

# AutoGLM：自主化的GUI基础代理

--智能人机交互新范式

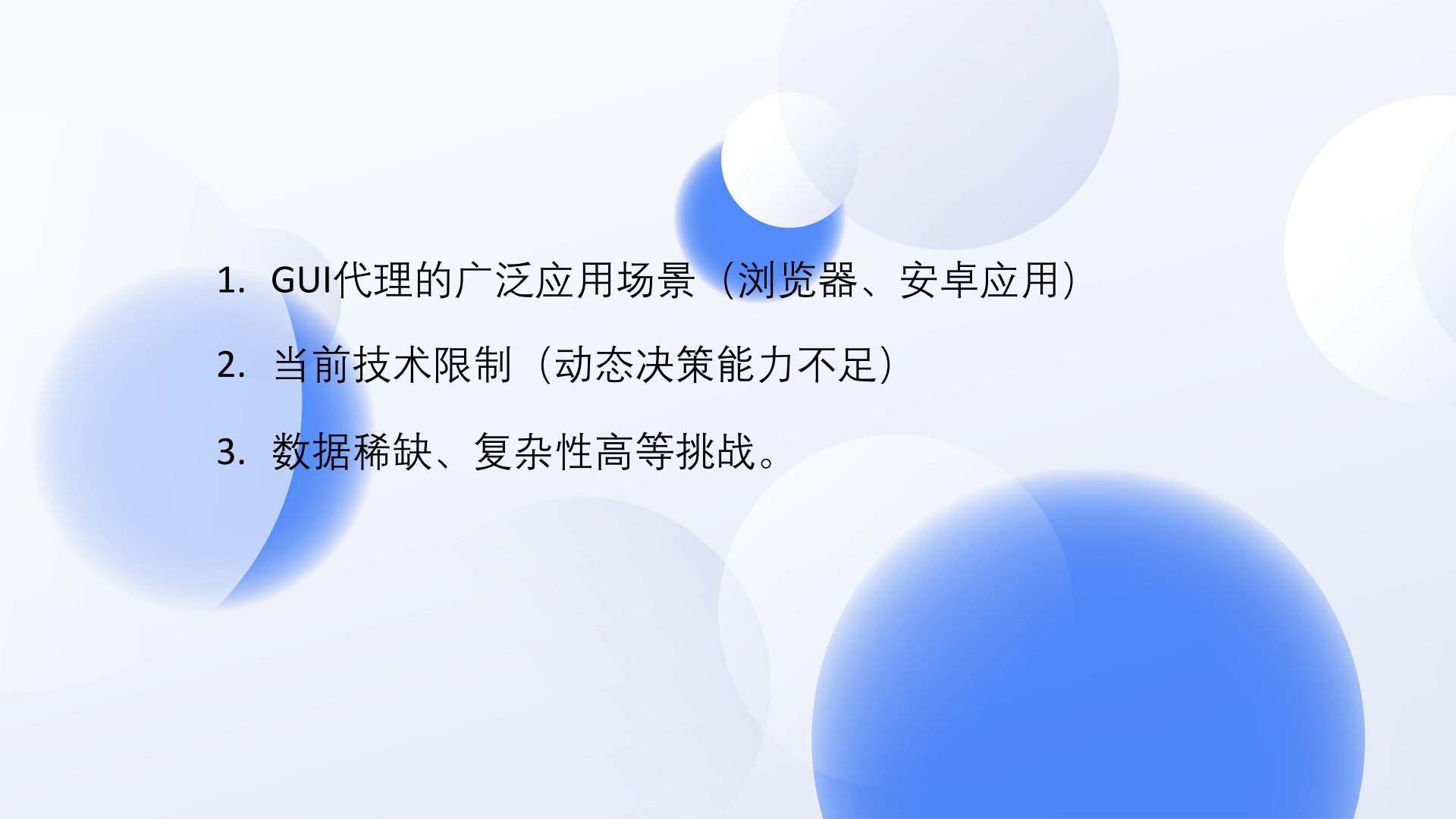
汇报人：龙卫平



01

# 背景与现状



- 
1. GUI代理的广泛应用场景（浏览器、安卓应用）
  2. 当前技术限制（动态决策能力不足）
  3. 数据稀缺、复杂性高等挑战。



02

# AutoGLM简介

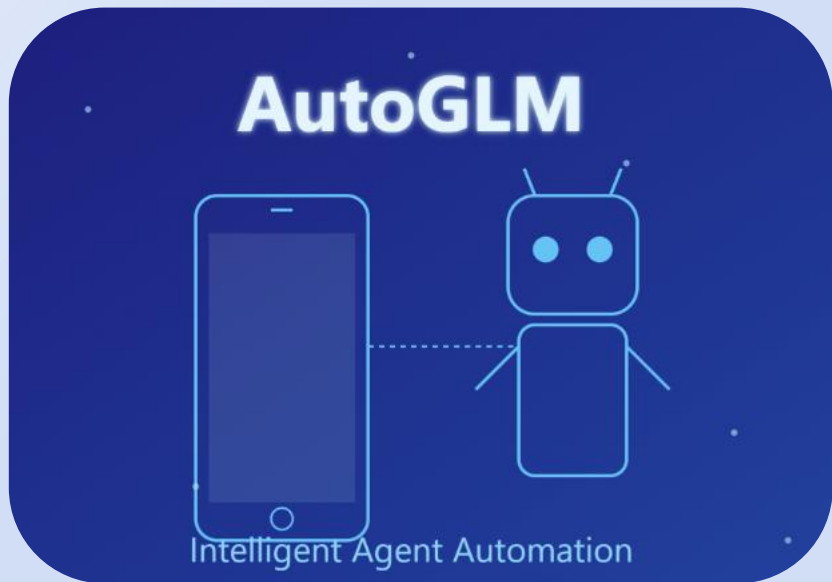




## 什么是AutoGLM?

产品定位：一种智能化GUI操作代理，专注于浏览器与安卓设备。

核心特点：动态环境中的任务决策与执行能力。





## 主要功能

### 💡 通用问答 —— 即时解答您的疑惑

唤起清言侧边栏，向清言提问，无需切换页面，让问题解答更加流畅。

### 📖 页面总结 —— 快速把握内容要点

清言能够智能总结网页内容，为您提供关键信息的概览，节省阅读时间，提升信息获取效率。

### 💬 页面对话 —— 互动式信息探索

通过清言侧边栏，与网页内容直接对话，提问并获得即时回答，让信息获取变得更有趣、更高效。

### 📄 划线助手 —— 网页内容的深度解读

选中感兴趣的内容，清言为您提供解释、总结、翻译，轻松掌握网页内容。

### ✍️ 写作助手 —— 文本创作的得力帮手

在浏览器中遇到写作需求？清言写作助手一键生成文章、评论、邮件，让创作变得轻松简单

### ✓ 勾选总结 —— 快速获取关键信息

在浏览网页时，选择您感兴趣的内容，清言为您提供深度总结，节省阅读时间。

### 🔍 高级检索 —— 信息的精准筛选

在知网、知乎、小红书等网站上，清言帮您执行复杂的搜索和筛选任务，快速获取您所需的信息。

# 手机实际使用示例

AUTOGLM 在中国安卓应用“美团”上的真实操作示例：执行指令“点一杯瑞幸咖啡的热椰椰拿铁，半糖”。



Step 1: Launch APP

Step 2: Choose delivery

Step 3: Tap Search Box

Step 4: Type Search Terms

Step 5: Confirm Shops

Step 6: Enter The Shop

Step 7: Enter The Shop



Step 8: Type Search Terms

Step 9: Choose The Item

Step 10: Confirm Flavors

Step 11: Select "Hot"

