesh Agrawala  
Stanford University

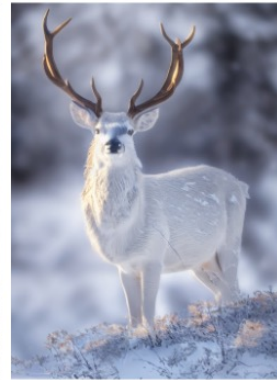
{lvmin, anyirao, maneesh}@cs.stanford.edu



Input Canny edge



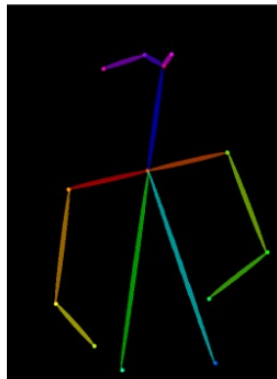
Default



“masterpiece of fairy tale, giant deer, golden antlers”



“..., quaint city Galic”



Input human pose



Default



“chef in kitchen”



“Lincoln statue”

Figure 1: Controlling Stable Diffusion with learned conditions. ControlNet allows users to add conditions like Canny edges (top), human pose (bottom), *etc.*, to control the image generation of large pretrained diffusion models. The default results use the prompt “a high-quality, detailed, and professional image”. Users can optionally give prompts like the “chef in kitchen”.



## Prof. Antti Oulasvirta

Antti Oulasvirta leads the Computational Behavior Lab ([cbl.aalto.fi](http://cbl.aalto.fi)) at Aalto University and the Interactive AI research program at FCAI (Finnish Center for AI). Prior to joining Aalto, he was a Senior Researcher at the Max Planck Institute for Informatics and the Cluster of Excellence on Multimodal Computing and Interaction at Saarland university. He received his doctorate in Cognitive Science from the University of Helsinki in 2006, after which he was a Fulbright Scholar at the School of Information in University of California-Berkeley in 2007-2008 and a Senior Researcher at Helsinki Institute for Information Technology HIIT in 2008-2011. During his postgraduate studies in 2002-2003, he was an exchange student at UC Berkeley's Neuropsychology Lab. He was awarded the ERC Starting Grant (2015-2020) for research on computational design of user interfaces. Dr. Oulasvirta serves as an associate editor for ACM TOCHI and has previously served International Journal of Human-Computer Studies, as well as served as a column editor for IEEE Computer. He frequently participates in the paper committees of HCI conferences, including the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI). His work has been awarded the Best Paper Award and Best Paper Honorable Mention at CHI fifteen times between 2008 and 2022. He has held keynote talks at NordiCHI'14, CoDIT'14, EICS'16, IHCI'17, ICWE'19, Chinese CHI'19, and IS-EUD'23. He is a member of ELLIS (European Laboratory for Learning and Intelligent Systems). In 2019, he was invited to the Finnish Academy of Science and Letters. He was a SICSA Distinguished Visiting Fellow in 2011 and in 2022.



### Research Lead

- [Computational Behavior Lab at Aalto University](http://cbl.aalto.fi)
- [Interactive AI research programme at FCAI](#)

