

# 智慧未來休憩港灣應用服務：以高雄亞灣為例建構複合休憩知識圖譜 Intelligent Future Recreation Harbor Application Service: Taking Kaohsiung Asia New Bay as an Example to Construct a Composite Recreational Knowledge Graph

吳典志 Dian-Zhi Wu<sup>1</sup>, 呂昱德 Yu-De Lu<sup>1</sup>, 董家銘 Chia-Ming Tung<sup>1</sup>, 黃柏洋 Bo-Yang Huang<sup>1</sup>  
黃薰慧 Hsun-Hui Huang<sup>2</sup>, 林謙德 Chien-Der Lin<sup>2</sup>, 盧文祥 Wen-Hsiang Lu<sup>1</sup>

資訊工程學系, 成功大學<sup>1</sup>

地方創生處, 資訊工業策進會<sup>2</sup>

{ P76101495, ne6101068, P78081065, P76091666 }@gs.ncku.edu.tw

{ beatrice, chiender }@iii.org.tw, whlu@ncku.edu.tw

## 摘要

有鑑於台灣目前針對港灣休憩的整體專屬規劃服務缺少，多樣的海洋休憩活動和海上景點觀光也未能統合市港週遭服務進行規劃發展，加上針對港灣休憩服務進行軟硬整合而發展的新型態產品及應用服務不多，以及台灣的港灣休憩服務相關產業正面臨數位轉型挑戰。因此資策會提出創新「智慧未來休憩港灣應用服務」計畫，並委託本團隊以高雄亞灣為主要實證場域，提出複合休憩知識建構模型，運用多來源知識圖譜建構與推論技術推薦適切的休憩服務資訊，讓遊客能夠享受最佳的擬真智慧人機感知互動服務體驗。

## Abstract

In view of the lack of overall specialized design services for harbour recreation in Taiwan nowadays, various marine recreational activities and marine scenic spots haven't yet been planned and developed in the integration of services around the city and harbour. As there are not many state-of-the-art products and application services, and Taiwan's harbour leisure services-related industries are facing the challenge of digital transformation. Institute for Information Industry proposed an innovative "Smart Future Recreational Harbour Application Service" project, taking Kaohsiung Asia's New Bay Area as the main field of demonstration. Using multi-source knowledge graph integration and inference technology to recommend appropriate recreational service information, as a result, tourists can enjoy the best virtual reality intelligent human-machine interactive service experience during their trip.

關鍵字：高雄亞灣、港灣休憩、知識圖譜、智慧人機互動服務體驗、複合休憩知識建構模型

Keywords: Kaohsiung Asia New Bay Area, Harbour Recreation, Knowledge Graph, Intelligent Human-Machine Interactive Service Experience

rience, Complex Rest Knowledge Construction Model

## 1 緒論

### 1.1 背景

本研究為因應經濟部技術處法人科專分項計畫「智慧未來休憩港灣應用服務」需要，本團隊提出複合休憩知識建構模型，擷取非結構化資料並建構港灣休憩知識圖譜，同時設計開發合適的知識圖譜建構框架於港灣休憩智慧應用服務的技術研究。

### 1.2 動機

「智慧未來休憩港灣應用服務」計畫主要有鑑於台灣目前針對港灣休憩的整體規劃服務缺少，多樣的海洋休憩活動和海上景點觀光也未能統合市港週遭服務進行規劃發展。此外，針對港灣休憩服務進行軟硬整合而發展的新型態產品及應用服務不多，以及台灣的港灣休憩服務相關產業面臨數位轉型挑戰，需要藉由新興科技的導入。

因此，本計畫預計將整合高效能、高附加價值及高創新應用之服務與產品發展之綜合考量，規劃發展以「港灣體感科技旅遊」為主軸、高雄亞灣為主要實證場域，並融入元宇宙虛擬體驗進行核心技術研發，以完成遊艇載具為主的實船串接模擬機連動與環境擬真體驗之即時擬真智慧人機感知系統，提供使用者線上線下遊艇休憩體感模擬與跨域共遊體驗。打造全台首創體感港灣休憩服務，讓民眾體驗遊艇航行環境的真實呈現和港灣休憩即時互動服務內容，滿足遊艇休憩活動的擬真互動感需求，對海洋活動不再陌生畏懼、近而產生興趣。

