Project 1

Group 12
施昱竹 104078502 莊翊鈞 104078514

Outline

- Goal
- Tool
- Framework
- Code
- Index of word cloud
- Word cloud of each department

Goal

- Analytics goal:

  Generating word cloud for each department.
  (47 departments in total)

- Business goal:

  For senior high students or researchers in different fields could have a big picture that what is the main topic in each department.

Tool

1. Python for text-mining
2. Word cloud using wordcloud2 website (https://timdream.org/wordcloud2.js/#les-miz)

Framework

Degree of completion: 100%

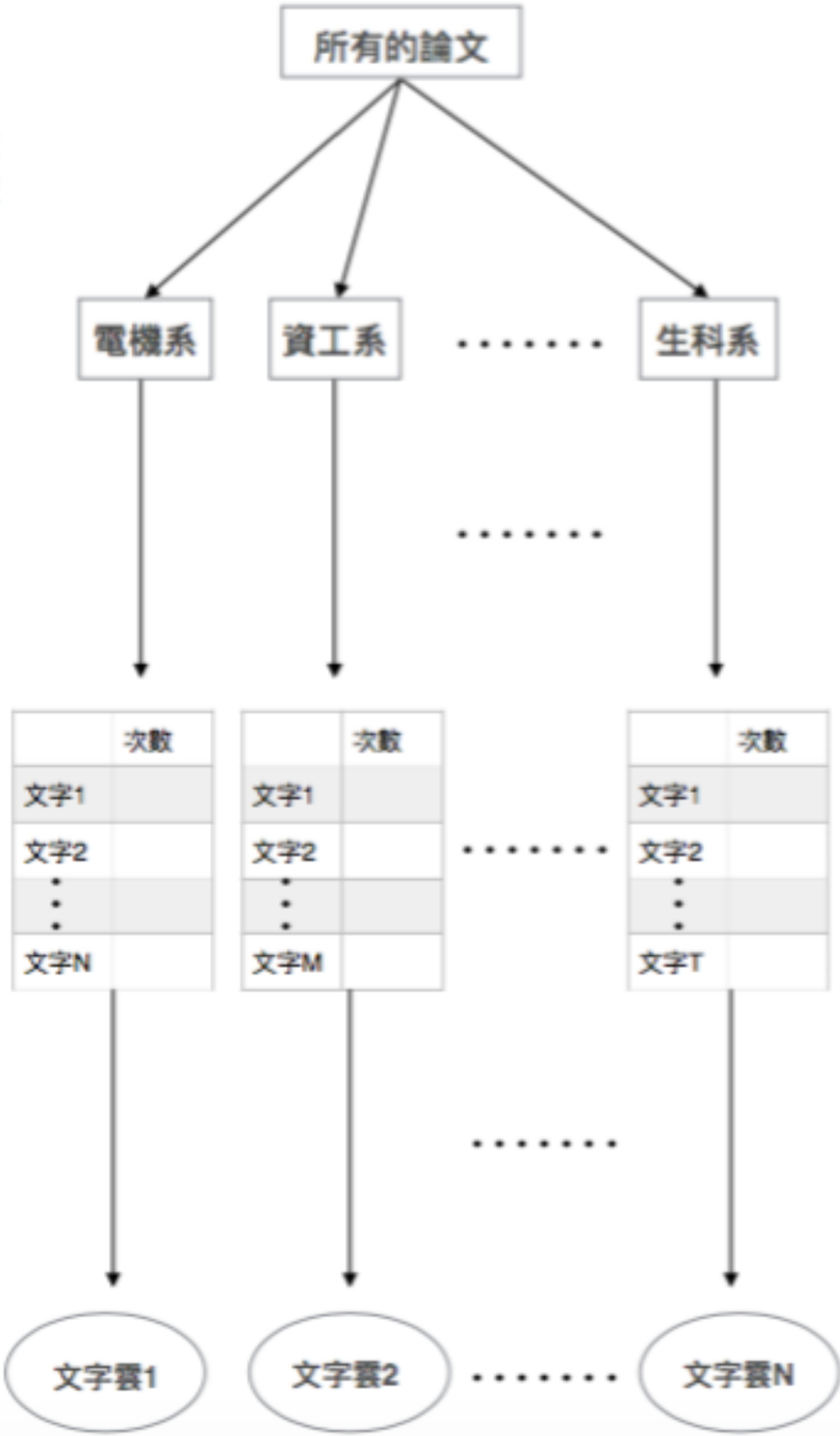
步驟說明

1.先將所有的論文，以"**系所名稱**"欄位做區分，個別分成小群集。目的是為個別系所建立文字雲；原因是以系所建立，將比較有參考價值，同類所研究較為相近。
*使用方法或套件:
using index ("系所名稱")

2. 因為論文多為中文撰寫，我們將以"**中文關鍵詞**"為目標，並計算該關鍵字出現在多少文件的次數(也就是Document Frequency)，來當作權重。
*使用方法或套件:
using nltk count the DF

3. 最後每個系所都會個別統計出該系所論文中，有哪些關鍵字，和各個關鍵子的次數統計，用這些次數統計，每個系所都將產生一個文字雲。
*方法或套件:
using wordcloud or
tool Timdream's wordcloud2

步驟示意



Code

```
import json
theses = json.load(open('nthu_thesis20170330.json'))
theses[0].index('系所名稱')
theses[0].index('中文關鍵詞')

#extract department and Chinese keyword
data_list = [(thesis[2], thesis[17].split()) for thesis in theses[1:]]

#store all the department into a list
index = []
i=0
while i < 5444:
    if data_list[i][0] not in index:
        index.append(data_list[i][0])
    i+=1
```

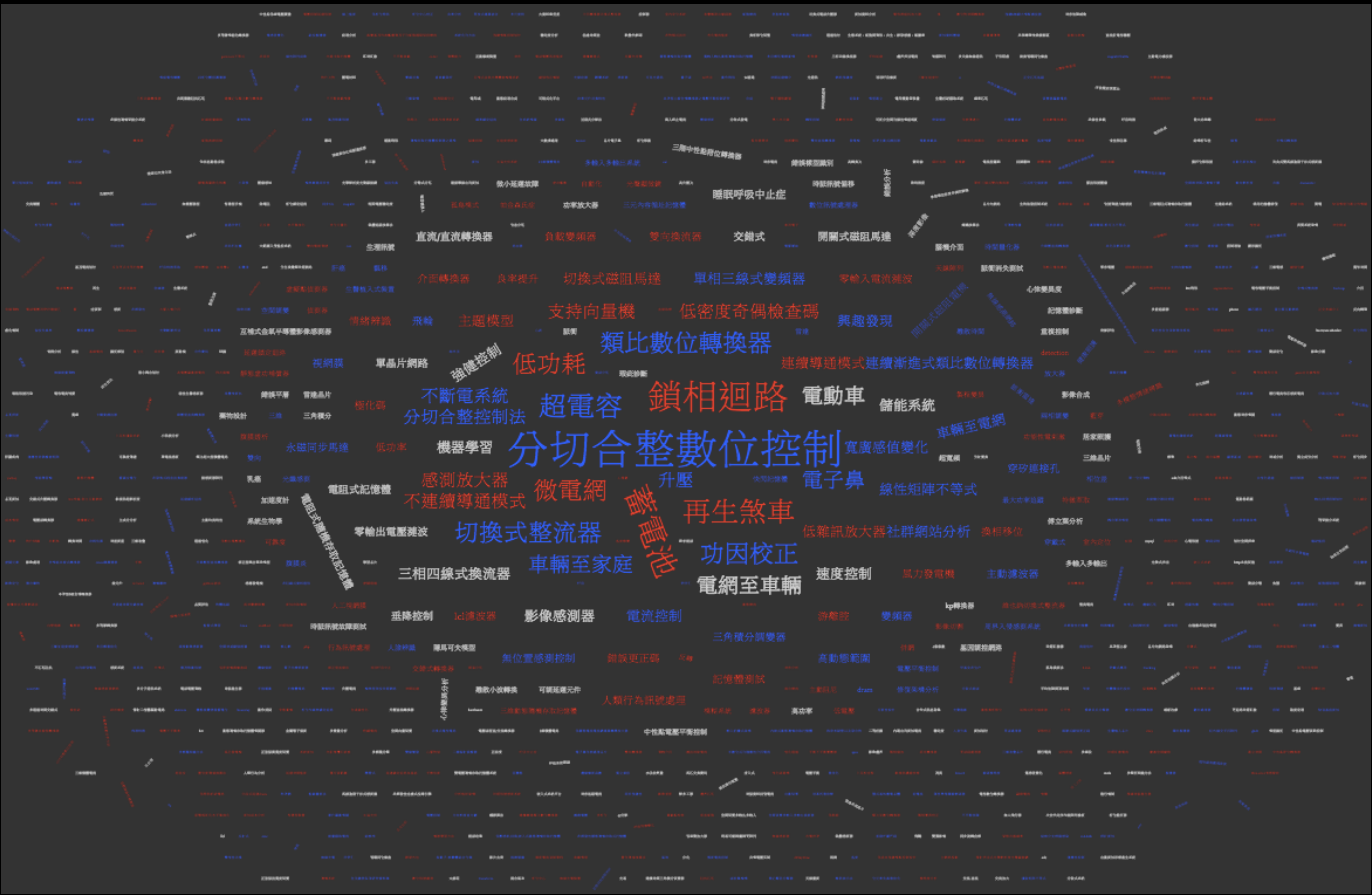
```
#Take EE as an example
ee = []

#store all the EE's keywords into a list
j=0
while j < len(data_list):
    if data_list[j][0] == index[0]:
        ee.append(data_list[j][1])
    j+=1

#remove list of list into one list
flat_ee = []
for sublist in ee:
    for val in sublist:
        flat_ee.append(val.lower())
```

```
#count the word and rank it
from collections import Counter
counter_ee = Counter(flat_ee)
sort_ee = sorted(counter_ee.items(), key=lambda x: -x[1])

#reverse the word's and number's position
m=0
while m < len(sort_ee):
    print (sort_ee[m][1], sort_ee[m][0])
    m=m+1

#copy the result and paste to wordcloud2 website
```

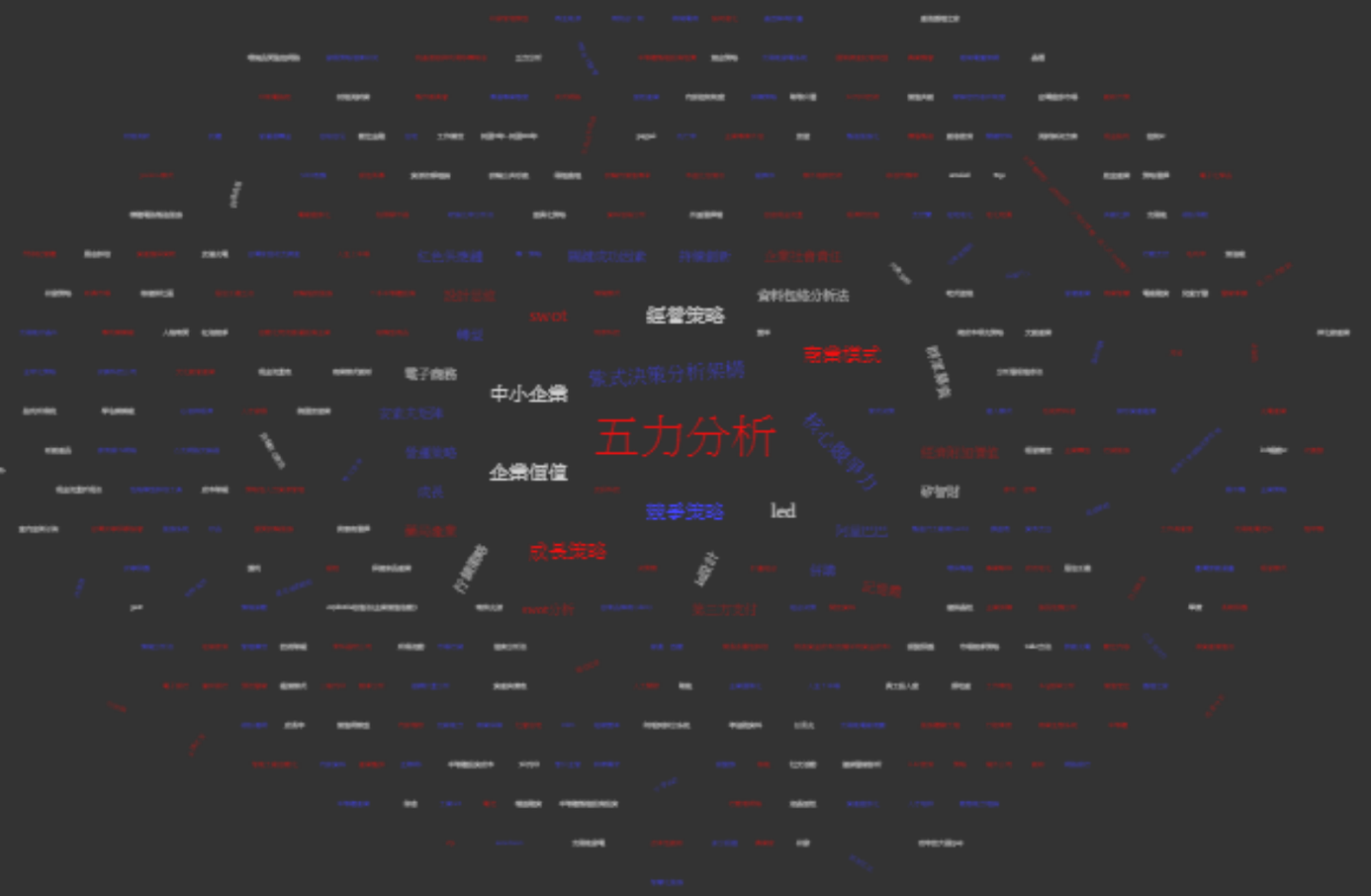
Index of word cloud

0. 電機工程學系
1. 高階經營管理碩士在職專班
2. 生醫工程與環境科學系
3. 材料科學工程學系
4. 分子與細胞生物研究所
5. 通訊工程研究所
6. 資訊工程學系
7. 經濟學系
8. 生物資訊與結構生物研究所
9. 工業工程與工程管理學系碩士在職專班
10. 歷史研究所
11. 台灣文學研究所
12. 電子工程研究所
13. 資訊系統與應用研究所
14. 奈米工程與微系統研究所
15. 物理系
16. 計量財務金融學系
17. 光電工程研究所
18. 數學系
19. 天文研究所
20. 動力機械工程學系
21. 外國語文學系
22. 工程與系統科學系
23. 化學系
24. 社會學研究所
25. 人類學研究所
26. 學習科學研究所
27. 服務科學研究所
28. 核子工程與科學研究所
29. 化學工程學系
30. 統計學研究所
31. 工業工程與工程管理學系
32. 分子醫學研究所
33. 中國文學系
34. 哲學研究所
35. 科技法律研究所
36. 生物科技研究所
37. 經營管理碩士在職專班
38. 科技管理研究所
39. 語言學研究所
40. 先進光源科技學位學程
41. 國際專業管理碩士班
42. 系統神經科學研究所
43. 臺灣研究教師在職進修碩士學位班
44. 生物醫學工程研究所
45. 半導體元件及製程產業研發碩士專班
46. 亞際文化研究國際碩士學位學程

