

初探數位旅遊服務整合平台 Taiwan Pass 營運定位之研究：結合文字探勘及 資料探勘進行關鍵因素萃取

A PRELIMINARY STUDY ON THE POSITIONING OF
TAIWAN PASS: COMBINING TEXT MINING AND
DATA MINING TO EXTRACT KEY FACTORS

熊正一 Zheng-Yi Shon ¹

吳東凌 Tung-Ling Wu ²

呂明穎 Ming-Ying Lu ³

張益城 Yi-Cheng Chang ⁴

葉 曜 Yao Yeh ⁵

(110年6月22日收稿，111年2月21日第一次修正稿，
111年7月13日第二次修正稿，111年8月15日接受)

摘 要

政府致力於推動符合時宜的新型態觀光旅遊消費機制與服務-Taiwan Pass 營運定位以及關鍵因素是很重要的議題。本研究利用焦點團體座談蒐集意見，以文字探勘技術建構關鍵字，萃取重要意見作為營運定位的關鍵因素，並利用資料探勘進行語意相似度分析，瞭解業者間意見異同程度。

-
1. 臺南應用科技大學國際企業經營系教授。
 2. 交通部運輸研究所運輸科技及資訊組組長。
 3. 東吳大學資料科學系助理教授 (聯絡地址:111 臺北市士林區臨溪路 70 號, 聯絡電話:02-23111531 轉 6513, E-mail: mylu@gm.scu.edu.tw)。
 4. 交通部運輸研究所運輸科技及資訊組副組長。
 5. 交通部運輸研究所運輸科技及資訊組助理研究員。

研究結果顯示，共有 5 個議題的產生，議題一（關於建置平台的架構、預算、時間與營運模式之考量）同質性最高、議題二（整合食宿遊購行，平台 API 資訊相互介接之共享、資安問題）次之、接續為議題四（公共運輸行動服務 Maas 經驗）和議題五（Umaji、好玩卡開發經驗），最低為議題三（平台服務數位化，開發統一規格技術的 QRcode）。本研究亦提出建構 Taiwan Pass 之營運定位方針作為未來推動參考。

關鍵詞： 數位旅遊服務、整合平台、營運定位、文字探勘、資料探勘

ABSTRACT

Due to the changes in tourism and consumer demand, the government is committed to promoting and mentoring the industry, in an attempt to launch a new type of digital tourism service -Taiwan Pass. Faced with challenges, the positioning of Taiwan Pass and the key factors for its success are very important issues. This research uses focus group to collect opinions and uses natural language processing to conduct text mining to extract keywords. Use word frequency analysis to extract important opinions as key factors, and use data mining for semantic similarity analysis. By understanding the degree of discrepancy of opinions participants, so the impact and influence of the industry can be reduced. The research results show that a total of five issues were generated, and the content of the issues was compared for homogeneity. The results show that issue 1 (considerations about the architecture, budget, time, and operation model of the platform) has the highest homogeneity, followed by issue 2 (integration of food, accommodation, travel, shopping, and transportation, platform API connected and security issues), topic 4 (MaaS experience) and topic 5 (Umaji, Taiwan Pass development experience), topic 3 (digitalization of platform services, development technical specifications of QRcode). This research also proposes the operating positioning policy of Taiwan Pass as a reference for the promotion of Taiwan Pass in the future.

Key Words : *Digital tourism service; Integrated platform; Positioning; Text Mining; Data Mining*

一、緒論

近年來，隨著科技和智慧觀光的发展，觀光產業的消費型態產生巨大改變，旅遊方式也產生了變化，自由行漸漸成為主要的旅遊方式之一。其中，新型態的目的地旅遊逐漸崛起，旅客可自行選擇合適的時間及交通工具，到達任意想要遊玩的城市或景點，並在到達後享受接送、導遊服務、遊覽、娛樂、當地交通、門票等輔助服務。此外，伴隨著消費習

性的改變以及數位轉型的推動，產品通路已由線下轉移至線上，或者是線上與線下的虛實整合的通路型態，甚至發展成全通路的型態，從供應商、通路、支付方式和搜尋等，都產生巨大的變革。面對旅遊型態與消費習性的改變，國內觀光產業面臨轉型的壓力，因此，建構符合消費者需求的觀光旅遊產品銷售機制和運營模式是重要的目標與挑戰。

根據市場消費型態的改變及現存急需改善之痛點，推動智慧觀光、數位轉型成為在觀光政策和產業發展上重要的推動方向，交通部運輸研究所於民國 109 年進行「Taiwan Pass 數位旅遊服務整合平台規劃案」委託研究計畫，期望藉由 Taiwan Pass 數位旅遊服務整合平台之建置，提供消費者更便利的介面，帶動相關產業創造更大的市場獲益，也為疫情後的觀光產業提供更完善的支援。然而，觀光旅遊市場的多樣性服務以及消費需求的轉變，Taiwan Pass 平台能否符合需求以及其營運定位勢必影響後續政策的成功與否，因此，找出 Taiwan Pass 營運的關鍵因素是很重要的議題。

Taiwan Pass 對於市場來說，是一個創新的政策和挑戰，整合產官學研的意見是很重要的基礎，營運關鍵因素和定位可透過系統性的方法來挖掘，針對議題深度的探索可透過訪談或座談的方式進行。訪談或座談等質性研究係用於蒐集對問題的深入見解或為研究產生新的想法，是一種在社會科學及教育學領域常使用的研究方法。其中，焦點團體座談是質性研究蒐集資料時，廣泛使用的一種方法，且對於意見的整合和收斂有不錯的效果 (Mishra^[1])。為確保後續 Taiwan Pass 數位旅遊服務整合平台之系統規劃建置、既有相關系統或平台之整合納入營運模式、策略研擬及後續移轉等之完善且符合需求，本研究透過焦點團體座談，就平台架構、平台使用對象/角色種類、功能需求、資訊系統及資安規劃、平台建置規劃、營運模式、票證載具、營運移轉、各利害關係人權利義務和相關法規等議題進行探討及釐清，邀請國內產官學研專家進行討論，進行相關資訊的蒐集，作為後續關鍵因素萃取的基礎。

焦點座談會等質性研究所獲取資料相較於問卷調查等量化研究屬非結構化的資料，不易驗證其信效度，為確保資料正確性及有效性，研究過程需設計資料檢核機制。透過焦點團體座談可以蒐集多種類型的數據，包括焦點小組參與者的錄音帶、主持人和助理主持人所做的筆記，以及主持人和助理主持人回憶的項目(Krueger and Casey^[2])。這些數據都是可以進行分析，並區分為基於轉錄的分析 (transcript-based analysis)、基於筆記的分析 (note-based analysis) 和基於記憶的分析 (memory-based analysis)，但在完成分析所需的時間和嚴格程度方面有所不同，其中，基於轉錄的分析是各種分析方法中最嚴格和最耗時的模式 (Onwuegbuzie, *et al.*^[3])。為了系統性的理解焦點團體座談中與會者感興趣的現象和重點部分，過往多採用人工進行語幹分析與萃取，再利用相互驗證確認其信度與有效性，然而，礙於分析方法的限制，常常無法大量處理文本資料。

焦點團體座談依會議時間的長短不同，通常會產生 50 到 70 頁的文本 (Krueger and Casey^[2])。實務上經常會採用基於磁帶的分析 (tape-based analysis)，這是一種分析焦點團體數據的方式，也就是研究人員聽取焦點團體座談討論的錄音，然後產生一個刪減內容的文本，為了可以加快處理速度，該文本資料通常比基於轉錄的分析中採用的完整文本短得

多，此外，也會受到研究人員的經驗影響，造成只專注在研究人員主觀認為有興趣的部分，因而可能會遺漏部份重要資訊 (Onwuegbuzie, et al. [3])。隨著文本資料的大量增加，資料中非結構化或半結構化數據的處理變得複雜且困難，自然語言處理技術的發展與應用，讓非結構化的文本資料透過斷詞處理再進行詞頻統計的文字探勘方式廣泛的被應用在各類型的質性研究中，也獲得不錯的效果 (Ming and Baumer [4]; Yu, Jannasch-Pennell, and DiGangi [5])。因此，本研究將採用文字探勘的方式來分析完整文稿，進一步萃取座談會的重要資訊，以作為後續生成焦點座談的主要意見或主要議題的基礎。

此外，大多數焦點團體研究人員使用小組作為分析單位 (Morgan [6])。研究人員對文本數據進行編碼並呈現主題，儘管這些主題可以產生重要且有趣的資訊，但僅分析和解釋文本可能會帶來問題，無法提供有關共識和異議程度的資訊 (Sim [7])。然而，產官學研專家間對於議題或意見的一致性也會牽動後續的發展，因此，與會者間對於意見表達的一致性將是另外一個需要關注的部分。本研究將利用相似性分析來確認與會者之間的意見異同，以作為後續意見同質與否的判斷依據，藉此瞭解意見的相似性。

綜合上述，本研究針對 Taiwan Pass 營運定位進行研究，以焦點座談方式蒐集意見和相關資料，進而利用文字探勘進行質性分析，藉此系統性且客觀的進行分析和收斂，再萃取出重要關鍵意見與議題，並透過相似性分析，進行議題一致性比對，藉此收斂議題並找出同質性高的議題。期望透過這樣的分析方法可以克服過往質性研究過多人為主觀和人力耗損大缺點，並提出更有效且全面的資訊萃取方法，達到推動 Taiwan Pass 的意見整合和營運定位的設定，藉此作為未來相關單位推動的參考依據。

二、文獻探討

2.1 Taiwan Pass 觀光旅遊平台

全球觀光旅遊產業及觀光活動的發展處於創新轉型的階段，一方面反映觀光旅遊價值提升的全球觀點，另一方面亦回應旅客對新的需求與期待。Travelport 亞太區總監馮克銘指出 (黃志偉 [8])，網路數位化時代，搜尋的便利性讓消費者的期望及要求愈來愈高，無論是搜尋引擎、線上旅行社 (OTA, Online Travel Agency) 或供應商網站搜尋或購買旅遊產品，都有多元化的選擇。因為智慧型手機或裝置的普及化，帶動了電子商務風潮與發展，旅遊商品的預訂在電子商務營收中占很大的一部分，2017 年全球線上旅遊預訂的收入達到 5,670 億美元。在亞太地區，消費者使用行動裝置購買旅遊商品，占亞太地區線上旅遊銷售的 50%，而印尼的千禧世代旅客更有 73% 透過手機應用程式或瀏覽器來進行預訂。這一再的顯示，民眾對於旅遊商品的消費習性已經傾向線上購買的方式來進行。

另外，隨著新興科技而發展出的「共享經濟」模式興起，業者為遊客提供不同的旅遊體驗，也在全球引起廣泛關注，各國紛紛採取行動，擬訂相關規定和政策來規範旅遊共享經濟的發展 (World Tourism Organization [9])。各國政府在大力推動智慧工程建設之際，大

量創新型態的技術也開始導入，許多相關的應用都率先在旅遊產業中推廣和應用。其中，創新的應用驅動旅遊業中的技術和管理模式進行改變，物聯網、機器人、混合實境 (MR, Mixed Reality)、Beacon 等技術及科技的影響是最為關鍵的部分 (World Tourism Cities Federation^[10])。這些創新應用的發展也意味著觀光旅遊的消費型態產生巨大變革，而且是快速的改變中，需有相對應的運營模式來滿足旅遊消費需求。