### 1.3 方法

本計畫之系統架構圖如圖1所示，首先蒐集選定景點官方網站(詳見3.1)的網頁資訊，同時結合高雄市政府公開資訊平台<sup>1</sup>以及高雄港務

<sup>1</sup>高雄市政府公開資訊平台

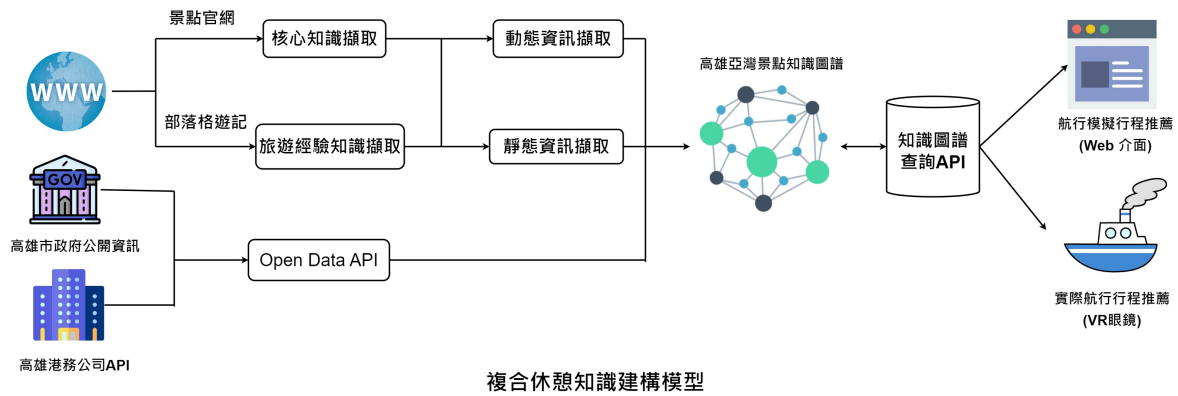


圖 1: 系統架構圖

公司<sup>2</sup>提供的港區船席現況 API 作為資料集，所蒐集的資料包含結構化、半結構化及非結構化資料，本研究提出複合休憩知識建構模型，擷取非結構化資料並建構港灣休憩知識圖譜，同時設計開發合適的知識圖譜建構框架於港灣休憩智慧應用服務的技術研究。

## 2 相關研究

### 2.1 知識圖譜建構

Google 於 2012 年提出知識圖譜的概念<sup>3</sup>，利用語意檢索從許多來源進行資訊蒐集並將其加入 Google Search 服務以優化其搜尋引擎。知識圖譜基本組成由「實體-關係-實體」三元組所建成，實體概念或詞彙代表節點 (node)、以關係代表邊 (edge)。透過自然語言處理 (Natural Language Processing, NLP) 的技術將文章內容進行分析以取出節點與邊的語意關係並建構出龐大的網路結構。知識圖譜的建構主要可以分成兩類，分別是通用型知識圖譜以及特定領域型知識圖譜，通用型知識圖譜最早可以追溯至 WordNet (Miller, 1992)，WordNet 是一個英文的語意網路，在 WordNet 中，名詞、動詞、形容詞和副詞各自被組織成一個同義詞的集合 (synsets)，每個同義詞集合都代表一個基本的語義概念，並且這些集合之間也由各種關係連接形成網路，後續基於 WordNet 的結構又衍生出許多不同的語言的語意網路甚至是多語言的語義網路，通用型知識圖譜除了常見的用於表達詞彙的定義、詞彙間的關係的語義網路以外，Vrandečić and Krötzsch (2014)，利用維基百科的資料建構了 WikiData 知識庫，藉由維基百科的共同協作性、多語性等特性，保證了知識圖譜的知識量會隨著時間持續擴增。另一方面，知識圖譜也被廣

<sup>2</sup>高雄港區停泊現況資訊

<sup>3</sup>Introducing the Knowledge Graph: things, not strings

泛應用在多個特定領域，例如醫療領域 (Gatta et al., 2017; Rastogi and Zaki, 2020)、教育領域 (Aliyu et al., 2020; Chen et al., 2018) 或是本計畫的主題旅遊領域。(Ling, 2020) 以中國的泰山作為核心景點並蒐集相關的景點資訊，例如文化背景、自然資源等等，建構出泰山旅遊導覽知識圖譜，DBtravel(Calleja et al., 2018) 利用擷取 Wikitravel<sup>4</sup>中的網頁資訊，並透過 Wikitravel 網頁中的特定 html 標籤結構建構出知識圖譜。目前現有之旅遊導覽知識圖譜大都多依賴於旅遊網站等結構化資料，這些客觀資訊較難反映出一個景點真實的旅遊經驗，本計畫將引入部落客遊記，當作使用者可參考的主觀資訊。

### 2.2 旅遊推薦系統

推薦系統近年來非常流行，最常見的應用在於電商平台，透過分析使用者的購買紀錄、瀏覽紀錄等資訊，用來預測並推薦消費者的偏好商品，促成交易 (Jain and Hegade, 2021; Zhao et al., 2022)。而隨著人們對於生活便利的強烈追求，更多領域引入了推薦系統，例如本計畫的旅遊領域，而使用知識圖譜進行旅遊行程推薦的系統更是少見，(Ling, 2020) 以泰山為核心景點建構知識圖譜並透過分析使用者的基本資訊、情境資訊、行為資訊去推薦使用者適合的泰山周遭景點或導覽資訊。本計畫則是以高雄亞灣為核心景點。

