1. **動機目的**

每個人的精力、記憶力和時間是有限的，農民在使用溫室設計專家知識系統時，在知道所要查詢的知識，卻不知道從技術知識庫中何處查詢起的農民，要如何從系統龐大的技術知識庫中查詢自己想查詢的、有效的知識？為此將透過「文字雲」展示技術知識庫中的知識內容與關鍵字出現頻率，根據目的性的不同，以視覺化的方式呈現。為農民提供更便捷的方式獲取知識內容，使農民在不需要閱讀所有技術知識庫內容的前提下，能夠快速聚焦在技術知識庫中最多人所查詢的知識，為此去尋找自己所要之內容。2

1. **文獻回顧**

**2.1 文字雲**

**2.1.1 什麼是文字雲?**

「文字雲(Word Cloud)」是由各種字詞組合成、如雲一般的圖形。“文字雲”是關鍵詞的圖像化、視覺化描述，用於匯總使用者生成的標籤或一個網站的文字內容。標籤一般是獨立的詞，常常按字母順序排列，其重要程度又能通過改變字體大小或顏色來表現，所以文字雲可以靈活地依照字序或熱門程度來檢索一個標籤。大多數標籤本身就是超連結，直接指向與標籤相聯的一系列條目[1][3]。

**2.1.2 文字雲的分類**

根據文字雲的作用，在應用中可以將其分成三大類。其中，第一類用於描述網站中的每個獨立條目，第二類著力於從整體上刻畫網站所有條目的標籤情況：

* 第一類標籤雲：每一個條目都有自己獨立的標籤雲，標籤字體越大，此條目中使用者使用過這個標籤的次數就越多，在頁面公開統計點擊且不要求精準數據的情況下十分適用。
* 第二類標籤雲：網站一般會有一個超大型標籤雲，標籤字體越大，網站裡使用過這個標籤的條目數就越多。第二類標籤雲可以顯示出標籤的熱門程度。
* 第三類標籤雲：在此類中，標籤作為一個數據項目的工具，用於表示在整個集合中各個項目的數據量的大小。

1. **方法及步驟**

**3.1 文字雲的產生及步驟**

針對溫室專家知識系統中技術知識庫之文字雲產生流程圖，如圖(一)。

****

**▲**圖(一) 文字雲產生流程圖

**3.1.1 抓取知識庫內容**

抓取掃描溫室專家知識系統中技術知識庫之內容，將知識內容進行文字分割整理。

**3.1.2 詞彙斷詞**

將所抓取之技術知識庫內容中能夠獨立運用並含有語義內容或語用內容（即具有表面含義或實際含義）的最小單位，進行詞彙斷詞。

**3.1.3 詞彙頻率統計**

統計技術知識庫中知識內容，每一個詞彙斷詞所出現的總次數。

**3.1.4 機器學習**

將每個詞彙所統計之次數，進行機器學習(例如：集群分析、智慧計算演算法)，以找出詞彙間的規則、關聯性、熱門程度等等。

**3.1.5 關鍵詞萃取**

根據機器學習的學習計算結果，排列出詞彙相關性、熱門、最常使用、最常出現等關鍵詞，進行關鍵詞萃取。

**3.1.65 資料視覺化－文字雲**

文字雲將依據關鍵詞萃取結果進行排序，排序結果的關鍵詞將通過不同的字體大小或顏色來表現其最常使用、最熱門、重要之程度。

1. **參考文獻**
2. 文字雲－維基百科

https://zh.wikipedia.org/wiki/%E6%A0%87%E7%AD%BE%E4%BA%91#cite\_note-1

1. Martin Halvey and Mark T. Keane, An Assessment of Tag Presentation Techniques, poster presentation at WWW 2007, 2007
2. KhaldoonDhou, Mirsad Hadzikadic ,and Mark Faust, “Typeface size and weight and word location influence on relative size judgments in tag clouds”, *Journal of Visual Languages & Computing*, Volume 44, Pages 97-105, February 2018.