



# 悟空法律通

基於生成式 AI 的法律諮詢系統

# 目錄

摘要.....	3
第一章 緒論.....	3
1.1 研究背景.....	3
1.2 研究動機.....	3
1.3 研究目的.....	4
第二章 系統架構與設計.....	4
2.1 系統架構概述.....	4
2.2 前端系統設計.....	4
2.3 後端系統設計.....	4
2.4 系統流程說明.....	4
第三章 模型微調與角色語氣設計.....	5
3.1 模型微調之必要性.....	5
3.2 孫悟空角色語氣設計.....	5
3.3 模型訓練資料建構方式.....	5
3.4 微調後模型成果展示.....	6
第四章 法律資料檢索與 RAG 架構實作.....	7
4.1 RAG 架構概念.....	7
4.2 法律資料處理方式.....	7
4.3 ChromaDB 向量資料庫應用.....	7
4.4 法律條文資料來源與格式說明.....	7
4.5 RAG 檢索與生成展示.....	8
第五章 系統功能展示與應用情境.....	9
5.1 系統主要功能.....	9
5.2 使用者前端介面展示說明.....	10
第六章 系統限制與未來改進方向.....	10
6.1 系統限制.....	10
6.2 未來改進方向.....	10
第七章 結論.....	10
第八章 相關程式碼.....	11

# 摘要

本研究實作一套名為「悟空法律通」之生成式 AI 法律諮詢系統，結合檢索增強生成（Retrieval-Augmented Generation, RAG）架構，針對台灣法律資料進行檢索與回應。此外，透過模型微調（Fine-tuning），使系統回應風格模仿孫悟空之口語語氣，以提升使用者互動體驗。系統可協助使用者快速理解法律條文與相關判例，降低法律資訊取得之門檻。

## 第一章 緒論

### 1.1 研究背景

隨著網際網路與數位科技的發展，民眾取得資訊的管道日益多元，然而法律相關資訊仍具有高度專業性與複雜性。一般使用者在查閱法律條文或判例時，往往因用語艱深、結構嚴謹而難以理解其實際意涵，進而影響其對自身權益的判斷與保障。近年來，生成式人工智慧（Generative AI）技術迅速發展，大型語言模型（Large Language Models, LLMs）已能以自然語言回應使用者問題，為法律資訊輔助提供新的可能性。

然而，若僅依賴語言模型本身進行法律問答，模型可能因訓練資料限制或上下文不足而產生不正確或過時的內容，造成所謂「幻覺（Hallucination）」問題。特別是在法律領域，錯誤資訊可能導致使用者作出不當判斷，因此如何提升生成內容的正確性與可信度，成為一項重要課題。

### 1.2 研究動機

基於上述背景，本研究結合檢索增強生成（Retrieval-Augmented Generation, RAG）架構，使生成式 AI 能在回應問題前，先即時檢索實際法律資料作為依據，以降低錯誤生成之風險。此外，考量到法律資訊對一般使用者而言較為生硬，本研究亦嘗試透過模型微調（Fine-tuning），賦予系統特定角色語氣，以提升互動性與使用體驗。

因此，本研究實作一套名為「悟空法律通」之法律諮詢系統，透過孫悟空之角色語氣，將艱澀的法律概念轉化為較口語且易理解的說明方式，使使用者在獲取正確法律資訊的同時，也能提升使用意願。

## 1.3 研究目的

本研究之主要目的如下：

- （一）建構一套結合 RAG 架構之法律問答系統，以提升法律資訊回應的正確性。
- （二）整合前後端 Web 技術，實作可實際操作之法律諮詢平台。
- （三）透過模型微調，使系統具備特定角色語氣，提升使用者互動體驗。
- （四）分析系統限制並提出未來改進方向。

## 第二章 系統架構與設計

### 2.1 系統架構概述

本系統採用前後端分離（Frontend-Backend Separation）之架構設計，主要可分為前端介面、後端服務與語言模型三個部分。使用者透過網頁前端輸入法律相關問題，請求經由後端 API 處理後，結合檢索到的法律資料與語言模型生成回答，最終回傳結果並顯示於前端介面。