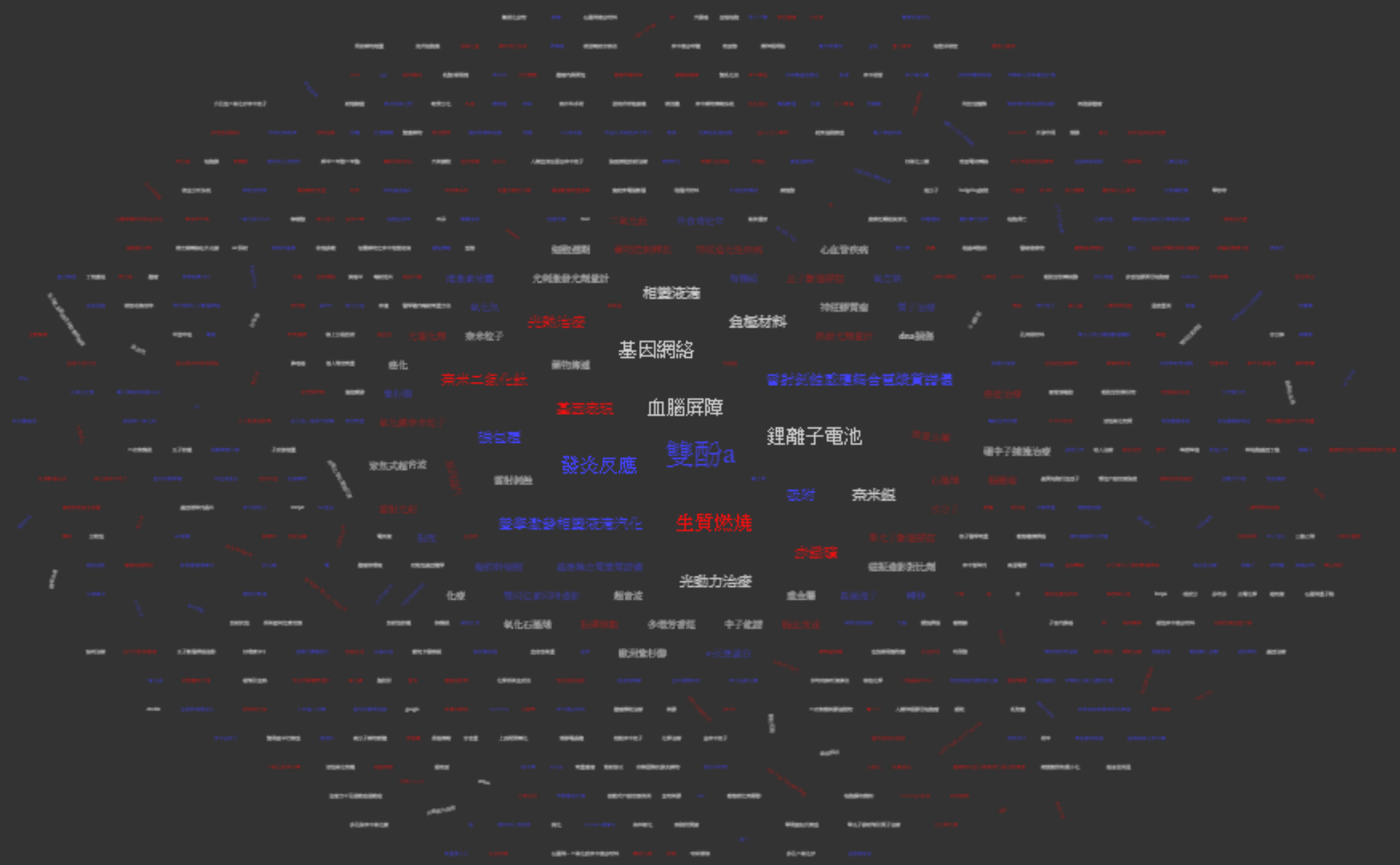
0.電機工程學系


分切合整數位控制
鎖相迴路
類比數位轉換器
超電容
蓄電池
再生煞車
功因校正
切換式整流器
微電網
低功耗
電動車
電子鼻
車輛至家庭
電網至車輛
不斷電系統
分切合整控制法
感測放大器
不連續導通模式
支持向量機
低密度奇偶檢查碼
升壓
強健控制
三相四線式換流器
速度控制
影像感測器
機器學習
儲能系統
寬廣感值變化
車輛至電網
線性矩陣不等式
連續導通模式
連續漸進式類比數位轉換器
低雜訊放大器
社群網站分析
零輸出電壓漣波
電阻式記憶體
永磁同步馬達
直流/直流轉換器
切換式磁阻馬達
單相三線式變頻器
開關式磁阻馬達
交錯式
雙向換流器
主題模型
興趣發現
睡眠呼吸中止症
電流控制
垂降控制
三角積分調變器
單晶片網路

1.高階經營管理碩士在職專班



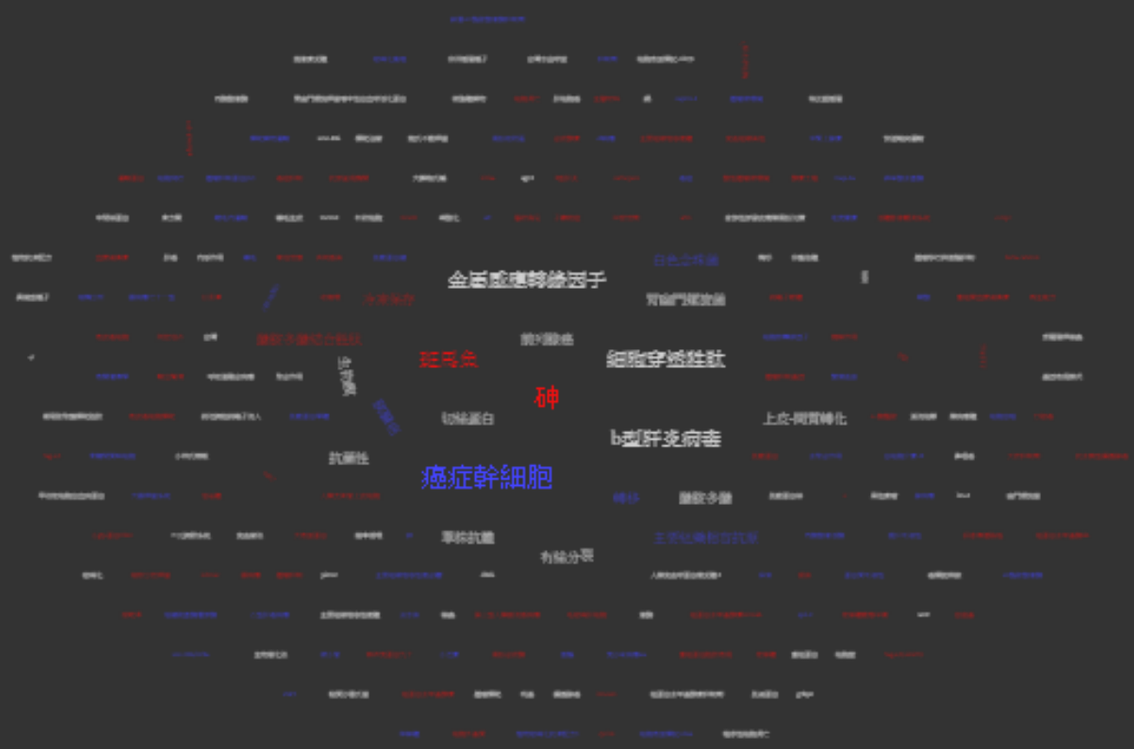
2.生醫工程與環境科學系



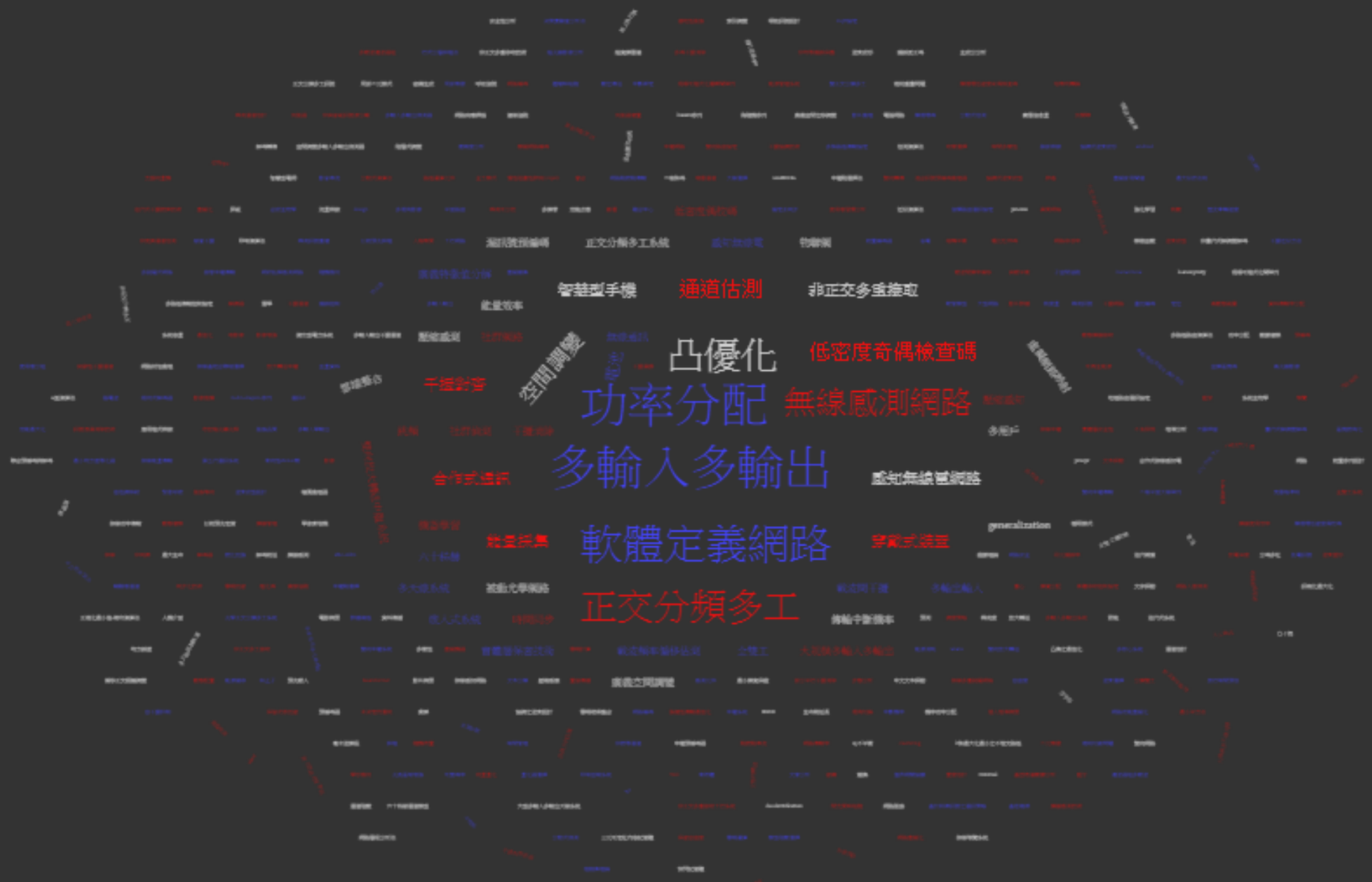
3.材料科學工程學系


有機發光二極體
奈米碳管
石墨烯
鋰離子電池
太陽能電池
氧化鋅
二氧化鈦
光觸媒
高熵合金
二維材料
奈米線
機械性質
有機薄膜電晶體
超級電容器
光電特性
散熱材料
光催化
熱電材料
碲化鉍
超合金
原子層沉積技術
超級電容
冷凍鑄造法
水熱法
共軛高分子
仿生
奈米銀線
合金設計
超硬合金
鋰電池
介電性質
奈米粒子
海水電池
電阻式記憶體
摻雜
薄膜
氧化物
熱穩定性
分子束磊晶
meh-ppv
奈米結構
鈮奈米粒子
電磁波屏蔽
反射式歐姆電極
溶膠凝膠法
超材料
紫外光
多階層結構
銅銦鎵硒
二硫化鉬
五苯環
複合材料
微結構
色心
砷化鎵
有機小分子
金奈米粒子
多層薄膜
氧化亞銅
高溫硬度
熱穩定性
氧化鋅奈米線
有機太陽能電池
耐火高熵合金
金屬氧化物
硬度
熱性質

