

Web3 數據分析與資料視覺化之研究

林億雄 *

助理研究員
國立高雄大學教學與發展中心
E-mail: yhslin@nuk.edu.tw

馬瀾嘉

副教授
國立成功大學統計學系暨數據科學研究所
E-mail: mcma@ncku.edu.tw

李育齊

助理教授
長庚科技大學通識教育中心
E-mail: jomireli@mail.cgust.edu.tw



摘要

Web3 數據分析和資料視覺化是數據科學領域一個重要探討議題。然，現今對於 Web3 數據分析處理幾乎沒有有效的系統化方法。因此，本研究引入 ELT 程序探討 Web3 數據分析的方法和呈現 Web3 資料視覺化的設計，並且實作於二個範例項目 OpenSea 次級 NFT 交易市場數據，與 STEP N 使用者數據。希冀透過現代新穎之網路科技技術與數據分析系統，將此類 Web3 資料分析之結論轉化為可供使用的知識與行動。

關鍵詞：數據分析、資料視覺化、第三代互聯網

* 為本文通訊作者

壹、緒論

我國政府於 2022 年 8 月 27 日正式成立數位發展部，規模為 1 部 6 司 6 處 2 署，屬於中央二級行政機關。其中，民主網絡司下成立多元宇宙科，該科核心業務為（1）全球 Web3 數位服務之研發、專案規劃與推動（2）參與全球分散式自治組織活動等。隨著政府部門日漸重視 Web3 創新發展，Web3 一詞迅速成為大家討論的對象。Web3 為指第三代互聯網，主要由分散式帳本技術所支援，一般常與區塊鏈去中心化、非同質性代幣（NFT）、元宇宙、加密貨幣、虛擬錢包、去中心化自治組織（DAO），及區塊鏈遊戲經濟（Gamefi）等同被討論。在 Web3 世界裡，所有權及掌控權是去中心化，這是基於區塊鏈的去中心化技術，也是驅動元宇宙的基礎建設技術。Web3 能否在可見將來成為主流，仍是未知之數。但，在區塊鏈去中心化技術熱潮帶動下，Web3 不再只是時下時髦流行用語，而可能是顛覆網路世界的科技革命。現行 Web2 在資料收集方面為依據主從架構資料庫設計，主要依賴機構或個人中心化伺服器。Web3 與 Web2 不同，在於 Web3 透過分散式帳本技術讓數據資料依存在開放式演算法與協議下，而不依賴機構或個人，形成所謂去中心化（Wood, 2014）。其中，用於產生 Web3 數據的應用程式稱為去中心化應用程式 DApp（Decentralized Application），其強調網路開放分散性。此類 DApp 允許任何使用者都可以參與，但不會任意使用個人數據。對使用者或內容創作者而言，Web3 的體驗與使用 Web2 差異不大，但在 Web3 下能對自己貢獻的內容保有所有權，還能獲得一定程度的回報。在隱私方面，使用者能清楚知道這些數據的用途，並且具有是否開放使用的決定權利。綜上所述，本文研究目的，主要在於探討 Web3 去中心化數據，包含 Web3 數據查詢，以及 Web3 數據分析與資料視覺化設計。

本文主要節次如下：第二節為文獻探討，主要介紹 Web3 數據與資料視覺化。第三節為引入 ELT 程序探討 Web3 數據分析及資料視覺化之研究方法；第四節為資料分析將以 OpeaSea 次級 NFT 交易市場資料與 STEPn 使用者數據，作為本文 Web3 數據研究之範例；最後，第五節將針對研究結果提出對於 Web3 數據分析與資料視覺化的相關建議，及未來研究方向。