## 3 方法

本研究整合網際網路資源、高雄市政府公開資訊平台及高雄港務公司船舶停靠資訊 API 以擷取亞灣休憩知識圖譜所需知識。

如圖 1 所示網際網路資源包含選定景點的官方網站資源為核心知識擷取的內容及景點相關之部落格遊記文章為旅遊經驗知識擷取的內容。

<sup>4</sup>Wikitravel

選定景點	選定景點
亞灣遊艇碼頭	高雄展覽館
高雄 85 大樓	高雄港埠旅運中心
流行音樂中心-鯨魚堤岸	流行音樂中心-音浪塔
駁二藝術特區	蓬萊商港區(棧庫群)
哈瑪星鐵道文化園區	新濱碼頭
鼓山輪渡站	鼓山漁港
哨船頭遊艇碼頭	西子灣風景區
打狗英國領事館文化園區	高雄燈塔
旗津輪渡站	旗津老街
旗后漁港	海軍第四造船廠
夢時代購物中心	

表 1: 亞灣 21 景點

- 核心知識:  
 官方網站將依資料特性區分為景點基本資訊的靜態資訊及最新消息、活動及展演等各類活動的動態資訊。動態資訊將另行開發排程定期更新，並依活動啓迄日期進行知識圖譜的內容更新，進行中的活動可因活動日期已結束而修改其活動類別，同時也會新增圖譜內不存在的新增活動，以維護知識圖譜內容的正確性。
- 旅遊經驗:  
 本研究提出複合休憩知識建構模型分析及擷取部落格與選定景點有關之遊記以擷取旅遊經驗，輔以高雄市政府及高雄港務公司之公開資訊分別可提供餐館類別、住宿等級及高雄港船舶停靠資訊，以用於知識圖譜查詢回應清單及未來應用於 VR 以提供遊船於港區航行之數位分身以呈現與現實船隻活動相同視覺。

### 3.1 資料集

本研究由資策會選定高雄亞灣週邊景點為知識圖譜建構範圍如表1所列 21 個景點。景點資訊主要蒐集官方網站、旅遊網站景點介紹資訊(交通、住宿、美食等)以及痞客邦中與表1有關之部落格遊記文章內容。除前述網站資源外，本研究另行整理附近景點、餐廳及住宿等三類詞庫，以供複合休憩知識建構模型提升實體詞標記之正確性。餐廳清單選自愛食記網站中所列位於高雄的餐廳及其分類；附近景點之詞庫則蒐集 Tripadvisor 中所列與高雄港相關但排除計畫選定之 21 個景點之亞灣附近景點；住宿的詞庫則是匯整高雄市政府資料開放平台提供之高雄市民宿資料、高雄市一般旅館資料及 Tripadvisor 中所列高雄住宿所匯整之詞庫。

實體詞	HTML 擷取規則
簡介	article
	div.entry-content
	div.ContentPlaceholder1
營業時間	div.day
	div.time
地址	div.location
	footer-address
電話	div.phone
	footer-address
	data.top.info
活動開始日期	div.starttime
	valid-dates
活動結束日期	div.endtime
	valid-dates
活動內文	thecontent
	entry-content
圖片	img
	img.BigImg

表 2: 景點核心知識擷取規則範例

### 3.2 景點核心知識建構

景點核心知識擷取的內容以各景點官方網站為主要知識來源，官網資料依其內容尚可分為靜態景點介紹之基本資訊及隨時間推移而有所變動的動態訊息，包含最新消息、展覽資訊及活動等訊息。擷取景點核心資訊所使用的規則依如表2所示，依其活動時間及查詢時間動態連結成為景點的歷史活動、現在活動及未來活動等知識。以下將以網站內容的形態區分靜態知識與動態知識擷取的方法：

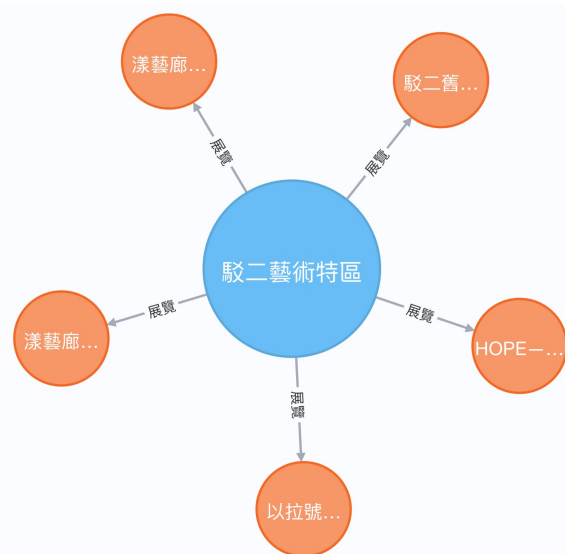


圖 2: 知識圖譜節點與關連範例

### 3.2.1 靜態知識擷取

知識圖譜的基本單元為「實體-關連-實體」，較常以句子中所擷取的名詞為實體詞來源，而動詞則是關連的來源，透過知識圖譜單元可建構出網狀如圖2的圖譜，在此範例中以駁二藝術特區串連各個展覽活動。