為了因應改變的浪潮，讓觀光產業可以跟上時代的潮流，交通部運輸研究所於民國 109 年進行「Taiwan Pass 數位旅遊服務整合平台規劃案」委託研究計畫，期望國內觀光產業可藉由 Taiwan Pass 數位旅遊服務整合平台之建置，提供國際旅客及國內旅客一站式觀光旅遊服務，以顧客為核心 (User-Centered) 的整合性服務，從顧客觀點出發，整合觀光旅遊需要的服務包括餐飲 (食)、飯店旅館 (宿)、觀光景點 (遊)、特色商品 (購)、交通 (行) 所有面向，將食、宿、遊、購、行服務數位化並於服務平台上進行整合與包裝，使用者可於平台上進行旅運規劃、服務預訂、完成付費及服務評價等。當觀光旅遊需求鏈集中在平台中進行，預期各類服務可有效整合，並帶動相關產業創造更大的市場獲益，消費者也可得到更高的運輸服務及觀光旅遊價值。此外，亦將透過旅客行為大數據之蒐集與分析，找出可能隱藏其中之珍貴資訊與趨勢，進一步轉換為營運管理、服務創新與政策決策支援之有效資訊，以做為後續公共政策、交通管理、增值服務、服務品質提升等之基礎。

在國內有許多業者也提供線上旅遊商品的銷售，有些業者分別提供餐廳的預定、飯店的預定、旅遊門票的預定和禮品的選購，甚至交通部觀光局也針對交通相關的整合與行銷推出好玩卡，企圖利用更便利的方式協助民眾解決觀光旅遊中的交通問題。然而，卻缺乏一站式的整合平台規劃，甚至有些節點無法突破，例如：預定旅遊行程，往往無法即時確認交通狀況等問題。因此，若能蒐集觀光相關業者的意見、政府單位的把關和學術研究單位的量能共同研擬 Taiwan Pass 的建置規劃，勢必能達到共贏的局面，也勢必能在疫情後，提供觀光旅遊市場不一樣的體驗。

2.2 質性研究

座談會是蒐集意見常採用的質性研究方法，與會人員可以透過交流互動，理解他人的意見，同時在相對無壓力的情況下充分表達自己的意見與看法。質性研究 (或稱質化研究、定性研究)，主要涉及資訊蒐集和分析非數字數據 (例如文本、視頻或音頻)，用以理解某特定議題的概念、觀點或經驗。它可用於蒐集某特定問題的深入見解或為研究產生新的想法，是一種在社會科學及教育學領域常使用的研究方法，相對量化研究而言，一般普遍認為，質性研究對於議題的資訊和資料探索較為深度。

其中，焦點團體座談是質性研究蒐集資料時，廣泛使用的一種方法，主要透過主持人的引導以及藉由團體的互動，蒐集團體成員對特定主題的想法、意見、知覺、態度與信念。焦點團體法最早出現於二次大戰期間，當時社會科學學者開始使用非指導式訪談方式於團體中進行訪談 (Rice and Ezzy^[11])。二次大戰後因商務蓬勃發展，市場調查研究人員為了

要瞭解消費者的看法，大量使用訪談方式於團體中進行訪談，以協助公司可以生產出暢銷的產品或是推出適當的服務。此外，1950年代起，美國大都市開始出現市場調查研究公司，提供與焦點團體有關的服務，協助找尋當地的團體參與者名單、召集前來參與討論、籌備規劃團體討論以及提供錄音與錄影服務，進行相關資料的蒐集 (Anderson^[12]; Denscombe^[13]; Patton^[14]; Wisker^[15])。然而，學術界在1980年代又開始重新正視焦點團體訪談，學術界乃擷取商界研究人員的部分策略並加以修改，使這類型的質性研究可被適當的用於其他研究對象。焦點團體方法可作為一種輔助或主要數據蒐集方法，且在社會和健康科學以及評估研究中變得越來越流行 (Kidd and Parshall^[16])。

焦點團體是一個謹慎規劃的系列討論，在一個舒適、包容的、無威脅性的情境進行一系列的討論，目的是瞭解人們對於一個特定的議題、產品或是服務項目的感受與意見。Krueger and Casey^[2]建議每個團體焦點應由一個技巧優良的主持人 (moderator) 帶領參與者進行討論，參與者均具有與該焦點團體的主題有關的某些特質或經驗。為了方便發掘主題的趨勢與規律性 (pattern)，可針對具有相似特性的參與者進行若干場次的團體討論。此外，研究者會運用縝密與系統的分析討論資料，最終提供對某一項產品、服務項目或是機會的研究結論與建議。在資訊蒐集方面，在調查問卷中，受限於題目，很少有受測者會花時間寫出詳細的答案，焦點團體座談和傳統調查問卷相較，焦點團體為完全開放式，焦點團體可以廣泛獲取資料與資訊，找出在對於參與者來說重要的問題。然而，基於成本限制因素，通常焦點團體座談不會有太多場次的討論或是較多的人員參與，因此，無法提供較完整的統計資訊。

從調查研究設計、資料蒐集過程、到資料整理分析，每個階段均會影響到資料的品質，故均須建立嚴謹的資料檢核工作與後續處理，以使所有蒐集回來的文字或數據資料，能符合研究所需並能獲得正確且合理的結果。一般而言，資料正確性的檢核方式包括人員稽核及資料檢核兩種。在人員稽核部份，則再分為現場作業人員的訓練及現場作業的品質管制。在分析方法的部分，常數比較分析法是很常用來分析焦點團體資料的方式，又稱常數比較法，最早用於紮根理論研究 (Glaser^[17]; Glaser and Strauss^[18]; Strauss and Corbin^[19])。常數比較分析法可分成三個階段，第一階段，數據被分成小單元，也就是分析單位，研究人員在這階段會將描述或代碼加到每個單元。然後，在第二階段，將第一階段的代碼進行分組。最後，在第三個階段，研究人員會發展出一個或多個主題來表達每個組的內容 (Strauss and Corbin^[19])。而資料檢核部份則主要在調查訪問工作結束後進行，主要目的在確認研究資料的正確性 (accuracy)，確保研究資料完整性 (completion) 與可用程度 (availability)。資料蒐集回來後，需進行初步資料檢核，再由專業人員進行覆閱，檢查項目包括答覆的完整性、內容是否清晰可辨。最後在分析時，亦會針對數據資料進行信效度分析，信度是用以衡量測驗的特性或測量結果的可靠性，透過衡量工具對同一受測群體進行重複性的測量，並加以驗證其測驗結果的一致性 (consistency) 及穩定性 (stability)。

一般來說在焦點團體座談中，助理主持人的職責包括記錄會議 (通過錄音還是錄影)、做筆記、創造一個有利於小組討論的環境 (例如確保每個人都有座位、安排茶點)，於座談

會後提供數據驗證，並幫助研究人員/主持人分析或解釋焦點團體數據 (Krueger and Casey^[2])。其中，分析內容容易限於助理主持人的主觀意識，隨著座談會的討論內容豐富程度增加，分析人員的負荷也跟著增加。因此，本研究採用文字探勘方法系統性且客觀的方式進行逐字稿內容的斷詞處理，來進行議題的生成，此外，在同意度與信度的計算上使用語幹為分析單位，以確保資料的信效度。

2.3 文字探勘與資料探勘

文字探勘是分析文字的過程，透過這過程可以獲得用於特定目的的有效資訊 (Gegick, Rotella, and Xie^[20])。文字探勘是資料探勘科學的一個分支，兩者的區別在於數據的形式。資料探勘主要是針對資料表等具有結構化的數據形式，文字探勘主要係針對像文本資料等非結構化的數據形式 (Imbar, Adelia, Ayub, and Rehatta^[21])。近年隨著自然語言處理技術的發展，文字探勘的相關應用越來越多廣泛，其可以系統性的處理非結構性的文本資料，提高資料處理的效率，也避免人為的失誤和主觀判斷的落差，其研究範圍包括將文本進行分類以及文字進行配對 (Nurdiana, Jumadi, and Nursantika^[22])；也有研究將此技術用於抄襲的檢測 (Priambodo^[23])；更可以針對字句進行情感分析，瞭解撰寫者的情緒反應 (Li and Wu^[24])；有些研究也利用自然語言處理來評估答案的優劣等 (Roy, Dandapat, Nagesh, and Narahari^[25])。這幾個類型的研究，過往多仰賴專家學者進行判斷，近年利用自然語言處理，不但可以提高處理的效率與效果，更可以避免掉主觀所造成的誤差和偏誤，在質性研究上都獲得不錯的效果。

其中，文本分類是自然語言處理領域中非常普遍的應用，有許多以自然語言進行文本分類的相關研究，像是情感分析、垃圾郵件分類和文章主題分類等。近年常被採用的文本分類方法是利用空間向量模型中生成的特徵，例如：Term Frequency-Inverse Document Frequency (TF-IDF; Zhang, Gong, and Wang^[26])或 Bag-of-words (Boulis and Ostendorf^[27])來表示文本的向量。TF-IDF 是一種進行資訊檢索與資料探勘常用的加權技術，它是基於統計方法來進行的。其中，TF-IDF 是兩個不同詞的組合，即詞頻和逆文檔頻率，TF 是指詞頻 (Term Frequency)，IDF 是指逆文字頻率指數 (Inverse Document Frequency)。其採行的概念係字詞重要性會因在檔案中出現的次數而成正比，但也會隨著它在語料庫中出現的頻率成反比下降，若一個詞頻在文章中出現的頻率高，且在其他文本中出現的情況很少，則就可以將這個詞視為有很好的類別區分能力。以 TF-IDF 來做文本分析都有不錯的結果，利用 TF-IDF 進行文本分類，解決傳統文本分類技術中忽略上下文語義鏈接和詞彙重要性不同等問題 (Liu, Sheng, Wei, and Yang^[28])；以自動文本分類技術幫助用戶查詢相關資訊，利用 TF-IDF 檢查關鍵詞與語料庫中的文檔相關性 (Qaiser and Ali^[29])；以 TF-IDF 來找熱門詞語，進而探索熱門話題 (Zhu, Liang, Li, Yu, and Liu^[30])。相關的研究顯示，TF-IDF 對於文本探索，有相當不錯的表現，也很適合來做資料探勘和資料檢索，可以有效地提升質性研究的效率。

三、研究方法

本研究透過焦點座談蒐集與會者的發言，焦點團體資料分析部分採用 Krueger and Casey^[2] 所提出的基於轉錄的分析概念，將座談會主持人和與會人員的發言以錄音的方式記錄，再透過人工轉譯逐字紀錄，並以完整逐字稿作為資訊的來源進行分析。為了有效從焦點座談會的資訊萃取出 Taiwan Pass 營運定位的關鍵因素與議題，本研究所採用的方法主要包含文字探勘和資料探勘兩個部分，整體架構如圖 1 所示。