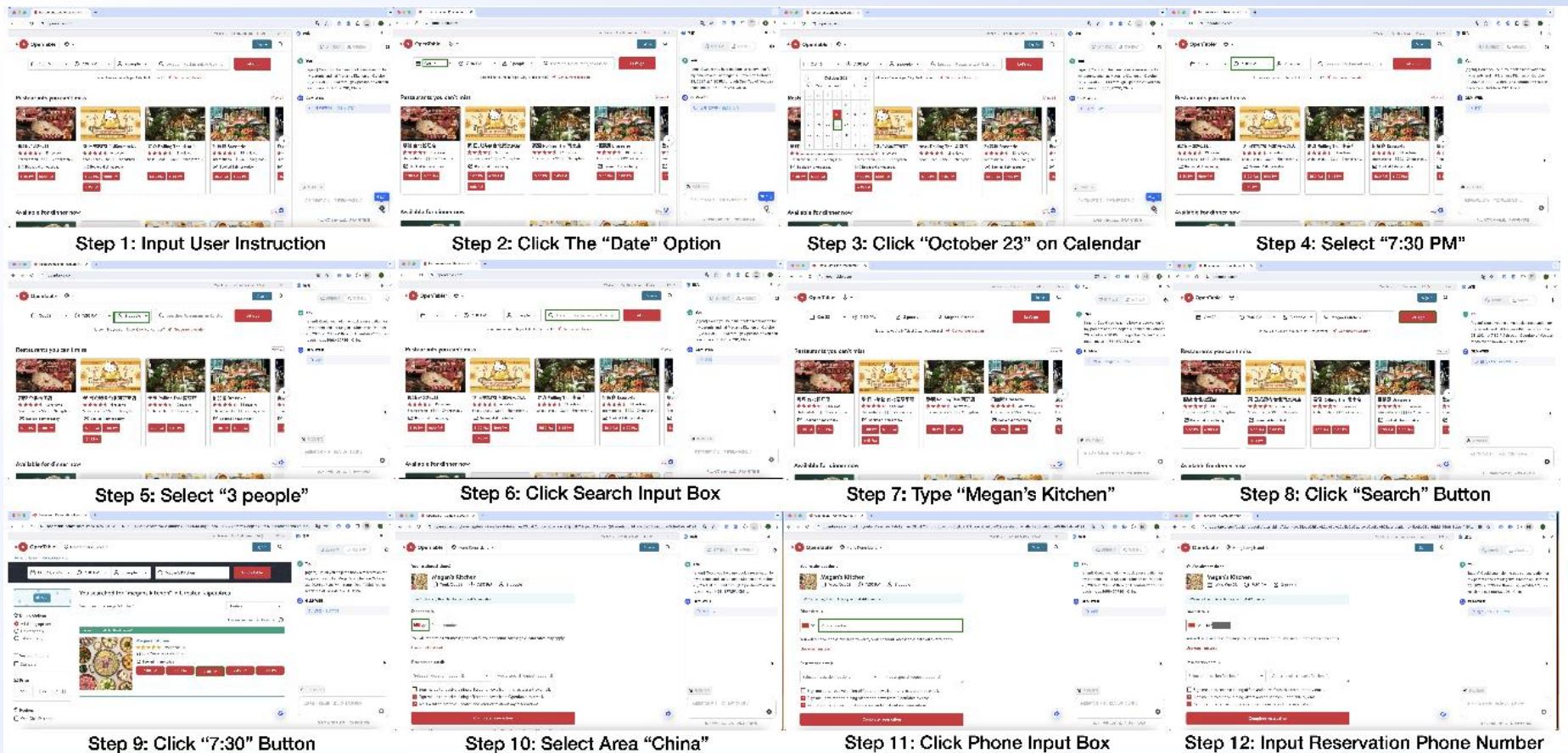
Step 12: Select "Less Sugar"