資料集所列景點官方網站這類型的内容其異動的頻率較低且能夠利用 HTML 標籤來擷取建構圖譜所需要的知識來源，並經由分析與擷取出文章句子中的實體詞及其關連以建立知識單元。

經過觀察各景點的網站原始碼，因各網站的開發與使用語言各不相同，因此本研究無法利用統一的 HTML 標籤來進行資料擷取，需依照各網站設定擷取資料的 HTML 使用規則，表2所列乃目前針對已蒐集的網站所建立的知識擷取規則。



圖 3: 駁二藝術特區最新活動截圖

<https://pier2.org/activity/info/1035/>

### 3.2.2 動態知識擷取

動態知識的來源為各官網的最新消息、展演及活動等有特定時效的知識，以圖3駁二藝術特區網站的最新活動消息為例，最新資訊的標題、簡介、開放時間，利用取出 HTML Tag 中所包含的 starttime 及 endtime 而獲得活動”駁遊路”的開始日期為 2022/09/03 與結束日期為 2022/09/04，加上網站 Title 取得活動名稱為「小人類 x 駁遊路 - 駁二蓬萊倉庫 | 駁二藝術特區」而建立活動名稱與開始與結束日期的知識。

### 3.3 旅遊經驗知識建構

本研究針對部落格旅遊文章内容的知識擷取重點定位在遊記中所提的延伸附近景點、餐飲美食及住宿，以供建立知識圖譜查詢應用的推薦清單。遊記的文章内容與景點官網的内容有較大的結構性差異，遊記主要以 Content 的内容為知識擷取來源而屬於處理非結構化文件。為提升知識擷取的效能，本研究提出複合休憩知識建構模型 (Complex Rest Knowledge

Construction Model, CRKC) 來分析遊記內文。CRKC 主要整合二大知識擷取方法，一為以自建的附近景點、餐廳及住宿三類詞庫搭配 CKIP CoreNLP 之複合休憩知識建構模型-實體詞辨識 (CRKC-NER) 以取得本研究感興趣的實體-關係-實體。另一方法則是應用 CKIP 中文句法剖析結果歸納出可識別主詞與受詞之語意角色集之複合休憩知識建構模型-句法剖析 (CRKC-Parser)。

#### 3.3.1 複合休憩知識建構模型-實體詞辨識 (CRKC-NER)

本研究為有效提升擷取旅遊經驗的知識，鎖定這類文章中提及的景點、餐飲及住宿等三類以供後續知識圖譜的推薦應用。因此，本研究以人工方式整理前述三類的正式名稱、同義詞、縮寫等建構出三類的詞庫，以提升實體詞辨識之準確率。以下分別說明每一類別詞庫建立的方法：

- 附近景點:  
此類詞庫的主要來源為 TripAdvisor 並加入遊記中所見的景點一般常用詞、簡稱等所組成的詞庫。例如在 TripAdvisor 中搜尋”駁二藝術特區”的景點，可以獲得”誠品書店-駁二店”、”北極殿”、”大銘高級腕錶”等 200 個附近景點。
- 餐廳: 愛食記網站中羅列台灣各地區最新、最流行的美食餐廳，並可了解網友對餐廳的評價及連結至部落客專業食記。因此本研究選定愛食記-高雄地區的餐廳為此類詞庫的資料來源。
- 住宿: 民宿與飯店的詞庫建立主要參考高雄市政府公開資料平台中所列的星級飯店、民宿清單及 TripAdvisor 而來。例如公開資料的高雄民宿可獲得民宿的登記編號、中文名稱、地址、電話及房間總數等資料，一般旅館資料可取得類別、星等、住宿名稱、地址、電話、房間數、電子郵件、網址、經緯度等資訊。

#### 3.3.2 複合休憩知識建構模型-句法剖析 (CRKC-Parser)

本研究利用 CKIP Parser<sup>5</sup>對部落格文章進行中文句法剖析，並經由觀察許多文章及分析句法語意角色内容以整合歸納出表3的内容來連結語意角色與主詞及受詞的關係。並應用於演算法1的主詞語意角色集 (Subject Semantic Role, SSR) 及受詞語意角色集 (Object Semantic Role,

<sup>5</sup>CKIP 中文剖析。

<http://parser.iis.sinica.edu.tw/>

OSR) 二項變數中。

如圖4所示。以「頂樓空中花園俯瞰高雄展覽館與高雄港」為例，可擷取出語意角色為 agent 的主詞片語「頂樓空中花園」及語意角色為 goal 的地方詞片語「高雄展覽館與高雄港」搭配剖析樹所得的動詞「俯瞰」而組成知識圖譜所需的 SVO 三元組。