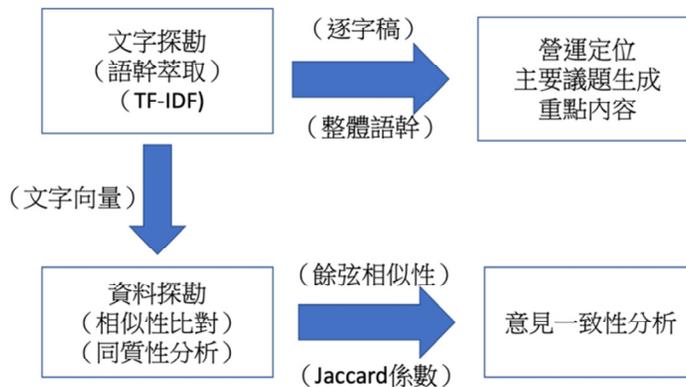


圖 1 研究架構

1. 文字探勘

在資料處理過程，利用文字探勘技術作為資料預處理的技術，利用自然語言處理進行逐字稿分析，透過斷詞以及關鍵詞的提取，作為語幹分析的基礎，除了可以系統化的理出重點外，更可以提升因子萃取的信效度。由於斷詞的精準度高度仰賴演算法能力，TF-IDF 是一種用於資訊檢索與資訊探勘的常用加權技術，可以有效的針對文本資料進行資料的檢索 (Qaiser and Ali^[29]; Kim, Seo, Cho, and Kang^[31])。其中，由中研院資訊所和語言所成立的中文語言小組所開發的 CKIP (Chinese Knowledge and Information Processing) 是針對繁體中文開發的工具，被認定為市場上繁體中文斷詞相對精準的工具，在研究上也獲得不錯的成果。整體來說，本研究透過文字探勘萃取語幹作為探討 Taiwan Pass 營運定位關鍵議題生成的基礎，並透過專家學者利用整體語幹與逐字稿對應生成主要議題，以及個別語幹與逐字稿對應生成重點內容。另外，透過文字探勘產生詞頻向量，作為後續資料探勘使用。

2. 資料探勘

在非結構性的文本資料進行相似性分析，可分成特徵 (Character based) 和字串 (Term based) 兩類型 (Hall and Dowling^[32])。其主要係利用斷詞獲得重要詞頻，轉換成向量後，再進行語意比對來進行驗證。因此，本研究採用兩階段的相似度比對進行參與人員間的意

見同質性比較，利用餘弦 (Cosine) 相似性進行內容一致性比對，再利用 Jaccard 相似係數進一步判斷異同程度的高低，藉此獲得關注議題的一致程度差異，達到焦點團體意見的收斂效果，瞭解產官學研的意見方向與異同。

詳細步驟包含：資料預處理、資料分析和相似性分析，透過資料預處理產生詞頻和整體研究技術架構如圖 2。

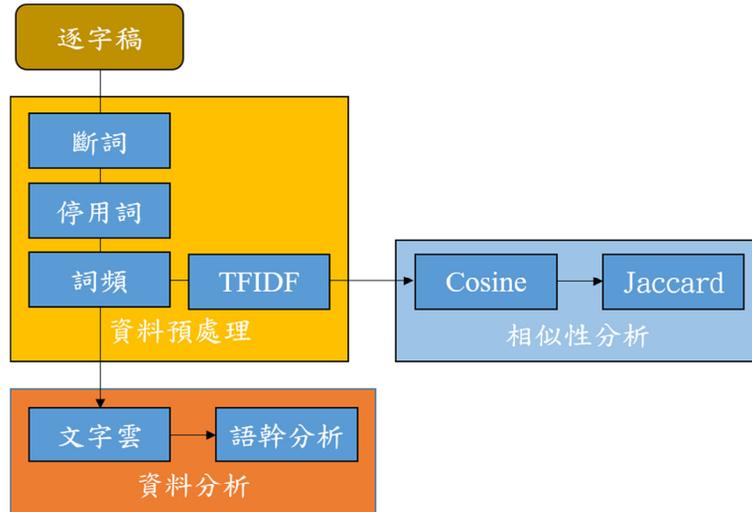


圖 2 研究技術架構

3.1 資料預處理

3.1.1 斷詞

詞是文章中有意義且可以自由使用的最小語言單位。通常語言處理機制中，系統必須能分辨出文章中的詞，才能有效的進行後續的語言分析或資訊抽取的處理。因此，斷詞的技術係成為語言處理不可或缺的工作之一。斷詞的方式，基本上係利用詞庫或辭典的詞和文本資料做比對，找出文章中可能包含的詞。然而，斷詞的切分可能造成詞句存在歧義，多數的斷詞程式都是利用演算法來解決斷詞歧義的狀況。

本研究斷詞的部分利用 CKIP Tagger 進行處理，將焦點座談的文字稿中所有字句透過 CKIP tagger 工具切割成有意義字詞，例如：專家 1 在一次發言中說道：「那些旅客是常來的，那些旅客是有興趣的，甚至於來過的我們可以再行銷給他類似這樣商品，那透過票卷的規格化，就如同我們剛剛講的，只要所有的服務他的票卷是同一個規格，他就可以通過我們未來再建置同一個票閘口，不管是我們剛剛講的，大眾運輸或者是公共運輸或場站，甚至於包括未來的飯店，包括未來的店家，他都可以用同樣的一套設備來完成」。一般來說，這類型的意見不易判斷重點，或是需要利用人工判讀，找出重點來記錄，費時耗工。

因此，本研究先利用斷詞，透過斷詞把文章中的句子做分割，分割成有意義的詞句，以利後續進行分析使用。

其中，上述的例子經過斷詞後將得到下列字詞列表：

「'那些'，'旅客'，'是'，'常'，'來'，'的'，'，'，'那些'，'旅客'，'是'，'有'，'興趣'，'的'，'，'，'甚至於'，'來'，'過'，'的'，'我們'，'可以'，'在'，'行銷'，'給'，'他'，'類似'，'這樣'，'商品'，'，'，'那'，'透過'，'票卷'，'的'，'規格化'，'，'，'就'，'如同'，'我們'，'剛剛'，'講'，'的'，'，'，'只要'，'所有'，'的'，'服務'，'他'，'的'，'票卷'，'是'，'同'，'一'，'個'，'規格'，'，'，'他'，'就'，'可以'，'通過'，'我們'，'未來'，'在'，'建置'，'同'，'一'，'個'，'票閘口'，'，'，'，'不管'，'是'，'我們'，'剛剛'，'講'，'的'，'，'，'大眾'，'運輸'，'或者'，'是'，'公共'，'運輸'，'或'，'場站'，'，'，'甚至於'，'包括'，'未來'，'的'，'飯店'，'，'，'包括'，'未來'，'的'，'店家'，'，'，'他'，'都'，'可以'，'用'，'同樣'，'的'，'一'，'套'，'設備'，'來'，'完成'，'。'。」

然而，斷詞的工作只是把文章或字句切割成有意義的詞，該列表中仍參雜許多贅詞以及標點符號等無意義，或是對於理解文章沒有幫助的資訊。因此，後續必須針對列表中無意義的符號和贅字進行分離或過濾，藉此達到重點詞萃取的目的。

3.1.2 過濾停用詞

文本預處理是準備文本數據的過程，以便後續可以使用它來執行分析、預測等任務。停用詞一般來說是任何語言中最常見的詞（如冠詞、介詞、代詞、連詞等），並且不會為文本添加太多資訊。停用詞在任何人類語言中都會被大量使用，從文本資料中刪除這些包含較低有用資訊的詞，便可以更加關注重要資訊。換句話說，刪除這些停用詞是不會對後續任務訓練的模型產生任何負面影響，反而去除停用詞肯定會減少數據量的大小，從而減少訓練時間。因此，過濾停用詞用途在於將無意義贅字或標點符號過濾，以免不重要的詞干擾主要詞頻的統計與產出，進而影響到後續的分析。

然而，停用詞並不是都需要被刪除，停用詞的刪除是基於執行的任務和實現的目標。例如，執行情感分析任務的模型，可能不會刪除過多停用詞，因為部分停用詞有其意義在，在去除停用詞後，負面評論可能變成了正面。停用詞通常需要被定義，而且沒有規則是通用的，也沒有停用詞列表是通用的，必須要單獨被建立。

因此，本研究將透過對文本內容進行檢視並建立停用詞庫，再將斷詞完成的資料進行停用詞庫比對，篩除不重要的符號和贅字。例如斷詞中所舉的範例，將該字詞列表經由建立的停用詞庫進行過濾，過濾結果為「'旅客'，'旅客'，'興趣'，'行銷'，'商品'，'票卷'，'規格化'，'服務'，'票卷'，'規格'，'建置'，'票閘口'，'運輸'，'公共'，'運輸'，'飯店'，'設備'」，過濾完的字詞只留下為較重要的字詞，以利後續分析使用。

3.1.3 計算詞頻

詞頻為字詞出現的頻率，也就是每個字詞出現的次數，透過計算詞頻可以初步了解逐

字稿中每個字詞被使用的狀況，更可以知道哪些有意義的資訊被重複或是多次提及。如同上例子，將過濾結果計算其詞頻，將得到表 1 資訊，此表也代表專家 1 在該次發言的詞頻結果。

表 1 詞頻結果範例

字詞	次數
旅客	2
興趣	1
行銷	1
商品	1
票卷	2
規格化	1
服務	1
規格	1
建置	1
票閘口	1
運輸	2
公共	1
飯店	1
設備	1

3.1.4 TF-IDF

隨著數據的大量增加，資料庫或資料表中結構化或半結構化數據的處理變得非常困難 (Trstenjak, Mikac, and Donko^[33]; Bafna, Pramod, and Vaidya^[34])。有許多技術或算法可用於協助有效的處理文本數據，TF-IDF 是一個數值統計，透過分析可以顯示出關鍵字與某些特定文檔資料的相關性，或者可以說，透過使用這些關鍵字可以識別或分類某些特定文檔。TF-IDF 演算法包含了兩個部分，詞頻率 (TF) 跟逆向文件頻率 (IDF)，透過兩個部分的執行，可以有效的做為檢查關鍵詞與語料庫中的相關性工作的前處理 (Qaiser and Ali^[29])。

詞頻指的是某一字詞在逐字稿中每次發言出現的頻率，假設第 t 個字詞出現在第 d 個發言的頻率記做 $tf_{t,d}$ ，舉例來說，如果第 d 個發言總共有 100 個字詞，而第 t 個字詞在這次發言出現的次數是 12 次，則 $tf_{t,d} = 12/100$ ，以頻率而不是次數來看待文字的重要性，讓發言與發言之間比較有可比較性。

逆向文件頻率則是用來處理常用字的問題。字詞 t 總共在 d_t 個發言中出現，則詞彙 t 的 IDF 定義成

$$idf_t = \log\left(\frac{D}{d_t}\right) \quad (1)$$

其中，

idf_t ：字詞 t 的逆向文件頻率

D ：語料庫中的總數發言數

d_t ：字詞 t 出現的發言數

假設第一個字詞總共出現在 25 個不同的發言中，則 $idf_1 = \log\left(\frac{D}{25}\right)$ 。如果字詞 t 在很多篇文章中都出現過，就代表 d_t 很大，此時 idf_t 就會比較小。

