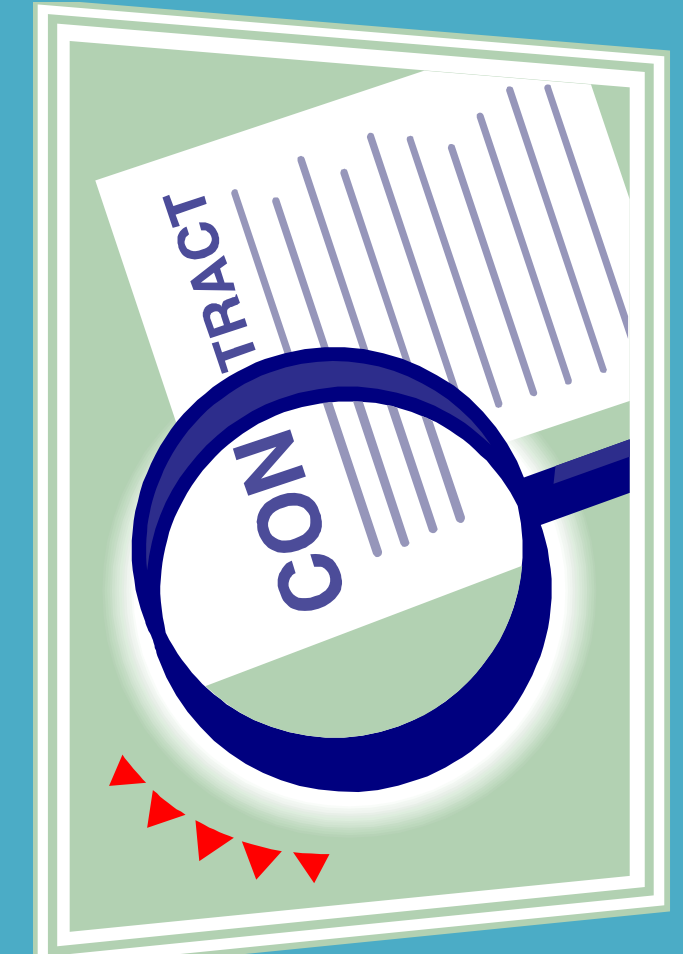




以過去閱讀經驗為基礎之文件重點提示系統開發：以新聞閱讀為例

