



Open Data

磐石行動

2019/6/29 總統盃 黑客松

優化開放資料平台中最重要的資料集分類和詮釋資料架構
提高平台可尋性讓使用者能找到需要的資料



磐石行動簡報大綱

01	主題	07	雛形產出
02	成員及提案單位	08	資料使用情形
03	議題痛點	09	預期產出
04	提案目標	10	預期效益
05	解決方案	11	可行性評估與複製性
06	理論基礎		

幫政府開放資料設計新家



Open Data
磐石行動



Open data, Open Mind!

資料門戶大改造
隨心所欲最好找
磐石行動一定要

查詢功能 Evolve

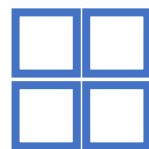
- 相關性排序
- 同義字
- 相關資料集

02

網站架構 Remodel

- 組織架構
- 分類架構
- 主題架構

01



03

圖表展示 Concise

- 資料預覽
- 網站統計資料

進階功能 Level Up

- 詮釋資料架構
- 資料應用展示
- 資料品質檢核
- 高價資料識別
- AI客服

04

成員及提案單位



Open Data
磐石行動



Open Data
磐石行動

關鍵字



中文 / En



江季洲
程式設計

工研院
巨量資訊科技中心



蘇志斌
程式設計

工研院
巨量資訊科技中心



趙家佑
團隊溝通、UI\UX

工研院
產業服務中心



林佳潔
插畫繪圖、UI\UX

驚嘆號星球人
插畫創作者

議題痛點 – 一般使用者

1. 分類架構不佳



一般使用者想找
流感接種的診所
名單



4. 分類架構不佳 導致統計無意義

服務分類	資料集總數	瀏覽次數
公共資訊	29,153	22,015,402
生活安全及品質	2,493	4,239,351
交通及通訊	1,718	3,140,658
投資理財	1,755	3,061,307

2. 查詢無相關性排序



3. 無資料預覽

資料資源欄位: 十碼章 醫院名稱 縣市別 行政區 郵遞區號 地址 連絡電話
檔案格式: CSV
編碼格式: UTF-8
資料量: 0
資料下載網址:
<http://opendataap2.e-land.gov.tw./resource/files/2019-01-03>
資料資源描述: 宜蘭縣公費流感抗病毒藥劑合約院所名冊
資料資源更新時間: 2019/06/26 00:00

宜蘭縣公費流感抗病毒藥劑合約院所名冊

資料集評分: ☆☆☆☆ No votes yet

資料集描述: 宜蘭縣公費流感抗病毒藥劑合約配置點一覽表

主要欄位說明: 十碼章 醫院名稱 縣市別 行政區 郵遞區號 地址 連絡電話 座標緯度 座標經度

資料下載網址:

- CSV
- 檢視資料** 宜蘭縣公費流感抗病毒藥劑合約院所名冊
- XML
- 檢視資料 宜蘭縣公費流感抗病毒藥劑合約院所名冊
- JSON
- 檢視資料 宜蘭縣公費流感抗病毒藥劑合約院所名冊

議題痛點 - 創業家



Open Data
磐石行動

1. 資料品質不佳

ISBN (020\$a) 978-986-01-6158-8	書名 (245 編著者 (C) 版本 (250 出版項 (C) 歐洲人權 司法院大法官書記臺北市; 孔孟荀 孟孟荀 蔡仁厚 著	
9.78957E+12 957130848X; 9789571308487	日本近代陳水逢著 初版 革命 :: 瑪查默斯 初版 淡江大學 淡江大學教育品質 ?2000	
9578158033 9573609134; 9789573609131	安撫哭聲 William S 初版 中庸 大身 楊洪注譯 初版 ?2004 中華臺北的競爭法與政策; Competitio	
9789868234031 (pbk.); 9868234034 (pbk) 9789577485595 : 9789862353424 (平裝) :	Decayed li Chung Wenyin ; trans Taipei Cit 為什麼是廣島?(18) 初版 ?2014 日本寫真大竹昭子 初版	
9.78986E+12 9789866006678 (平裝) :	保密局臺 許雪姬 主編 臺北市 :: 面對壓力 豐田圭一 初版 臺北市 :: 飢餓的軍 范軍 著 初版 臺北市 ::	
9.78957E+12 9789869350051; 9789869350082	胡適之先 胡頌平 編 初版; 增 ?2015 龍圖關 武 武 武 著 初版 臺北市 ::	