最後，計算字詞 t 在整篇逐字稿的發言中的 TF-IDF 權重以 $TFIDF_{t,d}$ 表示：

$$TFIDF_{t,d} = tf_{t,d} \times idf_t \quad (2)$$

其中，

$TFIDF_{t,d}$ ：字詞 t 在整篇逐字稿的第 d 個發言中的 TF-IDF 權重

$tf_{t,d}$ ：第 t 個字詞出現在第 d 個發言的頻率

idf_t ：字詞 t 的逆向文件頻率

當字詞 t 在第 d 個發言出現比較多次時 $tf_{t,d}$ 會比較高，同時字詞 t 較少出現在其他發言中，則 idf_t 也會比較大，整體的 $TFIDF_{t,d}$ 就會較大，也代表字詞 t 對於第 d 個發言是重要的。本研究透過 TF-IDF 演算法將每個單位發言的文字內容轉成向量形式，以利後續相似性分析使用。

3.3 相似性分析

為了分析與會專家學者以及業者對於意見的相似程度，本研究利用兩階段方式進行相似性的分析。首先，利用餘弦相似性作為發言方向的相似性分析，再利用 Jaccard 係數進行發言同質性的比較，藉此分析發言一致性的程度。

3.3.1 餘弦相似性

餘弦相似性 (Cosine) 為三角函數中的餘弦函數，其目的在於計算兩向量之間的相似度，若 Cosine 越小則代表兩向量之間的相適度是較高的，而 Cosine 值域為 $[-1,1]$ ，兩個向量為 0 度角的時候餘弦值是 1，而其他任何角度的餘弦值都不大於 1，最小值是 -1。利用餘弦值判斷相似性的原理係主要從而兩個向量之間的角度餘弦值確定兩個向量是否指向相同方向，當兩個向量有相同的指向時，餘弦相似度的值為 1，此時兩個向量夾角為 0 度角；兩個向量夾角為 90 度時，餘弦相似度的值為 0；兩個向量指向完全相反的方向時，餘弦相

似度的值為-1。然而，在資訊檢索的目的中，由於一個詞的頻率（公式 2 中的 $TFIDF_{t,d}$ ）不能為負數，所以兩個發言內容的餘弦相似性範圍為 0 到 1 之間，因此，兩個詞的頻率向量角度不能大於 90 度，其計算公式如下：

$$\cos(\theta) = \frac{\vec{P} \cdot \vec{Q}}{\|\vec{P}\| \cdot \|\vec{Q}\|} \quad (3)$$

其中，

$\cos(\theta)$ ：兩個給定向量餘弦值

P、Q：詞頻向量

本研究將各單位發言經由資料預處理後透過 TF-IDF 得到向量，再利用餘弦相似性計算各單位間發言的相似性，並找出相似性大於整體平均值的部分，探討其發言內容相似之處。

3.3.2 Jaccard

Jaccard 相似度指數（亦稱為 Jaccard 相似度係數）被拿來比較兩個集合的內容，查看哪些是同質的，哪些是不同的。它是衡量兩組數據相似性的指標，分數越高，表示兩個群體內容越相似。

本研究將兩單位間所有發言經由斷詞、停用自過濾後的所以字詞，並去除重複視作各自的集合，經由 Jaccard 係數的公式 (4) 進行計算，由於各集合皆是記錄關鍵字詞，因此，透過交集數除以聯集數的方式探討兩單位間發言的同質性。

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (4)$$

其中：

$J(A, B)$ ：給定的 AB 兩文字集合的 Jaccard 係數

A、B：文字集合

在上述公式產生 Jaccard 係數的步驟中，即先計算 AB 兩個集合之間共同詞的數量、再計算兩組（共同和非共同）中的詞總數，最後將共同的數量除以所有的詞總數。

四、結果與分析

本章節將說明本研究的文本數據蒐集來源，以及斷詞分析過程和分析結果，相關說明如後所述。

4.1 資料說明

4.1.1 資料來源與處理

本研究所採用的文本資料主要蒐集自「Taiwan Pass 數位旅遊服務整合平台規劃案」於 2021 年 2 月 4 日於國立臺灣師範大學 (臺北市和平東路一段 129 號) 圖書館校區教育大樓 2 樓《教 202 國際會議室》召開之「Taiwan Pass 數位旅遊服務整合平台規劃案」營運座談會議，座談會邀請對象包括旅行業者、票證公司、電信公司、臺高鐵路等相關單位，共計 46 位先進與專家出席該次會議。其中，主持人於會議中先行說明計畫對象、平台規劃概念及綜合討論等，會議中強調平台不會與民爭利，將透過導入公部門資源並研擬整合雙鐵 (高鐵、臺鐵)、四捷 (臺北捷運、桃園捷運、臺中捷運和高雄捷運)、二卡 (悠遊卡和一卡通) 服務，以此作為 Taiwan Pass 後續推動方向，並廣納各業者需求、意見後納入修正。主持人說明後，再由與會專家和各單位發表意見和想法，過程採用專人記錄和錄音的方式進行意見記載，最後由助理完成逐字稿的產生，並相互檢核逐字稿的正確性與完整性。其中，營運座談會議討論內容所產生的逐字稿有 38,222 個字，本研究以此份討論內容進行文本資料預處理，並進行文字探勘與資料探勘分析。

本研究先依據發言進行發言人標記，營運座談會議共有 4 位專家 (後續標記為專家學者 1~專家學者 4) 和 9 位業者 (後續標記為業者 1~業者 9) 進行發言，以此作為標記，標記識別係在資料探勘進行相似性分析時，作為識別不同業者的意見異同使用。

4.1.2 斷詞分析

首先，本研究利用 ckip tagger 將專家與業者的發言內容進行斷詞、過濾停用字及詞頻，詳細結果請參考附錄一。資料預處理後繪製文字雲，並依文字雲之關鍵字找出該會議探討的主要議題，圖 3 為整體逐字稿之文字雲，可以知道業者、系統、平台、整合、服務等關鍵字都是討論 Taiwan Pass 的重要用詞，也顯示本次斷詞內容有足夠的代表性。

另外，本研究透過詞頻分類、文字雲及整體逐字稿內容，由 3 位專家學者依據討論問題進行收斂和重複驗證後，生成五個關鍵議題。如下所示：

議題一：

關於建置平台的架構、預算、時間與營運模式之考量。

議題二：

整合食宿遊購行，平台 API 資訊相互介接之共享與資安問題。

議題三：

平台服務數位化，開發統一技術規格的 QRcode。

議題四：

MaaS (公共運輸行動服務) 經驗。

議題五：

Umaji、Taiwan pass (好玩卡) 開發經驗。



圖 3 整體逐字稿之文字雲

4.2 語幹萃取

語幹萃取的目的係快速收斂發言內容，主要透過各單位整體發言的內容，進行詞頻統計，再進一步繪製文字雲，針對文字雲與實際內容進行對應，利用文字雲的快速獲得關鍵詞的特點，有效的探討各單位於各議題回應之內容及方向。舉例來說，專家學者 1 的發言建構成圖 4 的文字雲，根據專家學者 1 發言的文字雲內容可以看到，業者、平台、資源、建置、政府、營運及服務等字詞數量較多，也代表較高的重要性。



圖 4 專家學者 1 之文字雲

此部分的目的，係可以系統性且有效率地將座談會中的發言內容進行收斂和重點萃取，因此，在語幹萃取後，透過與實際發言內容的對應比對分析，針對每個議題統整成以

下重點資訊：

議題一：關於建置平台的架構、預算、時間與營運模式之考量。

一、建置平台的架構：

- (1) 所有關於食宿遊購行的資訊，整合成一平台系統，服務給業者與消費者使用 (to B / C)。此平台需發揮各家業者的強項，且利用 API 串接政府機關與民間企業的資源，形成大資料庫給大家應用，讓觀光與交通運輸形成永續商業生態系。
- (2) 跨業界、民間資訊與政府資訊、票券與票閘口整合，平台如何規劃？
- (3) 尋求各業者對於投資此整合平台的服務之意願與看法。

二、平台建置預算、時間與營運模式之考量：

- (1) 是否影響既有市場平台與觀光旅遊業的經營？與民爭利？
- (2) 建置的經費與時間考量？
- (3) 平台資訊服務的營運模式以及 (政府 vs. 民間) 如何衡量？

議題二：整合食宿遊購行，平台 API 資訊相互介接之共享與資安問題。

資料串接之資安考量：

政府資源 (如國家基礎設施：臺、高鐵) 都會有此問題。

議題三：平台服務數位化，開發統一技術規格的 QRcode。

服務數位化：

目前交通運輸領域已由政府開發一版票卷規格 (QRcode)，希望未來其他業者也能沿用此規格技術，使票卷規格化、虛擬化，在所有的票閘口都能使用。

4.3 相似度分析

在相似度分析部分，主要係透過各單位的關鍵字詞頻，利用相似度分析之方法探討各單位間的相似性及同質性，相似性主要係瞭解發言內容是否同方向，採用餘弦相似度計算。確認其相似 (同方向) 性後，將相似性高的組合進一步採用 Jaccard 係數分析，以了解其發言內容的同質性，分析結果如後所示。

4.3.1 餘弦相似度

首先，本研究採用餘弦相似度來了解各單位間發言的相似程度，計算所有單位間 TF-IDF 向量的餘弦相似度，整體平均值為 0.0892，標準差為 0.0792，各單位間的相似度，如表 2、表 3 所示。進一步將各單位相互配對進行比較，共有 $13 \times 13 = 169$ 個配對，其中 62 個配對的餘弦相似度是高於整體平均值，與他人相似比例最高的為專家學者 2，達到 91.67%，其次為專家學者 1，達到 75%。

在業者中，與其他單位相似性比例最高的為業者 5 和業者 9，與他人的相似性比例為 41.67%。相似性高僅能表示討論的方向類似，然而，並無法表達出意見或內容的同質。因

表 2 各單位餘弦相似度列表 1

單位	專家學者 1	專家學者 2	業者 1	業者 2	業者 3	業者 4
專家學者 1	1	0.2239*	0.0554	0.0948*	0.1605*	0.108*
專家學者 2	0.2239*	1	0.1533*	0.2369*	0.2295*	0.1632*
業者 1	0.0554	0.1533*	1	0.1393*	0.011	0.0124
業者 2	0.0948*	0.2369*	0.1393*	1	0.0702	0.0765
業者 3	0.1605*	0.2295*	0.011	0.0702	1	0.0292
業者 4	0.108*	0.1632*	0.0124	0.0765	0.0292	1
專家學者 3	0.2197*	0.2289*	0.0636	0.1307*	0.3945*	0.0374
業者 5	0.1153*	0.3807*	0.1028*	0.0487	0.1611*	0.0257
業者 6	0.1385*	0.1392*	0.0393	0.0291	0.08	0.0336
業者 7	0.0892*	0.0896*	0.017	0.0136	0.0302	0.0467
業者 8	0.0528	0.234*	0	0.0663	0.0249	0.0466
業者 9	0.1453*	0.1921*	0.0445	0.0769	0.0856	0.089
專家學者 4	0.04	0.0511	0.0028	0.0835	0.0506	0.0183
相似數量	9	11	3	4	4	2
相似比例	75.00%	91.67%	25.00%	33.33%	33.33%	16.67%

