

从 GUI 到 NUI

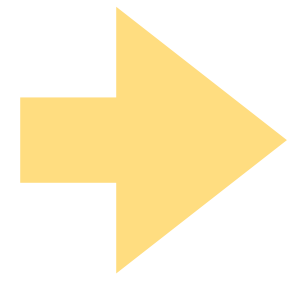
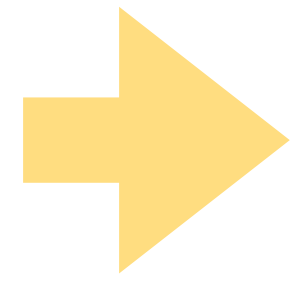
LLM 对交互模式的影响



谢亚东 / Lepton AI

Section 01

GUI / NUI 和 LLM



Command Line Interface
命令行界面

Graphical User Interface
图形用户界面

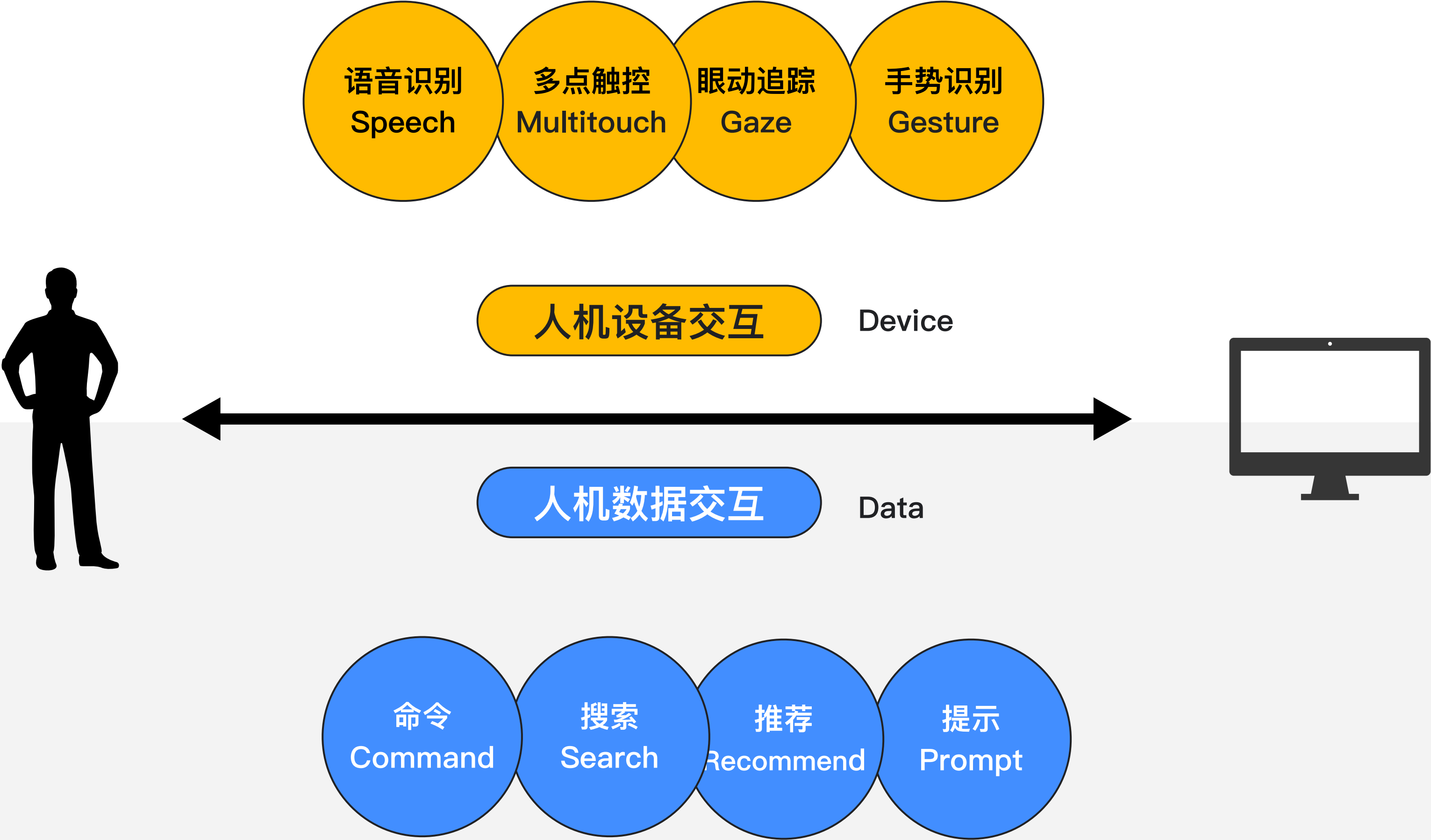
Nature User Interface
自然用户界面



“自然用户界面是一种旨在重复利用现有技能直接与内容交互的用户界面”

“An interface that is designed to reuse existing skills for interacting appropriately with content.”

