

Development of A Research Data Repository

研究資料寄存所的發展

2018 Road Ecology and Data Analysis Conference
Taipei, Taiwan

2018-06-20

Tyng-Ruey Chuang, Cheng-Jen Lee, Huang-Sin Syu, Wei-Ching Andrea Huang
莊庭瑞、李承鑫、許煌鑫、黃韋菁

Institute of Information Science, Academia Sinica
中央研究院 資訊科學研究所



Outline

- Collaborative Research
- Research Data Management
- data.depositar.io
 - Retooling an open data repository for a research data repository
 - Show Case: Place Names & Historical Maps
 - Show Case: Roadkill Observations
- Openness and FAIR Data

Collaborative Research

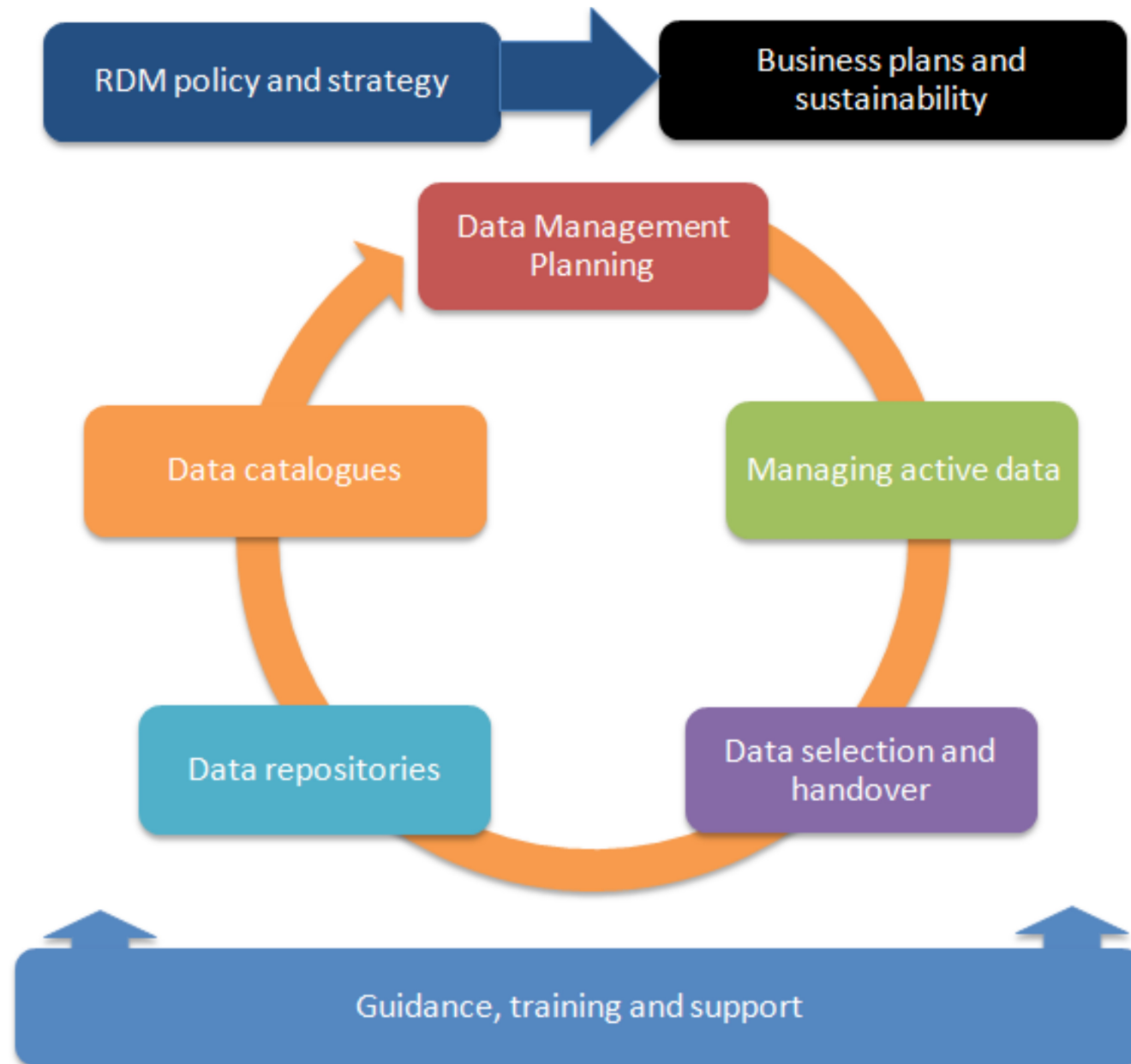
- Collaboration is the process of two or more people or organizations working together to realize or achieve something successfully.
 - *Wikipedia*
- For collaborative research, we set to make
 - the research project and
 - the research dataopen to project members and even to others.