*表示其相似度高於整體平均

表 3 各單位餘弦相似度列表 2

單位	專家學者 3	業者 5	業者 6	業者 7	業者 8	業者 9	專家學者 4
專家學者 1	0.2197*	0.1153*	0.1385*	0.0892*	0.0528	0.1453*	0.04
專家學者 2	0.2289*	0.3807*	0.1392*	0.0896*	0.234*	0.1921*	0.0511
業者 1	0.0636	0.1028*	0.0393	0.017	0	0.0445	0.0028
業者 2	0.1307*	0.0487	0.0291	0.0136	0.0663	0.0769	0.0835
業者 3	0.3945*	0.1611*	0.08	0.0302	0.0249	0.0856	0.0506
業者 4	0.0374	0.0257	0.0336	0.0467	0.0466	0.089	0.0183
專家學者 3	1	0.0829	0.1083*	0.0985*	0.0504	0.1083*	0.1209*
業者 5	0.0829	1	0.0075	0.1458*	0.0195	0.0226	0.0203
業者 6	0.1083*	0.0075	1	0.0364	0.0195	0.1216*	0.0346
業者 7	0.0985*	0.1458*	0.0364	1	0.0096	0.0468	0.0269
業者 8	0.0504	0.0195	0.0195	0.0096	1	0.1223*	0.0035
業者 9	0.1083*	0.0226	0.1216*	0.0468	0.1223*	1	0.0297
專家學者 4	0.1209*	0.0203	0.0346	0.0269	0.0035	0.0297	1
相似數量	8	5	4	4	2	5	1
相似比率	66.67%	41.67%	33.33%	33.33%	16.67%	41.67%	8.33%

*表示其相似度高於整體平均

此，產生並且確認各單位間的相似性後，本研究進一步分析相似性高的配對，將相似性高於整體平均的配對進行篩選，藉此瞭解各配對間對於議題和意見的同質性。

4.3.2 Jaccard 係數

計算完餘弦相似度後，為了瞭解相似性高的配對在每個議題的同質性或是內容的一致性。因此，本研究進一步針對餘弦相似度大於平均的兩單位計算 Jaccard 係數，詳細如表 4、表 5 所示。其中，表內值為 0 代表兩單位為自己或是餘弦相似度為小於平均數亦或是 Jaccard 係數為 0。其中，專家學者 1 在發言上，和專家學者 2 同質性最高，其次為業者 3。專家學者 2 在發言上，除了專家學者 1 外，和專家學者 3 同質性最高。專家學者 3，除了專家學者 2 外，和專家學者 1 的同質性最高。專家學者 4 和專家學者 3 的同質性最高。在業者部分，業者 1 和業者 7 與其他人的發言同質性都偏低，Jaccard 係數都未達到 0.1 以上。業者 2、業者 3、業者 4、業者 9 和專家學者 1 的同質性最高。業者 5 的部分和業者 3 的同質性最高，業者 6 和業者 8 與業者 9 的同質性最高。

計算 Jaccard 係數後，本研究進一步根據配對單位的文字雲及實際發言內容，統整各自針對的議題及內容。表 6 呈現各議題被相關性高且同質性高的配對單位提及的狀況。其中，相似性高且同質性高表示，大家討論的方向一致，且意見內容同質，以此作為重要度的判斷依據。其中，可以看到議題一為最主要相關探討之議題，相似性高且同質性高的配對次數達 23 次，與會者對於認為「建置平台的架構、預算、時間與營運模式之考量」相

表 4 各單位 Jaccard 係數列表 1

單位	專家學者 1	專家學者 2	業者 1	業者 2	業者 3	業者 4
專家學者 1	0	0.3258	0	0.2848	0.3072	0.1884
專家學者 2	0.3258	0	0.0513	0.2262	0.2439	0.1297
業者 1	0	0.0513	0	0.092	0	0
業者 2	0.2848	0.2262	0.092	0	0	0
業者 3	0.3072	0.2439	0	0	1	0
業者 4	0.1884	0.1297	0	0	0	0
專家學者 3	0.2842	0.3109	0	0.2516	0.2645	0
業者 5	0.1029	0.0847	0.0769	0	0.1591	0
業者 6	0.1727	0.1496	0	0	0	0
業者 7	0.084	0.069	0	0	0	0
業者 8	0	0.0431	0	0	0	0
業者 9	0.2252	0.1926	0	0	0	0
專家學者 4	0	0	0	0	0	0

表 5 各單位 Jaccard 係數列表 2

單位	專家學者 3	業者 5	業者 6	業者 7	業者 8	業者 9	專家學者 4
專家學者 1	0.2842	0.1029	0.1727	0.084	0	0.2252	0
專家學者 2	0.3109	0.0847	0.1496	0.069	0.0431	0.1926	0
業者 1	0	0.0769	0	0	0	0	0
業者 2	0.2516	0	0	0	0	0	0
業者 3	0.2645	0.1591	0	0	0	0	0
業者 4	0	0	0	0	0	0	0
專家學者 3	0	0	0.1119	0.0952	0	0.1603	0.1172
業者 5	0	0	0	0.0476	0	0	0
業者 6	0.1119	0	0	0	0	0.2195	0
業者 7	0.0952	0.0476	0	0	0	0	0
業者 8	0	0	0	0	0	0.1061	0
業者 9	0.1603	0	0.2195	0	0.1061	0	0
專家學者 4	0.1172	0	0	0	0	0	0

對重要。議題二「整合食宿遊購行，平台 API 資訊相互介接之共享與資安問題」次之，議題三「平台服務數位化，開發統一技術規格的 QRcode」則僅有專家學者 1 與業者 5 較著重探討，各單位間於相似議題的討論內容詳細於附錄二所示。

表 6 議題相關單位探討之狀況

議題	相似性高且同質性高單位	次數
議題一 關於建置平台的架構、預算、時間與營運模式之考量	專家學者 1 and 專家學者 2	23
	專家學者 1 and 業者 3	
	專家學者 1 and 業者 4	
	專家學者 1 and 專家學者 3	
	專家學者 1 and 業者 5	
	專家學者 1 and 業者 6	
	專家學者 1 and 業者 7	
	專家學者 1 and 業者 9	
	專家學者 2 and 業者 3	
	專家學者 2 and 業者 4	
	專家學者 2 and 專家學者 3	
專家學者 2 and 業者 6		

議題	相似性高且同質性高單位	次數
	專家學者 2 and 業者 7	
	專家學者 2 and 業者 8	
	專家學者 2 and 業者 9	
	業者 3 and 專家學者 3	
	業者 3 and 業者 5	
	專家學者 3 and 業者 6	
	專家學者 3 and 業者 7	
	專家學者 3 and 業者 9	
	業者 5 and 業者 7	
	業者 6 and 業者 9	
	業者 6 and 業者 8	
議題二 整合食宿遊購行，平台 API 資訊相互介接之共享與資安問題	專家學者 1 and 業者 2	5
	專家學者 1 and 業者 9	
	專家學者 2 and 高鐵	
	專家學者 2 and 業者 5	
	專家學者 2 and 業者 9	
議題三 平台服務數位化，開發統一技術規格的 QRcode	專家學者 1 and 業者 5	1
議題四 MaaS (公共運輸行動服務) 經驗	業者 3 and 專家學者 3	2
	專家學者 3 and 專家學者 4	
議題五 Umaji、Taiwan pass (好玩卡) 開發經驗	業者 1 and 業者 2	2
	業者 2 and 專家學者 3	

五、結論與建議

5.1 結論

Taiwan Pass 數位旅遊服務平台是為了提供國際旅客及國內旅客一站式觀光旅運服務，以顧客為核心 (User-Centered) 的整合性服務平台，從顧客觀點出發，整合觀光旅遊需要的服務包括餐飲 (食)、飯店旅館 (宿)、觀光景點 (遊)、特色商品 (購)、交通 (行) 所有面向。此一平台期望成為臺灣觀光整合性服務之開端，若能有妥善之營運暨建置規劃，可確保後續開發建置、服務推廣、營運移轉等得以符合相關規劃、使用者需求，以及營運永續性。本研究透過焦點座談方式，於座談會中蒐集 13 位產官學界專家的意見，從多角度檢視 Taiwan Pass 的規劃和定位，透過文字探勘和資料探勘方式，萃取重要關鍵意見和

議題，藉此了解產官學界對於平台定位的認知與建議。

整體來說，本研究所採用的文字探勘和資料探勘方法，可改善在質性研究中不易處理大量資料的缺點，有效地初探 Taiwan Pass 營運定位等關鍵議題。分析結果顯示，焦點團體座談的逐字稿內容透過文字探勘技術萃取出重要詞頻，利用詞頻分類組合，共有 5 個關鍵議題的產生，分別為：關於建置平台的架構、預算、時間與營運模式之考量；整合食宿遊購行，平台 API 資訊相互介接之共享、資安問題；平台服務數位化，開發統一規格技術的 QRcode；MaaS (公共運輸行動服務) 經驗；Umaji、好玩卡開發經驗。此外，本研究進一步利用文本相似度分析，了解參與焦點座談的學界或業界專家對於議題認同程度的異同，藉此收斂議題的歧見，達到共同意見萃取的目的，以作為未來 Taiwan Pass 推動的參考意見。

結果顯示議題一 (關於建置平台的架構、預算、時間與營運模式之考量) 同質性最高、其次為議題二 (整合食宿遊購行，平台 API 資訊相互介接之共享、資安問題)、再者為議題四 (MaaS (公共運輸行動服務) 經驗) 和議題五 (Umaji、好玩卡開發經驗)，同質性最低的為議題三 (平台服務數位化，開發統一規格技術的 QRcode)。這結果也意味著，與會人員中大家對於議題一的著墨和討論內容相似程度較高，對於議題三的著墨較少或共識較少，此結果可作為未來規劃和推動的參考。

最後，利用文本分析和資料探勘方法進行質性研究，透過資料預處理以及自然語言處理技術可以系統性的進行語幹的萃取，達到質性研究深度議題挖掘的目的。採用此方法可以減少人為的判斷失誤和主觀的偏好，同時也可以達到提升分析效率。另外，本研究進一步利用文本分析獲得的向量和詞頻進行相似性比對和同質性分析，可以用客觀且量化的方式了解發言者之間對議題的意見相似程度以及一致性，在研究中可以發現與會者對於學者 1 的發言有著高度的相似和同質，也可以了解在不同議題間與會者對議題的相似度和一致性。研究結果可以知道業者 1 不論在哪個議題間，和與會者的相似性和一致性都偏低。這些是過往在質性研究中不易達成的部分，未來在 Taiwan Pass 推廣的過程中，可以參考業者間的意見和對於議題的一致程度進行相關策略的調整以及政策說明，達到精準政策的目標。

5.2 建議

以下針對本研究的方法及結果提出相關建議，供未來研究以及相關單位推展 Taiwan Pass 參考。

首先，在本研究採用的方法部分，質性研究往往受限於資料處理成本高，因此，在相關研究中，往往只能採利用小型的焦點團體座談來蒐集資料。以下建議供未來研究參考：

1. 建議未來納入消費者端的意見

目前在 Taiwan Pass 的平台規劃只蒐集產官學的意見，現行以消費者為核心的世代，建議未來納入消費者端的意見參考，讓平台的規劃更為完善。

2. 資料蒐集方式可以更多元

本研究只以一次座談會作為資料來源，本研究已驗證可以利用文字探勘方式來進行質性研究分析，也獲得不錯的效果，建議未來可以納入更多的資訊，或是蒐集多個座談的資料來進行相關的意見萃取，以獲取更具代表性的意見。隨著社群網路的發達，有許多意見都會透過網路傳遞，未來在資料蒐集部分可以蒐集網路輿情資料，可以讓意見表達更加全面。