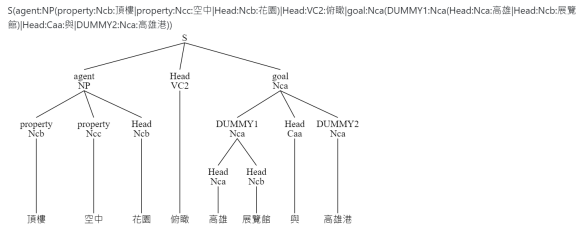


圖 4: CKIP 中文剖析樹

關係類型	語意角色標籤
主詞關係	Agent
	Causer
	Experiencer
	Topic
	Theme(句首)
受詞關係	Benefactor
	Goal
	Target
	Theme(句尾) Range

表 3: 語意角色之主詞及受詞關係

## 4 實驗

本章節將介紹初步實驗結果，4.1 說明已蒐集的資料集數量；4.2 說明景點核心知識圖譜的建構過程與結果；4.3 說明旅遊經驗知識圖譜的建置成果；4.4 則舉例說明提出方法的錯誤分析。

### 4.1 資料集

本研究依所列的 21 項景點之官方網站、旅遊網站之景點介紹、市政府旅遊觀光局處資料以建立靜態景點知識、動態展覽與活動及個人旅遊經驗及推薦等資料來建構出本研究所需的景點核心與旅遊經驗知識圖譜，目前已蒐集的各類文章如表4所示。

資料來源	資料類型	數量
景點資訊	靜態	11
	動態	228
部落格遊記	靜態	1100

表 4: 資料集分布

### Algorithm 1 剖析樹語意角色節點知識擷取法

**Input:** Sentence parsing result  $SP$

**Output:** Subject  $S$ , Extracted Object  $O$ , Extracted Object  $V$

- 1:  $SSR$ : Semantic role set of subject [agent, causer, experiencer, topic, theme (句首)]
- 2:  $OSR$ : Semantic role set of object [benefactor, goal, target, theme (句尾)]
- 3:  $V\_POS$ : Verb POS tagging set [VA, VAC, VB, VC, VCL, VD,...]
- 4:  $S \leftarrow []$
- 5:  $O \leftarrow []$
- 6:  $V \leftarrow []$
- 7: **for**  $token_i$  in  $SP$  **do**
- 8:     **for**  $token_i$  in  $SSR$  **do**
- 9:          $S \leftarrow token_i$
- 10:     **end for**
- 11:     **for**  $token_i$  in  $OSR$  **do**
- 12:          $O \leftarrow token_i$
- 13:     **end for**
- 14:     **if**  $token_i$  belong to "HEAD" && "V\_POS" **then**
- 15:          $V \leftarrow token_i$
- 16:     **end if**
- 17: **end for**

爲了優化亞灣休憩的知識圖譜建構，特別蒐集整理部落格文章內提到亞灣周遭並排除表1所列 21 景點的景點、餐廳及住宿之清單建構亞灣休憩延伸詞庫，以擷取出更多的知識，表5爲附近景點、餐廳及住宿三類詞庫的資料來源及人工整理之數量，未來將可透過人工方式新增詞庫內容。

資料標籤	資料來源	數量
附近景點	Tripadvisor	636
餐廳	愛食記-高雄	5813
住宿/飯店	Tripadvisor + 市府公開資料平台	342

表 5: 資料集分布

### 4.2 中文斷詞及詞性標記效能評估

中文斷詞與詞性標記的結果將影響後續知識擷取的成果，本研究隨機挑選部落格文章句子並使用史丹佛 CoreNLP<sup>6</sup>及中研院中文詞庫小組提供的 CKIP CoreNLP<sup>7</sup>、CKIP Transformers<sup>8</sup>套線上展示系統來比較各系統斷詞與詞性標記結果。

<sup>6</sup>Stanford NLP Group

<sup>7</sup>CKIP CoreNLP

<sup>8</sup>CKIP Transformers

圖5為 CKIP CoreNLP 斷詞結果，可見斷詞的結果可以比較貼近台灣的用詞，例如雞排、西子灣、狼腳色等詞彙。CKIP Transformers 所得的斷詞結果與 CKIP CoreNLP 相同。相同的句子使用史丹佛 NLP Group 的分析後可得到如圖6的結果，可以發現在斷詞部份就產出較多的錯誤。錯詞的斷詞包含林媽媽雞、排、西子、灣、個狼腳色。因此，本研究後續主要使用 CKIP CoreNLP 來進行中文斷詞及詞性標記。