Openness Benefits Research

- Help disseminate research findings.
- Help reproduce and re-purpose research results.
- Help encourage research collaborations.

Research Data Management (RDM)

<http://www.dcc.ac.uk/resources/how-guides/how-develop-rdm-services>



A Research Data Repository

- data.depositar.io
- Originally built for interdisciplinary regional study
 - A collaboration between Academia Sinica (Institute of Information Science and Center for GIS) and National Cheng Kung University
- Built upon CKAN
 - CKAN: software from OKI for “open data”
 - Adapt and extend CKAN for [research data](#) deposit, publishing, and reuse
 - Hosted at Academia Sinica

Search and Data Discovery

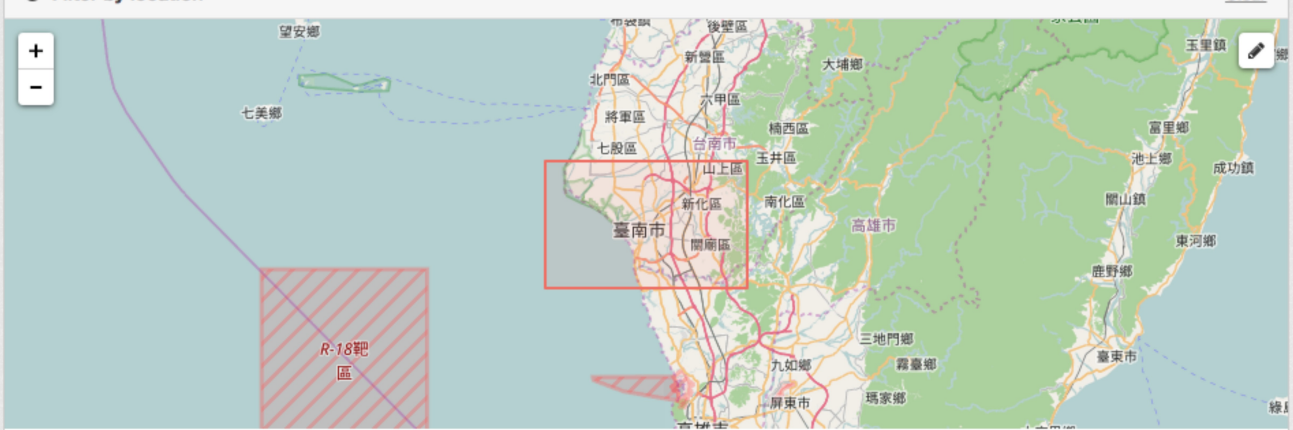
- free-text search
- multiple filters
- spatial and temporal constraints

跨領域區域研究資料集

Datasets Organizations Groups Help Search

Home / Datasets ?

Filter by location Clear



Map tiles & Data by OpenStreetMap, under CC BY-SA

Cancel

Search datasets...