3. 議題生成方式可以智慧化

本研究所採用的議題生成，係基於語幹萃取與逐字稿對應後生成，目前還是仰賴專家學者的協助，現行已經有文句生成的相關技術發展，未來可以考量利用語幹作為基礎進行文句生成，來達到更客觀的作法。在議題生成部分，可以考慮文字生成技術，透過人工智慧或機械學習演算法進行議題或內容的生成，可以達到更有快速的分析。

另外，在 Taiwan Pass 營運定位部分，未來若建置 Taiwan Pass 平台，國內業者可加入此平台，提供自身之商品或服務，供使用者進行查詢與預訂，並藉由金流清分系統，得到拆分後之收益。主管機關與業者則可透過此平台之分析，了解具時效性的觀光市場動態，以協助業者拓展商機，支援主管機關進行觀光決策等。因此，Taiwan Pass 數位旅遊服務整合平台的營運定位，對於規劃和建置至關重要，以下幾點作為未來推動的建議：

1. 關於建置平台的架構、預算、時間與營運模式之考量

(1) 平台建置定位

關於食宿遊購行的資訊，需整合成一個平台系統，並提供給服務業者與消費者使用，讓各家業者透過此平台發揮各自強項，且建構大數據資料庫給業者和政府做有效的利用，透過資料科學的應用讓觀光與交通運輸朝向永續的目標。跨業界、民間資訊與政府資訊、票券與票閘口必須在平台建置前完成整合，建置的經費與時間必須妥善考量，以提高業者對於投資平台與投入服務之意願。

(2) 營運模式定位

需要與既有市場平台與觀光旅遊業的經營訂出區隔，避免與民爭利，平台的資訊服務必須訂出政府和民間的使用定位。

2. 整合食宿遊購行，平台 API 資訊相互介接之共享與資安問題

Taiwan Pass 平台進行資料串接需考量資安相關議題，特別是與政府資源或是國家基礎建設（例如：臺、高鐵）都會有此問題產生，需避免資安對於相關單位的營運造成的衝擊與影響。

3. Taiwan Pass 數位旅遊服務整合平台服務數位化，開發統一技術規格的 QRcode

考量服務數位化的目標下，開發一版票卷規格（QRcode）是途徑之一，目前交通部已經有既定的規範，希望未來其他業者也能沿用此規格技術，使票卷規格化、虛擬化，在所有的票閘口都能使用。

4. MaaS（公共運輸行動服務）經驗

可參考過去整合客運的經驗，從了解業者實際需求出發，避免政府專案與系統商投入之後，以失敗或殭屍 APP 告終。此外，整合交通運輸，對於食宿遊購行、公部門、私部門整合來說，會面臨困難和挑戰。參考過去在 MaaS 推動上，遇到的困難與困境，都需要政府協助突破，因此，需要更妥善的規劃與設計。

5. Umaji、Taiwan pass (好玩卡) 開發經驗

過去串接 API 的經驗，針對一般訂票作業做相關開發或傳送交易資訊，面臨受到限制的問題且不夠彈性，像是團體優惠價的分票、退票都會產生問題。特別是與交通 (臺、高鐵) 整合不容易，Umaji 的經驗是值得參考的。

另外，本研究也針對未來研究提出相關的建議。首先，目前在 Taiwan Pass 的平台規劃只蒐集產官學的意見，現行以消費者為核心的世代，建議未來納入消費者端的意見參考，讓平台的規劃更為完善。其次，本次只以一次座談會作為資料來源，本研究已驗證可以利用文字探勘方式來進行質性研究分析，也獲得不錯的效果，建議未來可以納入更多的資訊，或是蒐集多個座談的資料來進行相關的意見萃取，以獲取更具代表性的意見。最後，在方法的部分，本研究議題生成的部分仍仰賴專家學者進行，現行已經有文句生成的相關技術發展，未來可以考量利用語幹作為基礎進行文句生成，來達到更客觀的作法。

參考文獻

1. Mishra, L., "Focus Group Discussion in Qualitative Research", *Techno Learn*, Vol.6, No. 1, 2016, pp.1-5.
2. Krueger, R. A., & Casey, M. A., *Focus Groups: A Practical Guide for Applied Researchers* 3rd ed., Thousand Oaks, CA: Sage, 2000.
3. Onwuegbuzie, A. J., Dickinson, W. B., Leech, N. L., & Zoran, A. G., "A Qualitative Framework for Collecting and Analyzing Data in Focus Group Research", *International Journal of Qualitative Methods*, Vol.8, Iss.3, 2009, pp.1- 21.
4. Ming, N., & Baumer, E., "Using Text Mining to Characterize Online Discussion Facilitation", *Journal of Asynchronous Learning Networks*, Vol.15, Iss.2, 2011, pp.71-109.
5. Yu, C. H., Jannasch-Pennell, A., & DiGangi, S., "Compatibility between Text Mining and Qualitative Research in the Perspectives of Grounded Theory, Content Analysis, and Reliability", *The Qualitative Report*, Vol.16, No.3, 2011, pp.730-744.
6. Morgan, D. L., *Focus Groups as Qualitative Research* 2nd ed., Thousand Oaks, CA: Sage, 1997.
7. Sim, J., "Collecting and Analyzing Qualitative Data: Issues Raised by the Focus Group", *Journal of Advanced Nursing*, Vol.28, No.2, 1998, pp.345-352.
8. 黃志偉，「旅業觀點」，*旅報 TTN Taiwan*，第 1041 期，民國 107 年，頁 9。
9. World Tourism Organization, *New Platform Tourism Services (or the so-called Sharing Economic) - Understand, Rethink and Adapt*, 2017.

10. World Tourism Cities Federation, 世界旅遊經濟趨勢報告, 2018。
11. Rice, P. L., and Ezzy, D., *Qualitative Research Methods: A Health Focus*, South Melbourne (Australia): Oxford University Press, 1999.
12. Anderson, G., *Fundamentals of Educational Research*, London: The Falmer Press, 1990.
13. Denscombe, M., *EBOOK: The Good Research Guide: For Small-Scale Social Research Projects*. McGraw-Hill Education (UK), 2017.
14. Patton, M. Q., *Qualitative Research & Evaluation Methods, 3rd Edition*, Sage Publications, Inc, 2002.
15. Wisker, G., *The Postgraduate Research Handbooks*, Great Britain: Palgrave, 2001.
16. Kidd, P. S., & Parshall, M. B., "Getting the Focus and the Group: Enhancing Analytical Rigor in Focus Group Research", *Qualitative Health Research*, Vol.10, No.3, 2000, pp.293-308.
17. Glaser, B. G., *Discovery of Grounded Theory*, Chicago: Aldine, 1999.
18. Glaser, B. G., & Strauss, A. L., *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Chicago: Aldine, 1967.
19. Strauss, A., & Corbin, J., *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, Thousand Oaks, CA: Sage, 1998.
20. Gegick, M., Rotella, P., and Xie, T., "Identifying Security Bug Reports via Text Mining: An Industrial Case Study", *2010 7th IEEE Working Conference on Mining Software Repositories (MSR 2010)*, IEEE, 2010, pp.11-20.
21. Imbar, V., Radiant. Adelia, Ayub, M., dan Rehatta, A., "Implementasi Cosine Similarity dan Algoritma Smith Waterman untuk Mendeteksi Kemiripan Teks". *Jurnal Informatika*, Vol.10, No. 1, 2014, pp.31-42.
22. Nurdiana, O., J. Jumadi, and D. Nursantika, "Perbandingan Metode Cosine Similarity Dengan Metode Jaccard Similarity Pada Aplikasi Pencarian Terjemah Al-Qur'an Dalam Bahasa Indonesia", *Journal Online Informatika*, Vol.1, No.1, 2016, pp.59-63.
23. Priambodo, J., "Pendeteksian Plagiarisme Menggunakan Algoritma Rabin-Karp Dengan Metode Rolling Hash", *Journal Informatika Universitas Pamulang*, Vol.3, No.1, 2018, pp.39-45.
24. Li, N. and D. D. Wu, "Using Text Mining and Sentiment Analysis for Online Forums Hotspot Detection and Forecast", *Decision Support Systems*, Vol.48, No.2, 2010, pp.354-368.
25. Roy, S., and S. Dandapat, A. Nagesh, and Y. Narahari, "Wisdom of Students: A Consistent Automatic Short Answer Grading Technique", *Proceeding 13th International Conference Natural Language Processing*, 2016, pp.178-187.
26. Zhang, Yun-Tao., Gong Ling., and Wang Yong-cheng., "An Improved TF-IDF Approach for Text Classification", *Journal of Zhejiang University-Science Applied Physics & Engineering*, Vol.6, No.1, 2005, pp.49-55.
27. Boullis, C., and Ostendorf, M., "Text Classification by Augmenting the Bag-of- Words Representation with Redundancy-Compensated Bigrams", *Proceedings of the SIAM International Conference on Data Mining at the Workshop on Feature Selection in Data*

- Mining (SIAM-FSDM 2005), Newport Beach, California, 2005.
28. Liu, C. Z., Sheng, Y. X., Wei, Z. Q., & Yang, Y. Q., “Research of Text Classification Based on Improved TF-IDF Algorithm”, In 2018 IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE), 2018, pp.218-222.
 29. Qaiser, S., & Ali, R., “Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents”, *International Journal of Computer Applications*, Vol.181, No.1, 2018, pp.25-29.
 30. Zhu, Z., Liang, J., Li, D., Yu, H., & Liu, G., “Hot Topic Detection Based on A Refined TF-IDF Algorithm”, *IEEE Access*, Vol.7, 2019, pp.26996-27007.
 31. Kim, D., Seo, D., Cho, S., & Kang, P., “Multi-Co-Training for Document Classification Using Various Document Representations: TF-IDF, LDA, and Doc2Vec”, *Information Sciences*, Vol.477, 2019, pp.15-29.
 32. Hall, P. A. V. and Dowling, G. R., “Approximate String Matching”, *Computing Surveys*, Vol.12, No.4, 1980, pp.381-402.
 33. Trstenjak, B., Mikac, S., & Donko, D., “KNN with TF-IDF based Framework for Text Categorization”, *Procedia Engineering*, Vol.69, 2014, pp.1356-1364.
 34. Bafna, P., Pramod, D., and Vaidya, A., “Document Clustering: TF-IDF Approach”, International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), Chennai, 2016, pp.61-66.