Joshua Blake



应用代表

Office /
Windows

Yahoo /
Google

Tiktok /
WeChat

Bard /
ChatGPT

What's
Next

人机设备交互

语音识别
Speech

多点触控
Multitouch

手势识别
Gesture

眼动追踪
Gaze

人机数据交互

命令
Command

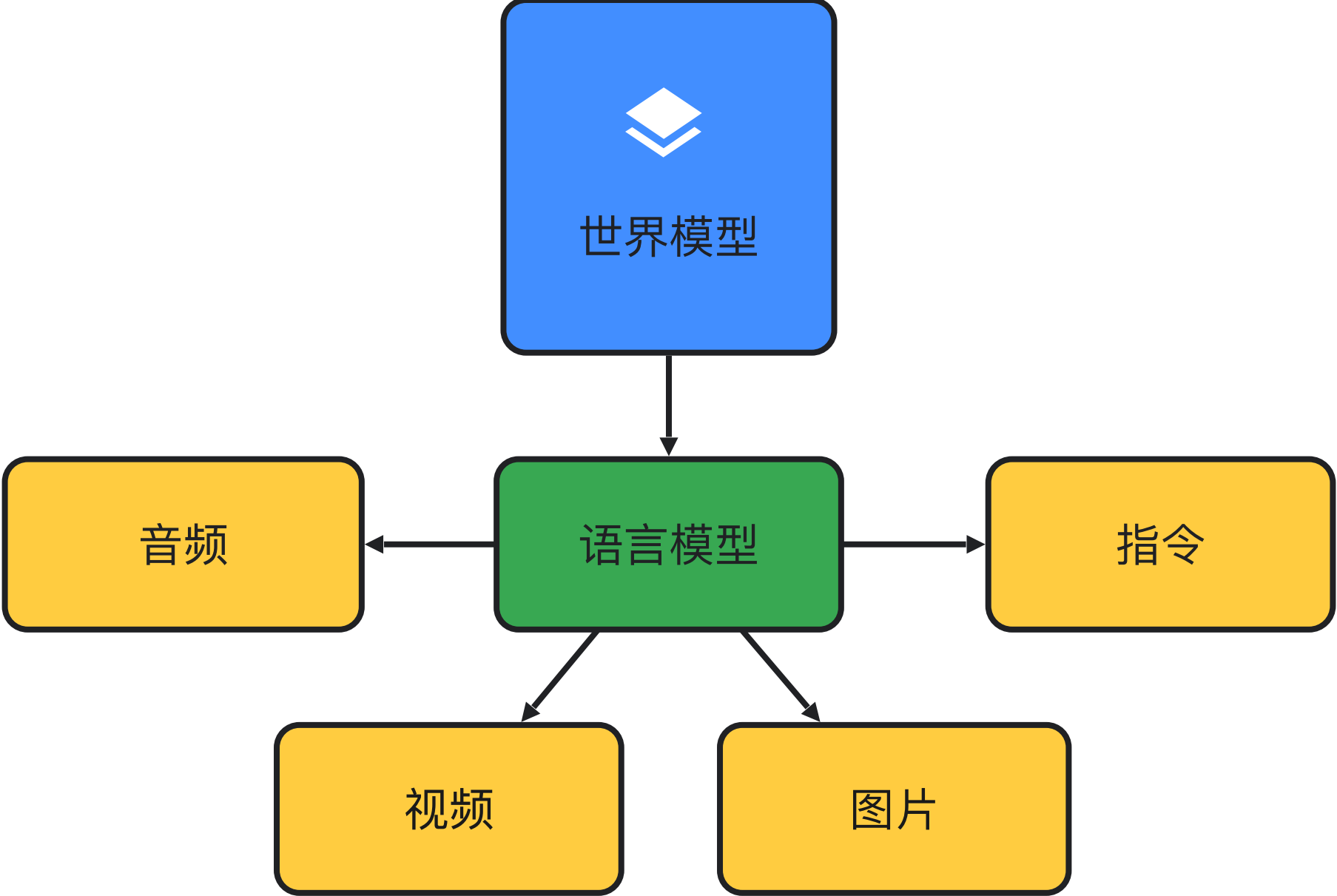
搜索
Search

推荐
Recommend

提示
Prompt



LLM 是 NUI 人机数据交互的基础



Section 02

发展趋势

NUI 在 LLM 下的发展趋势

1. 从不确定到确定

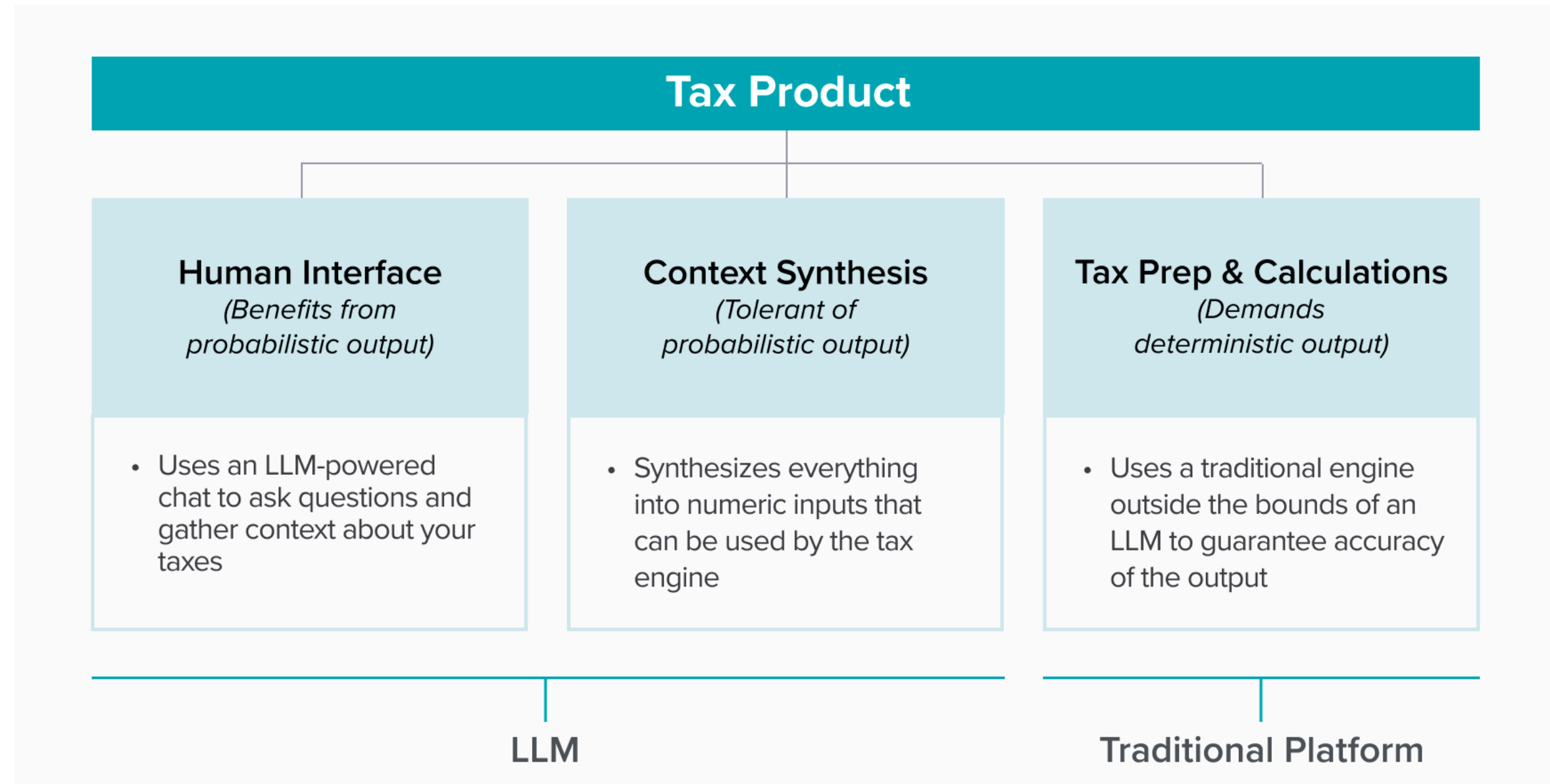
From probabilistic to deterministic

2. 渐进式升级

Progressive Updating

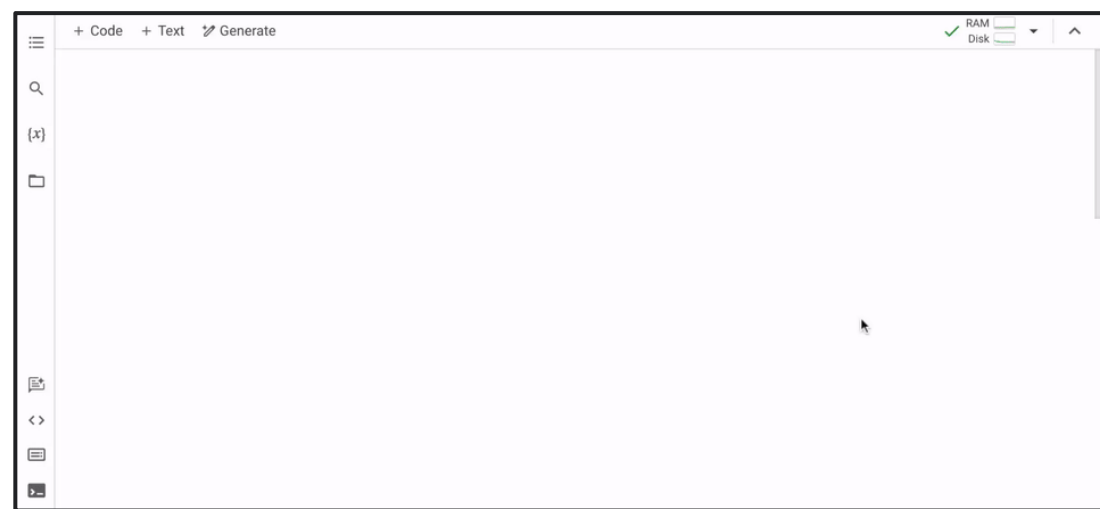
1. 从不确定到确定

- 可以从不确定性中获益的产品形态
- 可以忍受不确定性的产品形态
- 要求必须准确的产品形态



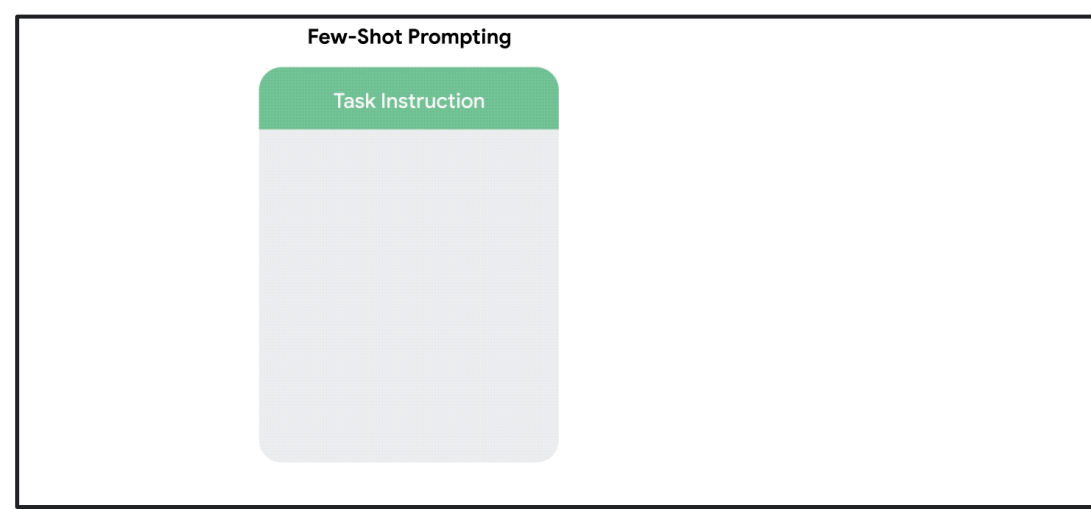
2. 渐进式升级

A. 保持现有交互



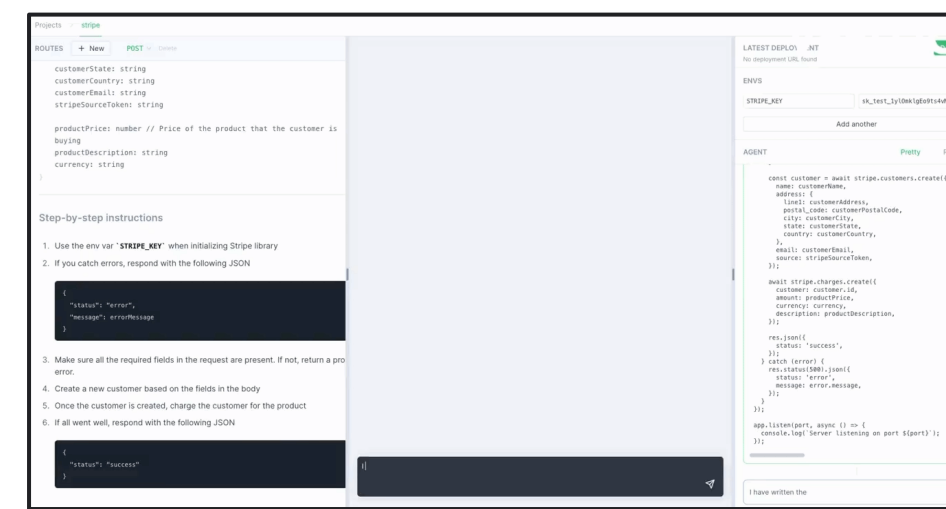
<https://blog.google/technology/developers/google-colab-ai-coding-features/>

B. 理解现有界面