51 datasets found Order by: Relevance

Temporal Search Clear

Or use time period shortcut

Organizations

昭和年間灌溉地給水區分圖：麻豆
1929年麻豆灌溉給水區分圖
TIF

昭和年間灌溉地給水區分圖：西港
1929年西港灌溉地給水區分圖

Dataset Visualization

臺南市里界圖

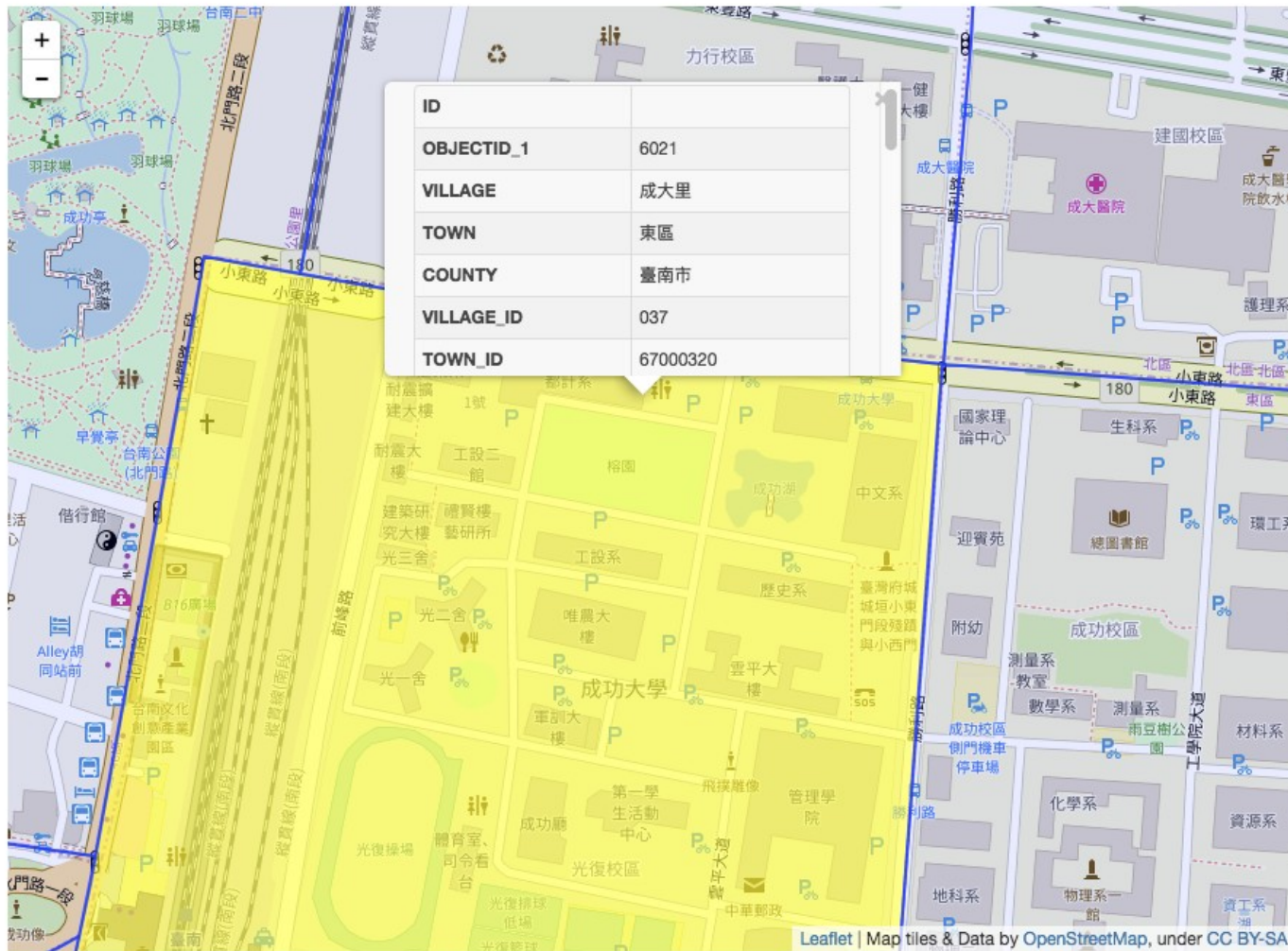
[Manage](#)[Go to resource](#)

URL: <https://data.depositar.io/dataset/76a232e5-a9c1-4dde-8d4c-6b58629f33ca/resource/c2bbe543-dc9d-41c2-82d5-d16a93b87835/do...>

屬性包含：「VILLAGE(里名)」、「TOWN(區名)」、「COUNTY(縣市名)」

📍 SHP

</> Embed



Metadata

- designed for cross-disciplinary research
- emphasizing spatial and temporal metadata authoring

Language:

Temporal Information (Time Period of Dataset)

Time Period Shortcut:

i This shortcut provides some historical periods for filling temporal information of the dataset.