學生姓名：劉松鑫、楊皓文、蘇詠盛 指導教授：侯建良 博士

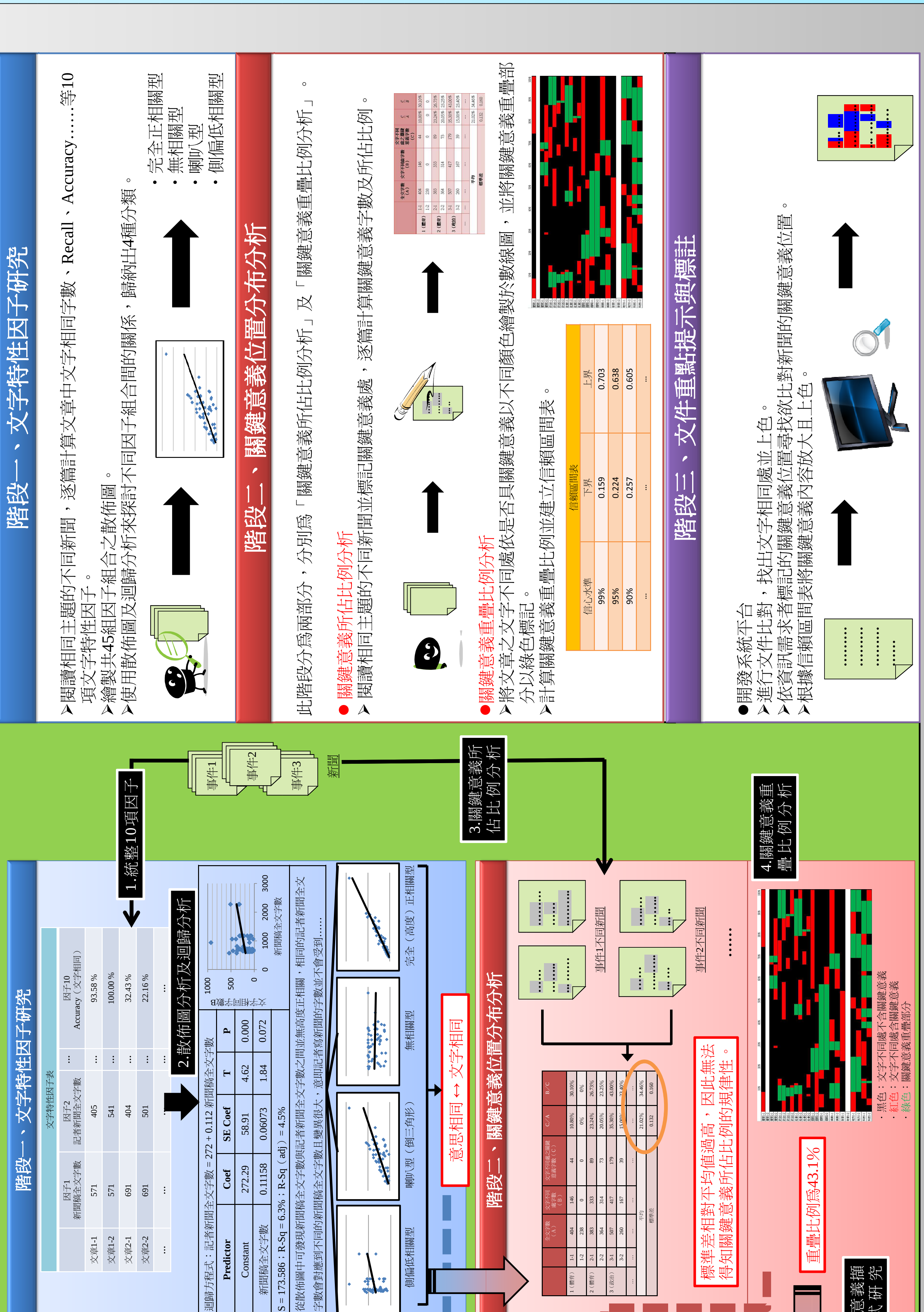