貳、文獻探討

王怡萱（2022）設計不同互動類型的數位教材，探究不同互動模式數位教材輔助學習影響。同時，根據眼動資料分析與訪談結果推論分析所開發的

數位教材對於使用者是否適切。對於該研究之數據資料區塊鏈數據分析可以提供具體的描述與洞見，幫助研究者理解數位教材的使用情況及其對學習者的影響。區塊鏈數據分析相關的應用，可以展示用於：一、學習行為分析：使用區塊鏈技術追蹤學習者在數位教材中的行為模式，例如點擊、觀看時間、回答問題的準確性等。透過分析這些數據，可以了解不同互動模式對學習者的吸引力和效果。二、數位足跡追蹤：將學習者在數位教材上的活動記錄在區塊鏈上，確保數據的不可竄改性和安全性。這些數據可以用於評估教材的使用情況，並作為改進的基礎。三、自評與眼動資料分析：將學習者的自評結果和眼動資料納入區塊鏈，建立可信的數據庫。透過比較自評和眼動資料，可以了解學習者的主觀感受和客觀行為之間的關聯性，進而評估數位教材的適切性。四、分散式評估系統：建立基於區塊鏈的分散式評估系統，讓多個評估者對數位教材進行評估並記錄評估結果。這有助於確保評估的客觀性和透明性，並提供更可靠的評估結果。總體而言，區塊鏈數據分析可以提供更準確、可靠的數據基礎，幫助研究者更適切地理解所開發的產品對學習者的影響，並進一步優化教學設計和評估流程。此外，徐新逸、何品萱（2021）探討國小學童使用水資源 VR 遊戲，區塊鏈數據分析可以提供用戶行為追蹤、數據安全性、用戶反饋收集及推廣效果評估等方面的支援和應用，幫助提升遊戲的吸引力和教育效果。同時，結合區塊鏈數據分析，陳致澄（2022）探討科技融入數學教學之成效，與魏鈺瑩、崔夢萍（2021）提出數位說故事與水資源保育影響等研究，可以透過區塊鏈數據分析幫助提升教學使用方法的吸引力和參與者的學習效果，並確保數據的安全性和可信度。

關於區塊鏈數據分析，林億雄、馬瀾嘉（2021）針對以太坊區塊鏈與去中心化金融數據進行分析，該文之以太坊區塊鏈數據資料就為 Web3 數據之一。該研究提到 Web3 資料數據都是公開的，節點間彼此共享與儲存數據副本。同時，該研究也說明 Web3 數據的資料結構是由區塊頭（Blockheader, H），已驗證交易列表（Transactions, T），及其他已知前交易區塊頭（Other Block Headers, U）所組合而成。截至 2022 年 8 月，根據 blockchair. 網站資料（<https://blockchair.com/ethereum/charts>），以太坊區塊鏈上的資料總計約為 468 GB，平均每個區塊大小約為 0.08 MB，每個區塊處理時間約為 12 秒，每秒可處理的交易數約為 15 筆交易（TPS）。研究人員對於此類 Web3 數據的資料分析，需先從全節點中提取相關的數據，再將數據轉換為電腦軟體可讀的格式；後續，將所有數據輸出到資料庫中並建立索引以實現最佳查詢性能。Web3 數據也是一種不可竄改只能追加的分散式資料庫，對於欲研究 Web3 數據資料的研究人員來說，同步下載節點數據是一個非常耗時且困

難的工作。研究人員除需先建立節點下載所有的交易區塊，還需將整個鏈上資料副本下載。此外，研究人員為進行對 Web 3 數據資料分析，仍需了解 Web3 鏈上儲存的數據，及其資料結構。

一、Web3 數據

Web3 數據是指在 Web3.0 環境下生成、存儲和流通的數據。Web3.0 是一種分散式的 Web 技術架構，它透過使用區塊鏈和去中心化應用（DApps）等技術，使網路更加開放、安全和透明。在 Web3.0 中，數據不再由中心化的審查者控制，而是由分散的節點共同管理和驗證。關於 Web3 數據之相關文獻通常有（一）、數據來源和生成方式、（二）、數據結構和格式、（三）、數據存儲和查詢、（四）、數據分析和應用，及（五）、隱私和安全等方面（Szabo, 1997; Buterin, 2014; Swan, 2015）。Web3 數據來源可分為鏈上和鏈下數據。關於鏈上數據主要包括：鏈相關資料（如雜湊、時間戳記等）、轉帳交易、錢包位址、智慧合約執行事件，及保存在緩衝區中的數據。此類數據被以區塊鏈技術的去中心化資料庫所維護，並以區塊鏈的共識機制來保證其正確性。此外，鏈下數據包括中心化交易所的數據、社交媒體數據、GitHub 數據，及一些典型的 Web2 數據等。因為鏈下數據缺乏確保其正確性的方法，所以本研究中所指 Web3 數據狹義定義為具有透過區塊鏈共識機制來保證其正確性的鏈上數據。