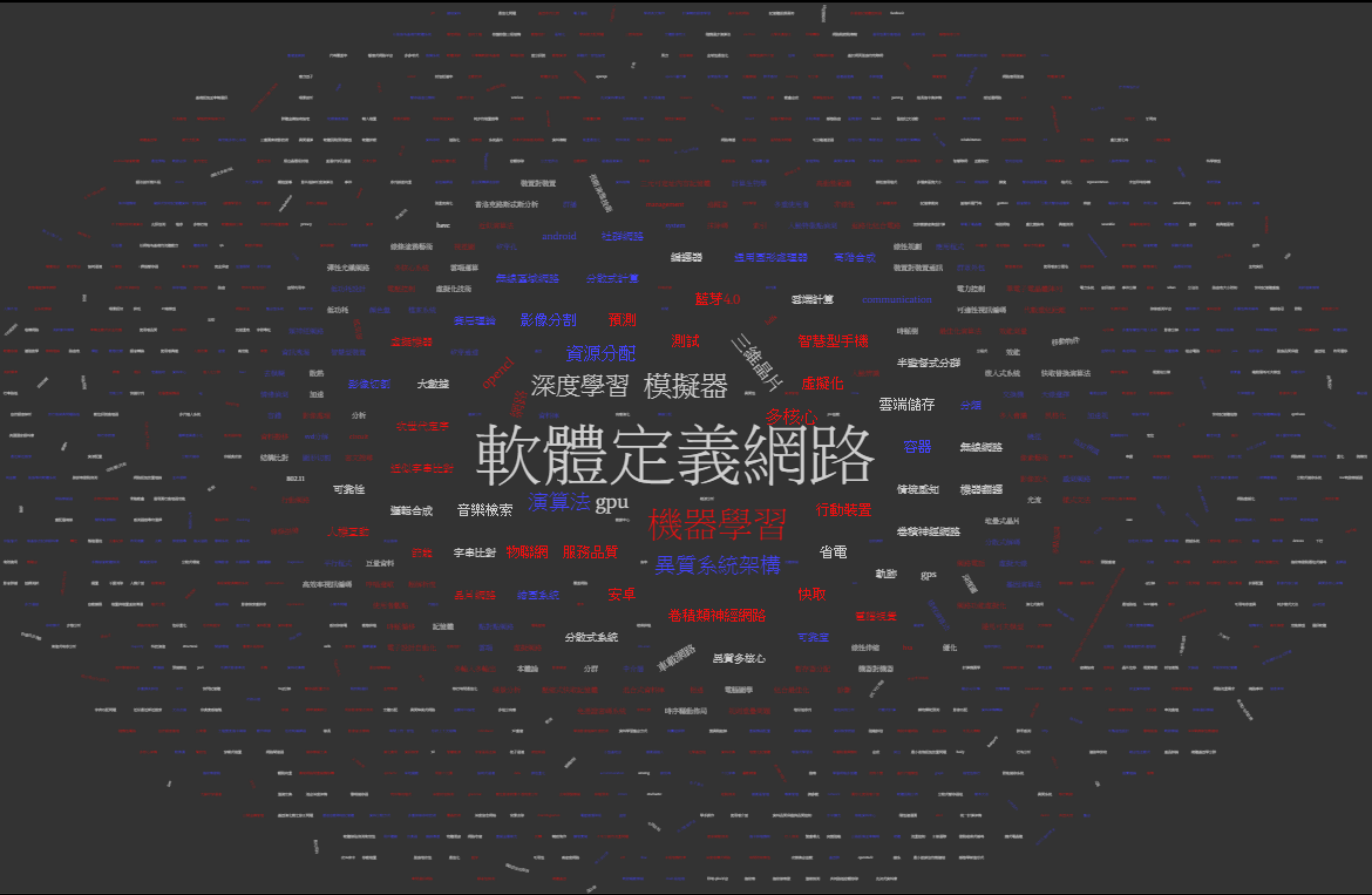
4.分子與細胞生物研究所



5.通訊工程研究所

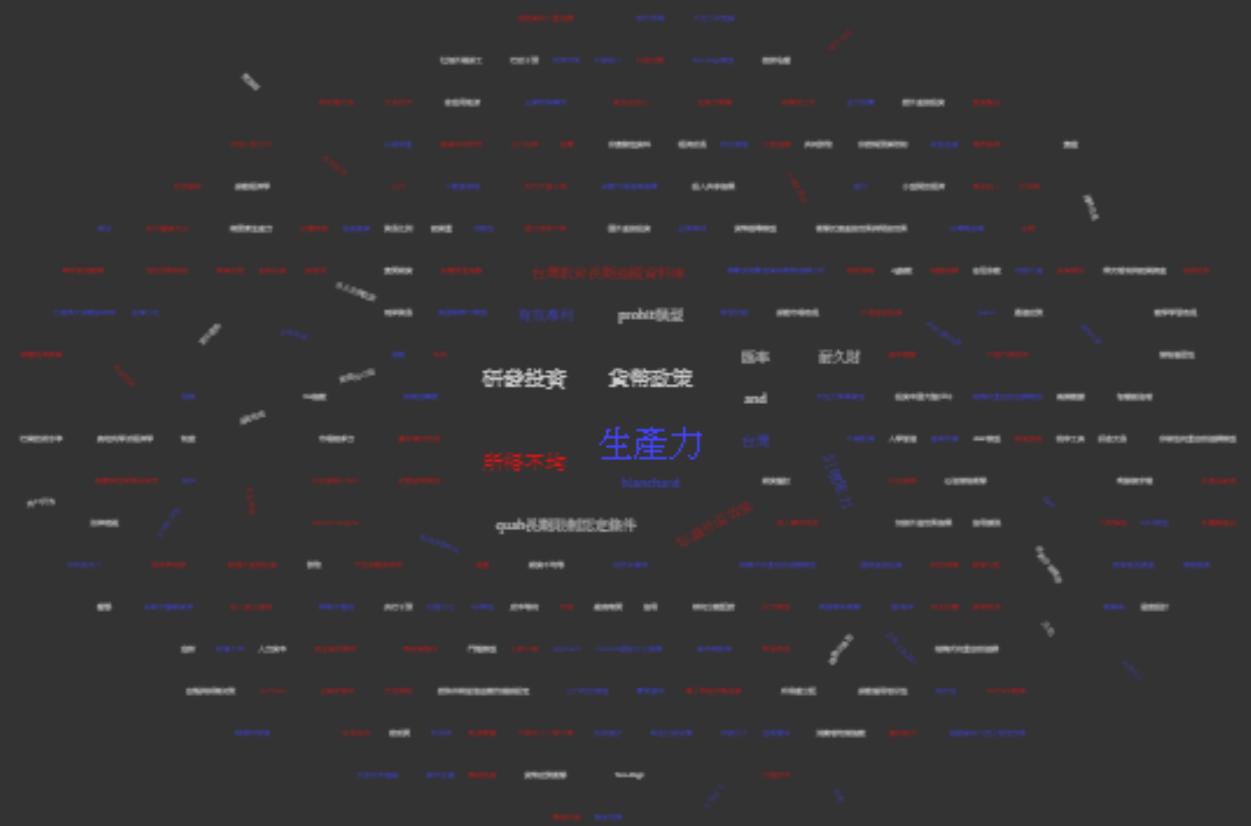

智慧型手機
通道估測
非正交多重接取
凸優化
低密度奇偶檢查碼
功率分配
無線感測網路
多輸入多輸出
軟體定義網路
正交分頻多工

6.資訊工程學系

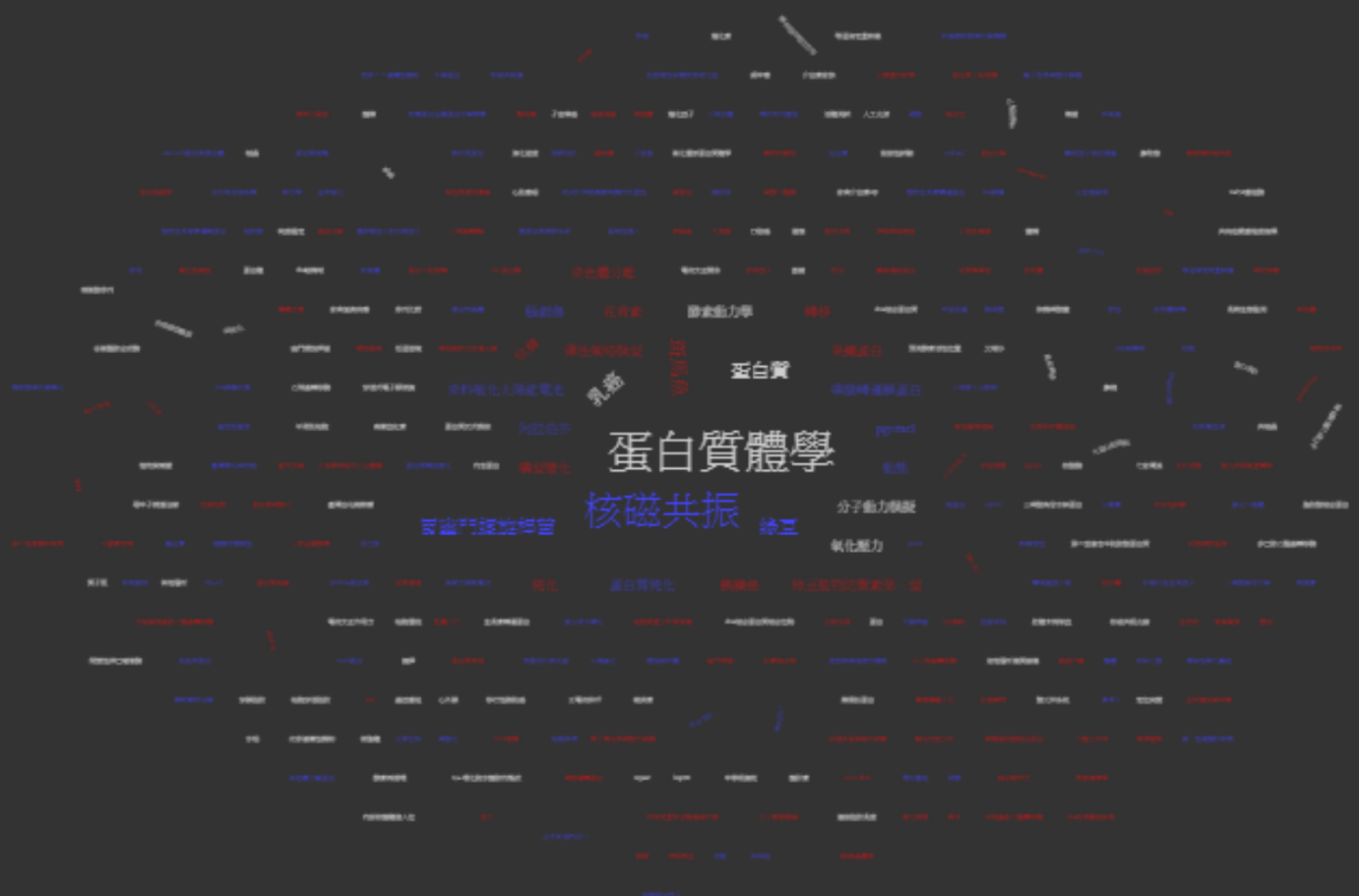


軟體定義網路
機器學習
深度學習
模擬器
異質系統架構
演算法
gpu
資源分配
影像分割
預測
測試
三維晶片
智慧型手機
虛擬化
多核心
雲端儲存
容器
行動裝置
卷積類神經網路
安卓
快取
物聯網
服務品質
音樂檢索
可靠性
省電
藍芽4.0
opencl
android
communication
分散式系統
字串比對
大數據
gps