附錄一

字詞	次數	字詞	次數	字詞	次數	字詞	次數
業者	128	食宿	16	遊購行	9	規範	6
系統	90	票卷	16	TPS	9	觀光局	6
平台	82	規劃	15	通路	9	APP	6
整合	77	時間	15	競爭	9	障礙	6
服務	76	臺高鐵	15	主導	9	票價	6
高鐵	72	運具	14	悠遊卡	8	公益	6
政府	61	kkday	14	日本	8	交易	6
旅行社	45	QRcode	14	乘車碼	8	限制	6
旅遊	44	klook	14	串接	8	遊購	6
資訊	41	功能	14	預算	8	通勤	6
客運	40	臺北	14	團體	8	供應商	6
MaaS	39	旅行	13	設備	8	面臨	6
介接	35	交通部	13	重要	8	經費	6
Taiwan pass	35	電子	13	公營	8	意願	6
票	34	標案	13	結合	7	出資	5
運輸	34	消費者	13	上架	7	IBM	5
臺灣	32	開放	13	民眾	7	Maa	5
營運	30	銷售	13	技術	7	永續	5
交通	30	投資	13	做法	7	支持	5
臺鐵	28	架構	12	業界	7	流量	5
產品	28	to C	12	車站	7	桃捷	5
部門	27	政策	12	國家	7	公權力	5
計劃	27	市場	12	公共	7	利益	5
廠商	26	相關	12	to B	7	mango	5
案子	26	國道	12	考量	7	國旅	5
遊程	26	發展	12	合作	7	困難	5
資源	26	旅客	12	經營	7	基礎	5
Umajj	25	需求	12	資安	7	資料	5
北捷	23	pass	11	核銷	7	賺錢	5
API	23	GDS	10	複雜	7	專案	5
建置	22	公總	10	客戶	7	住宿	5
票證	20	商品	10	成本	6	統一	5
捷運	20	景點	10	結構	6	理想	5
民間	18	觀光	10	商業	6	手機	5
好玩卡	17	興趣	10	成功	6	協調	5

字詞	次數	字詞	次數	字詞	次數	字詞	次數
產業	5	B to C	4	建置期	3	票種	2
中央	4	飯店	4	後端	3	規格化	2
北北基	4	顧問	4	得標	3	雙鐵	2
新北	4	旅運	4	票卡	3	數位化	2
壓力	4	法人	4	社區	3	實體卡	2
假期	4	民營	4	手續費	3	旅行業	2
公車	4	錢包	4	系統商	3	客制化	2
卡通	4	爭利	4	獨立	3	團體票	2
交換	4	特色	4	招標	3	電動車	2
人工	4	business model	4	採購法	3	宏碁	2
優惠	4	獲利	4	擴大	3	標準化	2
前端	4	解說	3	民宿	3	臺南	2
工程	4	國際	3	標準	3	刷卡	2
中華	4	訂房	3	日月潭	3	行政院	2
國外	4	國內	3	城際	3	英語化	2
營運權	4	設施	3	市長	3	赫爾辛基	1
行銷	4	讀卡機	3	紙本	3	逢甲	1
衙門	4	變化	3	署長	3	RFI	1
運研所	4	to E	3	行程	3	RFID	1
local trip	4	商務	3	學習	3	身份	1
票閘口	4	呼應	3	背後	3	藝術節	1
level	4	費用	3	官方	3	訂票	1
解釋	4	臺鐵局	3	事業	3	虛擬卡	1
運營	4	車票	3	聯合	3	TGO	1
遊戲	4	載具	3	網站	3	Uber Eats	1
OTA	4	便利	3	結案	3	遠鑫卡	1
積極	4	都會區	3	組裝	3	QR	1
退票	4	金額	3	市區	3	MaaS 卡	1
轉運站	4	銜接	3	Q&A	2	PDS	1
訂位	4	願景	3	TBS	2	OTM	1
資料庫	4	餐廳	3	亞太	2	48:28:00	1
線上	4	供需	3	TEXpress	2	高捷	1
美國	4	套票	3	google	2	馬來西亞	1
邏輯	4	地區	3	TDS	2	餘光黨	1
生意	4	世界大同	3	uber	2	G to	1
規格	4	籌組	3	transtaipei	2	非實體卡	1
B to B	4	科顧室	3	高鐵票	2	GTS	1

字詞	次數	字詞	次數	字詞	次數	字詞	次數
IT	1	船票	1	臺南市	1	利機	1
JR	1	市政府	1	臺北市	1	票卷商	1
JTA	1	屏東縣政府	1	印尼	1	東南投	1
KPI Check	1	主動	1	博物館票	1	東南亞	1
MOU	1	實體票	1	加拿大	1	三日卡	1
難度	1	客服	1	中興	1	中捷	1
芬蘭	1	五日卡	1	中端	1	三日通	1
郭台銘	1	五日通	1	中文化	1	故宮卡	1
foodpanda	1	國稅局	1	文化節	1	所有卡	1
公協會	1	噶瑪蘭	1	立法院	1	中共	1
ilook	1	臺灣版	1	科顧室	1	中國	1
念茲在茲	1	捷運卡	1				

附錄二

專家學者 1	專家學者 2
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
希望建置平台的架構：	
(1) 所有關於食宿遊購行的資訊，整合成一平台系統，服務給業者與消費者使用 (to B / C)。此平台需發揮各家業者的強項，且利用 API 串接政府機關與民間企業的資源，形成大資料庫給大家應用，讓觀光與交通運輸形成永續商業生態系。	(1) 系統後端介接數個業者 (住宿、遊程、交通、餐廳等)，而前端通路由旅行業或政府社區導覽遊程包裝成一產品銷售 (類似 GDS)，會比在各自平台銷售來得有效益。
平台建置預算、時間與營運模式之考量：	
(1) 建置的經費與時間考量？ (2) 是否影響既有市場平台與觀光旅遊業的經營？與民爭利？	(2) 系統會是一龐大且複雜的架構，且開發成本可能會超過好幾百萬。 (3) 此平台不去干擾業者原有營運規則，不跟既有的業者競爭。

專家學者 1	業者 2
議題二：整合食宿遊購行，平台 API 資訊相互介接之共享、資安問題	
(1) 政府資源 (如國家基礎設施：台、高鐵) 都會有此問題。	(1) 高鐵與臺鐵屬於國家一級設施，因資安問題而不開放第三方直接使用。

專家學者 1	業者 3
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台建置預算與營運模式之考量：	
(1) 是否影響既有市場平台與觀光旅遊業的經營？與民爭利？ (2) 建置的經費與時間考量？ (3) 平台資訊服務的營運模式以及 (政府 vs. 民間) 如何衡量？	(1) 未來整合方案要讓消費者有自主選擇權，而不是單純由提供固定方案。 (2) 整合方案的業者分潤問題，必須有公總的角色存在。

專家學者 1	業者 4
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
跨業界平台整合問題：	
(1) 跨業界、民間資訊與政府資訊、票券與票閘口整合，平台如何規劃？	(1) 住宿、local trip 等服務、種類變化較多，以住宿來說，就有分民宿、旅館飯店、青年旅

(2) 尋求各業者對於投資此整合平台的服務之意願與看法。	館等等，難以像交通（台鐵、高鐵、捷運、客運）一樣標準化。 (2) 住宿方面沒有比較大的機關可以負責主導小眾業者進入平台。
------------------------------	---

專家學者 1	專家學者 3
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
希望建置平台的架構：	
(1) 所有關於食宿遊購行的資訊，整合成一平台系統，服務給業者與消費者使用 (to B / C)。此平台需發揮各家業者的強項，且利用 API 串接政府機關與民間企業的資源，形成大資料庫給大家應用，讓觀光與交通運輸形成永續商業生態系。	(1) 此平台可以得到較多政府部門掌握的運輸資源，與食宿遊購整合一起，成為 business model，能自給自足。
平台建置營運模式之考量：	
(1) 是否影響既有市場平台與觀光旅遊業的經營？與民爭利？ (2) 平台資訊服務的營運模式以及（政府 vs. 民間）如何衡量？	(1) 平台靠介接是否能維持營運？ (2) 如何不與民爭利？

專家學者 1	業者 5
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台預算、時間、營運模式之考量？	
(1) 是否影響既有市場平台與觀光旅遊業的經營？與民爭利？ (2) 建置的經費與時間考量？ (3) 平台資訊服務的營運模式以及（政府 vs. 民間）如何衡量？	(1) 景點與交通的進出閘口裝置（刷統一票證 QRcode），中央是否有計畫輔助？ (2) 以公共服務角度來看，北捷認為此計畫對自己沒有利益。
議題三：平台服務數位化，開發統一規格技術的 QRcode	
(1) 目前交通運輸已開發一版票卷規格 (QRcode)，希望未來其他業者也能沿用此規格技術，使票卷規格化、虛擬化，在所有的票閘口都能使用。	(1) 北捷沒有 QRcode 系統，公總的交通 QR 試辦計畫與北捷系統無關。

專家學者 1	業者 6
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
跨業界平台整合問題：	

(1) 跨業界、民間資訊與政府資訊、票券與票閘口整合，平台如何規劃？ (2) 尋求各業者對於投資此整合平台的服務之意願與看法。	(1) 此平台如果成真，可解決旅行社面對不同層級的供應商與 API 串接。
平台預算、時間與營運模式之考量？	
(1) 是否影響既有市場平台與觀光旅遊業的經營？與民爭利？ (2) 建置的經費與時間考量？ (3) 平台資訊服務的營運模式以及(政府 vs. 民間)如何衡量？	(1) 案子的範圍訂定，涉及預算、時間等？ (2) 是否背後有願景政府未來可去主導？讓國內外旅客都願意使用。

專家學者 1	業者 7
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台營運模式考量：	
(1) 是否影響既有市場平台與觀光旅遊業的經營？與民爭利？ (2) 建置的經費與時間考量？ (3) 平台資訊服務的營運模式以及 (政府 vs. 民間) 如何衡量？	(1) 對於交通部很多計畫為了結案驗收，都與既有的單位平台結合 (如：kkday/klook)，未來如果計畫成熟要再轉型使用 Taiwan pass，可能又有成本支出。

專家學者 1	業者 9
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台營運模式考量：	
(1) 是否影響既有市場平台與觀光旅遊業的經營？與民爭利？ (2) 建置的經費與時間考量？ (1) 平台資訊服務的營運模式以及 (政府 vs. 民間) 如何衡量？	(1) 未來的型態、結構與法源根據？待與政府溝通。 (2) 此平台服務須符合 to B、E、C 需求。 (3) 簡單上架機制，完全競爭平台。
議題二：整合食宿遊購行，平台 API 資訊相互介接之共享、資安問題	
(1) 政府資源 (如國家基礎設施：台、高鐵) 都會有此問題。	(1) 資訊介接之困難度，API 介接越多，資安問題也越多。

專家學者 2	業者 1
無相同討論主題，只是都在探討 API	
因高鐵提出本身 API 有限制問題，台鐵應該也差不多，而熊正一老師對此話題有做呼	臺鐵發言：簡單介紹他們 API 串接的項目。

應。 (1) API 串接有限制問題與必須有公權力介入，希望公權力佔比維持一半以下，但需扮演維護資安、客戶資訊的角色。	
--	--

專家學者 2	高鐵
議題二：整合食宿遊購行，平台 API 資訊相互介接之共享、資安問題	
針對高鐵 API 限制問題回答： (1) API 串接有限制問題與必須有公權力介入，希望公權力佔比維持一半以下，但需扮演維護資安、客戶資訊的角色。	(1) 串接 11 支 API，而 API 串接有限制問題並非彈性，像是團體優惠價的分票、退票問題。

專家學者 2	業者 3
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台營運模式考量：	
(1) 平台客服要由提供商品的廠商做對應。 (2) 此平台不去干擾業者原有營運規則，不跟既有的業者競爭。	(1) 未來整合方案要讓消費者有自主選擇權，而不是單純由提供固定方案。 (2) 整合方案的分潤問題，必須有公總的角色存在。