創業家想找創業
可用的開放資料



童書繪本電商創業

全國圖書書目資訊網(NBInet)

資料集評分: ★★★★★
平均 3.3 (4 人次投票)

資料集描述: NBInet「圖書聯合目錄」收錄的資料包括國家圖書館及95所各類型合作單位陸續提供之書目及館藏資料、國際標準書號中心新書書目、漢學研究中心藏書目錄。

主要欄位說明: 標準書號(ISBN, ISSN等)、書名、編著者、版本、出版項、面頁數尺寸、集叢名、出版年、NBI紀錄號

資料下載網址:

- CSV 檢視資料 NBInet圖書聯合目錄(103年第1季).....
- CSV 檢視資料 NBInet圖書聯合目錄(103年第2季).....
- CSV 檢視資料 NBInet圖書聯合目錄(103年第3季).....
- CSV 檢視資料 NBInet圖書聯合目錄(103年第4季).....

2. 沒有所需資料

圖書館書籍借閱量

資料集評分: ☆☆☆☆
No votes yet

資料集描述: 新竹市公共圖書館每月書籍借閱量

主要欄位說明: 統計年月、館藏地、借閱人次、借閱冊數、續借人次、續借冊數、還書人次、數、取消預約取消預約人次、冊數、借閱續借人次、借閱續借冊數

統計年月	館藏地	借閱人次	借閱冊數	續借人次	續借冊數
108.03	新竹市文化局圖書館視聽室	1758	5232	505	1298
108.04	新竹市文化局圖書館一般閱覽室	10299	27694	2281	4646
108.04	新竹市文化局圖書館香山分館	5901	15199	1012	1777
108.04	新竹市文化局圖書館金山分館	5425	14067	1072	1958
108.04	新竹市文化局圖書館鹽水分館	4009	10083	795	1453
108.04	新竹市文化局圖書館南寮分館	2541	5682	446	704
108.04	新竹市文化局圖書館兒童閱覽室	5934	21659	777	1852
108.04	新竹市文化局圖書館視聽室	1557	4589	488	1296
108.05	新竹市文化局圖書館一般閱覽室	10503	28680	2374	5257
108.05	新竹市文化局圖書館香山分館	6378	16986	1009	1810
108.05	新竹市文化局圖書館金山分館	5771	15737	1163	2113
108.05	新竹市文化局圖書館鹽水分館	4079	10273	804	1438
108.05	新竹市文化局圖書館南寮分館	2760	6259	518	817
108.05	新竹市文化局圖書館兒童閱覽室	6351	24241	843	1936
108.05	新竹市文化局圖書館視聽室	1644	4878	466	1188

回應

D****/
2018/04/23
10:06

Re: ISBN欄位格式錯誤問題

敬啟者:

謝謝您的快速回覆，
經檢視後，發現無論以何種檔案格式開啟仍有以下問題：

議題痛點 – 開發者



Open Data
磐石行動

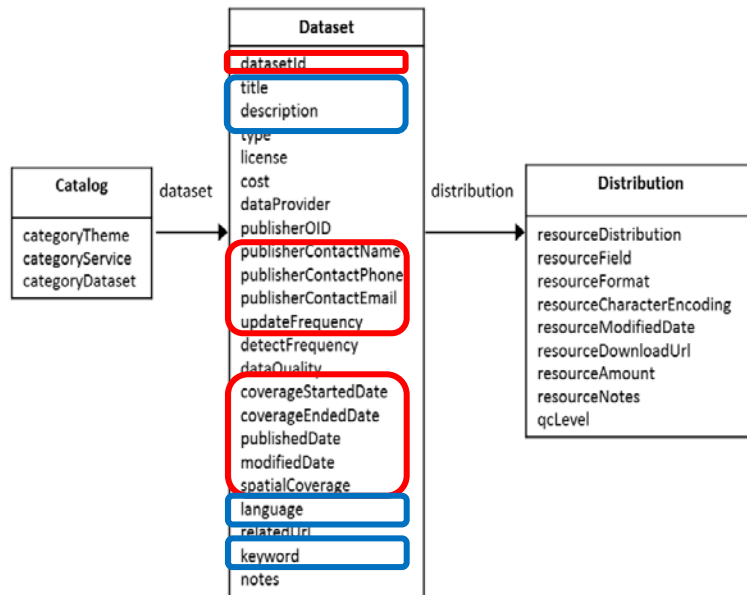
1. 詮釋資料品質不佳

中央研究院學術活動

資料集評分:	☆☆☆☆☆ No votes yet
資料集描述:	中央研究院學術活動
主要欄位說明:	id · subCatID · title · titleEn · pubDate · linkOut · linkUrl · linkUrlEn
資料下載網址:	JSON 檢視資料 中央研究院學術活動
提供機關:	資訊服務處
提供機關聯絡人姓名:	陳先生 (27899281)
更新頻率:	即時
授權方式:	政府資料開放授權條款-第1版
計費方式:	免費
上架日期:	2018/10/31
資料集類型:	原始資料
詮釋資料更新時間:	2018/10/31 00:00
關鍵字:	中央研究院



2. 詮釋資料與W3C不相容



文資局接受及支付補助金

資料集評分:	☆☆☆☆☆ No votes yet
資料集描述:	文資局接受及支付補助金相關資訊
主要欄位說明:	id(編號) · SiteId(站臺編號) · UnitId(單元編號) · TypeId(類別編號) · FileName(附件地址) · Url(連結) · BaseCategoryId(類別名稱) · Content(內容) · PublishTime(發佈日期) · PageUrl(連結回本站地址) · Mutilimage(多圖) · Mutilattach(多附件)
資料下載網址:	JSON 檢視資料 文資局接受及支付補助金
提供機關:	文化部文化資產局
提供機關聯絡人姓名:	葉小姐 (04-22177596)

- 名稱、定義與W3C一樣的只有4個
- 有許多欄位，定義相同，名稱卻不同
 - datasetId vs identifier
 - modifiedDate vs modified
 - spatialCoverage vs spatial

改造後享受開放資料的好處



Open Data
磐石行動

透過開放資料平台的優化及改造
讓資料經濟在台灣真的成形，讓每個人享受開放資料
帶來的好處



Open Data Value Chain

開放資料平台及
詮釋資料位置



議題痛點對應服務規劃列表

編號	議題痛點	服務規劃
1	分類架構不佳	重新規劃網站的組織、分類、主題架構
2	查詢無相關性排序	相關性排序、同義字、相關資料集
3	無資料預覽	XLS、CSV、XML、JSON、KML...資料預覽
4	分類架構不佳導致統計無意義	採用新的分類架構重新統計
5	資料品質不佳	資料品質檢核
6	沒有所需資料	資料申請小幫手*
7	詮釋資料品質不佳	詮釋資料品質檢核
8	詮釋資料與W3C不相容	導入W3C DCAT 資料模型

& MORE

資料應用展示、高價資料識別、AI客服



資訊架構(Information Architecture)設計 – 滿足使用者的資訊需求

三種資訊需求行為

1. 搜尋(search)
2. 瀏覽(browse)
3. 詢問(ask)



兩種資訊需求行為模式

1. 採莓模式(berry-picking)
2. 滾雪球模式(pearl-growing)



五大資訊架構體系

1. 組織(organization)
2. 導覽(navigation)
3. 搜尋(search)
4. 命名(labeling)
5. 詮釋資料(metadata)