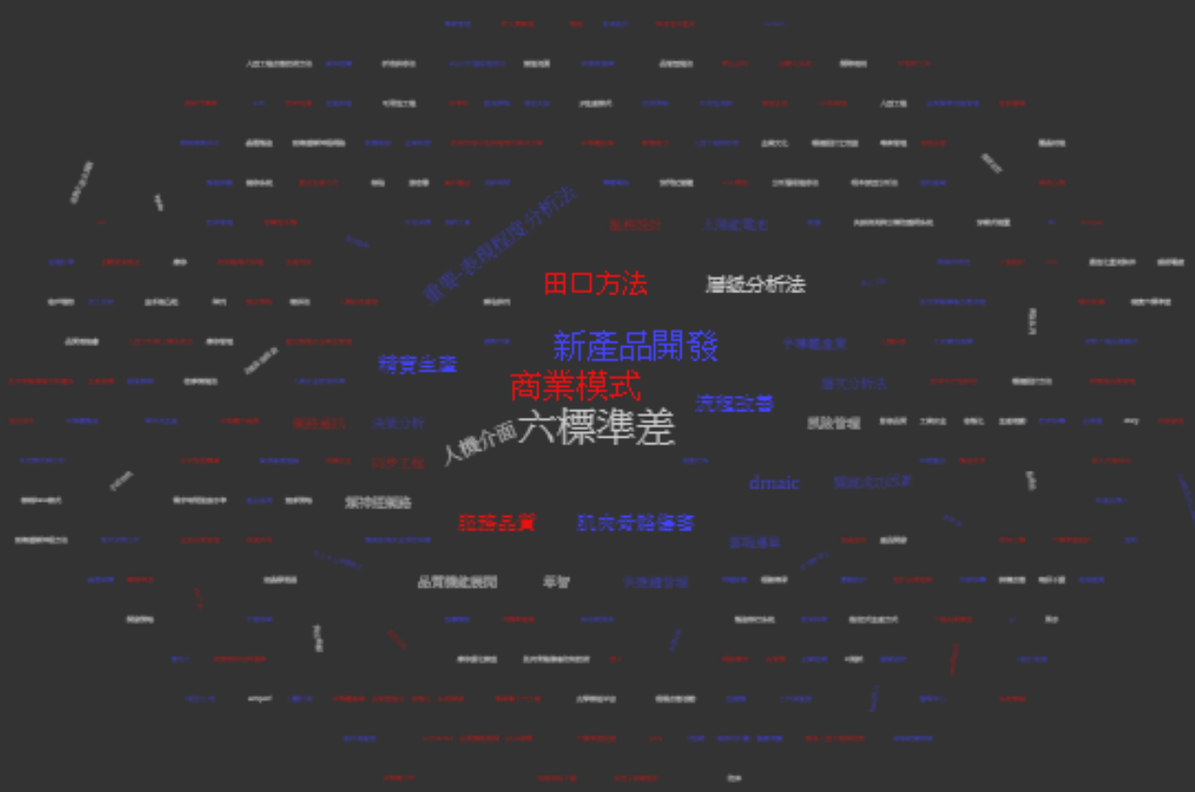
7.經濟學系



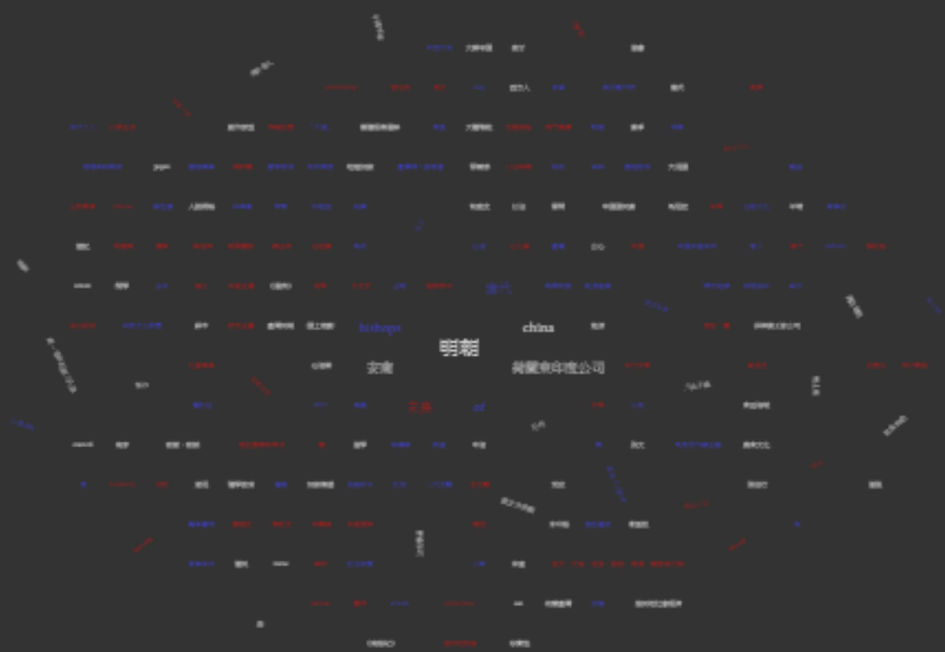
8.生物資訊與結構生物研究所



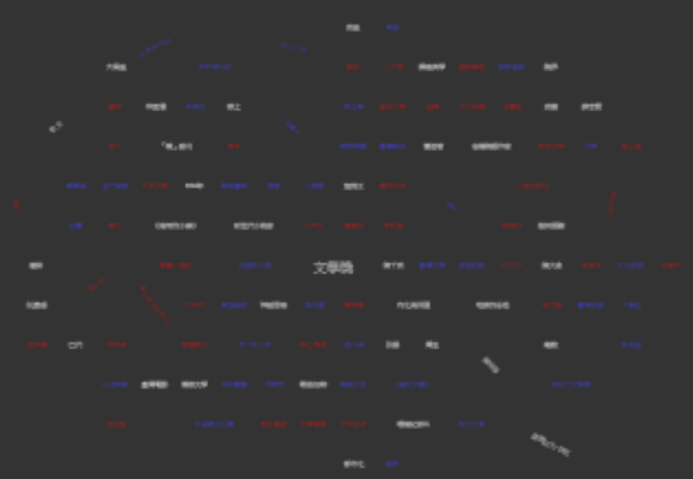
9.工業工程與工程管理學系碩士在職專班



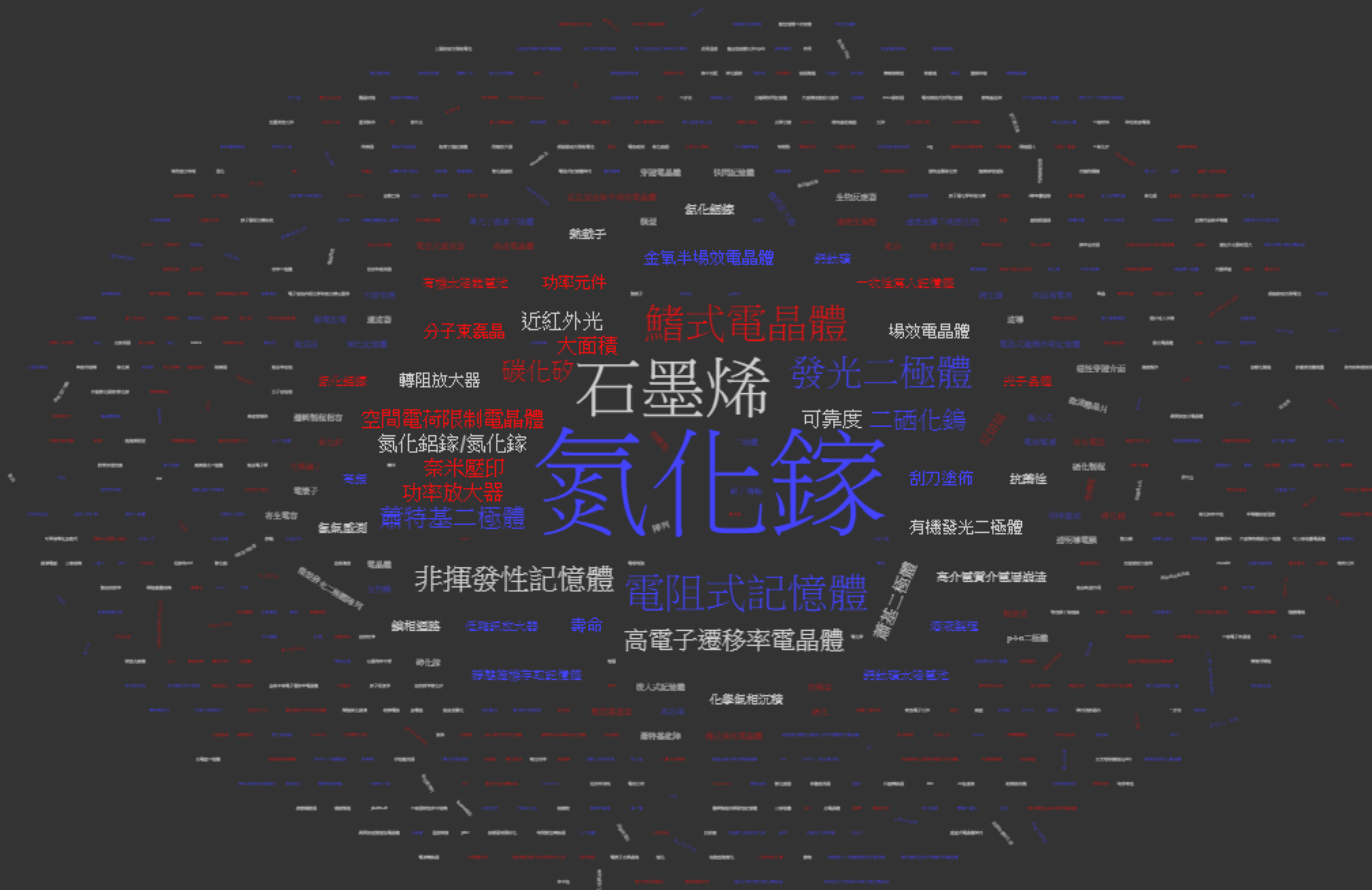
10.歷史研究所



11.台灣文學研究所

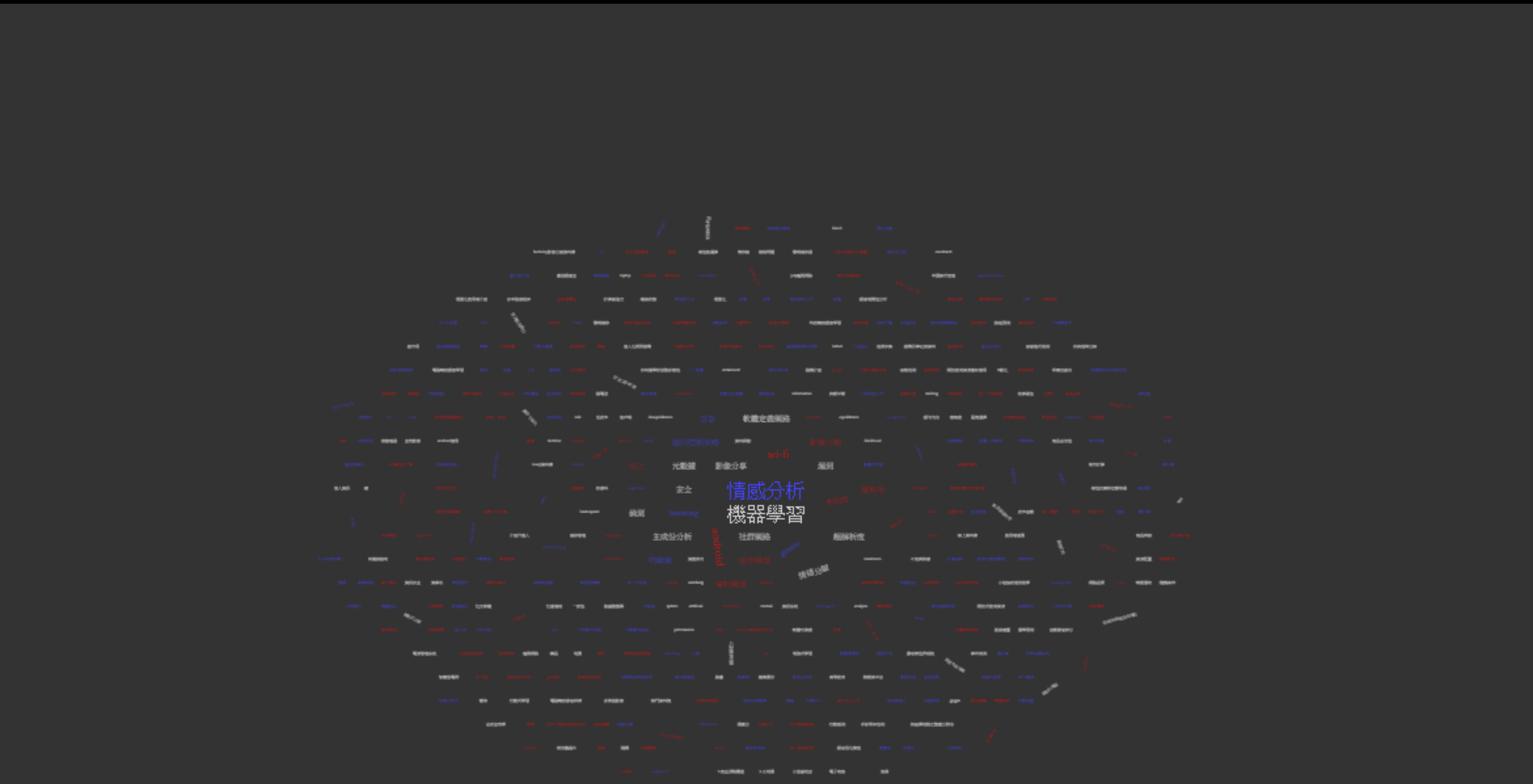


12.電子工程研究所

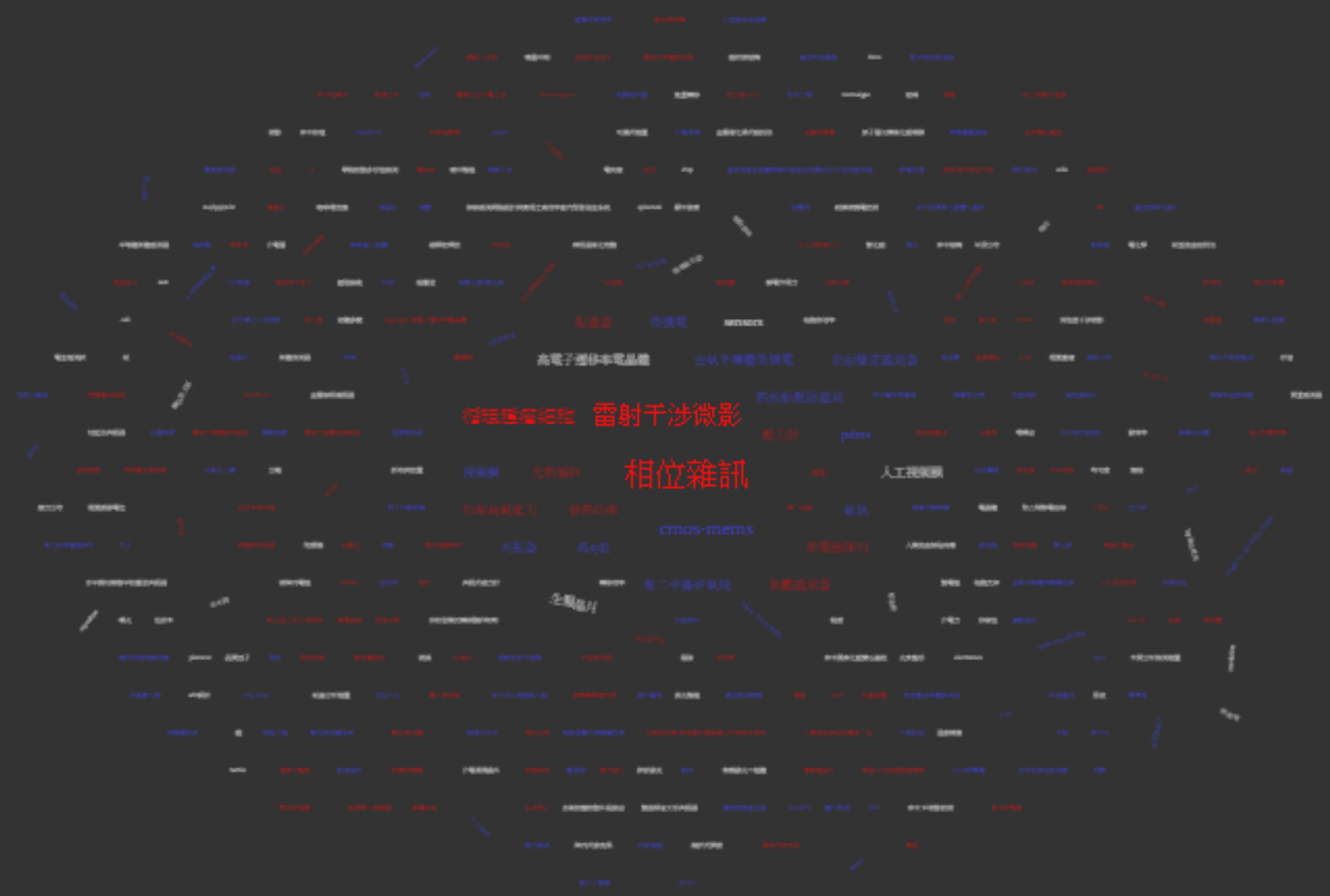


氮化鎵
石墨烯
鰭式電晶體
發光二極體
電阻式記憶體
非揮發性記憶體
高電子遷移率電晶體
蕭特基二極體
近紅外光
大面積
碳化矽
分子束磊晶
轉阻放大器
空間電荷限制電晶體
氮化鋁鎵/氮化鎵
奈米壓印
功率放大器
可靠度
二硒化鎢
場效電晶體
金氧半場效電晶體
功率元件
刮刀塗佈
有機發光二極體
蕭特基二極體

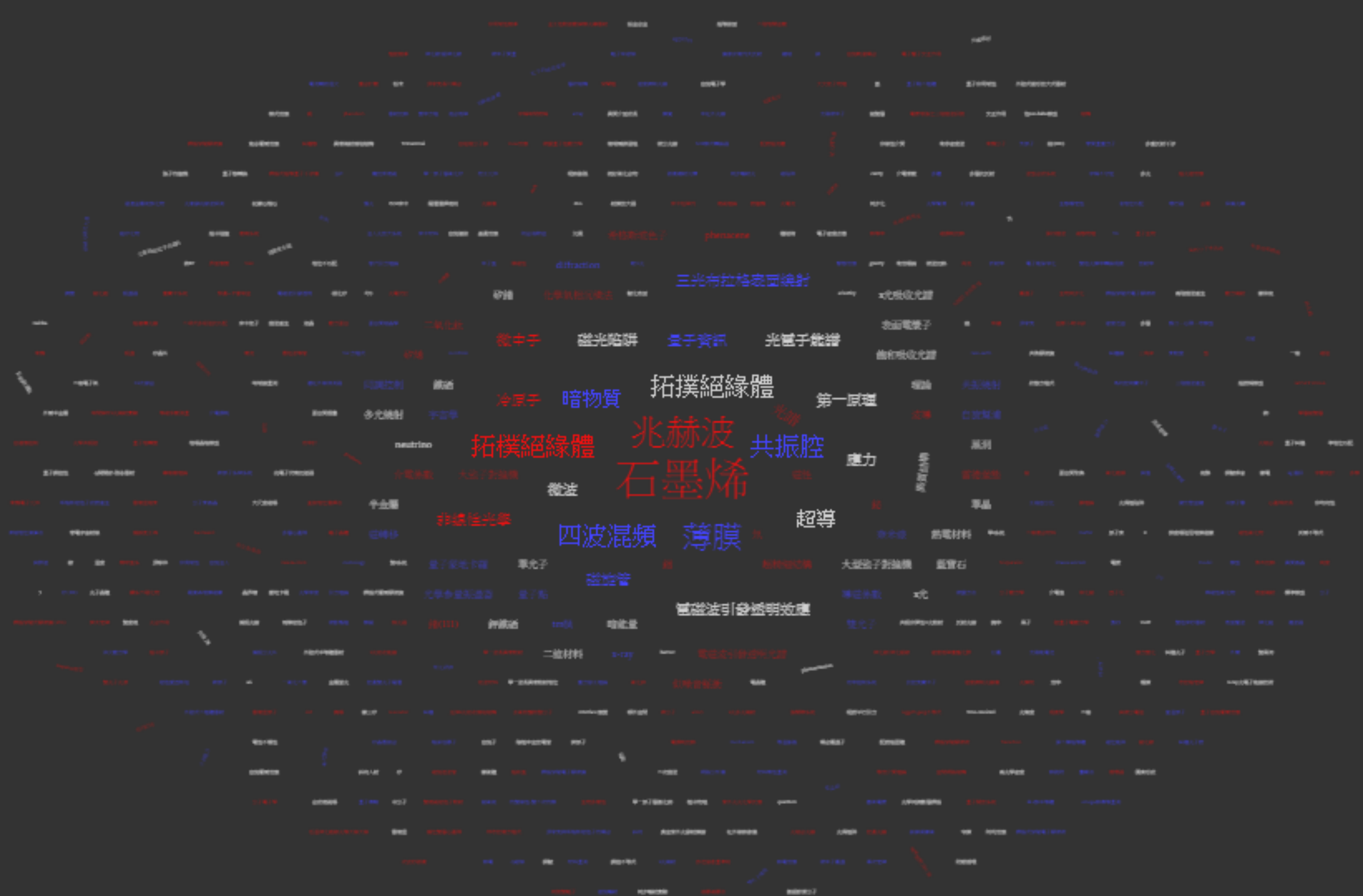
13.資訊系統與應用研究所



14.奈米工程與微系統研究所



15.物理系

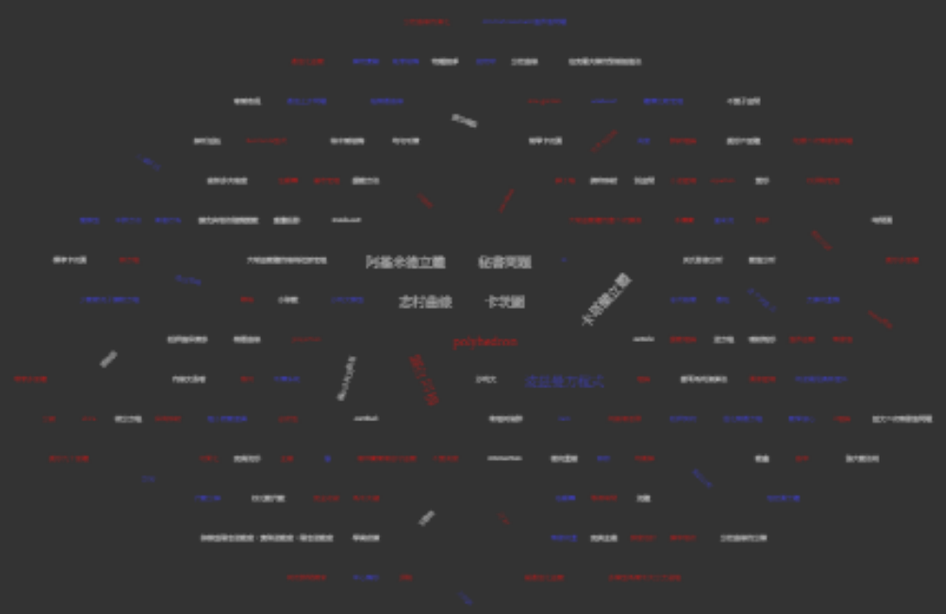


16.計量財務金融學系

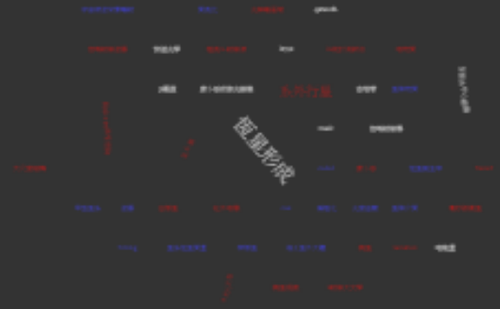
17.光電工程研究所


去氧核醣核酸
有機記憶體元件
超材料
脈衝量測
超快光學
兆赫波
正交分頻多工
二倍頻
雷射
光回饋
偏振態多工
半導體雷射
光偵測器
非線性動態
被動光學網路
機械損耗
光纖雷射
超短脈衝
光子晶體
三維