二、資料視覺化

資料視覺化領域的起源可以追溯到 1950 年代初的電腦圖學。當時，人們利用電腦繪製出首批圖表。後續，隨著電腦運算能力的迅速提升，研究人員建立起規模越來越大，複雜程度越來越高的數值模型，此舉造就各式各樣的大數據資料集。例如：研究者利用醫學儀器和顯微鏡之類的資料收集裝置產生巨量的醫學資料集；或研究者透過可以儲存文字、數值和多媒體資訊的資料庫伺服器來收集資料。故此，研究人員需要進階的電腦圖學技術方法來處理與視覺化此類規模龐大的大數據資料集。

數據分析和資料視覺化領域因大數據的熱門近期受到了極大的關注，透過數據分析與資料視覺化可將現代科技所產生的大數據轉化為可操作的知識。故此，數據分析和資料視覺化可視為是數據科學的基石之一。然而，現今對於大數據資料集的處理部分幾乎很少有效的方法，當中最大問題在於將大數據轉變為可以分析的數據，還包含有效的數據存儲、實現資料視覺化的

查詢和索引、大數據的視覺化呈現方式等。這些問題點是資料視覺化系統能夠有效提供處理兆位元等級大數據集，同時能將視覺回饋限制在幾毫秒內的關鍵技術。圖 1 為本研究針對臺灣 2021 年確診新冠肺炎回報人數所設計的臺灣確診地圖儀板表 (<https://reurl.cc/O1p7j9>)，該儀板表提供時間、確診人數與區域之資料視覺化，並將資料呈現於網站。

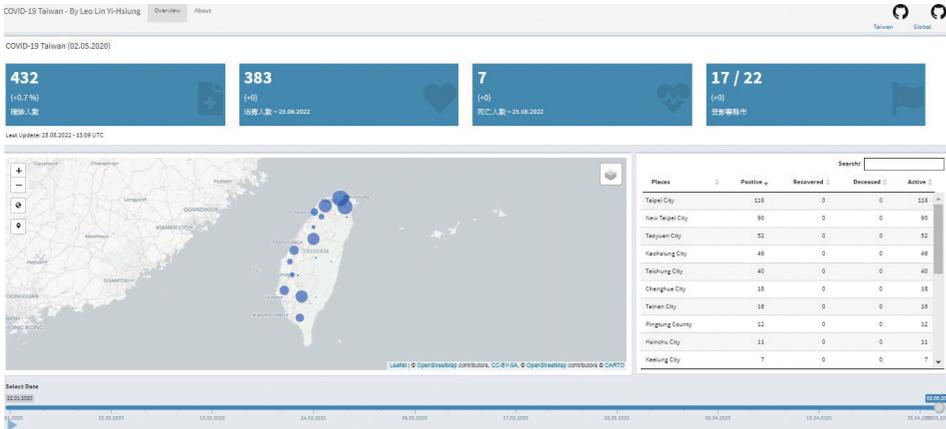


圖 1 2021 年臺灣新冠肺炎確診人數地圖之儀板表

在 Web3 的使用環境中可打破數據孤島，提升使用者的互動操作性，使用者只需要一個位址就可以瀏覽和使用各類 DApp，而這個位址所發生的每一筆鏈上互動數據，其對應的數據都可以被組合起來，無需任何許可。目前，研究 Web3 數據所使用的 DApp 大都為點對點 (Peer-to-Peer) 軟體，雖具有隨機存取功能應用程式介面 (Application Programming Interface, API) 提供類似檢查交易狀態、查詢錢包交易，及查詢錢包餘額等功能。但，API 端無法存取 Web3 鏈上所有的資料。所以，研究人員若想要分析 Web3 鏈上數據將需要建立節點、下載節點資料及建立索引查詢等技術性作業。為解決此一困難點，本文針對研究 Web3 數據提出使用大數據分析常用的 ELT 程序：提取、載入、轉換 (Extract、Load、Transform; ELT)。ELT 是一種數據整合過程，也是傳統 ETL (提取、轉換、載入; Extract、Transform、Load; ETL) 的一種變形。ELT 與傳統 ETL 之間的差異在於執行載入 (Load; L) 與轉換 (Transform; T) 的不同，最大的決定因素是數據轉換的執行方式、時間和地點 (Khine & Shun, 2017)。ELT 將原始數據從源頭服務器直接複製傳輸到目標服務器上的數據系統，然後使用者可以透過 SQL 語法查詢使用數據。以下將介紹 ELT 程序對數據執行三種不同的操作：