Step 13: Add To Cart

Step 14: Submit The Order



# 浏览器操作示例





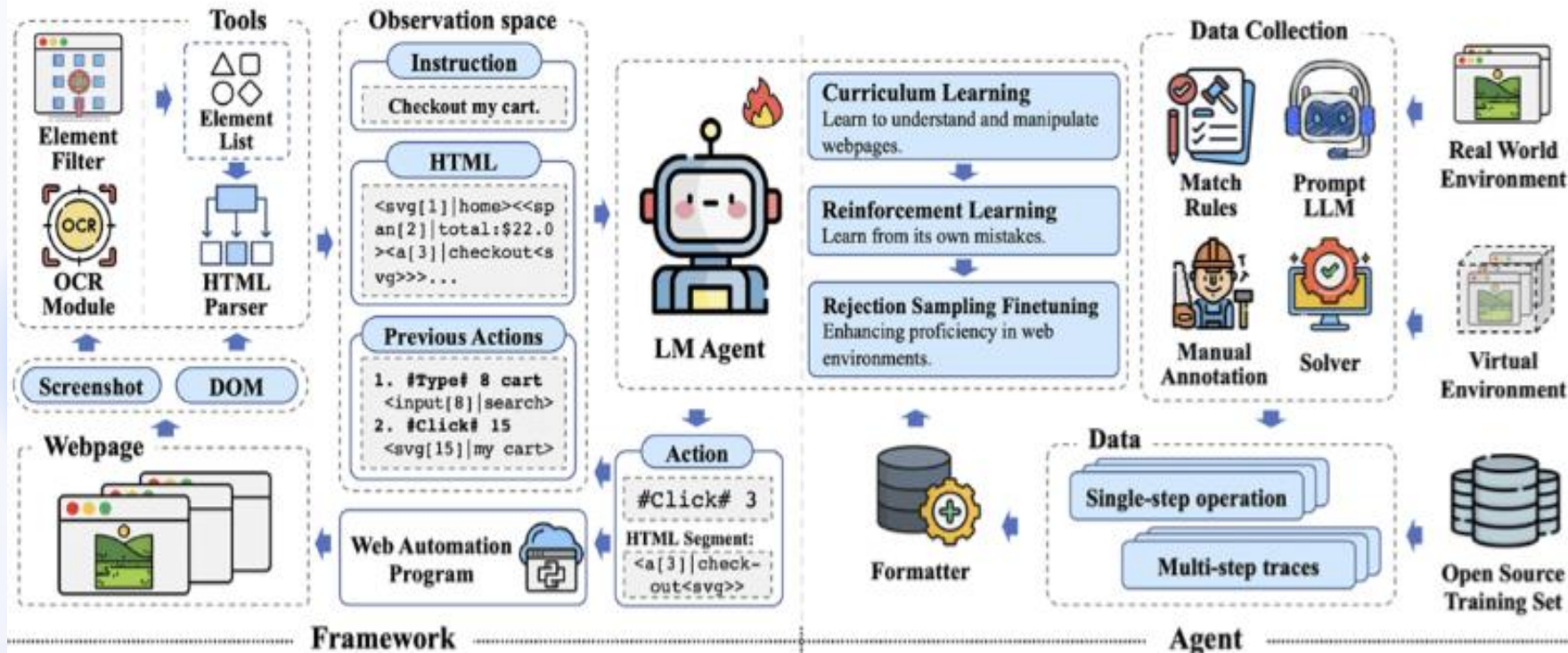


03

## 技术实现



# AutoGLM的核心技术架构





# 1.中间接口设计

## 规划和落地的分离

规划：生成任务意图（如“点击提交按钮”）。

落地：解析并找到对应元素的具体位置。

成效：显著提高了任务执行的准确性。

Table 1: Experiments on Intermediate Interface Design on VAB-WebArena-Lite [47; 23].

Observation Type	gpt-4o (text)	gpt-4o (visual)	gpt-4-vision-preview (visual)
End-to-End Agent	14.3%	18.2%	18.8%
Intermediate Interface Design	18.1% (+3.8%)	27.3% (+9.1%)	36.4% (+17.6%)

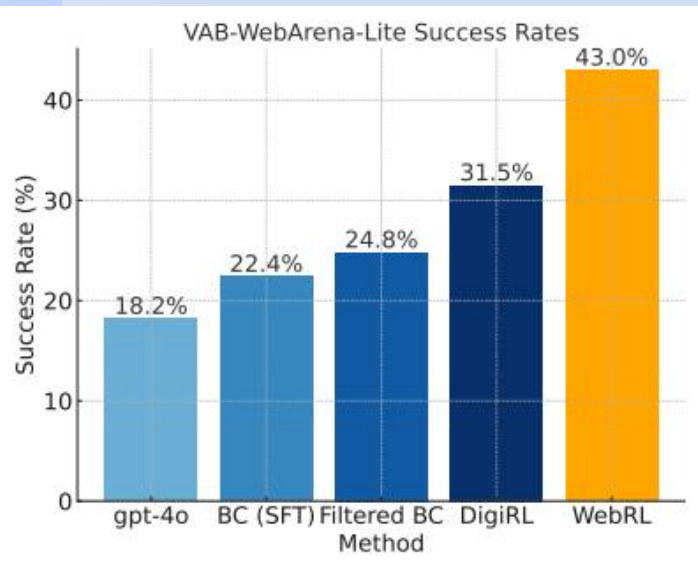


## 2.自我进化强化学习

### 在线强化学习框架

动态增强数据：生成多样化任务指令。

策略优化：解决任务复杂度递进与分布漂移问题。





04

## 实际任务操作



[https://xiao9905.github.io/AutoGLM/?utm\\_source=ai-bot.cn](https://xiao9905.github.io/AutoGLM/?utm_source=ai-bot.cn)







# 智能Agent能力

01

## 帮你筛-站内高级检索

只需要输入您的问题，讲根据您的指令识别出您的检索目的和筛选条件，清言将模拟人类动作，帮您再网站内输入问题，点击筛选条件，筛选相关信息，并打开前五条检索结果，并提供精确答案。

02

## 帮你看-量子速读

在任何网站上，选择您感兴趣的内容，清言插件能够一键生成深度总结，节省您宝贵的时间。  
如有特定需求，您可以在勾选后自定义提问，确保获取的信息更加精准匹配您的期望。

03

## 帮你写

当您在使用浏览器碰到需要写文章、写邮件、写周报等写作场景，清言写作助手可以根据您的需求，快速生成文本，一键插入输入框，省时省力。

<https://zhipu-ai.feishu.cn/wiki/ICTHwxNtjiol3RkvAFrc7mRYnVK>



05

## 学习启示





# AutoGLM对产品实现的启发

## 1. 技术实现：

任务分解（规划与执行分离）。

自适应学习（课程强化学习的引入）。

## 2. 产品设计：

用户体验为中心：界面友好，任务透明。

数据驱动优化机制：通过用户任务反馈持续改进模型。

## 3. 部署与应用：

小范围测试逐步扩展（如浏览器插件和安卓测试版）。

The background features several overlapping circles in various shades of blue and white, creating a soft, abstract pattern. The circles vary in size and opacity, with some appearing more prominent than others. The overall color palette is light and airy, with the blue circles providing a subtle contrast against the white and light blue background.

谢谢观看