介紹時提到本研究的重要性(傳統機器學習方式良率低...等)，在論文前段先針對中文wh疑問詞用語言學知識觀察，整理不同類型的語料，統整每種類型的語言學規則，在論文後段則將這些語言學規則作為實驗方法程式化，讓蒐集到的語料跑過這些程式模型，以證實中文wh疑問詞的語言學規則，並且將此次實驗和機器學習關於中文wh疑問詞的良率做比較。

你說的這些步驟與內容的確就是論文內該有的內容，對於你的論文我有幾個目標，你參考一下：

1. 焦點放在三種語意之一即可。

2. 時間不足的話，聚焦在「誰」或「什麼」之一即可。

3. 與機器學習的比較其實在我的文章內已經講蠻多的了，妳可以繼續延伸至 ChatGPT 3.5/4 上。

4. 文獻回顧除了wh詞本身的語言學文獻，NLP方面你可以先讀附檔中的文章，遇到不懂的不用太執著。

5. 務必要把處理的程序細節講清楚，這一點就是要你在現階段進入計畫的主因，從頭開始記錄自己從「搜集/整理語料」、「在Loki上互動微調」、「在Python中編碼中編碼」的操作細節，詳實記錄這些步驟細節與know-how會是很重要的貢獻，講得越仔細越好。

探勘整理語料、Articut/Loki互動以及閱讀文獻三個項目的處理過程，我都建議你建立做筆記的習慣，因為這三項筆記都會直接變成你碩論的材料，因此平時筆記詳實，到時候碩論內容就很容易生出來了。

第一，你的碩論目標是 (至少) 在 LOKI 上建立出一個 universal wh checker 歐！也就是說，使用者可以丟一個含有「誰」/「什麼」的任意中文句子進去，universal wh checker 可以告訴使用者這個句子的「誰」/「什麼」是否具有全稱語意。

第二，這個研究計畫的願景是要把上面的 universal wh checker 延展至 existential wh checker 以及 interrogative wh checker。

有了這三個 checkers，那麼使用者就可以丟任意含有「誰」/「什麼」的任意中文句子進去，程式會告訴使用者這個句子裡面的「誰」/「什麼」屬於何種語意，如下圖：



例如我們可以餵機器這個句子：「老王什麼都吃」

三個 checkers 整合的程式回應：

universal wh checker: True

existential wh checker: False

interrogative wh checker: False

如此我們可以藉由這三個checkers整合的一套判斷程式，朝向自然語言理解，建議你讀一下附件pp.34-38。

我在開學前有寄過這份檔案給你，現在你列為共同作者之一，下學期的語言與人類心智課程接近期末時，我會請你與我一起報告你在這個計畫中做的處理細節，當作你碩論資格考報告的練習，視情況在明年暑假開始前完成資格考口試。