- (1) 提取數據：從一個或多個來源系統中識別和讀取數據的過程。
- (2) 載入數據：將提取的數據載入到目標資料庫的過程。
- (3) 轉換數據：將數據從原來格式轉換為分析所需格式的過程。

Dune Analytics 是分析 Web3 數據重要工具資源，而且 Dune Analytics 的設計可視為是 Web3 數據執行 ELT 程序最佳平台之一。Dune Analytics 目前已經解碼部分 Web3 數據，並提供 PostgreSQL 介面來查詢數據集。此外，Dune 提供簡單介面來查詢結果並可創建簡單的數據儀板表。對於 Web3 數據分析而言，Dune Analytics 無疑是重要執行 ELT 程序的極佳平台。關於 Dune Analytics 鏈上數據分析平台的數據倉庫架構如下圖 2 所示，由下至上可以細分為 (1) 數據索引層：鏈上數據分析平台從區塊鏈的節點獲取原始鏈上數據。(2) 數據處理層：鏈上數據分析平台將原始數據以串流式處理方式進行數據 ELT 程序。(3) 數據存儲層：將處理後的數據依預先定義格式存取儲存入各數據表中以待後續使用。(4) 數據分析層：存儲的數據會被聚合運算，運算完成後的結果將即時被輸出。關於區塊鏈數據分析研究，林億雄、馬瀾嘉 (2021) 針對以太坊區塊鏈與去中心化金融數據分析提出二步驟研究方法：

- (1) 審視 DeFi 項目方的合約位址及其智能合約開源程式；
- (2) 於 Dune Analytics 撰寫提取分析 DeFi 項目方數據之 PostgreSQL 語法。



圖 2 Dune Analytics 鏈上數據分析平台的數據倉庫架構

林億雄、馬瀾嘉（2021）使用此二步驟方法研究特定 DeFi 項目方，該文發現可以透過 Dune 平台分析結果，運用來結合 DeBank、Zapper，及 Etherscan 等去中心化應用程式 DApps，深入研究用戶活動行為及其資產的轉移。本文在此架構下擴大探討研究對象至 NFT 交易轉移，此優點在於不需審閱項目方智能合約之程式碼，僅需在 PostgreSQL 查詢語法中指定 NFT 交易平台，或生成鑄造 NFT 的位址。本文二步驟研究方法，在方法上仍屬於使用 ELT 程序。故此，本文將原二步驟研究方法修改為：

- (1) 審視 Web3 項目 NFT 發行官方網站或文件，查詢該項目有關的重大訊息；
- (2) 於 Dune 撰寫提取分析 Web3 項目 NFT 生成鑄造數據之 PostgreSQL 語法。

肆、資料分析

本節將探討如何使用 Dune Analytics 提取 Web3 數據資料，及如何分析 Web3 項目之用戶使用情形；並將探討二個範例項目 OpenSea 次級 NFT 交易市場數據，與 STEPN 使用者數據實作於數據分析與資料視覺化設計上。

中華資訊與科技教育學會

一、Dune Analytics 介紹

Dune Analytics（以下簡稱 Dune 平台）為本文研究 Web3 數據資料提取分析的工具。Dune 平台（<https://duneanalytics.com/>）是一個基於網站開發的平台，其使用標準 PostgreSQL 結構化語言查詢，並從預先整理好的數據庫中查詢 Web3 數據。Dune 平台可視為研究 Web3 數據資料分析強大的工具，其核心為將區塊鏈上原始數據轉換為可查詢的結構化資料庫。Dune 平台具有可編程性，研究人員可針對 Web3 數據資料進行提取分析和視覺化。Dune 平台也是一綜合型的 Web3 數據平臺，可以對巨量的鏈上數據進行查詢、分析，以及視覺化。Dune 平台將鏈上數據進行解析，然後引入 PostgreSQL 的關聯性數據庫中。目前，Dune 平台支援 Ethereum、Solana、Polygon、Optimism 及 BSC 等鏈上原始與解碼數據。

二、OpenSea 次級 NFT 交易市場數據分析

OpenSea 是一數位貨幣收藏品和 NFT 市場，由共同創辦人暨執行長 Devin Finzer 建立於 2017 年。OpenSea 目前是世界最大 NFT 交易平台，該

