

---

# 專題研究

## 流程圖呈現方式對 LLM 流程檢查 的準確性探究

---

---

# 目次

03 LLM

---

04 研究動機

---

06 參考文獻

---

07 研究方法

---

09 PROMPT (提示詞)

---

10 資料說明

---

11 研究結果展示

---

12 結論

---

13 未來展望

---

# LLM(Large Language Model大型語言模型)

類似文字接龍

- 簡化舉例： $Y = w_1x_1 + w_2x_2$

$x_1$  是 "我"， $x_2$  是 "是"

然後要猜接下來的字  $Y$  是什麼

# 研究動機

提出問題：

LLM 擁有如此多的感官，那是否能讀懂流程？

對 LLM 的各種感官感興趣，先前 GPT-4o 釋出，讓 LLM 不只能讀文字 / 圖片，還具備了良好的聽覺感官。

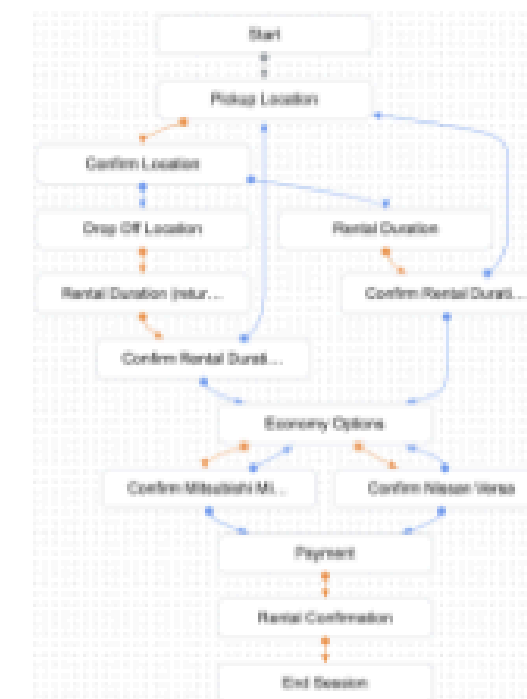
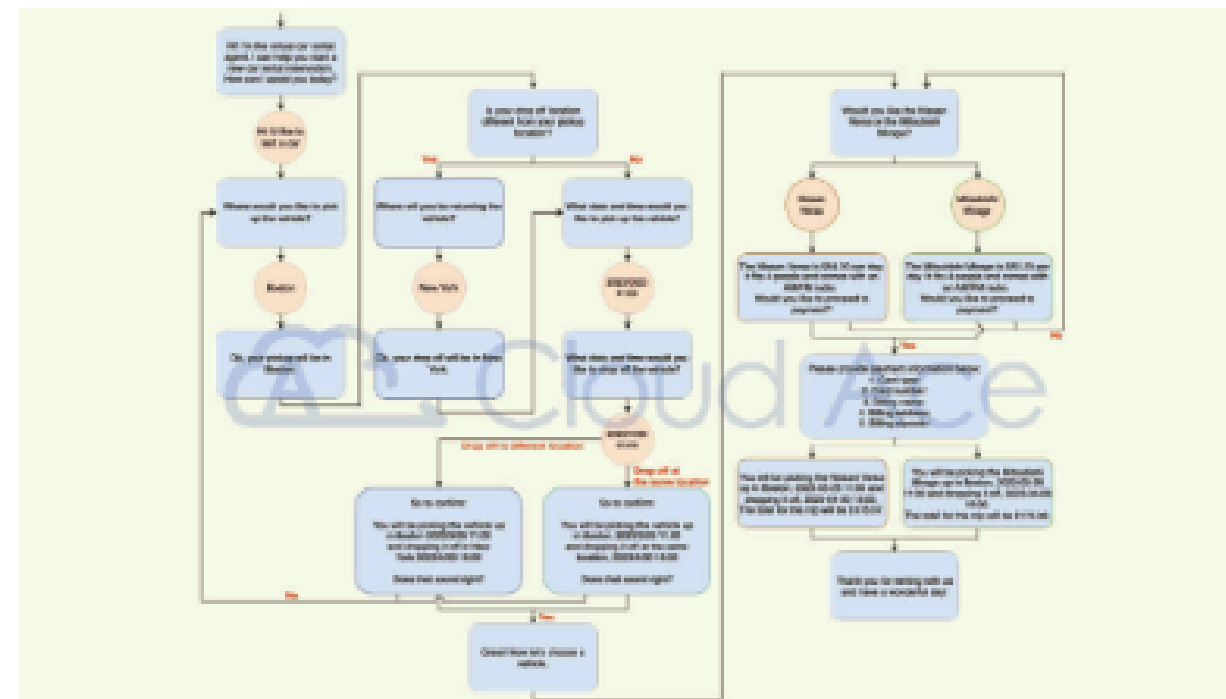
但根據我在業界的一些經驗，尤其是在程曦（台灣客服外包龍頭公司）兼職軟體開發，主要做生成式 AI 相關應用的研究與開發的經驗而言，LLM 對於流程的處理仍然有些薄弱。