<https://ai.googleblog.com/2023/05/enabling-conversational-interaction-on.html>

C. 根据需求生成界面



<https://twitter.com/mlejva/status/1641151421830529042>

+ Code + Text  Generate

✓ RAM
Disk ^



Few-Shot Prompting

Task Instruction

ROUTES [+ New](#) **POST** [Delete](#)

```
customerState: string
customerCountry: string
customerEmail: string
stripeSourceToken: string

productPrice: number // Price of the product that the customer is
buying
productDescription: string
currency: string
}
```

Step-by-step instructions

1. Use the env var `STRIPE_KEY` when initializing Stripe library
2. If you catch errors, respond with the following JSON

```
{
  "status": "error",
  "message": errorMessage
}
```

3. Make sure all the required fields in the request are present. If not, return a proper error.
4. Create a new customer based on the fields in the body
5. Once the customer is created, charge the customer for the product
6. If all went well, respond with the following JSON

```
{
  "status": "success"
}
```

||

LATEST DEPLOY .NET

No deployment URL found

ENVS

 [Add another](#)

AGENT

[Pretty](#)[Raw](#)

```
const customer = await stripe.customers.create({
  name: customerName,
  address: {
    line1: customerAddress,
    postal_code: customerPostalCode,
    city: customerCity,
    state: customerState,
    country: customerCountry,
  },
  email: customerEmail,
  source: stripeSourceToken,
});

await stripe.charges.create({
  customer: customer.id,
  amount: productPrice,
  currency: currency,
  description: productDescription,
});

res.json({
  status: 'success',
});
} catch (error) {
  res.status(500).json({
    status: 'error',
    message: error.message,
  });
}
});

app.listen(port, async () => {
  console.log(`Server listening on port ${port}`);
});
```

Section 03

GUI + NUI DEMO

Speech color changer

Tap/click then say a color to change the background color of the app. Try **aqua** azure beige bisque
black blue brown chocolate coral crimson cyan fuchsia ghostwhite gold goldenrod gray green indigo
ivory khaki lavender lime linen magenta maroon moccasin navy olive orange orchid peru pink plum
purple red salmon sienna silver snow tan teal thistle tomato turquoise violet white yellow .