網站不只擁有數位藝術品，還有加密收藏品、遊戲物品、虛擬土地和域名等。相較其他 NFT 交易平台 OpenSea 顯得較為友善，並沒有邀請制，任何人都可以免費建立和購買 NFT。本節將使用 Dune 平台提取分析 OpenSea 次級 NFT 交易市場數據資料，結果展示如網站 (<https://dune.com/linleo/open-sea>)，如圖 3。由下圖 3 可知 OpenSea 上單日使用者位址數，其中 2022 年 9 月 4 日當日使用者位址數有 40,668 個。該日買家位址數有 20,310、賣家位址數有 26,101，當日賣家人數的位址數量多於買家。（本研究紀錄時間 2022 年 9 月 4 日）研究步驟如下：

- i. 審視 Web3 項目 NFT 發行官方網站或文件，及查詢與該項目有關的重大訊息。研究對象：OpenSea 之官方網站 (<https://opensea.io/>)
- ii. 於 Dune Analytics 撰寫提取分析 Web3 數據之 PostgreSQL 語法。使用 Dune 平台上以太坊區塊鏈 NFT 交易資料庫 nft.trades，其中 PostgreSQL 語法內關於 NFT 交易平台選擇需填入 OpenSea (project = 'OpenSea')。同時，因為 OpenSea 買家與賣家可能為同一交易錢包位址，所以在搜尋總交易者位址數時，需排除買賣方位址重複計數情形。本範例之相關查詢語法，可參考表 1 內容。下表 1 中 block_time 為指區塊建立時間，透過查詢當日所有的交易區塊的區塊時間，並用此區塊時間來搜尋計算當日總交易人數位址。

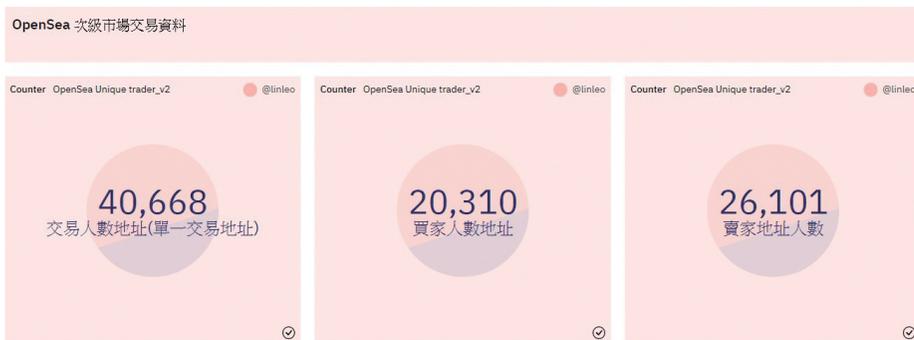


圖 3 OpenSea 上單日使用者人數、買家與賣家位址數

表 1 OpenSea 次級 NFT 交易市場當日總交易人數位址查詢語法

```

SELECT date_trunc('{{Time interval - Charts}}', block_time) AS time, NULL
      AS buyer, NULL AS seller, project, seller AS traders
FROM   nft.trades
WHERE  block_time > (date_trunc('{{Time interval - Charts}}', now()) - interval '1' day)
      AND block_time < (date_trunc('{{Time interval - Charts}}', now()))
      AND project = 'OpenSea'
      AND buyer != seller)

```

三、STEPN 使用者數據分析

STEPN 是一款融合區塊鏈技術的生活方式應用程式強調 Web3 生活型態，整合其社交經濟（SocialFi）和區塊鏈遊戲經濟（GameFi）元素。用戶透過在遊戲程式配備 NFT 運動鞋，然後在現實世界進行戶外散步、慢跑或跑步，藉此獲得加密貨幣賺取收入，達到遊戲 Move to Earn 的概念。STEPN 同時建於 Solana、BSC 及以太坊區塊鏈之上，這是一個促進碳中和，及邁向更健康生活方式的革命性概念。STEPN 由澳大利亞人 Jerry Huang 於 2021 年 8 月所創立，並由 Yawn Rong 進一步開發，該應用程式目前可在 Android 和 iOS 上使用。綜合上述，STEPN 成立宗旨在透過 GameFi 推動人們走向更健康的生活方式，並將用戶帶往 Web3 生活型態，同時依靠其 SocialFi 來構建一個持久的平台。因研究限制，本研究僅分析 Solana 鏈上 STEPN 使用者數據，其結果如網站（<https://dune.com/linleo/stepn-data-analysis>），如下圖 4。