自己用 LLM 嘗試過 SalesGPT 的做法，也有使用過 Google 釋出的 Agent Builder x DialogflowCX，所以我想利用這個想法做一個相關研究，研究如何讓 LLM 更好的處理複雜流程。探索要用什麼樣的流程輸入形式，才能確保 LLM 最好的了解流程內容，並遵守流程做出對應需求。

# 研究動機

# 業界實際應用

1. 客服人員通話流程品質檢查
2. 智能流程機器人可以更好的依照流程對話

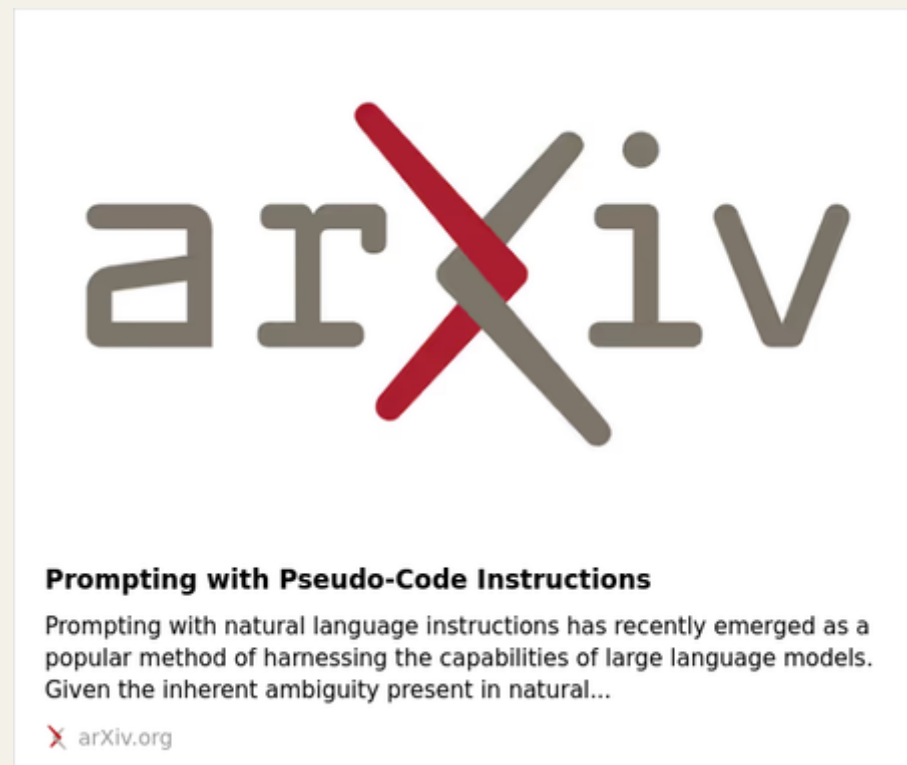


預約租車對話流程設計圖（左）；Dialogflow CX Console 對話流程設計圖（右）

截圖自：[Google Cloud Dialogflow CX Console](#) 頁面

# 參考文獻

IBM :  
Prompting with Pseudo-Code  
Instructions



Code Prompting Elicits Conditional  
Reasoning Abilities in Text+Code LLMs



# 研究方法

我們有大量的客戶進線通話紀錄 (目前先模仿一份通話紀錄，再根據標準答案想呈現的流程做細微修改)，這些紀錄的通話主題主要為電商客服人員處理發貨/退貨。並有對應的發貨/退貨客服人員對答流程圖，我們要讓 LLM 檢查對話的客服人員是否依照流程進行。

具體操作為：

1. LLM 回復在哪裡的時候客服人員沒有照流程走，或是完全遵循流程。
2. 回復格式例如：4, 3.1, Pass，Pass 表示完全遵循流程，數字表示客服人員遵守到的流程節點。

## 五種流程圖比較

我們將比較五種不同的流程圖呈現方式給 LLM，觀察其回答的準確率：

1. 把流程圖換成虛擬碼 (例: if ... else if ... else) --> LLM 強大的生成/理解程式碼能力，對 if else 應有良好處理
2. 把流程圖轉成專門製圖的 Mermaid 語法 --> 經典的製圖語法，LLM 生成圖表多是使用此語法做轉換
3. 自然語言描述流程圖 (例: 首先，客服人員要 ...，然後要 ...，他這邊有可能可以選擇 ... 或是選擇 ...)
4. 使用可以識圖的 LLM (例: GPT-4o)，直接呈現流程圖片給 LLM
5. 將流程寫成 Excel 表 (第一層流程會在第一 Column，第二層流程會在第二 Column，一層流程可能會有分支，所以在不同 row 就是不同分支)



# 研究方法

因為這次是 Demo 一個初版的關係，故先利用單筆通話紀錄擴充製作 6 筆後做初次試驗

流程圖比較僅先對前兩種做比較

1. 把流程圖換成虛擬碼 (例: if ... else if ... else)
  - > LLM 強大的生成/理解程式碼能力，對 if else 應有良好處理
2. 把流程圖轉成專門製圖的 Mermaid 語法
  - > 經典的製圖語法，LLM 生成圖表多是使用此語法做轉換



# Prompt (提示詞)

```
def call_aied(chart, mess):
```

```
    prompt = f"""
```

```
    以下在一段python流程偽代碼，用於處理發貨問題
```

```
    {chart}
```

```
    以下是客服員工和客戶的對話，請仔細分析，並以JSON格式輸出偽代碼流程中的過程變量及變量值，最後輸出流程中的返回值
```

```
    {mess}
```

```
    """
```

```
    prompt_mer = f"""
```

```
    以下在一段 流程 Mermaid code，用於處理發貨問題
```

```
    {chart}
```