圖 5: CKIP CoreNLP 斷詞範例

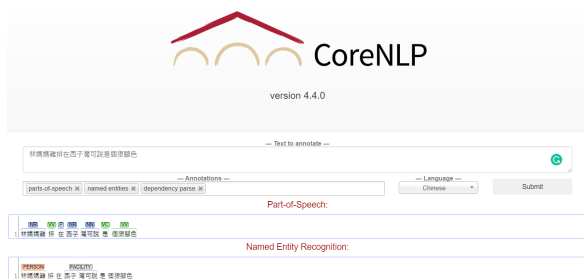


圖 6: Stanford CoreNLP 斷詞範例

### 4.3 旅遊經驗知識建構效能評估

#### 4.3.1 實體詞擷取效能評估

實體詞擷取的結果將可影響知識擷取的結果，本研究評比 Stanford CoreNLP、CKIP Transformers、CKIP CoreNLP 及本研究提出之 CRKC-NER，由表6可了解 CRKC-NER 可獲得最佳的實體詞擷取結果。

方法	Precision	Recall	F1
Stanford CoreNLP	0.56	0.41	0.47
CKIP Transformers	0.67	0.93	0.78
CKIP CoreNLP	0.71	0.87	0.78
CRKC-NER	0.94	0.73	0.82

表 6: 實體詞擷取方法效能比較

舉例”新濱駁前將高雄的舊三和銀行改建成咖啡店，是高捷西子灣/輕軌哈瑪星站很受歡迎

迎的高雄咖啡廳”應該要能正確辨識出的實體詞有”新濱駁前”、”高雄”、”三和銀行”、”高捷西子灣/輕軌哈瑪星站”、”高雄咖啡廳”等實體詞。

利用前述四種方法比較實體詞擷取結果，表7可以看到 Stanford CoreNLP 僅能正確辨識一個實體詞，CKIP Transformers 與 CKIP CoreNLP 二者辨識實體詞的數量一樣，但實體詞有些許落差。例如 Transformers 可以正確辨識出”三和銀行”，但 CoreNLP 則無法辨識出相同的實體詞，但 CoreNLP 可以正確辨識出”高雄咖啡廳”而 Transformers 則只能辨識出”高雄”，CRKC-NER 則因多加詞庫的影響，僅有”新濱駁前”未識別出。”新濱駁前”實體詞正確的名稱是”新濱·駁前”，此為一家新開幕的餐酒館。經由維護詞庫將”新濱·駁前”及其同義詞”新濱駁前”加入餐廳類別的詞庫後 CRKC-NER 即可辨識出五個實體詞。

方法	正確	類別錯誤	未識別
Stanford CoreNLP	1	3	1
CKIP Transformers	3	1	1
CKIP CoreNLP	3	1	1
CRKC-NER	4	0	1

表 7: 例句”新濱駁前將高雄的舊三和銀行改建成咖啡店，是高捷西子灣/輕軌哈瑪星站很受歡迎的高雄咖啡廳”實體詞擷取方法的正確詞彙比較

#### 4.3.2 「實體-關連-實體」關係擷取效能評估

為驗證本研究提出的 CRKC-Parser 於旅遊經驗知識擷取的成果，使用 Stanford CoreNLP 提供的 Dependency Parser 所擷取到的實體詞及其關連來取得「實體-關連-實體」的關係進行效能評比。本研究隨機選取資料集中的 50 個句子來進行效能評比，比較結果如表8所示。可發現當作對照組的 Stanford CoreNLP 的方法對於實體的擷取結果較不理想，例如「駁二藝術特區原本是舊倉庫」經由 Stanford Core NLP 處理後取得一個 PERSON 的實體詞”駁二藝術特區”，而 CKIP 則取得一個行政區的實體詞”駁二藝術特區”，相較之下 CKIP 較能貼近在地使用體驗。

方法	Precision	Recall	F1
Stanford CoreNLP	0.26	0.18	0.21
CRKC-Parser	0.63	0.35	0.45

表 8: 旅遊經驗知識建構之「實體-關連-實體」關係擷取效能評估

### 4.4 亞灣休憩知識圖譜建構與系統展示

本研究目前初步利用駁二藝術特區、高雄展覽館等景點來蒐集官網、遊記文章以建立知識圖

譜的雛形系統並利用 Neo4j 建構出本系統的核心可視化知識圖譜，可視化的呈現可提供人員可以方便快速的查閱知識圖譜所有節點與關係，圖7即為以哈瑪星鐵道文化園區為中心所展示的圖譜節點與關係，系統維護人員可容易的查看目前知識圖譜的內容及可透過 API 進行人為的圖譜知識節點與關係的新增、刪除與修改等維護作業。目前進度已蒐集高雄展覽館、駢二藝術特區、哈瑪星鐵道文化園區等 5 個主要的景點進行初期建構與驗證。