圖 4 STEPN 每日、每週、新加入及連續使用者位址數

圖 4 為 STEPN 每日與每週活躍使用者位址，其中在 2022 年 9 月 4 日當日活躍使用者位址共有 3,708 個，當週活躍使用者位址計有 21,247 個。STEPN 當日新加入者位址共有 966 個，與連續使用超過 60 天位址可視為具長期連續穩定使用者有 2,706 個。（本研究紀錄時間 2022 年 9 月 4 日）

研究步驟如下：

(1) 審視 Web3 項目 NFT 發行官方網站或文件，及查詢與該項目有關的重大訊息。

研究對象：STEPN 之官方網站 (<https://stepn.com/>)

(2)於 Dune Analytics 撰寫提取分析 Web3 數據之 PostgreSQL 語法。

使用 Dune 平台 Solana 鏈上 NFT 交易資料庫 solana.transactions。其中，STEPN 官方鑄造鞋子的位址為”STEPNq2UGeGSzCyGVr2nMQAzf8xuejwqebd84wcksCK”。所以，在查詢中需要指定由此位址生成的 NFT 鞋子作為判斷是 STENP 項目使用者。相關查詢語法可參考表 2。表 2 為查詢 STEPN 當日與當週使用者位址數的語法，透過 account_keys 指定生成 NFT 的位址，及利用 now 函數與 block_time 的時間差來計數當日、或當週使用者位址數。

表 2 STEPN 當日與當週使用者位址數查詢語法

```

SELECT
  COUNT(DISTINCT signer) filter (where (now() - block_time) <= interval '24' hours)
  AS Active_Users_24H,
  COUNT(DISTINCT signer) filter (where (now() - block_time) <= interval '1' week')
  AS Active_Users_7D
FROM solana.transactions
WHERE array_contains(account_keys, "STEPNq2UGeGSzCyGVr2nMQAzf8xuejwqebd84wcksCK")
  AND block_date >= (CURRENT_DATE - '7 days'::INTERVAL)
  AND error is NULL

```



2022 年 09 月 15 日以太坊區塊鏈主網合併完成，區塊鏈共識機制由最初的 PoW 挖礦機制，轉成 PoS 權責驗證機制。從此開始，以太坊區塊鏈將不再靠礦工電腦和顯卡運作，而是由驗證者協助驗證區塊。以太坊區塊鏈聯合創始人 Buterin, V. (2022) 指出以太坊區塊鏈主網合併會讓以太坊網路消耗電力下降 99.988%，同時為全世界電力消耗降低 0.2%，這可能是歷史上最大的減碳努力之一。另一方面，由於驗證工作時間大幅縮短，鏈上交易手續費大幅下降，同時交易時間變更快，估計以太坊區塊鏈每秒可完成 10 萬筆交易 (TPS)。故此，以太坊區塊鏈的 PoS 權責驗證機制讓 Web3 數據更加容易產生與儲存，而這樣的 Web3 數據將更融入人們真實的生活情境中。所以，對於 Web3 數據分析與視覺化的研究顯得更加重要。目前對於 Web3 數據研究仍屬於初期，但相信隨著投入 Web3 研究時間增加，及投入研究 Web3 的機構與專家越來越多，研究人員越清楚研究項目的實質內容，及其重要事件等。研究人員透過整合 Dune 平台，與 PostgreSQL 語法可以查詢提取更加適當數據，以供研究人員進行高階統計分析使用，與設計出更直觀合宜資料視覺化的儀板表。