Temporal Resolution:

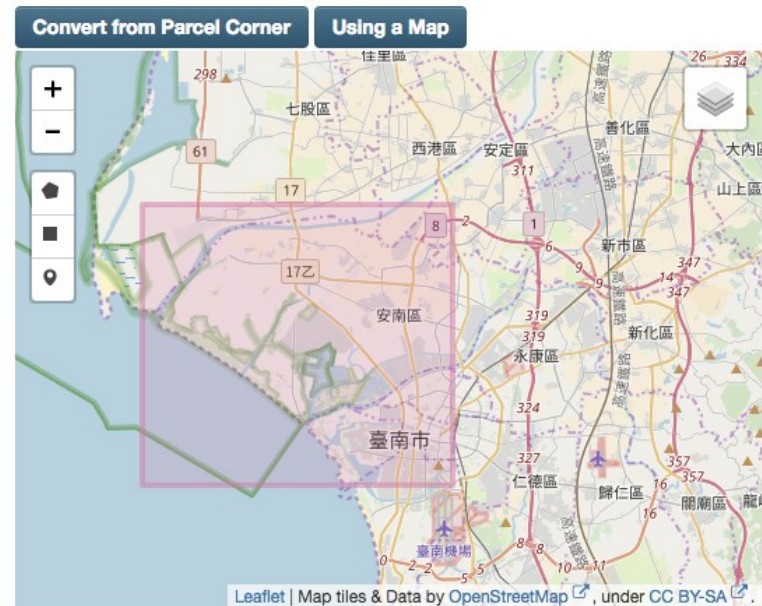
Start Time:

i Depend on the temporal resolution setting.
Acceptable formats: 'YYYY', 'YYYY-MM', or 'YYYY-MM-DD'.

End Time:

i Depend on the temporal resolution setting.
Acceptable formats: 'YYYY', 'YYYY-MM', or 'YYYY-MM-DD'.

Spatial Fields



Spatial:

```
{ "type": "Polygon", "coordinates": [[ [120.05825042724611, 22.969773432534776], [120.05825042724611, 23.09678621760499], [120.21068572998048, 23.09678621760499], [120.21068572998048, 22.969773432534776], [120.05825042724611, 22.969773432534776]] ] }
```

This column is used for geo-indexing the dataset. Click for details.

Multilingual metadata

▼ Data Type
Books (455)
Vector (144)
Statistics (103)
Pictures (Spatial) (34)
Pictures (Non spatial) (25)
▼ Language
Chinese (730)
Dutch (22)
English (8)
Spanish (Castilian) (1)
▼ Historical Material
Local chronicles in Qing dynasty (36)
Maps (20)
Japan officials (19)
Newspapers (16)
Inscriptions (9)
Taiwanese Governor Office Files (9)
Dutch Formosa (5)
Officials in postwar period (1)

▼ 資料類型
文獻書籍 (455)
向量資料 (144)
統計資料 (103)
圖像 (空間類) (34)
圖像 (非空間類) (25)
▼ 語言
華語/中文 (730)
荷蘭語 (22)
英語 (8)
西班牙語 (1)
▼ 使用史料
清代方志 (36)
地圖 (20)
日治出版品與官方資料 (19)
報紙 (16)
碑文 (9)
台灣總督府檔案 (公文類纂、府報、調查報告、職員錄、衛生法規) (9)
荷治時期檔案 (5)
戰後官方資料與鄉鎮志 (1)

Generalized and Multilingual Metadata

- One set of simple metadata for all datasets; three categories:
 - **Basic metadata:** title, description, data type, ...
 - **Descriptive metadata:** language, temporal, and spatial information, keywords, ...
 - **Management metadata:** license, author, created time, organization, maintainer, ...

Wikidata-powered Keywords

- Keywords
 - controlled vocabularies for tagging datasets
- Define your own list(s) of keywords
 - a never-ending process ...
- Use [Wikidata](#) as the source for keywords
 - 37M+ entries
 - multilingual
 - semantic relations enable data inference
 - e.g. Tainan is part of Taiwan



Wikidata Entries for Keywords

Search and select keywords:

Keywords: tainan

nagement

ID	Label	Description
Q140631	Tainan	city in southern Taiwan
Q477601	Tainan County	county of Taiwan
Q702035	Tainan Airport	airport

* License:

Keywords (as Wikidata IDs) are stored.