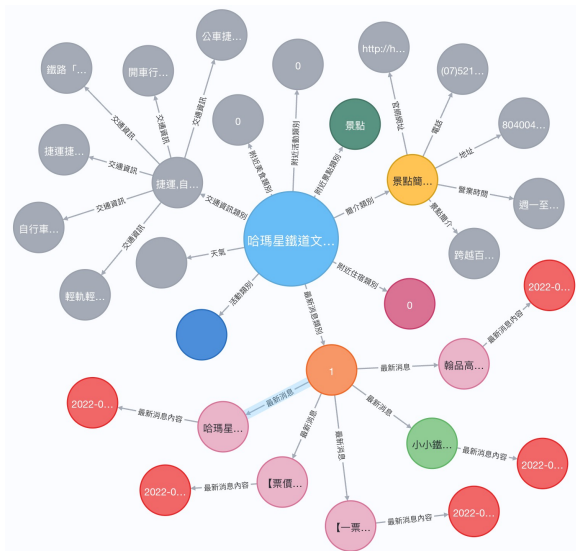


圖 7: 可視化知識圖譜

本系統設計一套三階知識圖譜查詢的 API 邏輯以降低前端應用程式查詢時需要回傳大量資料，以加速系統的回應時間。詳細的 API 設計邏輯如下：

- L1-景點基本資訊: 本階層目的在提供遊客該景點基本簡介及連結的知識，包含景點交通、住宿、餐飲及活動、展覽等類別供遊客深入探索。

- 輸入: 高雄展覽館
- 輸出:

- \* 高雄展覽館 (KEC) 位於亞洲新灣區核心，是台灣第一座多功能的臨港會展中心，由經濟部投資興建，安益國際展覽集團營運，致力打造一個國際商業與貿易的平台，推動產業聚落永續發展，引領會展城市經濟邁向另一個高峰。多功能的室內、戶外以及水岸活動空間及會議中心，可激發創意盡情馳騁，全方位滿足會展需求。2021 年導入 5G 專頻專網建設與創新會展應用，轉型為高

科技多功能場域，以「場館即平台、科技即服務」的新目標定位，透過會展商務推廣，推動產業數位轉型，發展成為 5G AIoT 科技新創應用商機拓展的大平台。

- \* 官網: <https://www.kecc.com.tw/zh-tw/kec>
- \* 地址: 高雄市前鎮區成功二路 39 號
- \* 電話: 07-2131188
- \* 營業時間: 週一至週日: 10:00 - 18:00
- \* 活動: 現在活動、歷史活動、未來活動
- \* 交通: 汽車、捷運、公車
- \* 美食類別: 中式料理、日本料理、韓式料理、美式料理、義式料理、泰式料理、港式料理
- \* 附近住宿: 一星級、二星級、三星級、四星級、五星級、其他

- L2-景點 + 類別: 本階層主要回應使用者於 L1 點選感興趣的類別資訊，例如使用者想了解高雄展覽館附近五星級住宿的相關資訊，前端應用程式將以高雄展覽館 + 五星級來進行第二階知識查詢。

- 輸入: 高雄展覽館 + 五星級
- 輸出:

- \* 名稱: 英迪格酒店、漢來大飯店

- L3-景點行程項目詳細資訊: 第三階層顯示內容為使用者所選定景點及特定項目的完整資訊。

- 輸入: 高雄展覽館 + 五星級 + 漢來大飯店
- 輸出:

- \* 官網: <https://www.grand-hilai.com/>
- \* 房間數: 540 間
- \* 簡介: 漢來大飯店座落於高雄繁華的商業中心，樓高 186 公尺，可俯瞰大高雄全景，擁有三千多件隨處可見的古董藝術珍品、540 間舒適客房、13 家中西餐廳及 1000 個停車位。距離高鐵站只要 20 分鐘車程；步行至捷運站也僅需 15 分鐘。館內除了住宿、餐飲、會議、宴客的功能之外，還結合了精品雲集的漢神百貨，並符合一次購足的便利。充滿新古典主義風格的大廳，古董藝術品

遍佈全館，呈現壯麗典雅的設計風格。房內採用穩重大方的原木材質，加以現代感的簡潔線條設計與歐洲古典紋飾，呈現出尊貴優雅的氣質。打開窗簾，高雄港的美麗景緻盡收眼底，不但令人驚艷，更造就了漢來大飯店的無價！