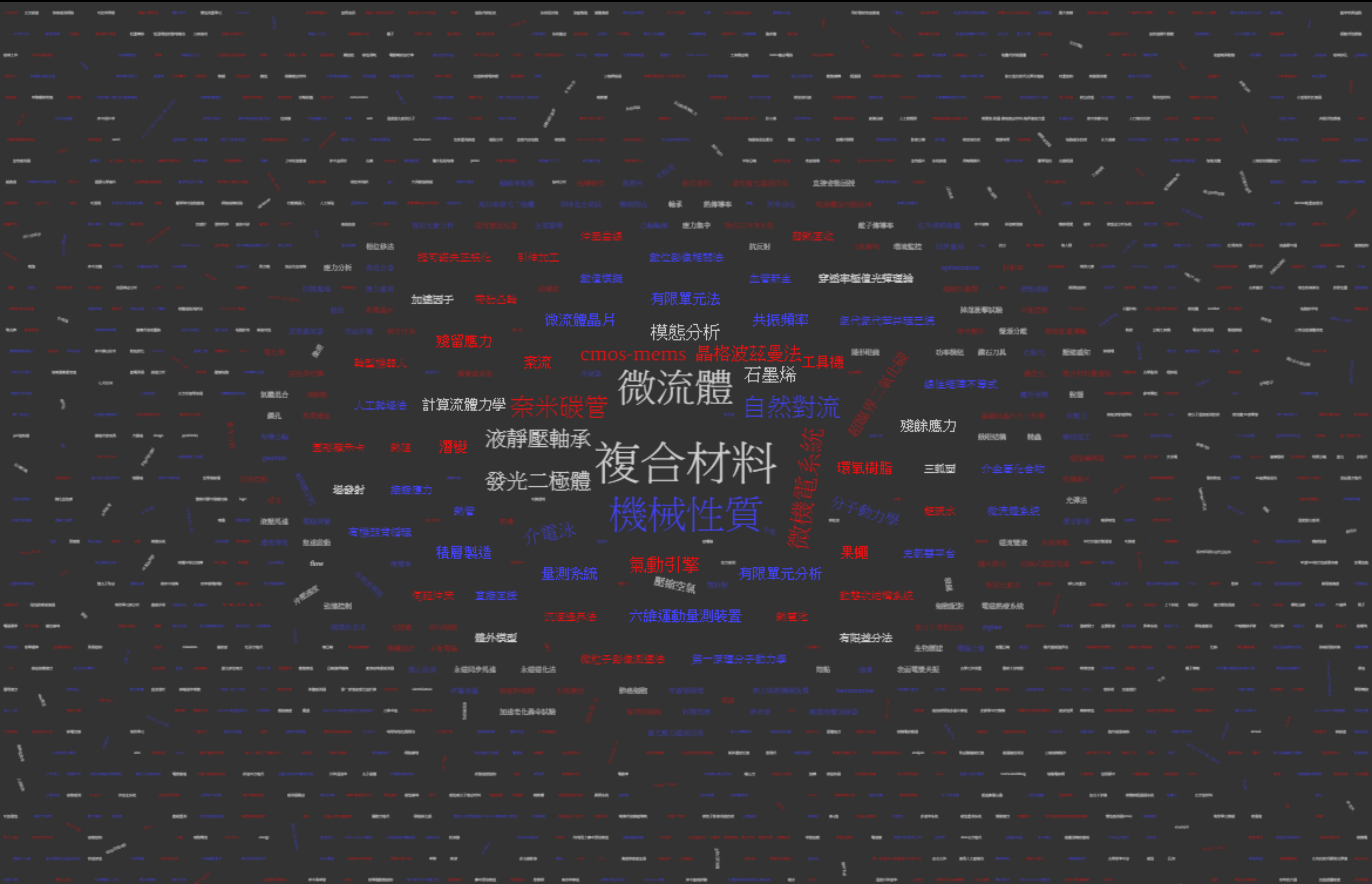
18.數學系



19.天文研究所

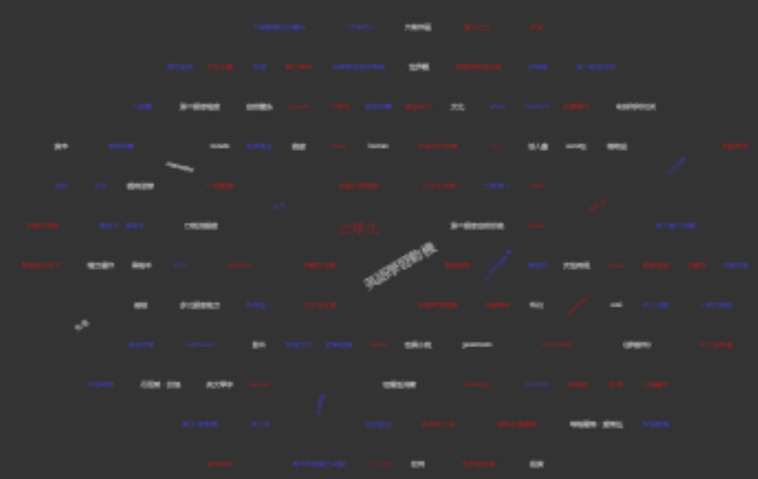


20.動力機械工程學系

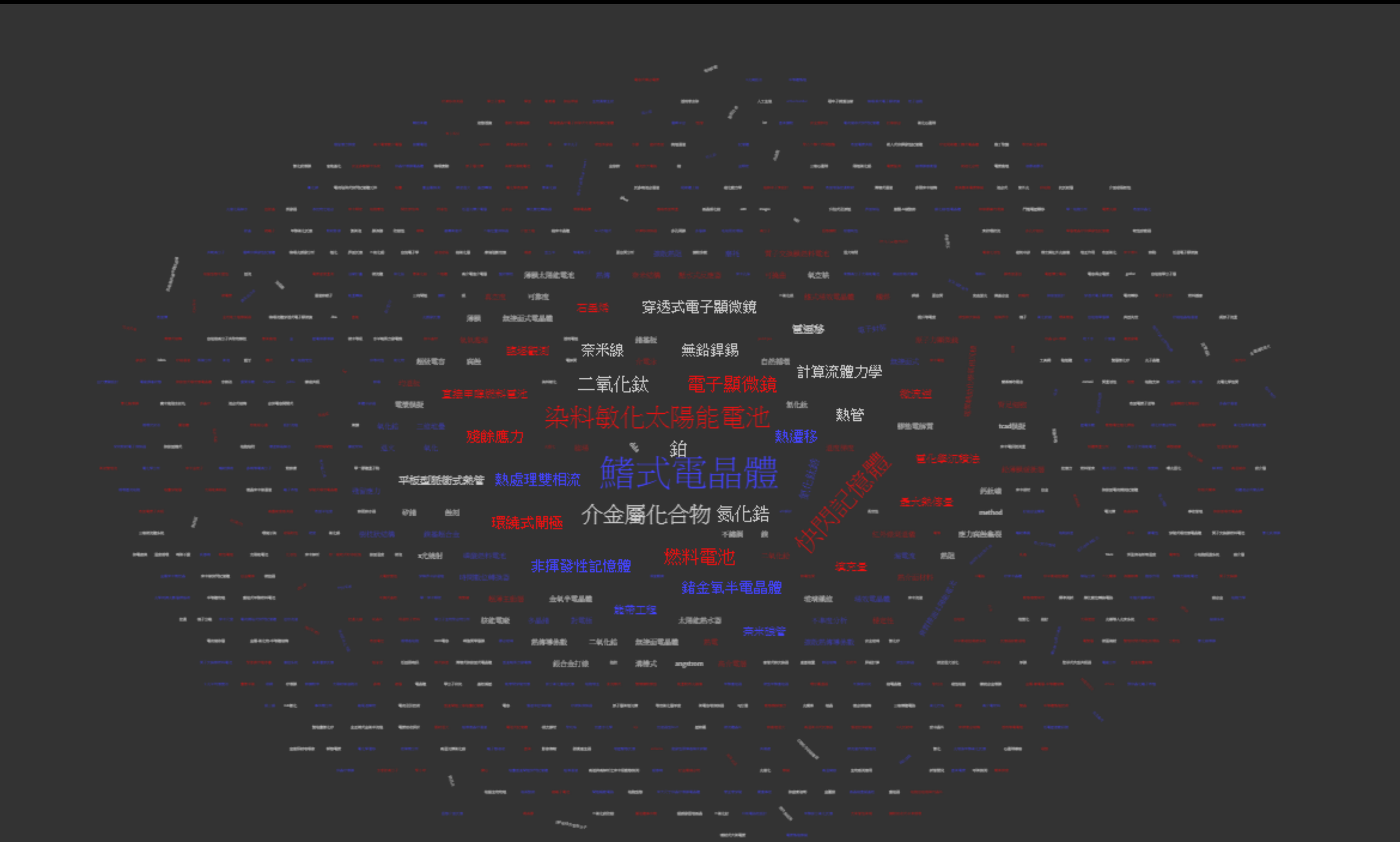


複合材料
微流體
機械性質
奈米碳管
自然對流
微機電系統
cmos-mems
晶格波茲曼法
工具機
石墨烯
模態分析
液靜壓軸承
發光二極體
計算流體力學
有限單元法
微流體晶片
共振頻率
殘留應力
紊流
潛變
介電泳
分子動力學
積層製造
量測系統
氣動引擎
有限單元分析
六維運動量測裝置
殘餘應力
環氧樹脂
果蠅
有限差分法

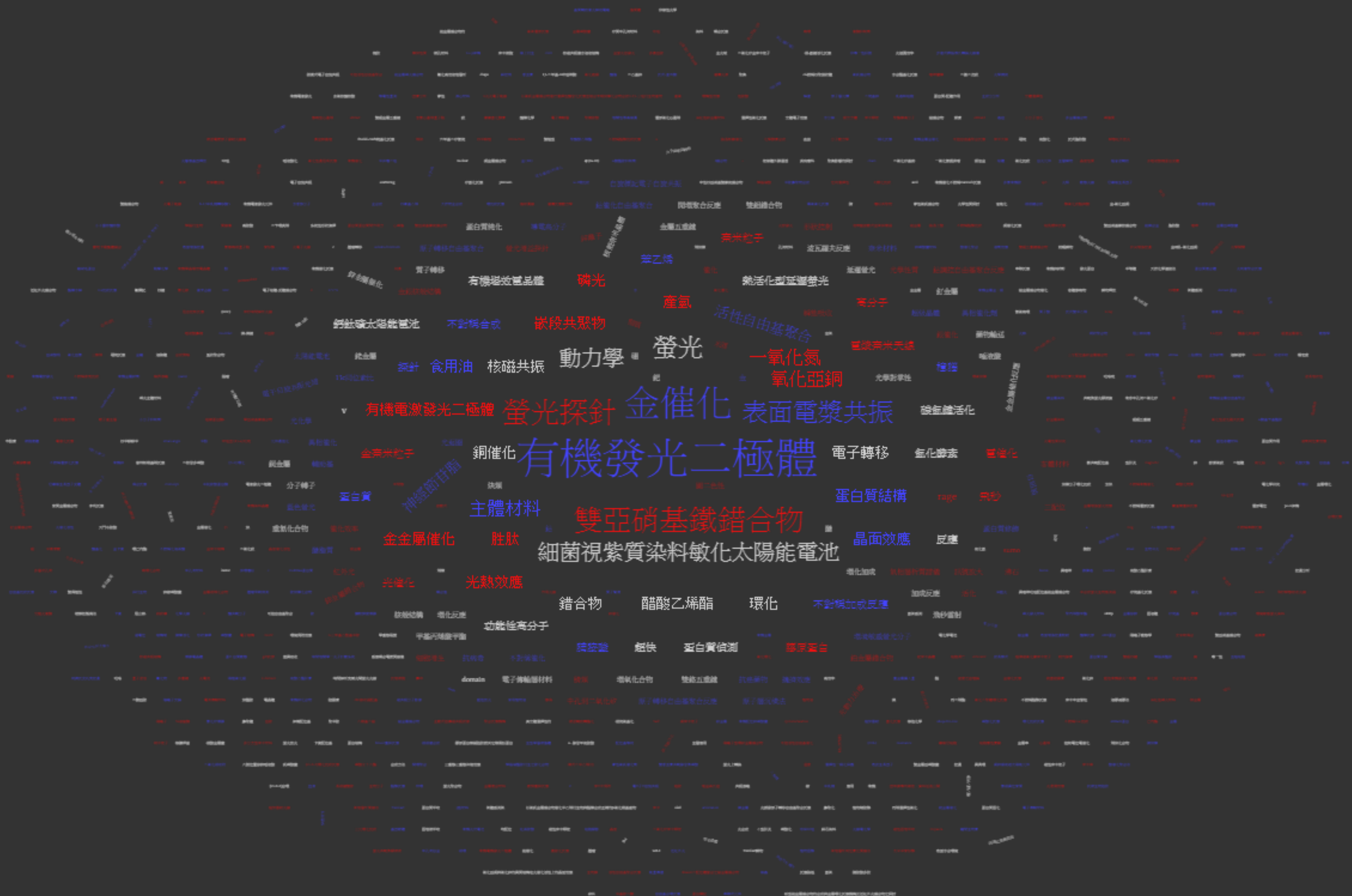
21.外國語文學系



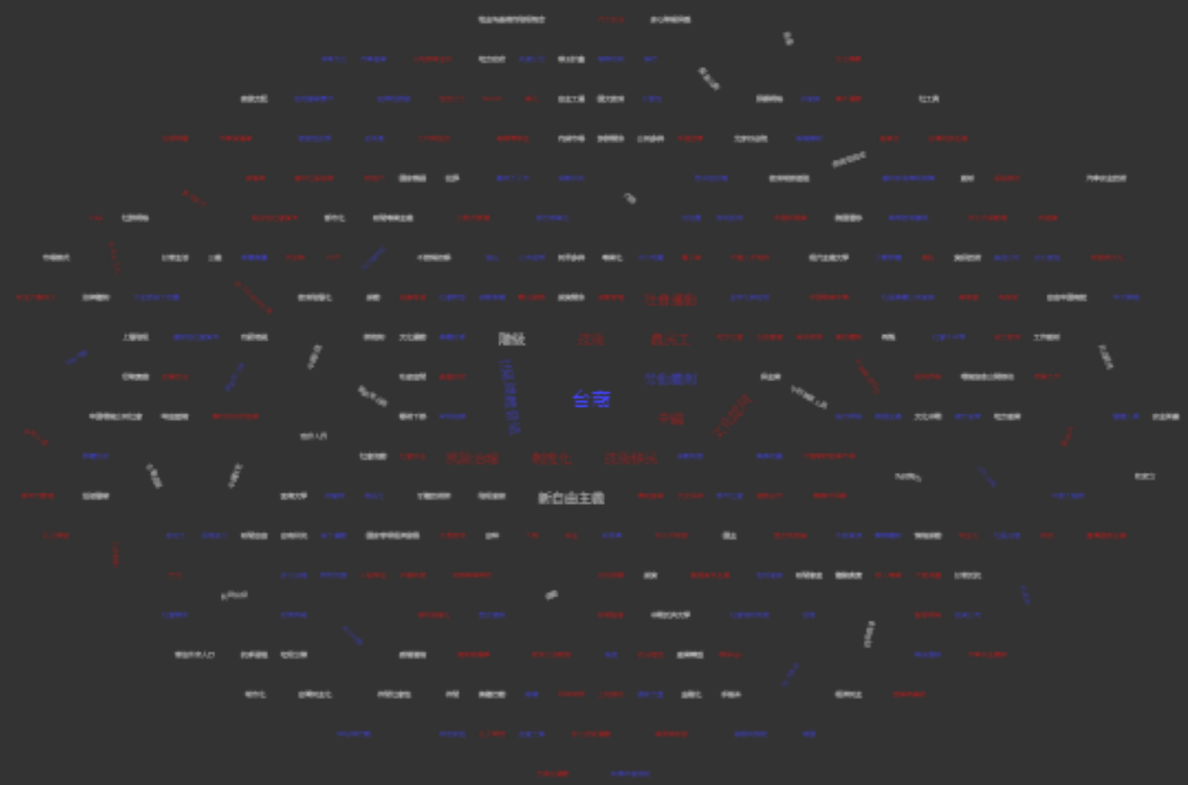
22.工程與系統科學系


穿透式電子顯微鏡
奈米線
無鉛銲錫
二氧化鈦
電子顯微鏡
計算流體力學
染料敏化太陽能電池
殘餘應力
熱管
熱遷移
鉑
鰭式電晶體
平板型脈衝式熱管
熱處理雙相流
介金屬化合物
氮化鋯
環繞式閘極
快閃記憶體
燃料電池
非揮發性記憶體
鍺金氧半電晶體
能帶工程
奈米碳管