...diagnostic messages

The download may take a few minutes, only for the first run. The subsequent refreshes and runs will be faster.

Chat Demo

The chat demo is based on [vicuna-7b-v1.1](#) model and [RedPajama-INCITE-Chat-3B-v1](#) model. More model supports are on the way.

vicuna-v1-7b-q4f32_0

[System Initalize] Finish loading on WebGPU - apple

Hi

Hello! How can I help you today?

what is your name

As an AI language model, I go by the name Vicuna and was trained by researchers from Large Model Systems Organization (LMSYS).

Enter your message...

Send

Reset

prefill: 14.3821 tokens/sec, decoding: 21.2188 tokens/sec

I want to check out



```
{  
  "name": "checkout",  
  "description": "when user want to checkout"  
}
```



I want to buy iPhone

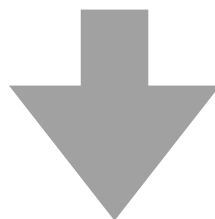


```
{  
  "name": "buy",  
  "description": "when user want to buy something"  
}
```

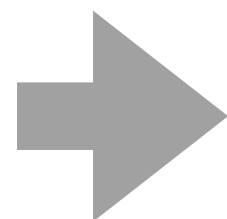




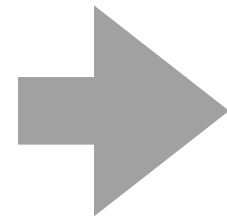
Voice



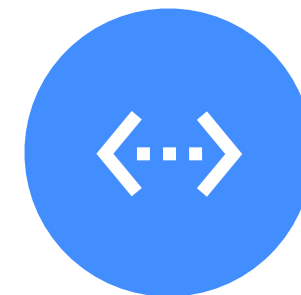
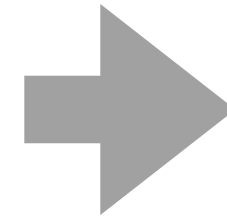
Speech
Recognition