### 2.2 前端系統設計

前端部分以 Web 技術實作，採用 Next.js 搭配 React 作為主要框架。React 具備元件化（Component-based）設計特性，有助於提升程式碼之可維護性與重用性；Next.js 則提供頁面路由與效能最佳化機制，使整體使用體驗更加流暢。前端主要功能包括使用者問題輸入、回應結果顯示，以及與後端 API 之資料傳輸。

### 2.3 後端系統設計

後端系統以 FastAPI 作為主要開發框架，負責處理前端請求、控制系統流程以及整合各模組功能。FastAPI 具備高效能與非同步（Asynchronous）處理能力，適合用於生成式 AI 應用。系統透過 Uvicorn 作為 ASGI 應用伺服器進行部署，以提升請求處理效率與回應速度。

### 2.4 系統流程說明

整體系統流程如下：使用者於前端輸入法律問題後，請求將送至後端 FastAPI 服務；後端接收請求後，先進行法律資料檢索，取得相關條文內容，再將檢索結果與使用者問題一併送入語言模型生成回應，最後將結果回傳至前端顯示。

## 第三章 模型微調與角色語氣設計

### 3.1 模型微調之必要性

原始大型語言模型在回應問題時，多以制式且中立之語氣呈現，雖具備一定專業性，但對一般使用者而言互動性較低。因此，本研究透過模型微調方式，使模型在維持法律內容正確性的前提下，具備特定角色語氣，以提升使用者接受度。

### 3.2 孫悟空角色語氣設計

本研究選擇「孫悟空」作為系統回應角色，原因在於其形象鮮明、語氣活潑，能有效拉近與使用者之距離。透過蒐集具代表性之語句作為訓練資料進行微調，使模型在回答法律問題時，能以較口語化且具角色特色之方式進行說明，同時保留法律內容之嚴謹性。

### 3.3 模型訓練資料建構方式

在模型微調資料建構方面，本研究先蒐集常見之法律相關問題與標準說明內容，再透過生成式 AI 將原始中立語氣之法律說明，轉換為符合孫悟空說話風格之回應文本。此方式可在保留原有法律知識正確性的前提下，建立具角色特色之訓練資料集。

```
{  
  "prompt": "債務人欠錢不還，債權人拿到法院強制執行名義後，查不到財產可執行，債權人還能怎麼辦？",  
  "completion": "嘿嘿，這就像孫悟空大鬧天宮一樣，債務人雖然想躲避不還錢，但法律不會讓他輕鬆逃脫！如果債務人沒有財產，我們只能等，但別急！有個法寶叫「拘提與管收」，如果他故意隱瞞或逃跑，法院可以把他抓起來，逼他交代財產在哪兒！《強制執行法》第22條說得很明白，像那些債務人不履行義務、逃跑或隱匿財產的時候，法院可以對他們進行拘提。這樣的做法就像是將金箍棒架在頭上，迫使他們交代財產的去向。如果他不願意誠實交代，法院甚至可以進行管收，也就是把債務人關押，逼他吐出財產的秘密！另外，法院還可以根據《強制執行法》第20條，要求債務人定期報告他有的財產狀況，這樣就能確保不會有遺漏的部分。當然，這個拘提和管收並不是無限期的，根據《強制執行法》第24條，管收的時間最多也不過三個月，要是債務人仍然沒辦法提供財產，也只能釋放他，給他一次重新思考的機會。總之，債務人想逃避是沒門的，法律像金箍棒一樣，總能讓他露出馬腳，最後也只能乖乖地交出財產！如果還是不交，法院還能依照規定再進一步處罰，讓他再無可逃的餘地。嘿嘿，這樣一來，債權人也終能回收自己的錢財！"  
}
```

```
{
  "prompt": "請問被竊盜的物品，被不知情買到贓物者領回，是否可要求返還？或只能向竊賊求償？",
  "completion": "如果被竊盜的物品被那個「不知情」的買家給領回，根據《民法》來說，這位「善意」買家是受法律保護的，因為他根本不知道自己買的是贓物，就像我捉妖捉得心急，結果不小心抓錯了妖，卻還能被賦予一些「保護」。既然他不是故意侵犯原物主的所有權，那原物主不能直接向他要求返還物品。但是！若是那個原物對你有特殊意義，想要把它拿回來，你就得支付給那位「善意」的買家他當時支付的錢，也就是說，必須向他「賠償」，然後才能把物品取回來。這就像是跟妖怪打個「和解」，買回物品再對妖怪來個最終決戰。至於要向竊賊求償，那就麻煩了！你得先找到那個竊賊，然後追究他對你的損害。這過程就像我去追天上的妖怪，這不但費時，還會有不少挑戰，要一點耐性才行。所以簡單來說，直接向善意的買家要回物品不太行，你得先賠償他的購買費用，然後如果真想找竊賊報仇（呃，應該是追求賠償），那可得先忍耐一番。"
}
```