23.化學系

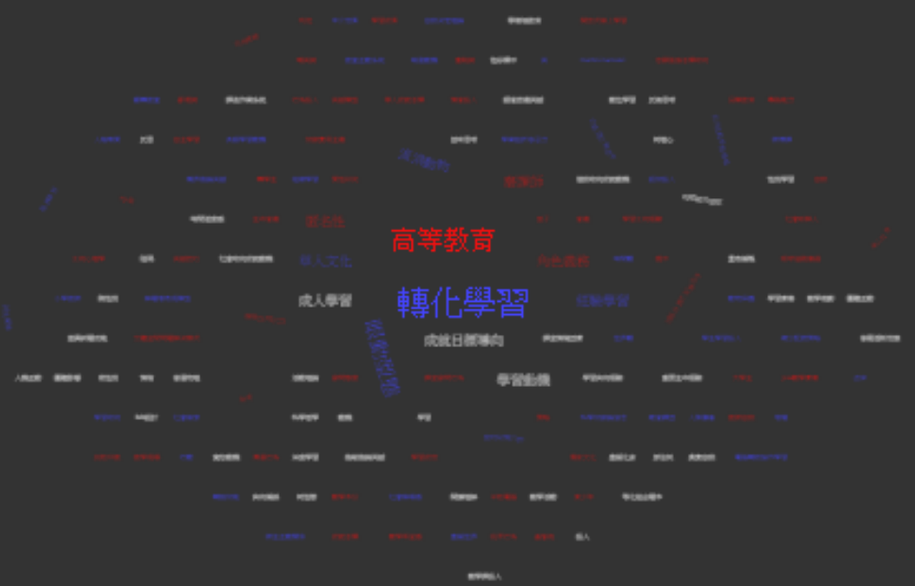

有機發光二極體
金催化
螢光探針
表面電漿共振
雙亞硝基鐵錯合物
細菌視紫質染料敏化太陽能電池
螢光
動力學
一氧化氮
氧化亞銅
有機電激發光二極體
核磁共振
食用油
銅催化
電子轉移
主體材料
蛋白質結構
晶面效應
金金屬催化
胜肽
光熱效應
錯合物
醋酸乙烯酯
環化
嵌段共聚物
磷光
產氫
活性自由基聚合