最後，本研究使用二個範例，分別為探討 OpenSea 次級 NFT 交易市場數據，與 Solana 鏈上 STEPn 使用者數據，期能吸引更多研究人員投入關於 Web3 數據分析與資料視覺化設計等開發。OpenSea 數據分析之 PostgreSQL 全文語法可由網站 (<https://dune.com/linleo/opensea>) 查閱，STEPn 使用者數據分析之 PostgreSQL 全文語法亦可由網站 (<https://dune.com/linleo/stepn-data-analysis>) 查閱。Web3 數據是分散式資料庫，其具有去中心化的特性，其指數據存儲在多個節點上，而不是集中在單一中心伺服器上。這種分散性和去中心化使 Web3 數據變得更加安全、不易遭受單一點的故障，並且具有防範篡改的能力。然而，這種特性也帶來了一些挑戰，例如數據的存取和分析變得更加複雜，需要新的方法和工具來應對。在 Web3 環境中，數據分析和資料視覺化變得至關重要，因為它們可以幫助解釋和理解這些分散式數據的複雜性。這些方法和設計包含去中心化數據分析，由於 Web3 數據存儲在多個節點上，傳統的中心化數據分析方法可能不再適用。研究人員需要探索新的去中心化數據分析方法，例如分散式機器學習，以實現對數據的協作分析而無需集中式數據存儲；此外，智能合約數據分析因 Web3 環境中的智能合約產生大量的交易和數據，需要適當的方法來分析和視覺化合約的執行和效能，用以幫助合約開發者更好地理解其合約的運行情況。同時，因為 Web3 數據的特性使得分析和視覺化變得更加複雜，研究人員需要透過開發適當的方法和直觀的資料視覺化設計，用以幫助使用者在這個新興的去中心化數據環境中發現更多商機，並獲得更多有用的知識。這不僅對個人用戶有益，也對企業和整個 Web3 生態系統都具有潛在價值。

參考文獻

- 林億雄、馬瀾嘉 (2021)。以太坊區塊鏈與去中心化金融數據分析之研究，*智慧科技與應用統計學報*，19 (1)，1-19。
- 王怡萱 (2022)。探究不同互動模式數位教材輔助學習影響，*教育科技與學習*，17，39-70。
- 徐新逸、何品萱 (2021)。國小學童使用水資源 VR 遊戲之研發與推廣，*教育科技與學習*，16，79-95。
- 陳致澄 (2022)。師培生設計科技融入數學教學活動表現之研究，*教育科技與學習*，17，1-38。
- 魏鈺螢、崔夢萍 (2021)。數位說故事融入專題導向學習對國小五年級學生水資源保育學習成效之影響，*教育科技與學習*，16，1-26。
- Buterin, V. (2014). A next-generation smart contract and decentralized application platform. *white paper*, 3(37), 2-1.
- Buterin, V. (2022). *Vitalik Buterin Says Ethereum Merge Cut Global Energy Usage by 0.2%, One of Biggest Decarbonization Events Ever*. Retrieved from <https://www.coindesk.com/business/2022/09/15/vitalik-buterin-says-ethereum-merge-cut-global-energy-usage-by-02-one-of-biggest-decarbonization-events-ever/>
- Khine, P., & Shun, W. (2017). Big Data for Organizations: A Review. *Journal of Computer and Communications*, 5, 40-48.
- Swan, M. (2015). *Blockchain: Blue print for a new economy*. O'Reilly Media, Inc.
- Szabo, N. (1997). *The idea of smart contracts*. Retrieved from https://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart_contracts_idea.html
- Wood, G. (2014). Ethereum: A secure decentralised generalised transaction ledger. *Ethereum project yellow paper*, 151, 1-32.

A Research on Web3 Data Analysis and Data Visualization

Yi-Hsiung Lin*

Assistant Researcher
Center for Teaching and Learning
National Kaohsiung University
Kaohsiung City, Taiwan
E-mail: yhslin@nuk.edu.tw

Mi-Chia Ma

Associate Professor
Department of Statistics Institute of Data Science
Center for Innovative FinTech Business Models
National Cheng Kung University
Tainan City, Taiwan
E-mail: mcma@ncku.edu.tw

Yu-Chi Li

Assistant Professor
Center for General Education
Chang Gung University of Science and Technology
Taoyuan City, Taiwan
E-mail: jomireli@mail.cgust.edu.tw



Abstract

Web3 data analysis and data visualization are important topics in the field of data science. However, there is currently a lack of effective systematic methods for handling Web3 data analysis. Therefore, this study introduces an ELT (Extract, Transform, Load) process to explore methods for Web3 data analysis and the design of Web3 data visualization. It is implemented in two example projects: OpenSea secondary NFT market data and STEP N user data. The goal is to leverage modern and innovative internet technologies and data analysis systems to transform the findings from Web3 data analysis into actionable knowledge.

Keywords: *Data analysis; Data visualization; Web3*

* Corresponding author



CACET
中華資訊與科技教育學會