- \* 地址: 高雄市前金區成功一路 266 號
- \* 電話:(07)215-7266

#### 4.5 知識擷取錯誤分析

針對本研究利用 CKIP 中文剖析樹搭配演算法進行 SVO 擷取的過程目前仍有些問題待解決。

- 單一句子，多重剖析樹:  
 首先是中文剖析系統會自動利用標點符號進行字串的分割，因此一句較長的句子會先依標點符號斷開後解析成多個剖析樹，在未使用自然語言處理的搭配詞 (Collocation) 的方法前會得到較差的知識擷取成果。表9為錯誤的例句。以第一句範例「走到底就會抵達駁二藝術特區，經過高雄港牌樓，倉庫群就在眼前」所取得的剖析樹只能擷取語意角色為 goal 的受詞「駁二藝術特區」，該句並無法擷取到屬於主詞的語意角色，而「經過高雄港牌樓」及「倉庫群就在眼前」分別只能得到動詞片語及名詞片語的剖析樹。
- 缺少明顯主詞或主詞不明確:  
 第二類錯誤是句子沒有明顯的主詞或是主詞不明顯時也容易造成擷取 SVO 時的錯誤，第二句範例「房間能看見 85 大樓與高雄流行音樂中心」，人工分析時會將房間當成主詞、看見是動詞而 85 大樓及高雄流行音樂中心則為受詞，但 CKIP 剖析樹得到的是一個動詞片語，房間被辨識為地方詞而非主詞，動詞及受詞則與人工標記結果相同。因此需考量將地方詞加入演算法判斷主詞及受詞的規則中並以該詞出現的位置來決定標記成主詞或受詞才可擷取到預期的知識。

## 5 結論

本研究提出複合休憩知識建構模型，結合 HTML 標籤及整合 CKIP 中文剖析樹、語意角色之旅遊經驗知識擷取演算法以建構符合計畫需求之高雄港灣休憩知識圖譜。目前仍在持續蒐集各景點官網、旅遊網站及遊記經驗等資

例句	正確結果	分析結果
走到底就會抵達駁二藝術特區，經過高雄港牌樓，倉庫群就在眼前	倉庫群 (s)/就 (v)/在眼前 (o)	抵達 (v) 駁二藝術特區 (o)
房間能看見 85 大樓與高雄流行音樂中心	房間 (s)/看見 (v)/85 大樓與高雄流行音樂中心 (o)	85 大樓與高雄流行音樂中心 (o)

表 9: SVO 剖析知識擷取錯誤分析範例

料並持續建構中，以達期末成果展示之所需。複合休憩知識建構模型未來仍有優化、改進之處：

- 引入搭配詞:  
 遊記經驗的文章中的景點與內文較易出現字句過長的現象，目前的建構方法尚無法擷取出適當的知識單元，未來將加入搭配詞 (Collocation) 的機制來串連文章前後句的關係，提高主詞擷取成功的機率。
- 動態資訊定期更新與維護:  
 景點官網的動態資訊需要定期更新與維護，需要排程式定期取得各網站的最新消息、展演及活動等資訊並更新知識圖譜內容，以確保圖譜內知識的正確性。
- 知識圖譜維護機制:  
 本研究已設計一套可經由人工維護知識圖譜內容的方法，待景點資訊建構完成後將接續完成此一維護介面，使得知識能夠建立、應用及解釋，對於錯誤學習的知識也可經由人工校正，以健全知識圖譜內容。

## References

- Ismail Aliyu, A. F. D. Kana2, and Salisu Aliyu. 2020. Development of knowledge graph for university courses management. volume arXiv:2004.00071, pages 1–4. International Journal of Education and Management Engineering.
- Pablo Calleja, Freddy Priyatna, Nandana Mihindukulasooriya, and Mariano Rico. 2018. *DBtravel: A Tourism-Oriented Semantic Graph: ICWE 2018 International Workshops, MATWEP, EnWot, KD-WEB, WEOD, TourismKG, Cáceres, Spain, June 5, 2018, Revised Selected Papers*, pages 206–212.



- Penghe Chen, Yu Lu, Vincent W. Zheng, and Xiyang Chen. 2018. Knowedu: A system to construct knowledge graph for education. volume 6, pages 1–4. IEEE Access.
- Roberto Gatta, Mauro Vallati, and Lenkowicz. 2017. Generating and comparing knowledge graphs of medical processes using pminer. pages 1–4. K-CAP.
- Sourabh Jain and Prakash Hegade. 2021. [E-commerce product recommendation based on product specification and similarity](#). In *2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, pages 620–625.
- Feng Ling. 2020. [Design of tourism intelligent recommendation model of mount tai scenic area based on knowledge graph](#). pages 241–244.
- George A. Miller. 1992. Wordnet: a lexical database for english. volume *Speech and Natural Language: Proceedings of a Workshop Held at Harriman, New York, February 23-26, 1992*, pages 23–26.
- Nidhi Rastogi and Mohammed J. Zaki. 2020. Personal health knowledge graphs for patients. volume arXiv:2004.00071, pages 1–4.
- Denny Vrandečić and Markus Krötzsch. 2014. [Wiki-data: A free collaborative knowledgebase](#). *Communications of the ACM*, 57:78–85.
- Kai Zhao, Yukun Zheng, Tao Zhuang, Xiang Li, and Xiaoyi Zeng. 2022. [Joint learning of e-commerce search and recommendation with a unified graph neural network](#). WSDM '22, page 1461–1469, New York, NY, USA. Association for Computing Machinery.