24.社會學研究所

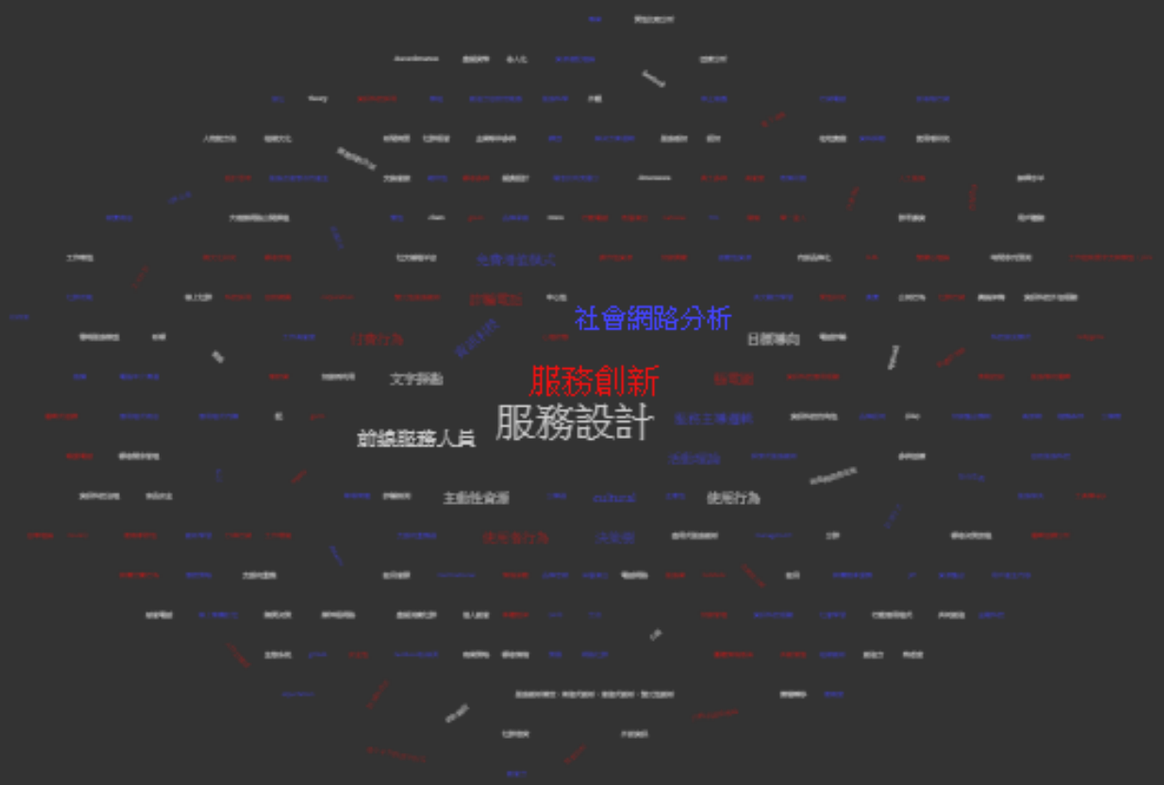


25.人類學研究所

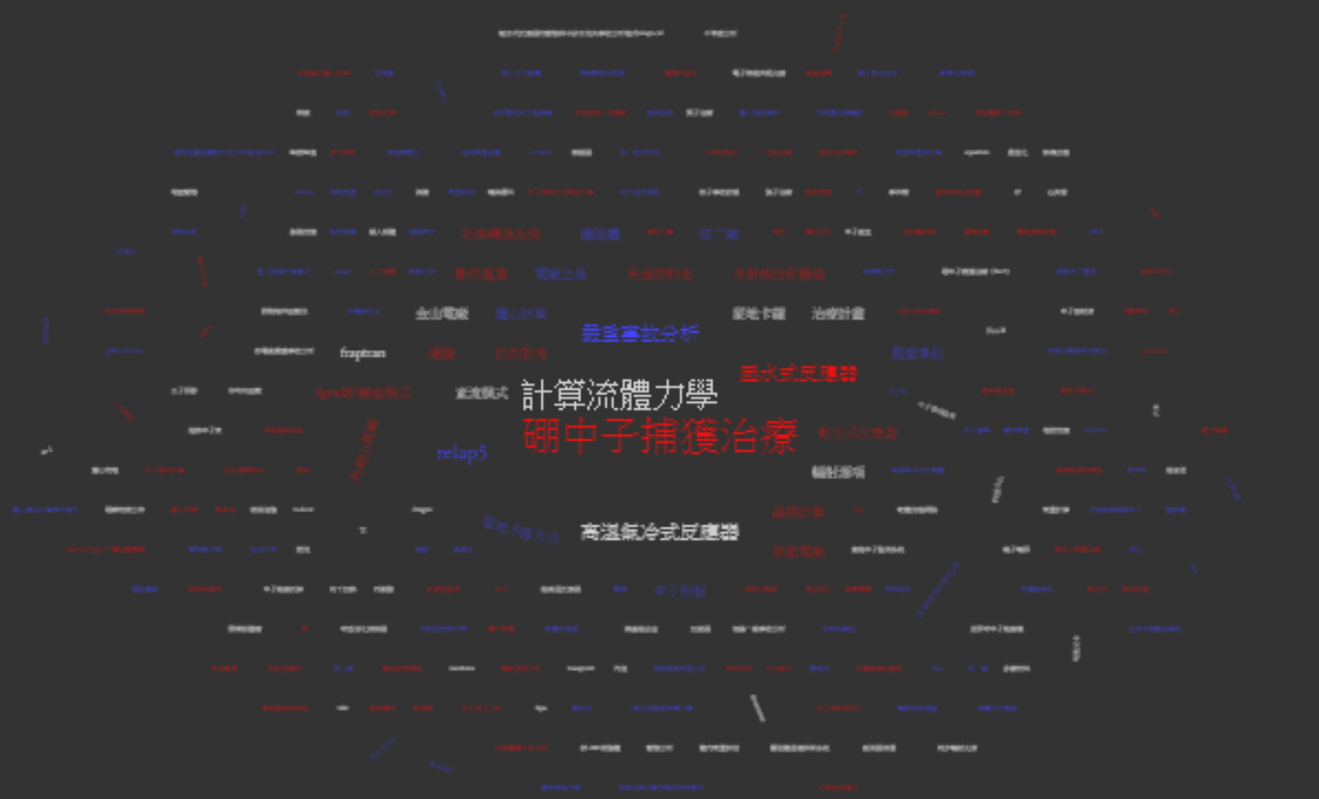
26.學習科學研究所



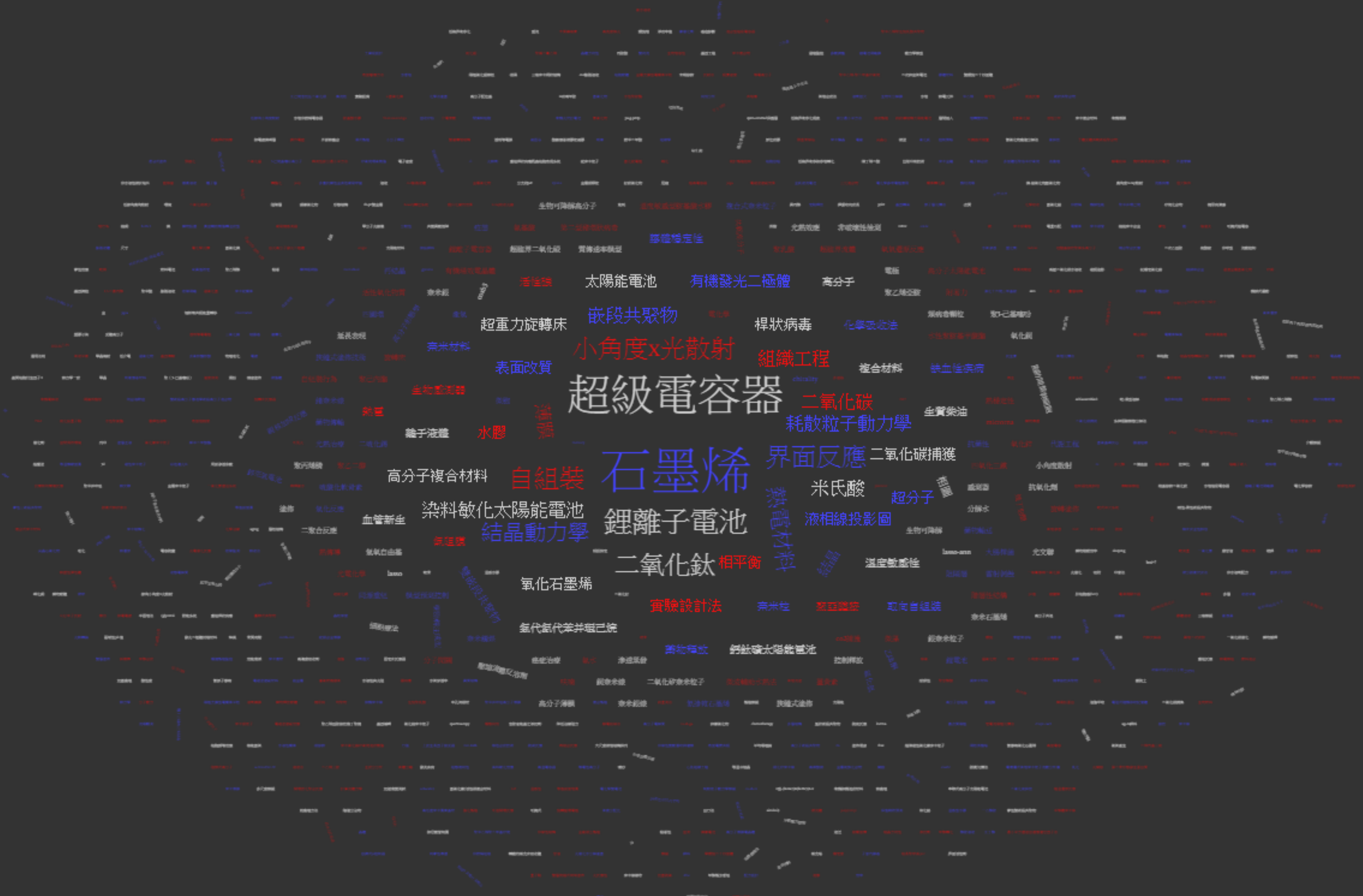
27.服務科學研究所



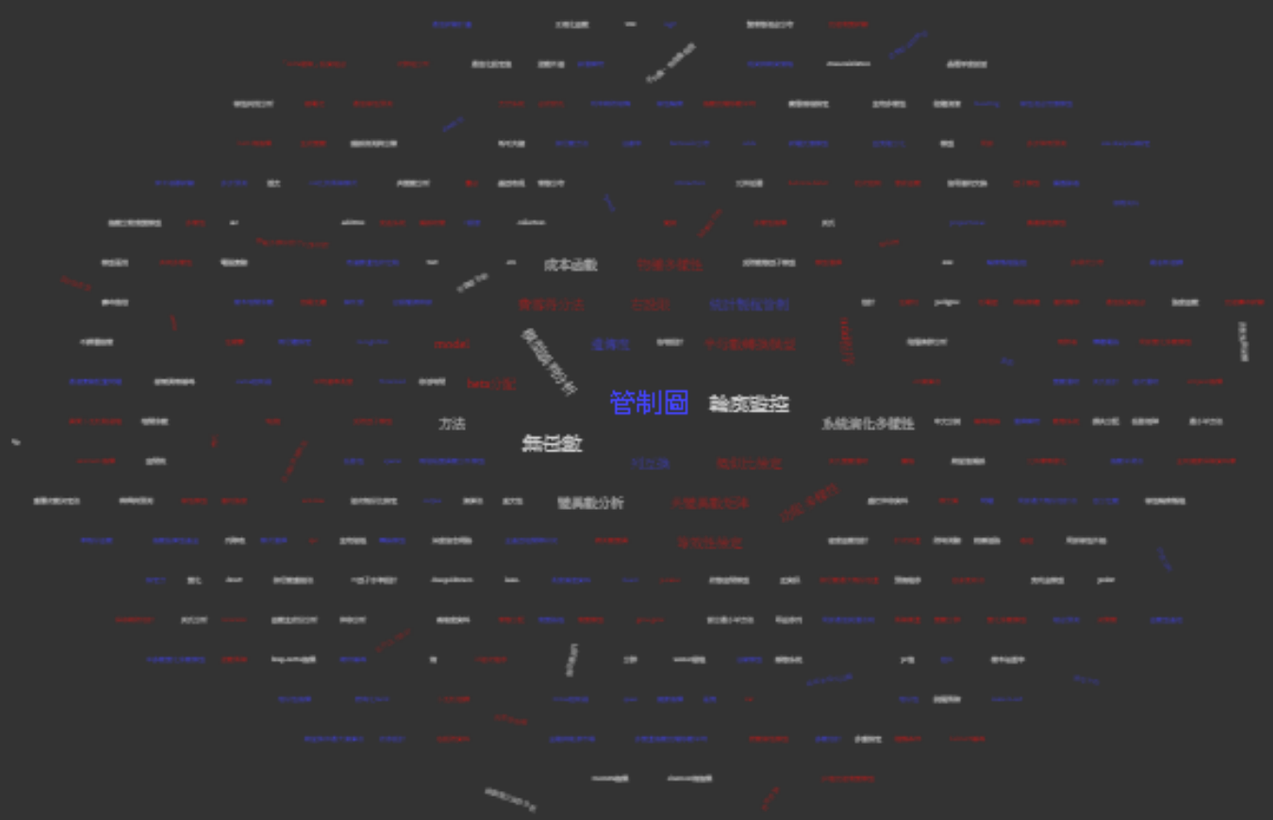
28.核子工程與科學研究所



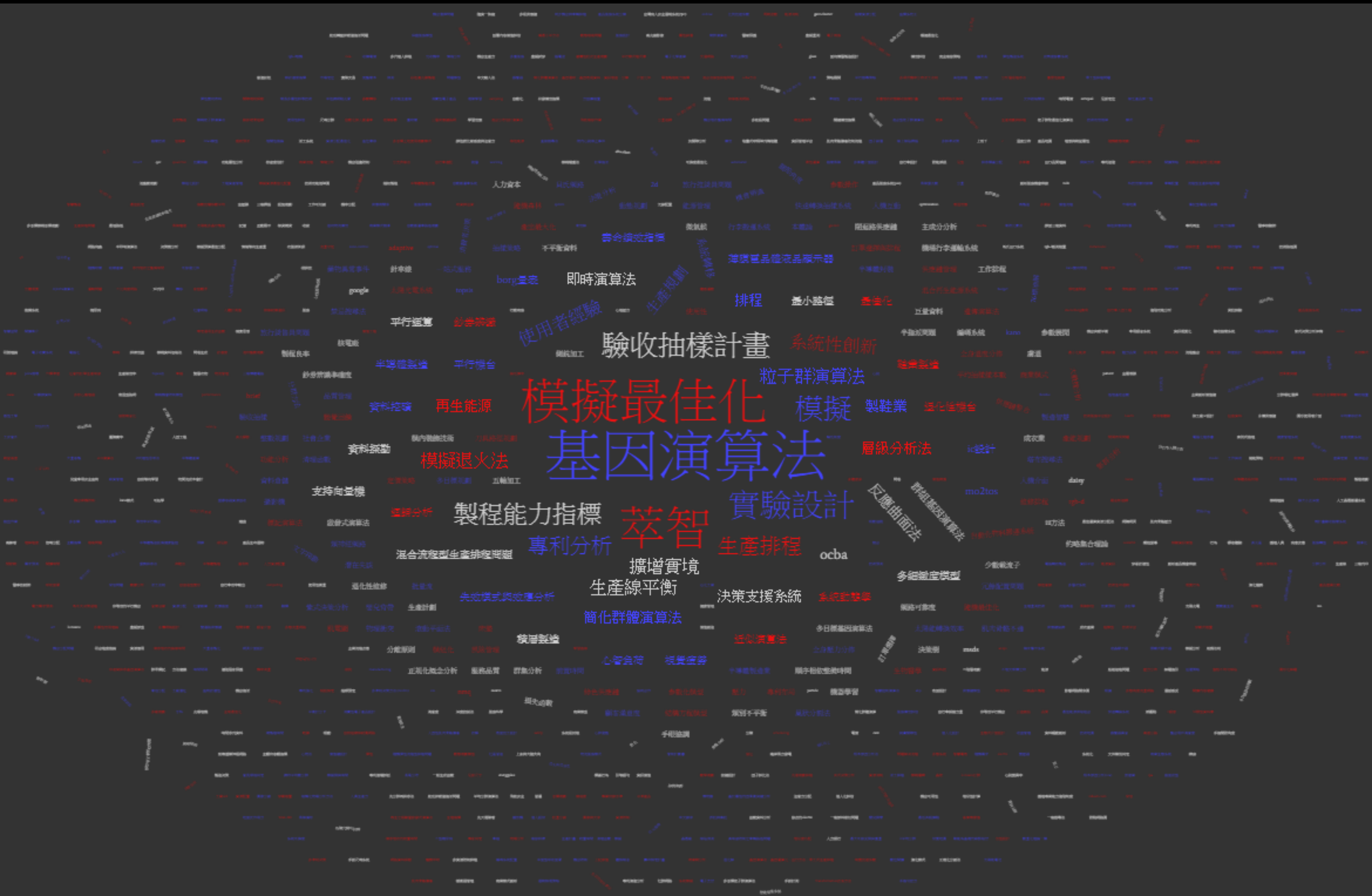
29.化學工程學系


石墨烯
超級電容器
鋰離子電池
二氧化鈦
界面反應
小角度x光散射
自組裝
組織工程
二氧化碳
耗散粒子動力學
嵌段共聚物
結晶動力學
熱電材料
染料敏化太陽能電池
太陽能電池
有機發光二極體
表面改質
米氏酸
超分子
液相線投影圖
高分子複合材料
氧化石墨烯
實驗設計法
相平衡
超重力旋轉床
桿狀病毒
二氧化碳捕獲
水膠
高分子
複合材料
生質柴油