Viewed in English:

Keywords

Tainan

Viewed in Chinese:

關鍵字

臺南市

Place Names & Maps – A Showcase

<https://data.depositar.io/dataset/place-names-in-west-central-district-of-tainan>

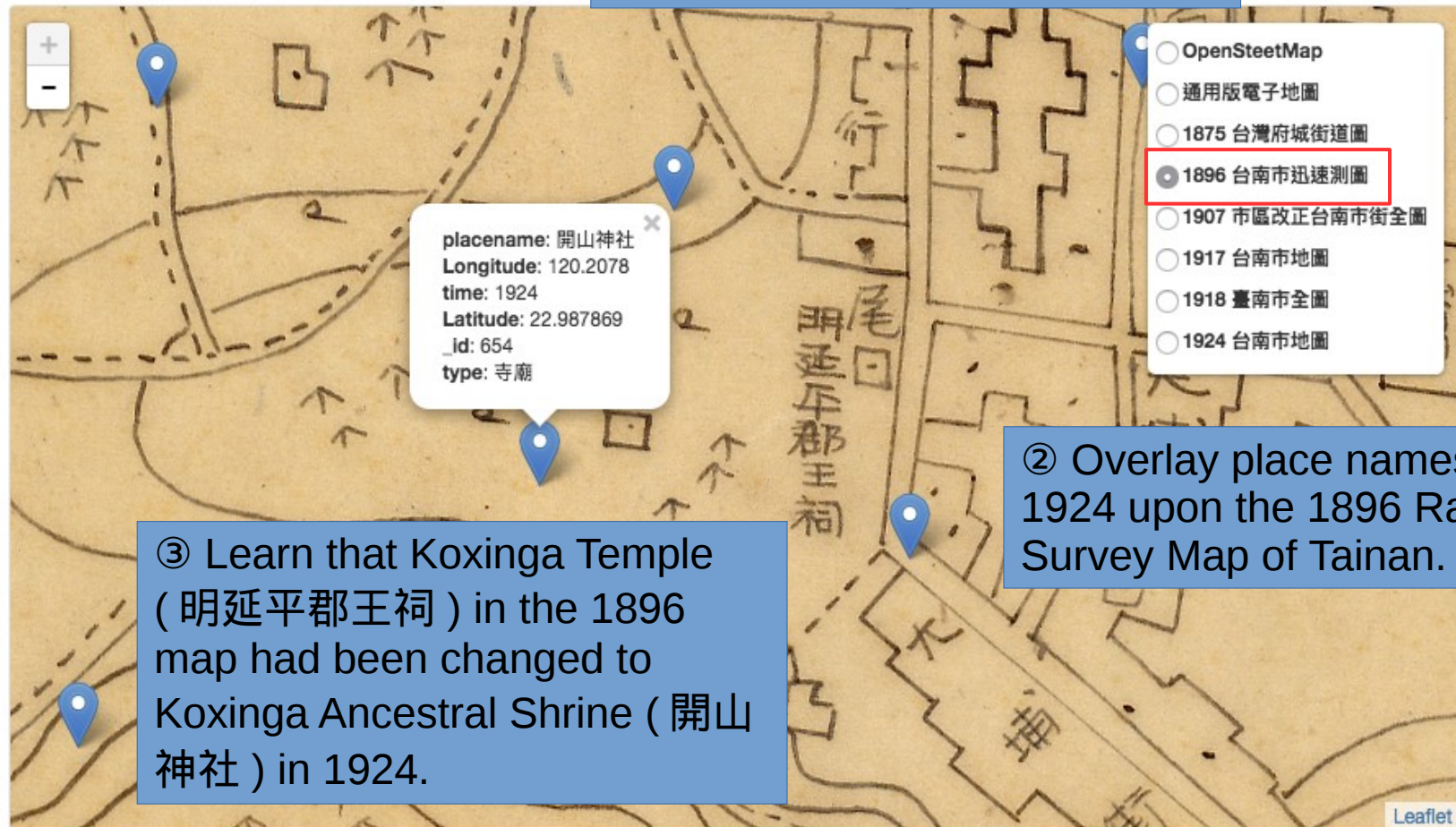
Source: Place Names in West Central District of Tainan

📍 All 📍 1875 📍 1896 📍 1907 📍 1917 📍 1924 🗄 資料瀏覽器

① Place names extracted from a map of Tainan in 1924 (displayed as blue place marks).