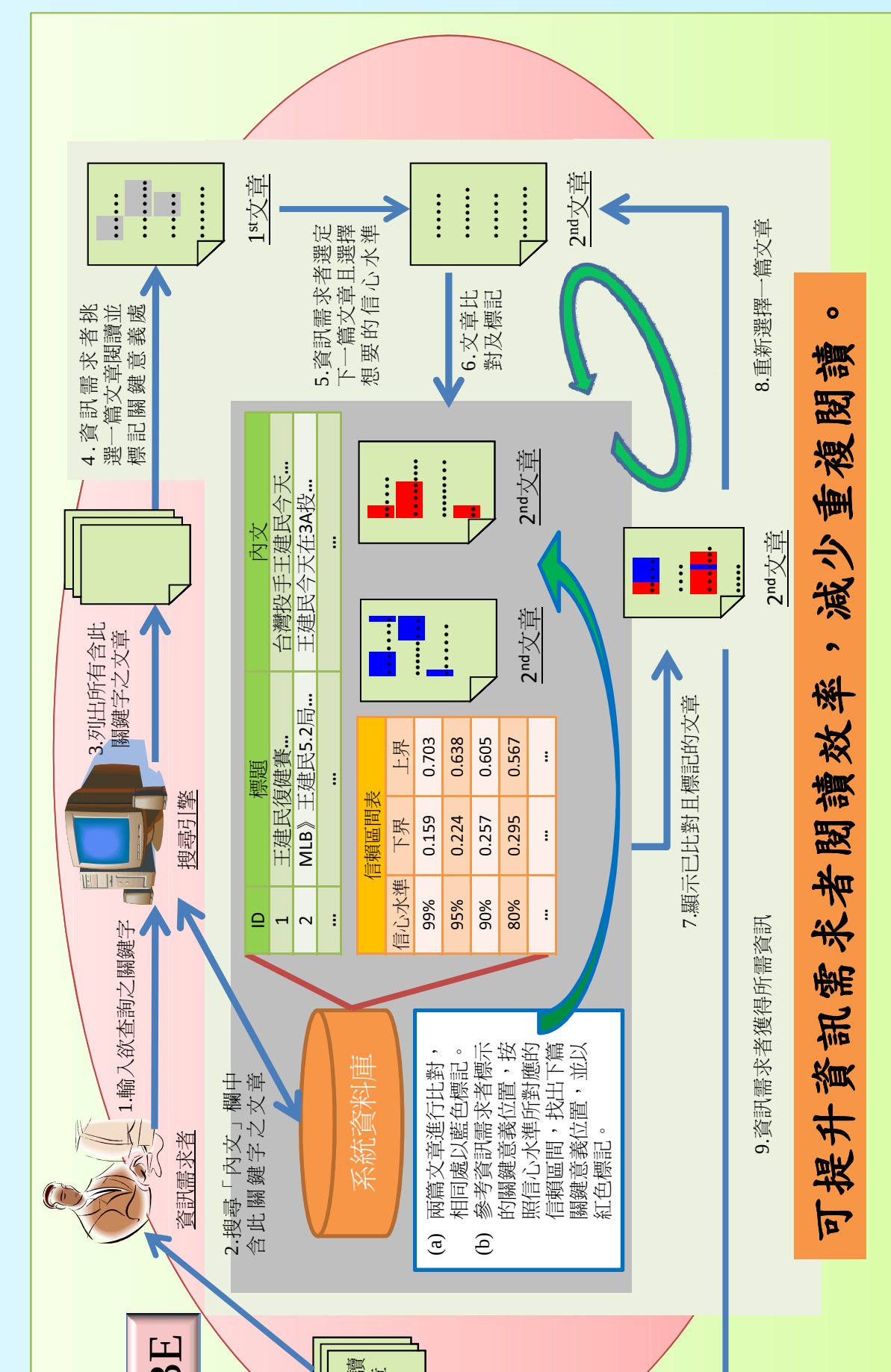
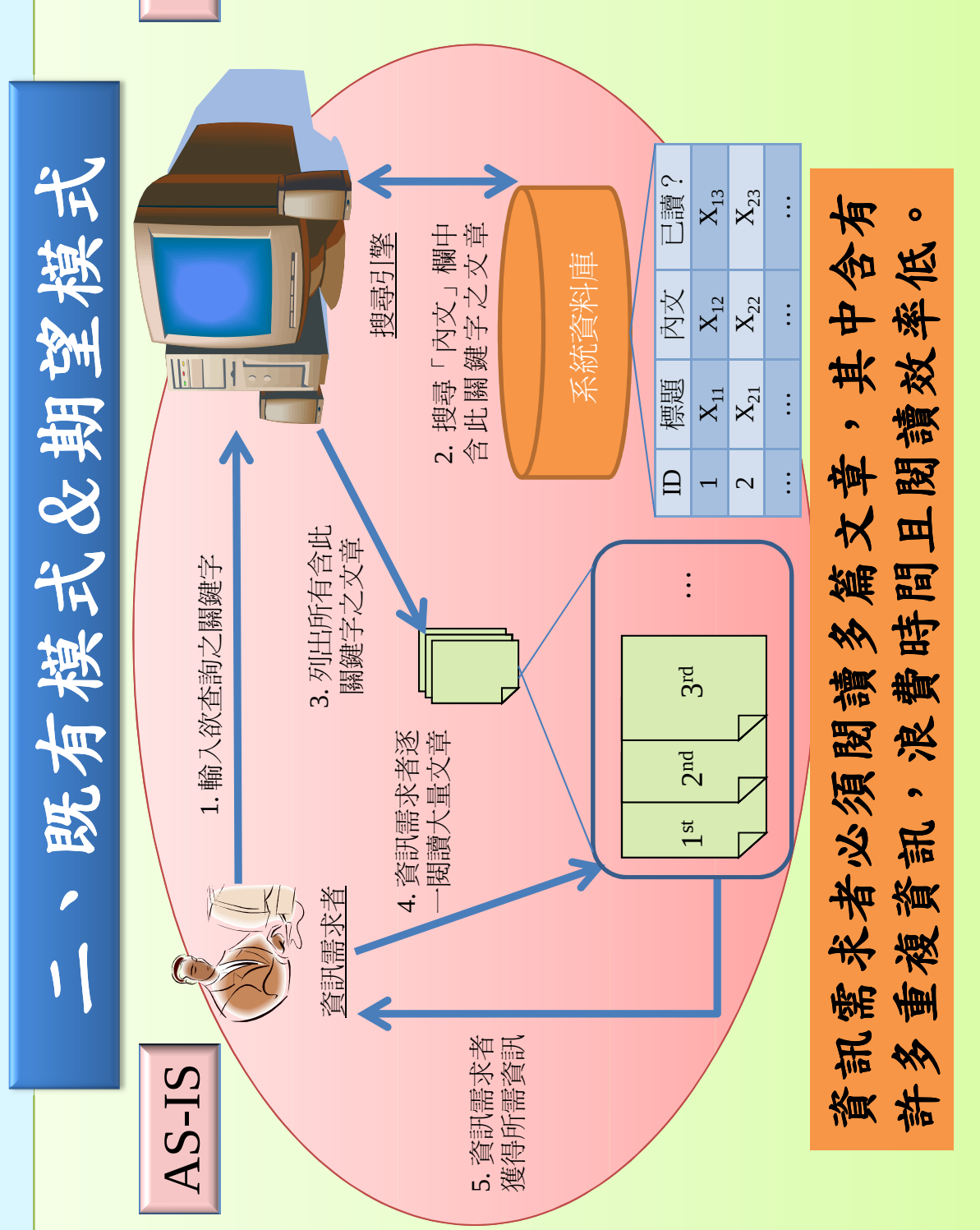