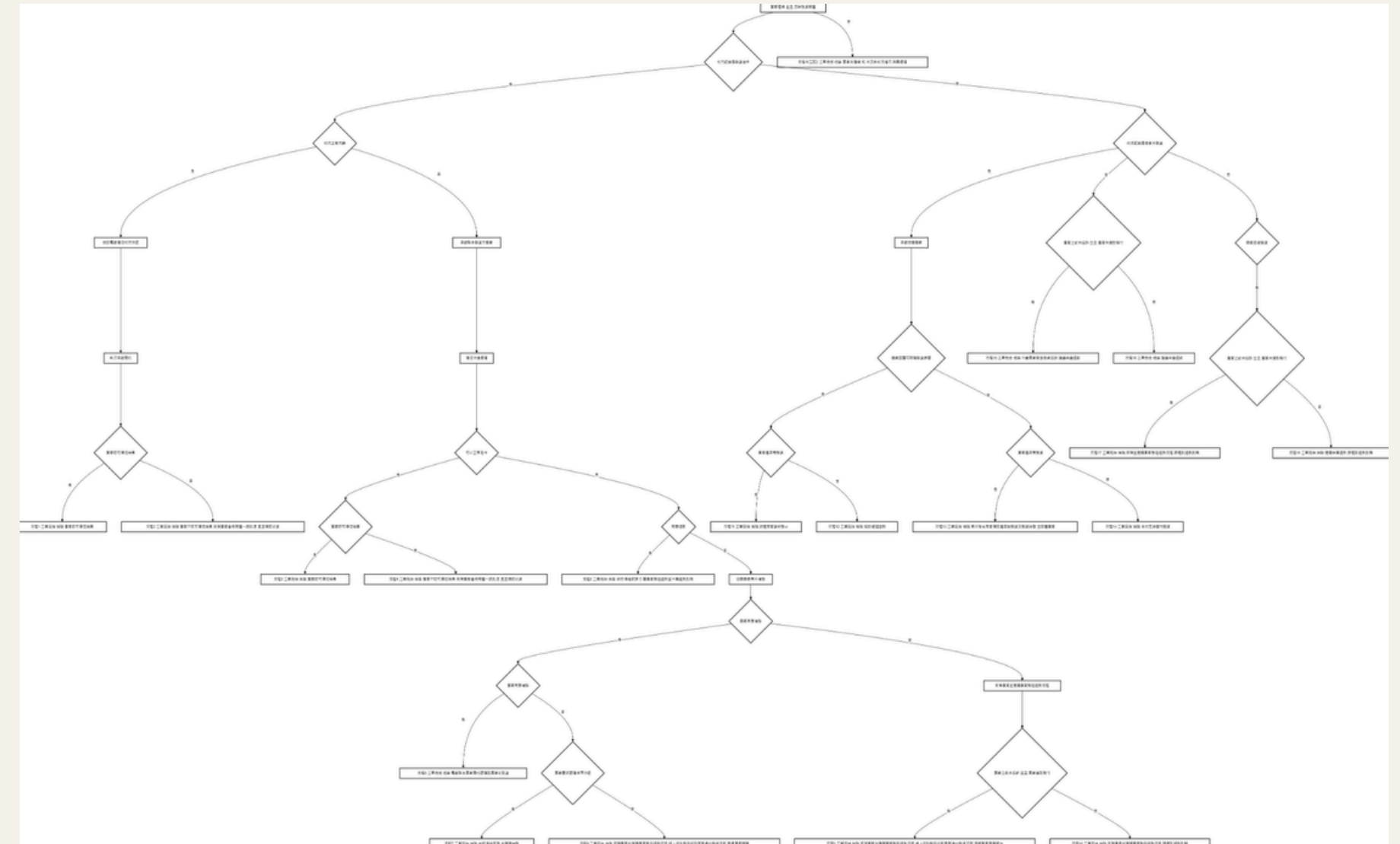
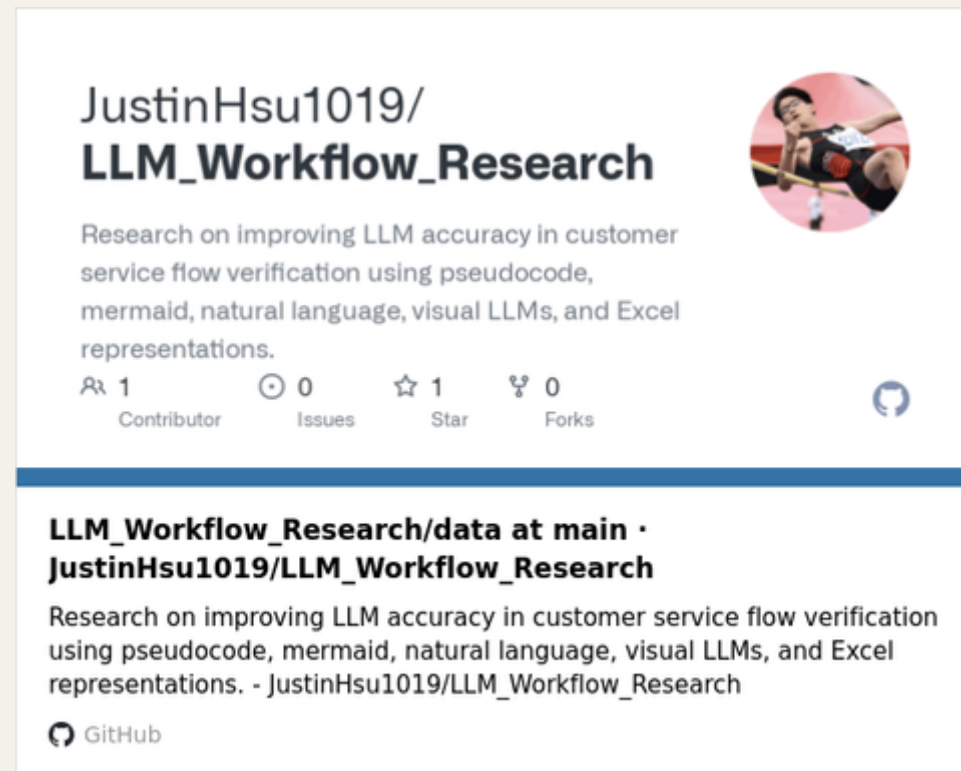
```
    以下是客服員工和客戶的對話，請仔細分析，並以JSON格式輸出偽代碼流程中的過程變量及變量值，最後輸出流程中的返回值
```

```
    {mess}
```

```
    """
```