四種資訊需求

1. 最佳解答(已知項目尋找)
2. 一些就好(探索式尋找)
3. 一網打盡(地毯式尋找)
4. 再次尋找



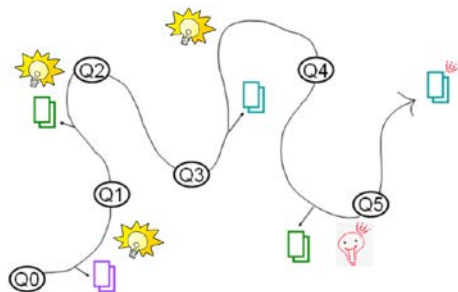
兩種資訊需求行為面向

1. 整合(integrate)
2. 反覆(iterate)



六種數位空間階層

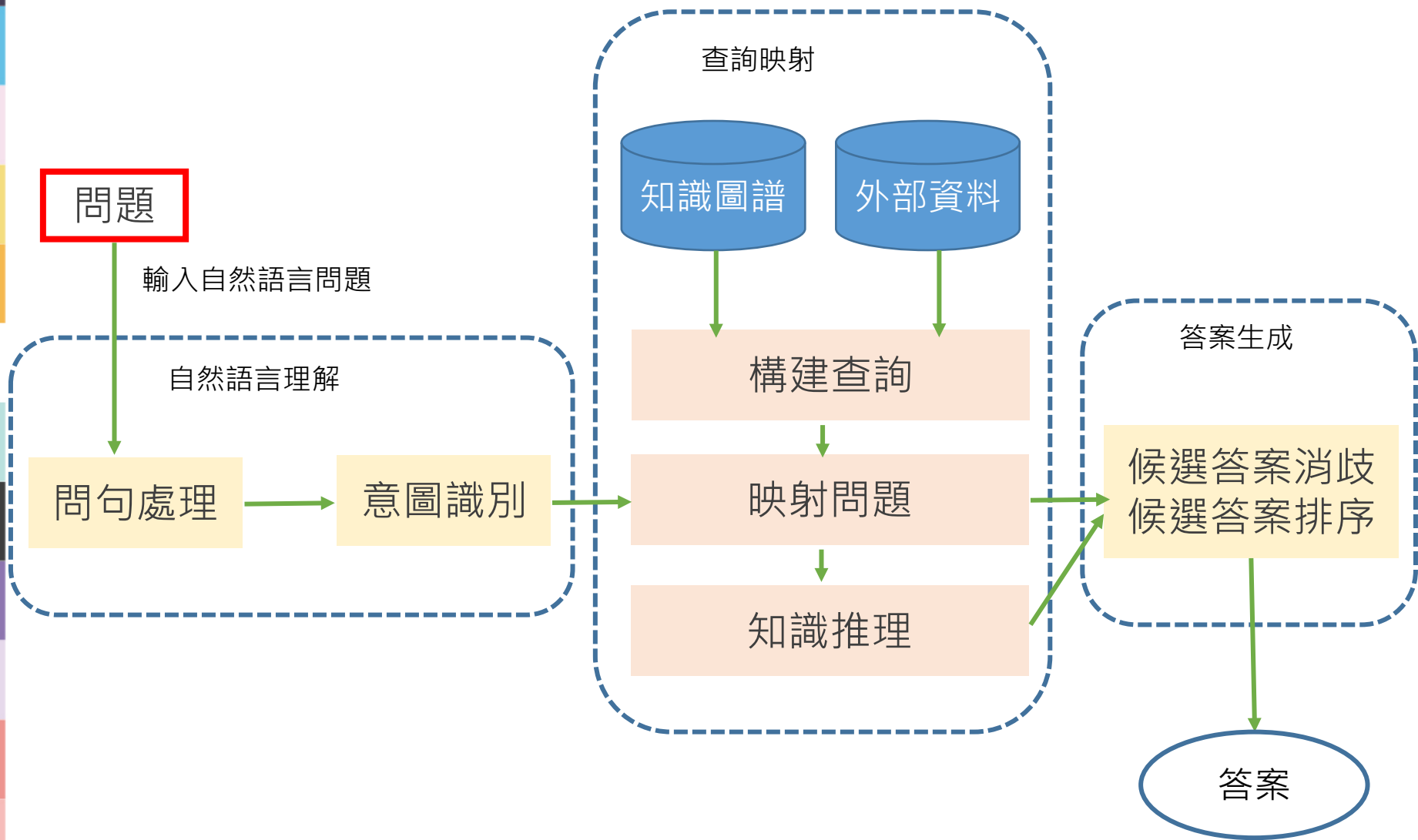
1. 場址(site)
2. 結構(structure)
3. 外表(skin)
4. 服務(service)
5. 空間格局(space plan)
6. 物品(stuff)