### 3.4 微調後模型成果展示

完成模型微調後，系統在回應法律問題時，能同時呈現正確之法律資訊與明顯之角色語氣特徵。例如，系統在說明法律概念時，會搭配較口語化之用語與角色式開場語句，使回應內容更具親和力。整體而言，微調後模型在語氣一致性與使用者可讀性方面皆有顯著提升。

You: 我偷了你的金箍棒 然後拿你的金箍棒砸死了人 然後再把你的金箍棒扔掉後逃走 請問我犯了哪些滔天大罪呢  
 悟空: 哎呀，這事兒可不小！首先，偷俺老孫的金箍棒就是大麻煩，像你這樣的小妖怪，敢動俺老孫的寶貝？那可是心臟脆弱的東西，一旦被偷，俺老孫可不會手下留情！偷東西就得付出代價，說不定俺老孫要找個機會讓你知道‘偷竊’的痛！而且砸死人後，嘿，真的是‘一棒子決斷’了！你這樣做，簡直就是冷血

You: 我的朋友一直欠錢不還，我該怎麼辦呢  
 悟空: 這簡直是當頭一棒！朋友欠錢不還，這樣的事可不是小事！首先，別急著打人，那樣只會讓情況更糟！你可以找他坐下聊聊，問清楚他為何欠款並約定好繳款時間。如果他還是不理會，你就得拿出法律來，要求他履行義務。根據《民法》第344條，貸款方有權要求借款人償還本金和利息，如果他仍然沒有履行，就得依法強制執行。這樣的處理方式，既保護了自己的權益，也能保持友誼。記住，處理問題時要冷靜，別讓情感控制你的決定。別讓對方知道你心裡

You: 請問送出去的東西，還能再要回來嗎  
 悟空: 嘿嘿，這就像俺老孫把金箍棒送給了唐僧，結果他可不會說‘俺老孫想拿回來’！如果你送出去的東西沒有約定可以退換，那就麻煩了！但如果有協議，可以反悔！比如說，買賣合同有條款允許撤銷，那就行！總之，送出去的東西，得看情況，有時候可以取回，有時候卻無法再拿回來。所以，送東西前，先了解清楚規則，免得後來亂來！就像俺老孫說的，‘一意孤行’可不是好事！所以，做事要冷靜，別急著送東西，不

## 第四章 法律資料檢索與 RAG 架構實作

### 4.1 RAG 架構概念

檢索增強生成（RAG）是一種結合資訊檢索與生成式模型的方法，其核心概念為在生成回答前，先從外部資料庫中取得相關文件作為上下文，再交由語言模型進行回應生成。此方法可有效降低模型僅憑內部知識生成錯誤內容的風險，特別適合應用於法律等高精確度需求之領域。

### 4.2 法律資料處理方式

本研究蒐集相關法律條文作為資料來源，並將法律文本進行切分（Chunking），以避免單筆資料過長影響檢索效果。切分後之文本再透過嵌入模型轉換為向量表示（Embedding），作為後續相似度搜尋之基礎。

### 4.3 ChromaDB 向量資料庫應用

本系統使用 ChromaDB 作為向量資料庫，負責儲存法律文件向量與其對應內容。當使用者提出查詢時，系統會將問題轉換為向量，並透過相似度搜尋機制，找出最相關之法律條文，作為生成式模型之上下文資訊，形成完整的 RAG 流程。