# 資料說明

- 客戶進線通話紀錄 (發貨/退貨)
- 發貨/退貨標準流程圖



# 研究結果展示

各自分析結果 (詳細見 GitHub)

C	D
虛擬碼_分析結果	Mermaid_分析結果
{	{
"過程變量": {	"過程變量": {
"買家進線": true,	"process_variables": {
"反映發貨問題": true,	"買家進線": true,
{	{
"對話變量": {	"變量": {
"客戶訴求": "收貨",	"買家進線": true,
{	{
"買家進線": true,	"過程變量": {
"反映發貨問題": true,	"買家進線": true,
{	{
"買家進線": true,	"variables": {
"反映發貨問題": true,	"買家進線": true,
{	{
"買家進線": true,	"variables": {
"反映發貨問題": true,	"買家進線": true,

六筆通話紀錄的正確性檢查

F	G	H
正確答案	虛擬碼_檢查	Mermaid_檢查
反映發貨問題		
發貨途中	100	100
反映發貨問題		
發貨途中	100	100
發貨途中		
物流正常流轉	100	100
商家未發貨		
商家回覆可明確發貨時間	100	100
商家未發貨		
商家未明確發貨時間   聯系	100	100
反映發貨問題		
商家拒絕發貨	100	100

# 結論

可以看出 LLM 對虛擬碼和 Mermaid 語法在 LLM 對流程圖方面的辨識效果是不錯的

目前因樣本數稀少，故得到的結果僅得知六筆通話紀錄在兩者間都是 100% 的正確率，不過這次僅是一個初版 Demo，後續會擴大樣本數及流程圖呈現方式做正式的試驗。

# 未來展望

## 1. 完整的完成簡報第七頁的整體研究計畫

- 利用 LLM 擴充通話紀錄，針對不同的流程遵守都要有足夠的通話紀錄樣本數做試驗
- 將流程圖轉換成 5 種不同的表達形式，目前僅針對我認為正確率/表現結果會最好的虛擬碼和 Mermaid 做測試，但後續還須加入 自然語言 / 直接識圖 / Excel 3 種流程表示法。