# Trip to Stanford University

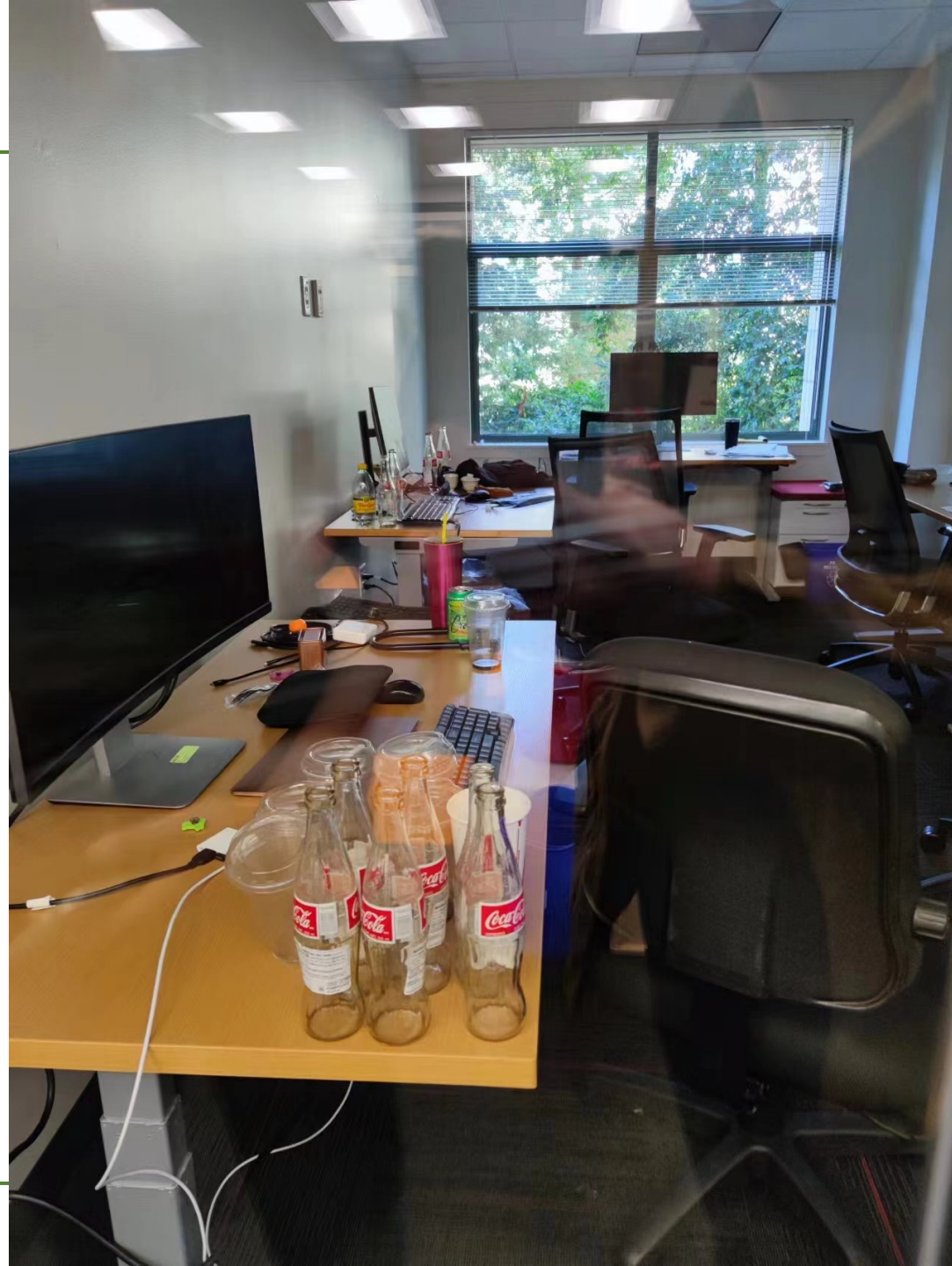


胡佛塔





# 博士/博后办公室





# 校园长廊







Fei-Fei Li的办公室





# 周日集市

下雨了怎么办?



# Trip to UC Berkeley



萨瑟塔  
(钟楼)



# 肠粉店...



# Trip in SF













# Trip in SF



# NBA Game



# HCI elements in NBA Game



# Other Learned Things

---



Learn from a chat with Tommy (Taewook Kim)

- From Prof. Juho Kim
  - It is an important thing that let supervisors not worry about you.
  - Expectation management is important in the student-supervisor relationship.
  - You should also get into the HCI community. The community is a family that helps students grow up. Everyone should also volunteer to help the community.
- From Prof. Xiaojuan Ma
  - Don't easily give up. Think of your co-authors. Everyone is helping you.