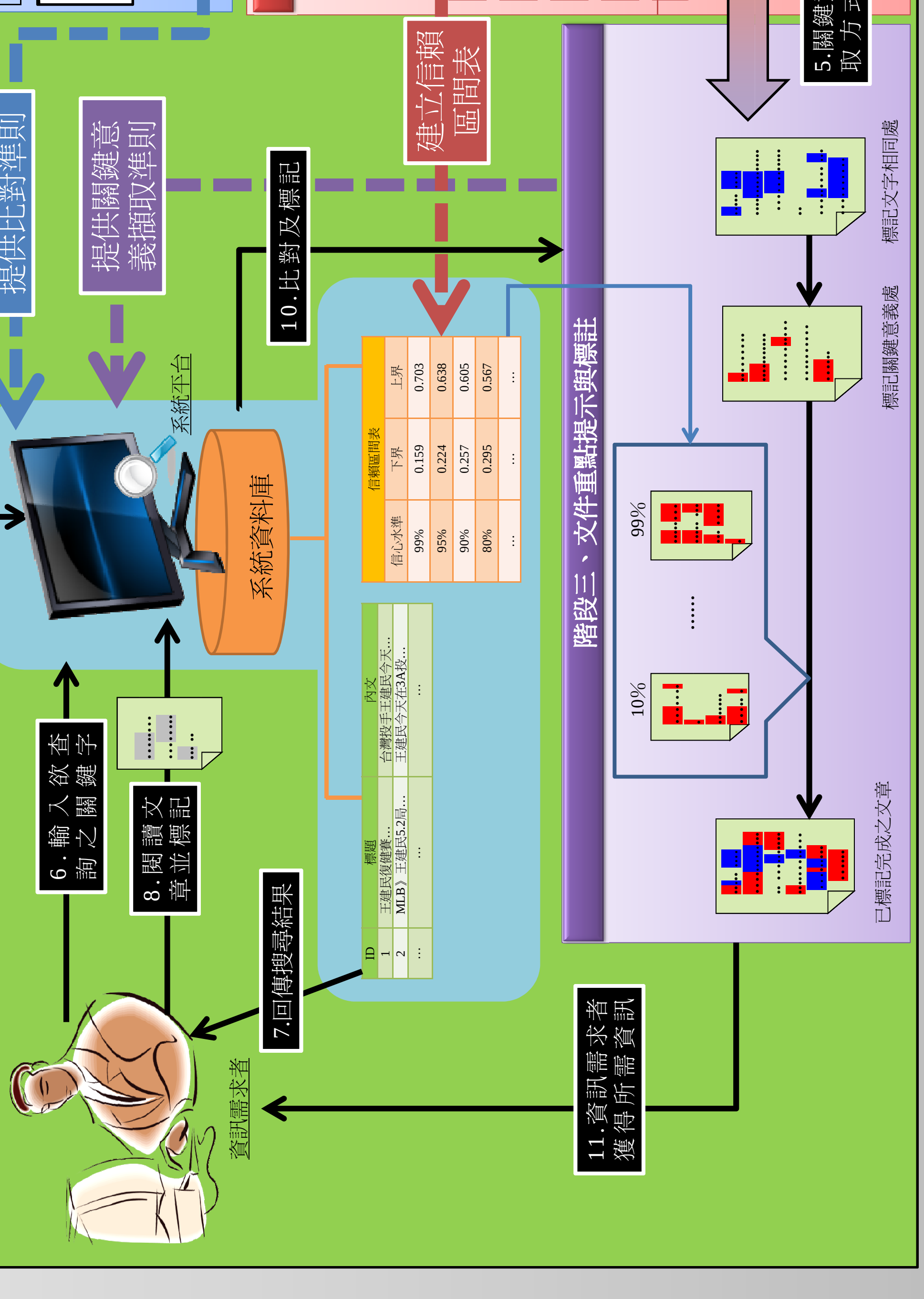
一、研究動機與目的