</> Embed

Add Filter



Roadkill Observations – A Showcase

<https://data.depositar.io/dataset/2016>

跨領域區域研究資料集

Datasets Organizations Groups Help

Search

Log In

Datasets / 路殺社範例資料集-2016

路殺社範例資料集-2016

6

Followers

0

Social

Google+

Twitter

Facebook

License

License Not Specified

Dataset extent

Map files & Data by OpenStreetMap, under CC BY-SA.

Other Access

The information on this page (the dataset metadata) is also available in these formats:

JSON

via the CKAN API

Dataset Groups Activity Stream History

路殺社範例資料集-2016

透過群眾所提野生動物遭受車輛撞擊或輾壓致死的現象，並以結構化方式紀錄相關資訊

Data and Resources

路殺社範例資料集

Explore

Tags

roadkill

Basic Information

Data Type Statistics

Language Chinese

Temporal Information (Time Period of Dataset)

Start Time 2016

Spatial Fields

Spatial show more

X.min 118.35888862609865

X.max 122.68355369567873

Y.min 21.45913969982141

Y.max 26.44352480769591

Management Information

Author 路殺社

About 跨領域區域研究資料集

Contact us: data.contact AT depositar.io

CKAN API | Site Statistics | Site Status | Documentation

Powered by ckan

Source code available on GitHub

Language: English

跨領域區域研究資料集

Datasets Organizations Groups Help

Search

Log In

Datasets / 路殺社範例資料集-2016 / 路殺社範例資料集

路殺社範例資料集

Go to resource

Data API

URL: <https://data.depositar.io/dataset/b78e1f21-1f5d-4fb6-95b5-6193cb44797/resource/0158ed01-1989-4a64-b67c-e7962c353d57/download>

From the dataset abstract

透過群眾所提野生動物遭受車輛撞擊或輾壓致死的現象，並以結構化方式紀錄相關資訊

Source: 路殺社範例資料集-2016

map

Embed

Add Filter

Grid Graph Map 3218 records 1 - 20 Search data ... Go Filters

_id	id	科名	科名 (...)	物種名	物種名 (...)	紀錄時間	地點	經度	緯度
1	24933	Bufo	Bufo	Rana	Rana	2016-01...	宜蘭縣員...	121.636...	24.754443
2	24934	Musca	Musca	Turdus	Turdus	2016-01...	基隆市中...	122.08166	25.62861
3	24935	Muridae	Muridae	Bandicota	Bandicota	2016-01...	高雄市茄...	120.195...	22.894975
4	24944	Talpidae	Talpidae	Mogera	Mogera	2016-01...	台中市北...	120.788...	24.173821
5	24971	Colubridae	Colubridae	Lycodon	Lycodon	2016-01...	新竹縣橫...	120.98417	24.698021
6	24972	Viperidae	Viperidae	Trimeres	Trimeres	2016-01...	苗栗縣頭...	120.974...	24.702631
7	25004	Bufo	Bufo	Bufo	Bufo	2016-01...	台北市士...	121.538...	25.127258
8	25006	Potamidae	Potamidae	Geothel	Geothel	2016-01...	高雄市大...	120.408...	22.725732
9	25008	Laniidae	Laniidae	Lanius	Lanius	2016-01...	台南市柳...	120.349...	23.261861
10	25009	Muridae	Muridae	Rattus	Rattus	2016-01...	台中市潭...	120.724...	24.231545
11	25023	Rallidae	Rallidae	Amauror	Amauror	2016-01...	金門縣金...	118.412...	24.508996
12	25025	Columbi...	Columbi...	Streptop...	Streptop...	2016-01...	金門縣金...	118.319...	24.471015
13	25026	Bufo	Bufo	Duttaphr...	Duttaphr...	2016-01...	台北市文...	121.538...	25.008142
14	25027	Columbi...	Columbi...	Columba	Columba	2016-01...	新北市貢...	121.984...	24.99522
15	25028	Rhacoph...	Rhacoph...	Rhacoph...	Rhacoph...	2016-01...	基隆市暖...	121.734...	25.091321
16	25029	Rhacoph...	Rhacoph...	Rhacoph...	Rhacoph...	2016-01...	基隆市暖...	121.734...	25.091321
17	25030	Rhacoph...	Rhacoph...	Rhacoph...	Rhacoph...	2016-01...	基隆市暖...	121.734...	25.091321
18	25032	Colubridae	Colubridae	Ptyas	Ptyas	2016-01...	台中市太...	120.7971	24.10225
19	25033	Strigidae	Strigidae	Strigidae	Strigidae	2016-01...	台北市文...	121.586...	24.982227
20	25034	Sturnidae	Sturnidae	Acridoth...	Acridoth...	2016-01...	金門縣金...	118.413...	24.507908