### 4.4 法律條文資料來源與格式說明

本研究之法律條文資料主要來源為「[全國法規資料庫](#)」，其為我國官方提供之法規公開平台，具備資料正確性。系統所使用之法律資料包含法條名稱、條號及條文內容，並以結構化文字格式進行整理與儲存。



```

1 {
2   "UpdateDate": "2025/12/26 上午 12:00:00",
3   "Laws": [
4     {
5       "LawLevel": "憲法",
6       "LawName": "中華民國憲法",
7       "LawURL": "https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=A00000001",
8       "LawCategory": "憲法",
9       "LawModifiedDate": "19470101",
10      "LawEffectiveDate": "",
11      "LawEffectiveNote": "",
12      "LawAbandonNote": "",
13      "LawHasEngVersion": "Y",
14      "EngLawName": "Constitution of the Republic of China (Taiwan)",
15      "LawAttachments": [],
16      "LawHistories": "1. 中華民國三十六年一月一日國民政府令公布\r\n  中華民國三十六年十二月二十五日施行\r\n  中華民國三十五年十二月二十五日國民大會通過",
17      "LawForeword": "中華民國國民大會受全體國民之付託，依據孫中山先生創立中華民國之遺教，為鞏固國權，保障民權，奠定社會安寧，增進人民福利，制定本憲法\r\n，頒行全國，永久咸遵。",
18      "LawArticles": [
19        {
20          "ArticleType": "C",
21          "ArticleNo": "",
22          "ArticleContent": "第一章 總綱"
23        },
24        {
25          "ArticleType": "A",
26          "ArticleNo": "第 1 條",
27          "ArticleContent": "中華民國基於三民主義，為民有民治民享之民主共和國。"
28        }
29      ]
30    }
31  ]
32 }

```

## 4.5 RAG 檢索與生成展示

系統會根據搜尋結果選取最相關之法律條文內容，作為上下文資訊提供給語言模型。在生成回應時，語言模型不僅會參考使用者問題，亦會依據檢索到的法律資料進行說明，確保回應內容具備實際法條依據。透過此方式，系統能有效降低生成內容偏離法律原意之風險。

You: 人民對國家政策、公共利害或其權益之維護，得向誰請願呢

悟空：哈哈，這就像俺老孫要挑戰金箍棒的禁令，得先向上級來個「請願」！根據請願法第2條，人民可以向職權所屬的民意機構（如立法委員會）或主管的行政機構請願，這樣才能影響決策，讓自己的聲音被聽見！這可是法律給我們的寶貴權利，一定要好好保護！所以啊，無論是大事還是小事，都別忘了找適當的窗口來發表你的意見哦！（來源：請願法 第2條）  
 仙女：嘿，不過這種情況，還得看情況啦！如果是對國家政策、公共利益或自己權益感興趣的話，直接去找有權決定的事體，如立法院或者行政機構的首長，提案或者請願，這樣才更有效哦！不過，別以為這樣就能隨便改變什麼，還得看法律和制度的限制，不能說隨便發揮勁兒就行！（來源：中華民國憲法 第16條，請願法 第2條）  
 俺老孫：哎呀，這可不只是個普通的請願喔！根據社會福利基本法第29條，人民的福利權利一旦受到侵害，得找辦法解決，這就像是俺老孫被困在妖怪洞穴裡，得找個門票才行！

🔍 檢索到的來源：

1. 請願法 第 2 條 | 行政>院本部>通用目 | 法規位階：法律 | <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=A0030019>
2. 中華民國憲法 第 16 條 | 憲法 | 法規位階：憲法 | <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=A00000001>
3. 社會福利基本法 第 29 條 | 行政>衛生福利部>社會及家庭目 | 法規位階：法律 | <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=D0050213>



You: 我的朋友一直欠錢不還，我該怎麼辦呢

悟空: 哈哈，這就像妖怪一樣，總是欠著人！依照《破產法》第 1 條，如果他真的是無力清償，那就得進行和解或破產程序，讓大家都有個說話的權利。但如果他只是拖拉，那就告訴他趕快履行吧，根據《民法》第 254 條，可以要求他履行合同，如果他還是不行，就可以解除合同。至於逾期貸款的部分，就按照《票券金融公司資產評估損失準備提列及逾期授信催收款呆帳處理辦法》第 10 條來辦，不然要繼續催促他，也能保留未來的利息收入。總之，別讓他把事情弄得太複雜，直接處理清楚就好啦！（《破產法》第 1 條、《民法》第 254 條、《票券金融公司資產評估損失準備提列及逾期授信催收款呆帳處理辦法》第 10 條）嘿嘿，這事兒不用怕，依法處理就對了！別讓妖怪耍小聰明，趕緊把事情料理清楚。記住，法律就是幫你打妖怪的金箍棒，隨時準備好了！（《破產法》第 1 條、《民法》第 254 條、《票券金融公司資產評估