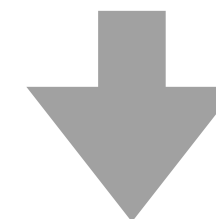
Text



LLMs



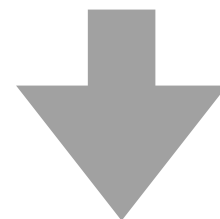
Structured
data

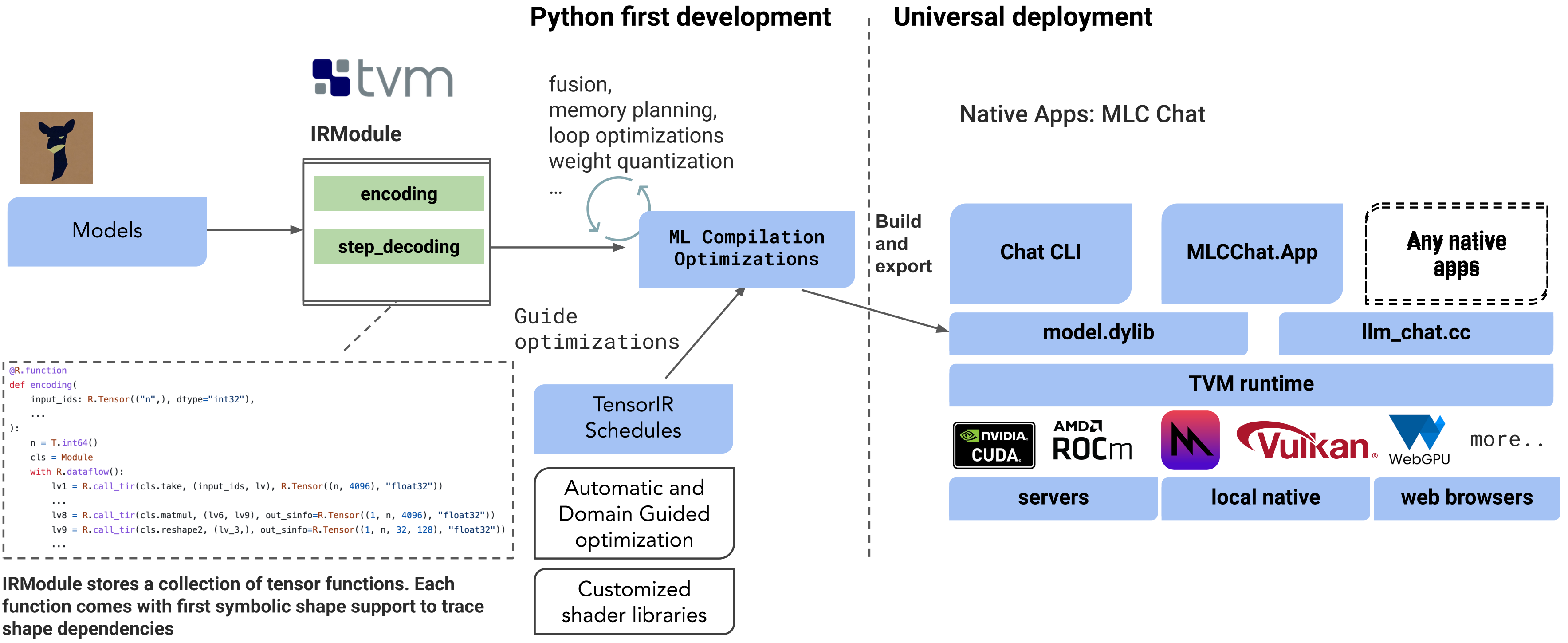


Render



Prompt







Chrome ships WebGPU

After years of development, the Chrome team ships WebGPU which allows high-performance 3D graphics and data-parallel computation on the web.

Published on Thursday, April 6, 2023

Table of contents

[A new dawn for web graphics](#)

[Browser support](#)

[Library support](#)

[Resources](#)

[Acknowledgments](#)

Filter

Web Speech API

▼ Guides

[Using the Web Speech API](#)

▼ Interfaces

[SpeechGrammar](#)[SpeechGrammarList](#)[SpeechRecognition](#)[SpeechRecognitionAlternative](#)[SpeechRecognitionErrorEvent](#)[SpeechRecognitionEvent](#)[SpeechRecognitionResult](#)[SpeechRecognitionResultList](#)[SpeechSynthesis](#)[SpeechSynthesisErrorEvent](#)[SpeechSynthesisEvent](#)[SpeechSynthesisUtterance](#)[SpeechSynthesisVoice](#)

Using the Web Speech API

The Web Speech API provides two distinct areas of functionality — speech recognition, and speech synthesis (also known as text to speech, or tts) — which open up interesting new possibilities for accessibility, and control mechanisms. This article provides a simple introduction to both areas, along with demos.

Speech recognition

Speech recognition involves receiving speech through a device's microphone, which is then checked by a speech recognition service against a list of grammar (basically, the vocabulary you want to have recognized in a particular app.) When a word or phrase is successfully recognized, it is returned as a result (or list of results) as a text string, and further actions can be initiated as a result.

The Web Speech API has a main controller interface for this — [SpeechRecognition](#) — plus a number of closely-related interfaces for representing grammar, results, etc. Generally, the default speech recognition system available on the device will be used for the speech recognition — most modern OSes have a speech recognition system for issuing voice commands. Think about Dictation on macOS, Siri on iOS, Cortana on Windows 10, Android Speech, etc.

Note: On some browsers, such as Chrome, using Speech Recognition on a web page involves a server-based recognition engine. Your audio is sent to a web service for

In this article

[Speech recognition](#)[Speech synthesis](#)

Section 04

总结

- LLM 是工具，而不是目的
- 学习成本与效率的平衡
- 离真正的 NUI 还有距离，但已经在路上

Google  Extended

Thank You