- 研究動機**
- 背景：近年來由於網路愈趨大眾化，資訊流通量迅速增加，若能**有效擷取文章重點**，便能增加閱讀效率。
 - 現況：資訊需求者**必須閱讀所有文章**才能獲取所需之完整資訊。
 - 問題：資訊需求者所閱讀的文章中可能含有**內含重複性高或不具關鍵意義**的文章，如此將使閱讀效率不彰。

- 研究目的**
- 發展「**以過去閱讀經驗為基礎之閱讀重點標記**」方法論。
 - 以上述方法論為基礎建置一套「**文件比對及標註系統**」，此系統乃依據資訊需求者已閱讀之文章之**關鍵意義位置**回饋，標示即將閱讀之文章之**關鍵意義部分**，避免資訊需求者閱讀重複或非**關鍵意義**內容，進而提升資訊閱讀效率。

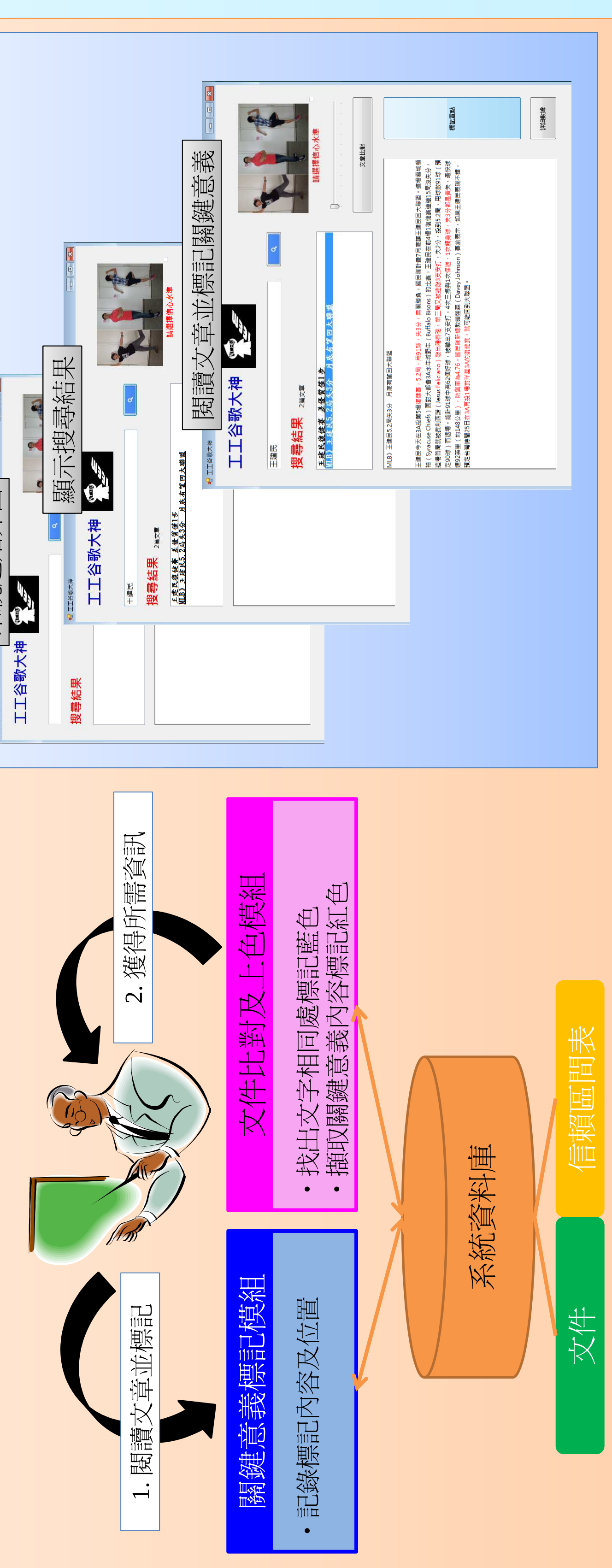
三、研究主軸

本研究乃發展一套**以過去閱讀經驗為基礎之閱讀重點標記方法論**，透過**文字特性因子研究**、**關鍵意義位置分布分析**、**文件重點提示與標註三階段方法**以分析網路新聞文字特性、建立後端資料庫、擷取資訊需求者標記之**關鍵意義位置**並與後端資料庫中的**信賴區間表**相互連結，找出後續閱讀文章之**關鍵意義位置**以增進閱讀效率。本方法論的三階段方法如圖所示：



四、系統開發