理論基礎 – AI 客服機器人



Open Data
磐石行動



沒有IA，就沒有AI



Open Data
磐石行動

Information Architecture enables Artificial Intelligence



source: <https://thewest.com.au/sport/surfing/gilmore-wins-seventh-world-surfing-title-ng-b881033914z>

“AI requires machine learning, machine learning requires analytics, and analytics requires the right data and information architecture (IA). In other words, **there is no AI without IA.**”

— Rob Thomas, general manager at IBM Analytics.

AI就好像沖浪選手般是世界注目的焦點；資料和IA就好像是海浪，沒有資料和IA的配合，沖浪選手只有滅頂的命運！

雛型產出(1/4)



Open Data
磐石行動

Before



資料集服務分類



After



資料集 | 最新消息 | 應用 | 互動專區 | 諮詢小組 | 關於平台



資料集



小臉的嘴型喊著
“喔” = Open Data的
“O”，像是在輕鬆愉
快的唱歌，希望這個平
台大家使用上是很簡易
輕鬆的。

雛型產出(3/4)



Open Data
磐石行動

Before

The screenshot shows the '政府資料開放平臺' (DATA.GOV.TW) interface. It features a search bar at the top with the text '輸入您想要搜尋的關鍵字(資料集)'. Below the search bar are several filter buttons: '資料集 (40621)', '最新消息 (8663)', '活化應用 (78)', '我有話要說 (6293)', and '我想要更多 (1297)'. On the left side, there is a navigation menu for '中央機關' (Central Government) and '地方機關' (Local Government), listing various agencies like '國家安全會議', '司法院', '監察院', etc., with their respective dataset counts. The main content area displays a list of datasets, with the first entry being '108年度行政院農業委員會花蓮區農業改良場會計報告'.

After(尚未套版)

The screenshot shows the new CKAN portal interface. The top navigation bar includes 'ckan', 'Datasets', 'Organizations', 'Groups', and 'About', along with a search bar. The main heading is '/ Organizations'. Below this, there is a section titled 'What are Organizations?' explaining that CKAN Organizations are used to create, manage and publish collections of datasets. A prominent 'Add Organization' button is visible. A search bar for organizations is present, followed by the text '36 organizations found' and an 'Order by: Name Descending' dropdown menu. The main content area displays a grid of organization cards, each with a logo, name, and dataset count. The organizations shown include '飛航安全調查委員會', '金融監督管理委員會', '農業委員會', '衛生福利部', '花蓮縣政府', '臺中市政府', and '經濟部'.

雛型產出(4/4)



Open Data
磐石行動

1. 預覽 CSV、XLS

Organizations / 國家發展委員會 / 5-4 國家重要濕地 / 國家重要濕地(CSV)

國家重要濕地(CSV) Manage

URL: <https://ws.ndc.gov.tw/Download.sash?u=LzAwMS9RZG1pbnZhdH9hGyLzEwL3JhZGZpbGUwT01MS2MkZl.2024>