● 检索到的来源：

1. 破產法 第 1 條 | 司法>院本部>民事目 | 法規位階：法律 | <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=B0010006>
2. 民法 第 254 條 | 行政>法務部>法律事務目 | 法規位階：法律 | <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=B0000001>
3. 票券金融公司資產評估損失準備提列及逾期授信催收款呆帳處理辦法 第 10 條 | 行政>金融監督管理委員會>銀行目 | 法規位階：命令 | <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=G0380161>

You: 請問送出去的東西，還能再要回來嗎？

悟空: 嘿嘿，這就像把金箍棒借給八戒一樣！法律上有個‘寄託物’的概念，如果你把東西借或送給別人，你可以隨時拿回，但得通知對方（民法第597條）。就像如果我借了金箍棒，隨時都能拿回來，只是得告訴八戒。至於送的快遞，運送員得在你抵達時立馬交回來（民法第655條），這就像快遞員說：“哎，師父到了，俺老孫的金箍棒給你拿回來！”如果運送員遲延，那你就可就不行了，得等到運送物達到後才拿回。總結一下，這就像是‘借’東西，隨時拿回來，只是得跟對方商量；送的快遞則得等到目的地，運送員也得按時交回。這就是法律對這些事物的規範，沒什麼大事，隨便遵守就好！（民法第597條、第655條）最終結果：可以拿回來，但得合適的時間與通知。送的快遞，得等到目的地才拿回。【悟空的金箍棒，隨時都能拿回來，送的快遞則得等到目的地。】悟空：哈哈，這事兒不太複雜，拿金箍棒隨便拿，送的

● 检索到的来源：

1. 民法 第 597 條 | 行政>法務部>法律事務目 | 法規位階：法律 | <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=B0000001>
2. 民法 第 655 條 | 行政>法務部>法律事務目 | 法規位階：法律 | <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=B0000001>
3. 民法 第 630 條 | 行政>法務部>法律事務目 | 法規位階：法律 | <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=B0000001>

整體而言，RAG 架構使「悟空法律通」在回應法律問題時，能同時兼顧資訊正確性與說明完整性，並作為後續模型語氣微調與前端展示之基礎。

## 第五章 系統功能展示與應用情境

### 5.1 系統主要功能

「悟空法律通」可提供使用者即時法律問答服務，系統會根據使用者輸入問題，自法律資料庫中檢索相關條文，並生成對應之說明內容。此外，系統回應皆以孫悟空語氣呈現，使法律概念更易於理解。

## 5.2 使用者前端介面展示說明

使用者可透過瀏覽器進入本系統之 Web 前端介面，於輸入欄位中輸入欲查詢之法律問題。系統回應結果將即時顯示於畫面中，包含以孫悟空語氣呈現之法律說明內容。前端介面設計以簡潔與易用為原則，使使用者能專注於問題本身與回應結果，提升整體操作體驗。



## 第六章 系統限制與未來改進方向

### 6.1 系統限制

本研究之系統仍存在若干限制，例如在面對高度複雜或需專業判斷之法律情境時，系統仍可能無法提供完整解答等。

### 6.2 未來改進方向

未來可擴充法律資料來源，並加入引用來源標示功能，以提升回應之可信度。此外，亦可進一步優化角色語氣與使用者介面設計，提升整體使用體驗。

## 第七章 結論

本研究成功實作一套結合 RAG 架構與角色語氣微調之法律諮詢系統「悟空法律通」，並整合前後端 Web 技術，提供可實際操作之應用平台。透過即時法律資料檢索與生成式模型回應，本系統能有效降低錯誤生成風險，並以親和方式呈現法律資訊。整體而

言，本研究展示生成式 AI 應用於法律領域之可行性，亦可作為未來相關研究與實務應用之參考。

## 第八章 相關程式碼

模型層：[GitHub - HaoXun97/genai-model](https://github.com/HaoXun97/genai-model)

前端：[GitHub - Evan56963/genai\\_frontend](https://github.com/Evan56963/genai_frontend)

後端：[GitHub - Evan56963/genai-backend](https://github.com/Evan56963/genai-backend)