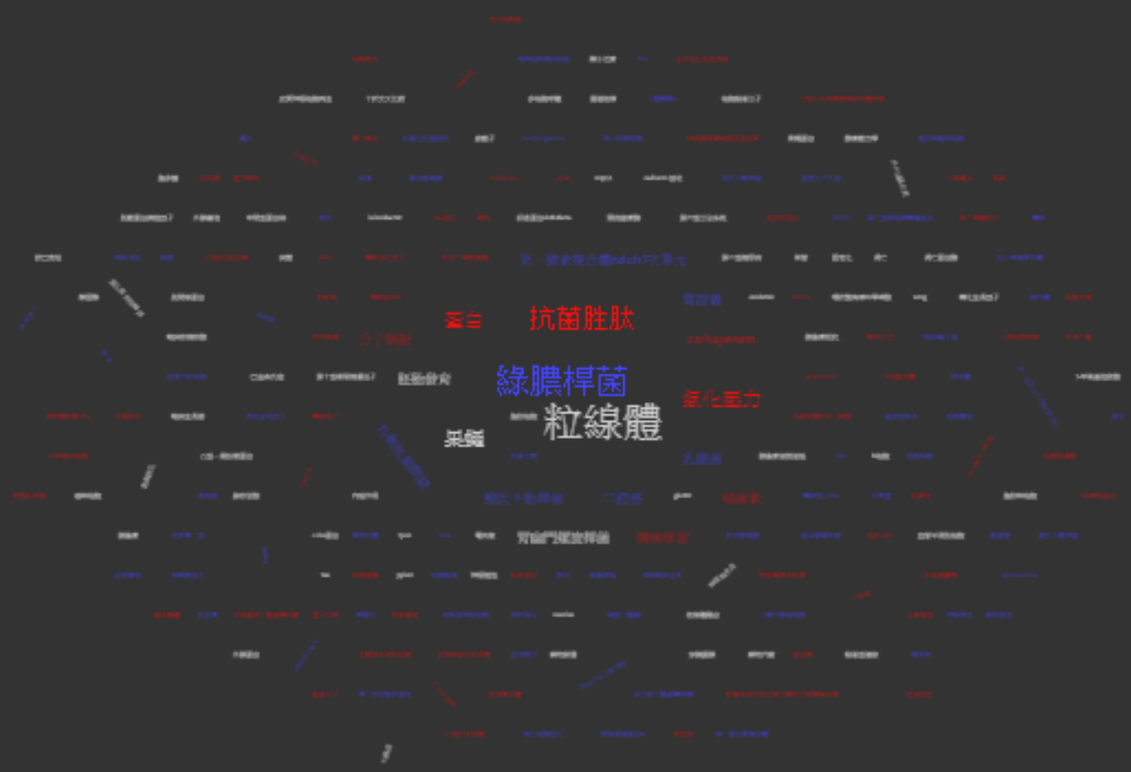
30.統計學研究所



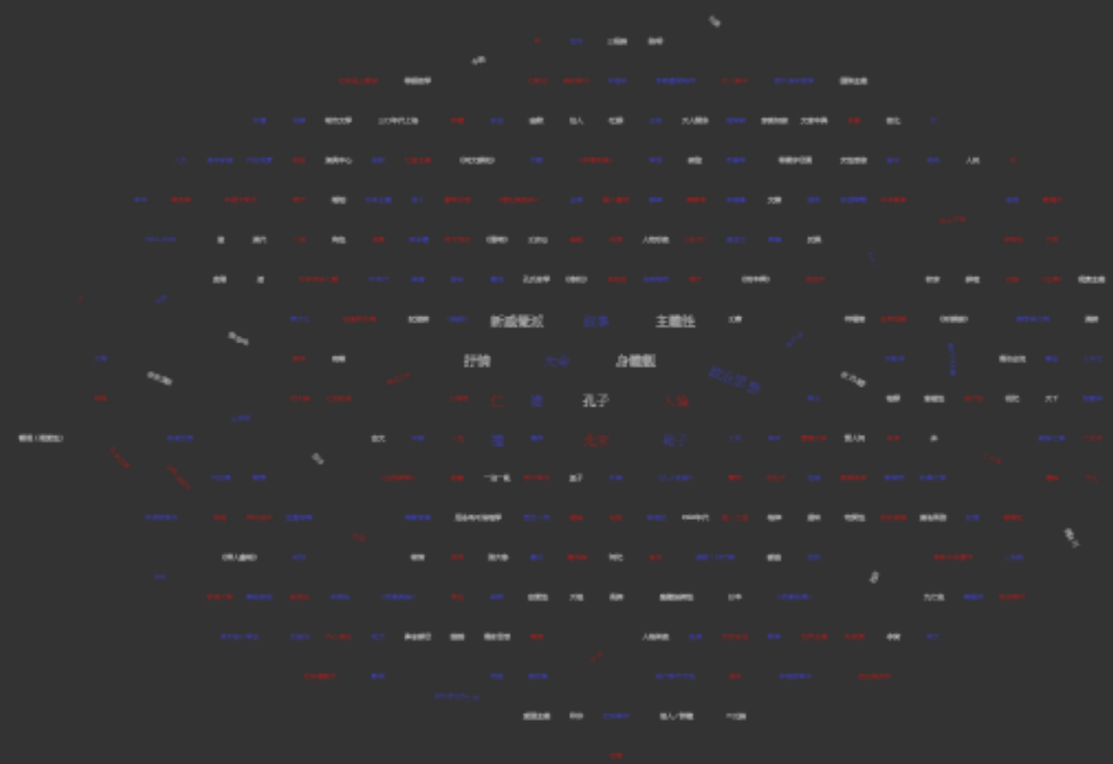
31.工業工程與工程管理學系



32.分子醫學研究所

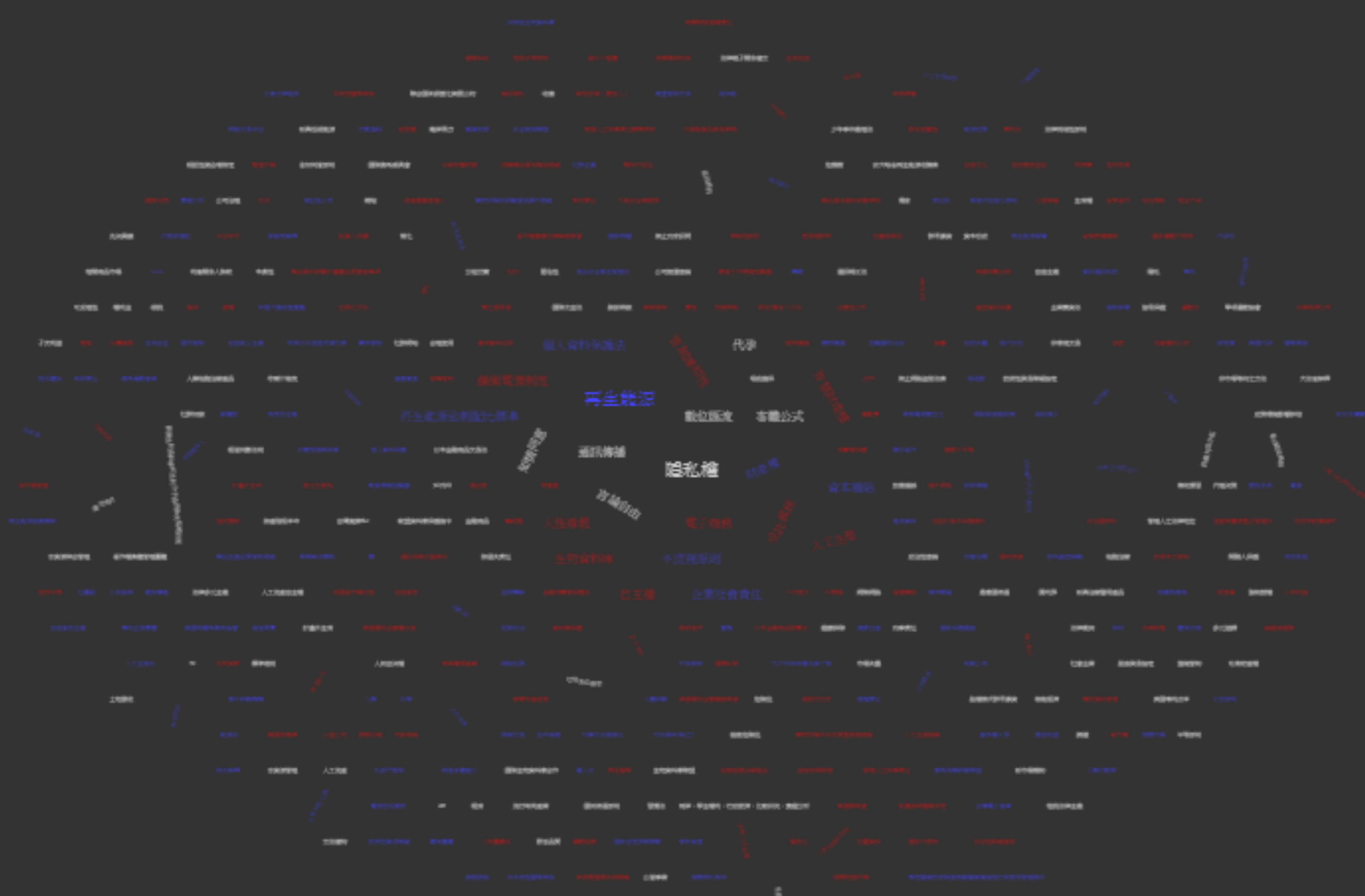


33.中國文學系

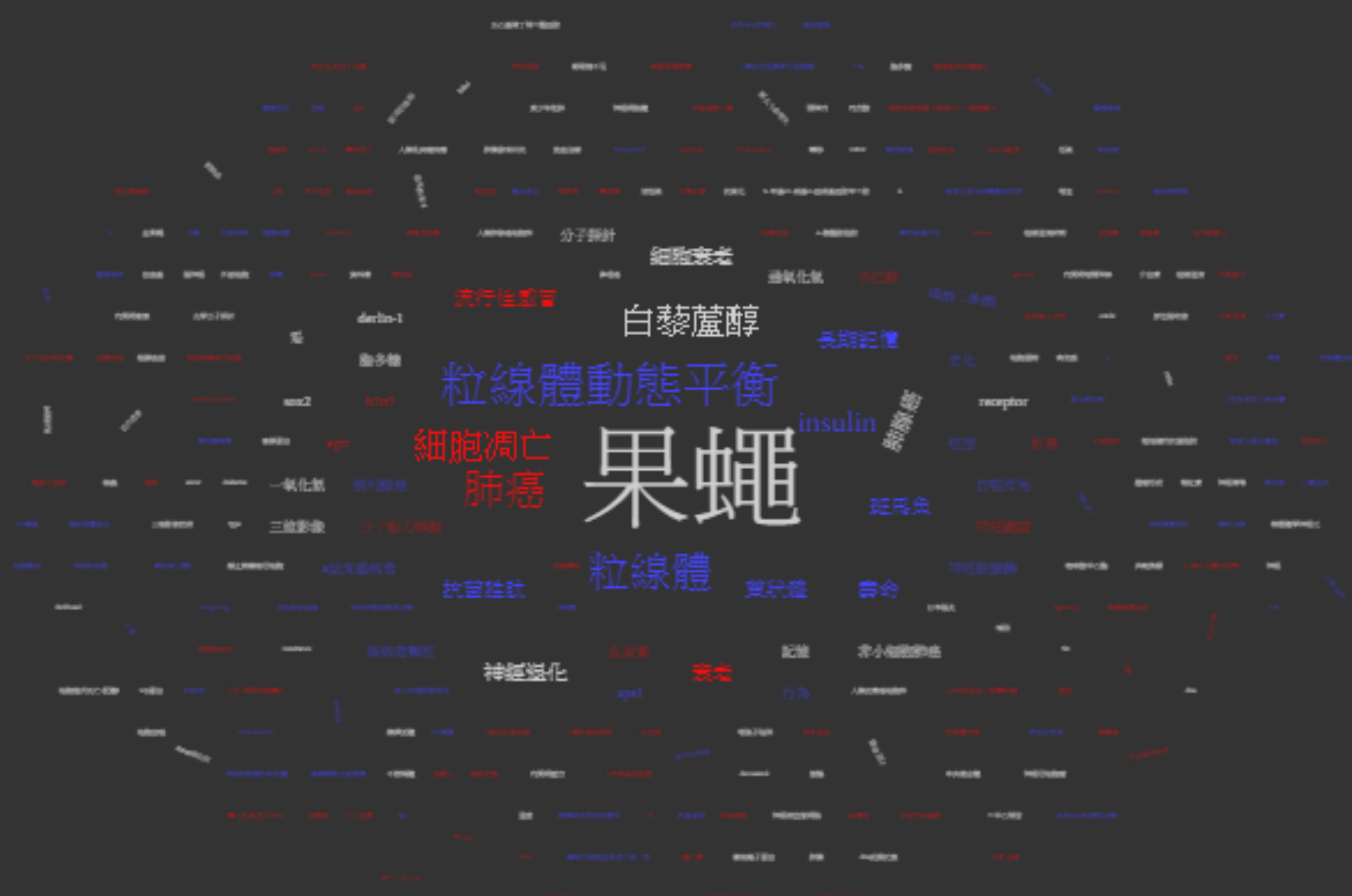


34.哲學研究所

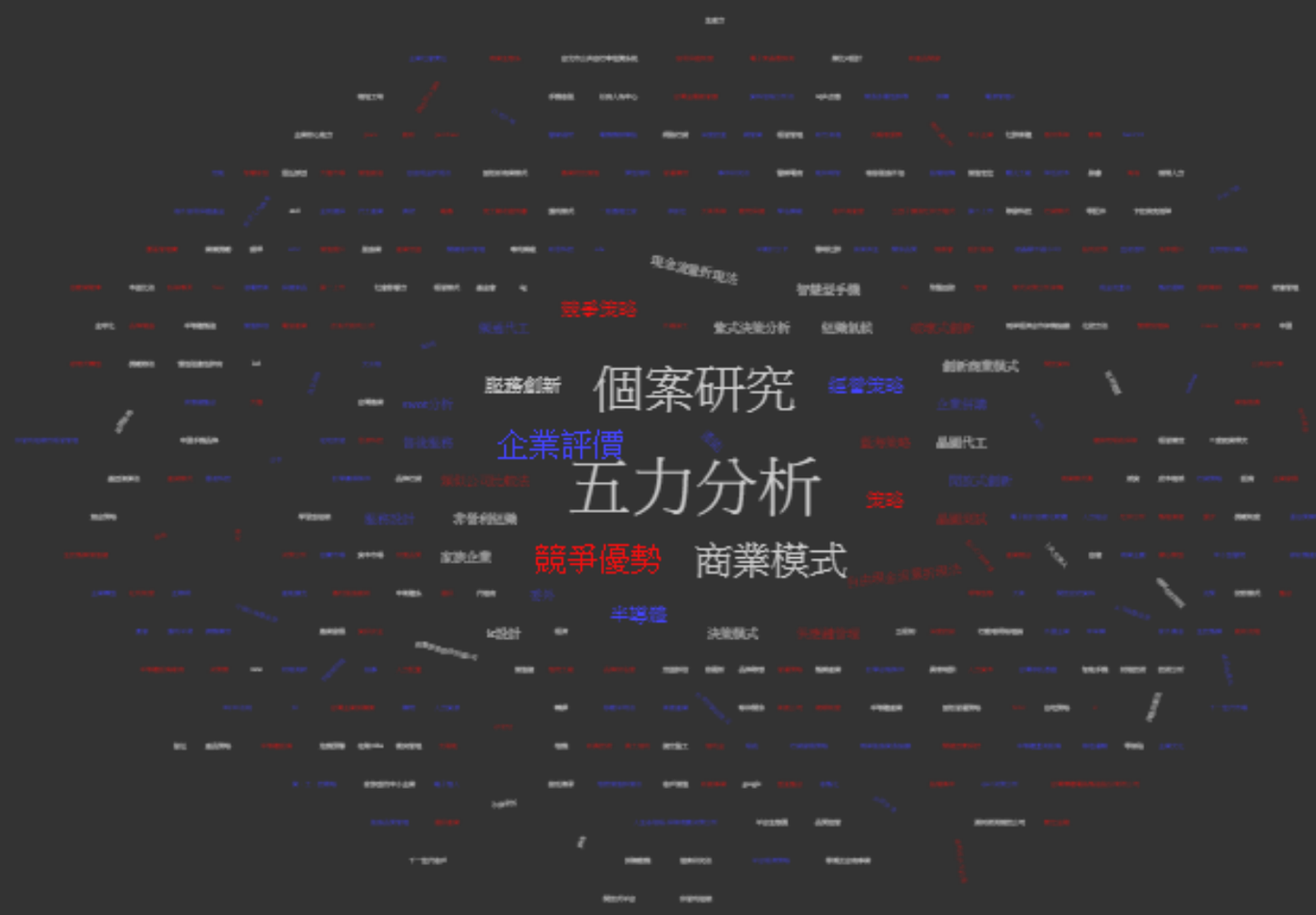
35.科技法律研究所



36.生物科技研究所

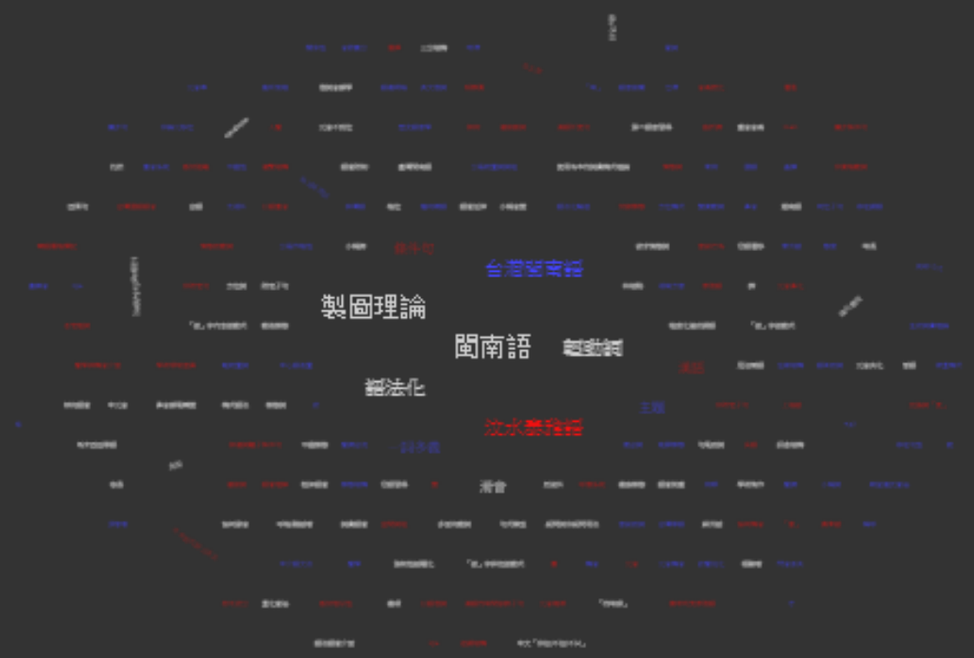


37.經營管理碩士在職專班

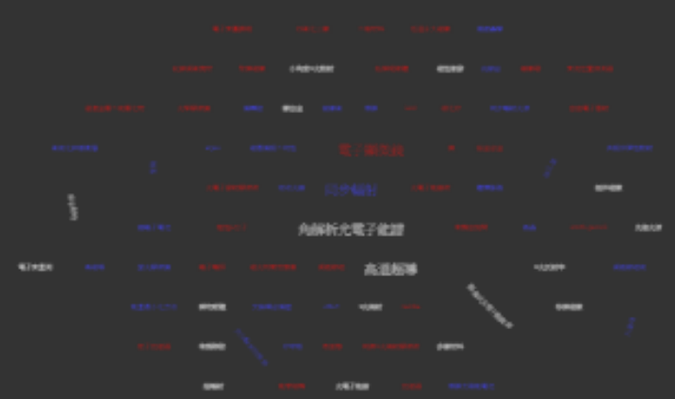


38.科技管理研究所

39.語言學研究所

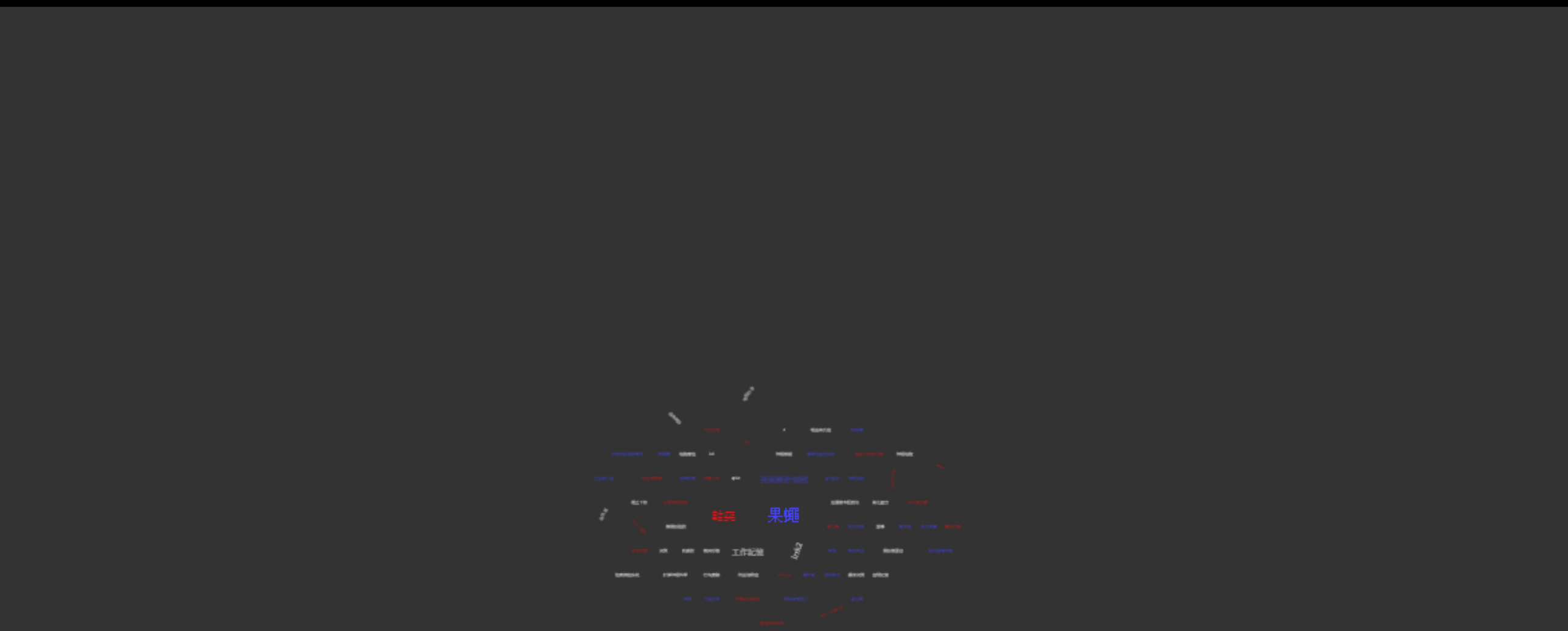


40.先進光源科技學位學程

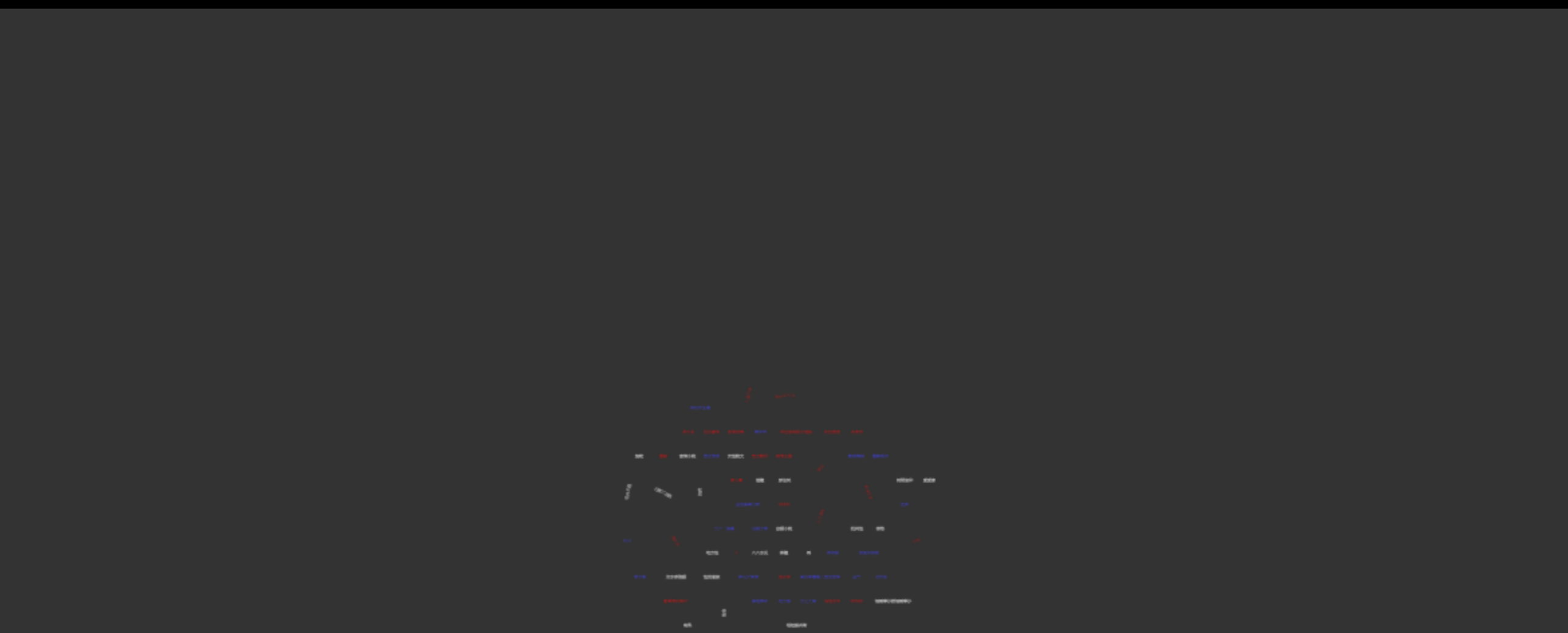


41.國際專業管理碩士班

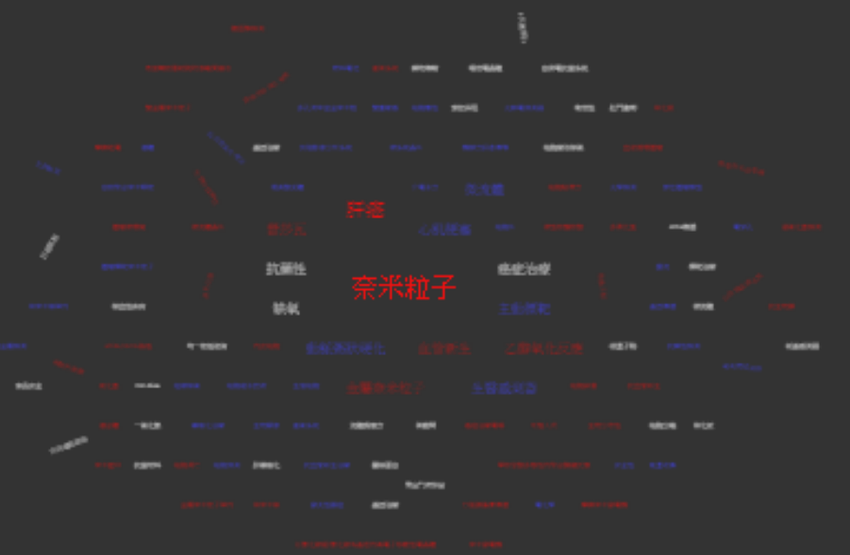
42.系統神經科學研究所



43.臺灣研究教師在職進修碩士學位班



44.生物醫學工程研究所



45.半導體元件及製程產業研發碩士專班

46.亞際文化研究國際碩士學位學程