專家學者 2	業者 4
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
跨業界平台整合問題：	
(1) 系統後端介接數個業者 (住宿、遊程、交通、餐廳等)，而前端通路由旅行業或政府社區導覽遊程包裝成一產品銷售 (類似 GDS)，會比在各自平台銷售來得有效益。	(1) 住宿、local trip 等服務、種類變化較多，以住宿來說，就有分民宿、旅館飯店、青年旅館等等，難以像交通 (台鐵、高鐵、捷運、客運) 一樣標準化。 (2) 住宿方面沒有比較大的機關可以負責主導小眾業者進入平台。

專家學者 2	專家學者 3
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
希望建置平台的架構：	
(1) 系統後端介接數個業者 (住宿、遊程、交通、餐廳等)，而前端通路由旅行業或政府社區導覽遊程包裝成一產品銷售 (類似 GDS)，會比在各自平台銷售來得有效益。	(1) 此平台可以得到較多政府部門掌握的運輸資源，與食宿遊購整合一起，成為 business model，能自給自足。

平台營運模式考量：	
(1) 平台客服要由提供商品的廠商做對應。 (2) 此平台不去干擾業者原有營運規則，不跟既有的業者競爭。	(1) 平台靠介接是否能維持營運？ (2) 如何不與民爭利？

專家學者 2	業者 5
議題二：整合食宿遊購行，平台 API 資訊相互介接之共享、資安問題	
<ul style="list-style-type: none"> 與 (交通) 捷運整合問題： (1) 銷售產品基本上希望綁定四都捷運業者，未來是否都可接受非實體化票卷的 QRcode？ (2) 是否可把捷運一日卡、三日卡、五日卡上到平台去賣？ 	<ul style="list-style-type: none"> (1) 北捷沒有 QRcode 系統，公總的交通 QR 試辦計畫與北捷系統無關。 (2) 一日票、二日票、三日票是實體卡，不具電子票證功能。

專家學者 2	業者 6
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
跨業界平台整合問題：	
<ul style="list-style-type: none"> 與 (交通) 捷運整合問題： (1) 銷售產品基本上希望綁定四都捷運業者，未來是否都可接受非實體化票卷的 QRcode？ (2) 是否可把捷運一日卡、三日卡、五日卡上到平台去賣？ 與 (交通) 臺鐵、高鐵有關： (1) API 串接有限制問題與必須有公權力介入，希望公權力佔比維持一半以下，但需扮演維護資安、客戶資訊的角色。 	<ul style="list-style-type: none"> 與 (旅遊) 整合問題： (1) 此平台如果成真，可解決旅行社面對不同層級的供應商與 API 串接。
平台預算、時間、營運模式之考量？	
<ul style="list-style-type: none"> (1) 系統會是一龐大且複雜的架構，且開發成本可能會超過好幾百萬。 (2) 系統後端介接數個業者（住宿、遊程、交通、餐廳等），而前端通路由旅行業或政府社區導覽遊程包裝成一產品銷售（類似 GDS），會比在各自平台銷售來得有效益。 (3) 平台客服要由提供商品的廠商做對應。 (4) 此平台不去干擾業者原有營運規則，不 	<ul style="list-style-type: none"> (1) 案子的範圍訂定，涉及預算、時間等？ (2) 是否背後有願景政府未來可去主導？讓國內外旅客都願意使用。

跟既有的業者競爭。	
-----------	--

專家學者 2	業者 7
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台營運模式考量：	
(1) 系統會是一龐大且複雜的架構，且開發成本可能會超過好幾百萬。 (2) 此平台不去干擾業者原有營運規則，不跟既有的業者競爭。	(1) 對於交通部很多計畫為了結案驗收，都與既有的單位平台結合（如：kkday/klook），未來如果計畫成熟要再轉型使用 Taiwan pass，可能又有成本支出。

專家學者 2	業者 8
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台營運模式考量：	
(1) 平台客服要由提供商品的廠商做對應。 (2) 此平台不去干擾業者原有營運規則，不跟既有的業者競爭。	(1) 後端問題：客戶遇到問題，客服是誰要出來對應？ (2) 在平台銷售、包裝自家產品，是否影響自己本身的生意？

專家學者 2	業者 9
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台營運模式考量：	
(1) 系統後端介接數個業者（住宿、遊程、交通、餐廳等），而前端通路由旅行業或政府社區導覽遊程包裝成一產品銷售（類似 GDS），會比在各自平台銷售來得有效益。	(1) 未來的型態、結構與法源根據？待與政府溝通。 (2) 此平台服務須符合 to B、E、C 需求。 (3) 簡單上架機制，完全競爭平台。
議題二：整合食宿遊購行，平台 API 資訊相互介接之共享、資安問題	
<ul style="list-style-type: none"> 與（交通）整合有關： (1) API 串接有限制問題與必須有公權力介入，希望公權力佔比維持一半以下，但需扮演維護資安、客戶資訊的角色。	(1) 資訊介接之困難度，API 介接越多，資安問題也越多。

業者 1	業者 2
議題五：Umaji、Taiwan pass (好玩卡) 開發經驗	
(1) 串接 16 支 API，針對一般訂票作業做相關開發，傳送交易資訊。	(1) 串接 11 支 API，而 API 串接有限制問題並非彈性，像是團體優惠價的分票、退票問

	題。
--	----

業者 1	業者 5
沒有相關連主題	
(1) 臺鐵說明自己 API 串接情況。	(1) 北捷說明此案跟自己本身沒有什麼利益連結。

高鐵	專家學者 3
議題五：Umaji、Taiwan pass (好玩卡) 開發經驗	
(1) 串接 11 支 API，而 API 串接有限制問題並非彈性，像是團體優惠價的分票、退票問題。	(1) 變成城際型態。 (2) 與交通 (臺、高鐵) 整合不容易。

業者 3	專家學者 3
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台營運模式考量：	
(1) 未來整合方案要讓消費者有自主選擇權，而不是單純由提供固定方案。 (2) 整合方案的分潤問題，必須有公總的角色存在。	(1) 平台靠介接是否能維持營運？如何不與民爭利？
議題四：MaaS (公共運輸行動服務) 經驗	
<ul style="list-style-type: none"> 整合客運過去經驗： (1) 由於不了解客運業者，政府專案與系統商投入多以失敗、殭屍 APP 告終。 (2) 自行開發 TBS 系統，整合 11 家客運公司。 	<ul style="list-style-type: none"> MaaS 推動： (1) 整合交通運輸，對於食宿遊購行、公部門、私部門整合來說，有困難。

業者 3	業者 5
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台經費與營運模式考量：	
(1) 未來整合方案要讓消費者有自主選擇權，而不是單純由提供固定方案。 (2) 整合方案的分潤問題，必須有公總的角色存在。	(1) 景點與交通的進出閘口裝置 (刷統一票證 QRcode)，中央是否有計畫輔助？ (2) 以公共服務角度來看，北捷認為此計畫對自己沒有利益。

專家學者 3	業者 6
--------	------

議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台經費、時間與營運模式考量：	
(1) 平台靠介接是否能維持營運？如何不與民爭利？	(1) 案子的範圍訂定，涉及預算、時間等？ (2) 是否背後有願景政府未來可去主導？讓國內外旅客都願意使用。
希望建置平台的架構	
(1) 此平台可以得到較多政府部門掌握的運輸資源，與食宿遊購整合一起，成為 business model，能自給自足。	(1) 此平台如果成真，可解決旅行社面對不同層級的供應商與 API 串接。

專家學者 3	業者 7
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
希望建置平台的架構	
(1) 此平台可以得到較多政府部門掌握的運輸資源，與食宿遊購整合一起，成為 business model，能自給自足。	(1) 對於交通部很多計畫為了結案驗收，都與既有的單位平台結合（如：kkday/klook），未來如果計畫成熟要再轉型使用 Taiwan pass，可能又有成本支出。

專家學者 3	業者 9
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台營運模式考量：	
(1) 平台靠介接是否能維持營運？如何不與民爭利？	(1) 未來的型態、結構與法源根據？待與政府溝通。 (2) 此平台服務須符合 to B、E、C 需求。 (3) 簡單上架機制，完全競爭平台。

專家學者 3	專家學者 4
議題四：MaaS (公共運輸行動服務) 經驗	
(1) 整合交通運輸，對於食宿遊購行、公部門、私部門整合來說，有困難。	(1) 日本國土交通省來觀摩臺灣的 MaaS 兩次。代表做得不錯。

業者 5	業者 7
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台經費、營運模式考量：	
(1) 景點與交通的進出開口裝置（刷統一票證 QRcode），中央是否有計劃輔助？	(1) 對於交通部很多計畫為了結案驗收，都與既有的單位平台結合（如：kkday/klook），未

	來如果計畫成熟要再轉型使用 Taiwan pass，可能又有成本支出。
--	-------------------------------------

業者 6	業者 9
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台經費、時間與營運模式考量：	
(1) 案子的範圍訂定，涉及預算、時間等？ (2) 是否背後有願景政府未來可去主導？讓國內外旅客都願意使用。	(1) 未來的型態、結構與法源根據？待與政府溝通。 (2) 此平台服務須符合 to B、E、C 需求。 (3) 簡單上架機制，完全競爭平台。

業者 6	業者 8
議題一：關於建置平台的架構、預算、時間與營運模式之考量	
平台整合、營運模式考量：	
(1) 此平台如果成真，可解決旅行社面對不同層級的供應商與 API 串接。	(1) 未來的型態、結構與法源根據？待與政府溝通。 (2) 此平台服務須符合 to B、E、C 需求。 (3) 簡單上架機制，完全競爭平台。

跟團旅遊怎麼

縮水了!

1

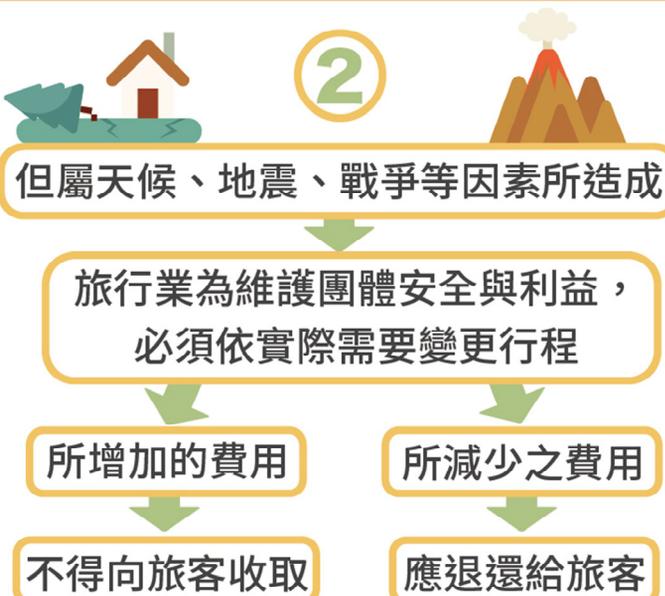
旅行業若未依旅遊契約所訂定的旅程、交通、食宿或遊覽等項目執行，消費者可以要求賠償**2倍**差額違約金。



國外旅遊定型化契約
應記載及不得記載事項



國內旅遊定型化契
約應記載及不得記載事項



行政院消費者保護處 廣告