本研究開發之「**文件比對及標註**」系統包含**記錄資訊需求者標記關鍵意義位置之關鍵意義標記模組**，及**擷取文字不同處之關鍵意義並上色之文件比對及上色模組**，系統運作架構及操作介面如下圖所示：



五、結論與展望

結論

- 透過「**文字因子特性分析**」可確認「**意思相同**」與「**文字相同**」能夠互相推得，因此**文字相同處可忽略不看**，文字不同處則必須先留下，此結論能**做為不同文章比對時的準則**。
- 建立「**文字不同處關鍵意義重疊比例**」之**信賴區間表**，資訊需求者可以選定想要的信心水準，**做為系統擷取關鍵意義內容多寡的依據**。
- 所開發之「**文章比對及標記系統**」可透過資訊需求者過去的閱讀經驗，找出即將閱讀之文章的**關鍵意義部分**，藉此**提升資訊需求者的閱讀效率**。

展望

- 本研究的**關鍵意義擷取方式**為不遺漏所有**關鍵意義資訊**而擷取過多不重要的資訊，因此本**研究期望找出另一種更為精準的擷取方式**，增進資訊需求者的閱讀效率。
- 建立詞庫比對方式**，解決系統將**表相同意義的不同詞彙**視為相異的問題。
- 持續精進系統，使系統能**針對同一事件多篇文章**，隨資訊需求者閱讀的文章數目越多而**增加系統記憶性**，進而更精準的找出後續文章的**關鍵意義位置**。
- 更新系統介面中文章重點的標記方式**，讓使用者可更具**直觀和性**地進行重點標記。