From the dataset abstract

1. 都市及區域發展計畫(彙編)中都市及區域發展計畫之都市及區域發展計畫統計資料；2. 本項資料來源：行政院農委會

Source: 5-4 國家重要濕地

Data Explorer

年	重要濕地名稱	位置	等級	面積	備註
2017	1-綠之溪-山頂池	臺中市	國家級	3001	1.96 年評選 2,104 年...
2017	2-口車潭池	臺南市	國家級	557	1.96 年評選 2,104 年...
2017	3-綠之溪濕地	臺北市	國家級	1	1.96 年評選 2,104 年...
2017	4-1臺北港北堤濕地	新北市	國家級	357	1.100 年評選 2,104 年...
2017	4-2松子尾濕地	新北市	國家級	66	1.96 年評選 2,104 年...
2017	4-3沙洲河-江樹林濕地	新北市	國家級	109	1.96 年評選 2,104 年...
2017	4-4寶島濕地	臺北市	國家級	379	1.96 年評選 2,104 年...
2017	4-5芝蔴濕地	新北市	國家級	175	1.96 年評選 2,100 年...
2017	4-6七溝新舊濕地	臺北市	國家級	559	1.96 年評選 2,104 年...
2017	4-7新港人工濕地	新北市	國家級	31	1.96 年評選 2,100 年...
2017	4-8海山人工濕地	新北市	國家級	42	1.100 年評選 2,104 年...

2. 預覽 XML、JSON

Organizations / 經濟部 / 新創圓夢網-找資金 / 新創圓夢網-找資金(JSON)

新創圓夢網-找資金(JSON) Manage Go to resource

URL: https://sme.moeasmea.gov.tw/startup/upload/opensource/gov_funding_opendata.json

From the dataset abstract

新創圓夢網-找資金

Source: 新創圓夢網-找資金

Text

```
{
  "資金名稱": "臺北市產業發展獎勵補助",
  "資金類型": "研發補助",
  "資金領域": "已創業, 創新研發投資",
  "創業階段": "已創業",
  "身分別": "在學學生, 從業青年, 原住民族青年, 青年青年, 選拔位官兵, 其他弱勢青年",
  "申請方式": "申請簡章或向其他機關身, 應參閱下列文件並向各級就業服務中心: chr /1/, 申請書 chr /2/, 投資計畫書 chr /3/, 公司或商標",
  "對象": "1. 於臺北市完成登記之公司或商標。 chr /2/. 創業計畫書其他政府單位補助。",
  "資金額度": "ep/創業計畫書每案補助最高400萬元, 研發計畫書每案補助最高500萬元, 品牌建立計畫書每案補助最高500萬元, 創新育成計畫書每案補助",
  "諮詢窗口": "本計畫採隨時受理, 請洽政府服務諮詢窗口「臺北市產業發展獎勵補助專案辦公室」(地址: 臺北市信義區信和路1號北區1樓工商服務中心)",
  "網址": "https://www.industry.gov.tw/live/taipei/index.aspx",
  "建立時間": "2015-02-28 09:08:08",
}
```

3. 預覽 KML、GeoJson

Organizations / 臺中市政府 / 台中市自行車道位置 / 自行車道位置

自行車道位置 Manage Download

URL: <http://127.0.0.1:8080/dataset/d275f71f-440b-4a18-9433-c8143729321b/resource/ba5c5e0c3ce1-438c-a21c-250760ef694/download/1...>

From the dataset abstract

台中市自行車道位置

Source: 台中市自行車道位置

Map viewer

Before

資料資源欄位: 十碼章 醫院名稱 縣市別 行政區 郵遞區號 地址 連絡電話 座標緯度 座標經度 備註 建立日期 更新日期

檔案格式: CSV

編碼格式: UTF-8

資料量: 0

資料下載網址:
<http://opendataa2.e-land.gov.tw//resource/files/2019-01-03/b3cdd447cc5758ca888ef9986d137d5.csv>

資料資源描述: 宜蘭縣公費流感抗病毒藥劑合約院所名冊

資料資源更新時間: 2019/06/26 00:00

宜蘭縣公費流感抗病毒藥劑合約院所名冊

資料集評分: ☆☆☆☆ No votes yet

資料集描述: 宜蘭縣公費流感抗病毒藥劑合約院所名冊

主要欄位說明: 十碼章 醫院名稱 縣市別 行政區 郵遞區號 地址 連絡電話 座標緯度 座標經度

資料下載網址:

- CSV 精選資料 宜蘭縣公費流感抗病毒藥劑合約院所名冊
- XML 檢視資料 宜蘭縣公費流感抗病毒藥劑合約院所名冊
- JSON 檢視資料 宜蘭縣公費流感抗病毒藥劑合約院所名冊

無資料預覽

資料使用情形



Open Data
磐石行動

資料來源

政府資料開放平台

40382筆資料集註釋資料



- Wikipedia
- 部會組織、地方政府開放資料平台



資料產出

- 組織架構
- 分類架構
- 主題架構
- 資料集
- 資料來源

編號	分類中文	分類英文
1	法律	Law
2	公共服務	Public Service
3	環境	Environment
4	文化&教育	Culture & Education
5	交通	Transport
6	財政	Finance
7	社會&安全	Society & Safety
8	地方政府	Local Government
9	科技	Science & Technology
10	經濟	Economy
11	農業	Agriculture
12	健康	Health

預期產出



Open Data
磐石行動



預期效益



Open Data
磐石行動

運作良好的開放資料平台， 就是資料經濟的溫床



一般使用者可以容易找到所需的資料

創業家可以利用開放資料經營企業

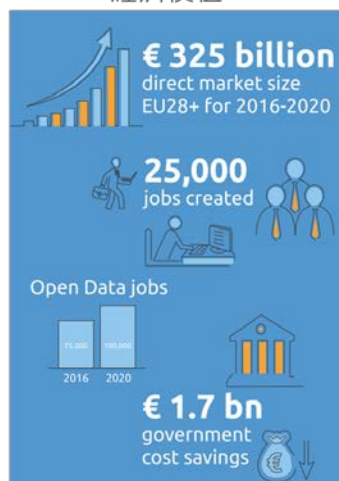
開發者可以快速開發應用，串連不同
資料平台

平台營運者透過資料使用情形分析獲得Insight



歐盟報告指出開放資料可創造出價值估計

經濟價值



社會福利價值



可行性評估與複製性(1/2)



Open Data
磐石行動

政府部門
可行性評估



民間單位
複製性評估

單位	合作項目	合作內容
國發會	政府資料開放平台	1.雛形網站供改版參考 2.建議報告供改版參考 3.AI客服可洽談合作
科技部	政府開放資料彙整平台	團隊整合，共同開發

單位	開放資料平台項目
中研院	研究資料寄存所
宏華環境保護與數位未來基金會	大肚台地計畫資料開放平台
g0v	Vote Taiwan投票指南

顯示民間資料中心需求存在
全球使用CKAN資料中心達200個以上

可行性評估與複製性(2/2)



Open Data
磐石行動

科技部政府開放資料彙整平台 Integrated Platform for Government Open Data (IPGOD)

1. 說明

政府開放資料彙整平台提供一個不分資料屬性類別的雲端政府資源資料倉儲，提供使用者更快速、更精確的方式找到資料以及更方便的方式取得資料，用更無痛、更直覺的方式操作與學習使用資料，讓政府提供的開放資料更有感，更快速與方便被應用。



2. 動機

行政院國發會所建置之《政府資料開放平台》擁有大量政府開放資料，但依然有可加強的部分：

- 如僅提供連結、無資料預覽、不提供分析主機...

更加強化 data.gov.tw 網站不足之處：

問題	解法
網站僅提供連結，不真正存放資料	統一存放於彙整平台中
各地資料主機服務不穩定	平台服務主機將放置於國家高速網路與計算中心機房
無過往資料	資料抓取後依時間存放
編碼格式不統一	標準化、轉碼
檔案格式眾多	分類存放、資料過濾
無資料預覽	提供資料預覽與分析試算
不提供分析主機	快速將資料匯入國網中心大資料平台



- 直接引用組織、分類架構沒調整就會出現有1萬多個組織，每個組織只有少數資料集的問題
- 雙方可合作互補，共同開發



Open Data

磐石行動

Thanks for your attention.