Resources

路殺社範例資料集

Social

Google+

Twitter

Facebook

Additional Information

Last updated June 4, 2018

Created June 4, 2018

Format CSV

Encoding UTF-8

Coordinate Systems EPSG: 4326

License License Not Specified

About 跨領域區域研究資料集

Contact us: data.contact AT depositar.io

CKAN API | Site Statistics | Site Status | Documentation

Powered by ckan

Source code available on GitHub

Language: English

Roadkill Observations – A Showcase

<https://data.depositar.io/dataset/2016>

Add Filter

Grid Graph Map 3218 records « 1 – 20 » 🔍 Search data ... Go » Filters

_id	id	科名	科名 (...	物種名	物種名...	紀錄時間	地點	經度	緯度
1	24933	Bufo	蟾蜍科	Rana lon...	長腳赤蛙	2016-01...	宜蘭縣員...	121.636...	24.754443
2	24934	Muscica...	鶺鴒科	Turdus p...	白腹鶺鴒	2016-01...	基隆市中...	122.08166	25.62861
3	24935	Muridae	鼠科	Bandicot...	鬼鼠	2016-01...	高雄市茄...	120.195...	22.894975
4	24944	Talpidae	鼯鼠科	Mogera i...	台灣鼯鼠	2016-01...	台中市北...	120.788...	24.173821
5	24971	Colubridae	黃頰蛇科	Lycodon...	白梅花蛇	2016-01...	新竹縣峨...	120.98417	24.698021
6	24972	Viperidae	蝮蛇科	Trimeres...	赤尾青竹絲	2016-01...	苗栗縣頭...	120.974...	24.702631
7	25004	Bufo	蟾蜍科	Bufo ban...	盤古蟾蜍	2016-01...	台北市士...	121.538...	25.127258
8	25006	Potamidae	溪蟹科	Geothel...	黃綠澤蟹	2016-01...	高雄市大...	120.408...	22.725732
9	25008	Laniidae	伯勞科	Lanius c...	紅尾伯勞	2016-01...	台南市柳...	120.349...	23.261861
10	25009	Muridae	鼠科	Rattus n...	溝鼠	2016-01...	台中市潭...	120.724...	24.231545
11	25023	Rallidae	秧雞科	Amauror...	白胸苦惡鳥	2016-01...	金門縣金...	118.412...	24.508996
12	25025	Columbi...	鳩鴿科	Streptop...	珠頸斑鳩	2016-01...	金門縣金...	118.319...	24.471015
13	25026	Bufo	蟾蜍科	Duttaphr...	黑眶蟾蜍	2016-01...	台北市文...	121.538...	25.008142
14	25027	Columbi...	鳩鴿科	Columba...	岩鴿	2016-01...	新北市貢...	121.984...	24.99522
15	25028	Rhacoph...	樹蛙科	Rhacoph...	台北樹蛙	2016-01...	基隆市暖...	121.734...	25.091321
16	25029	Rhacoph...	樹蛙科	Rhacoph...	台北樹蛙	2016-01...	基隆市暖...	121.734...	25.091321
17	25030	Rhacoph...	樹蛙科	Rhacoph...	台北樹蛙	2016-01...	基隆市暖...	121.734...	25.091321
18	25032	Colubridae	黃頰蛇科	Ptyas dh...	過山刀	2016-01...	台中市太...	120.7971	24.10225
19	25033	Strigidae	鴞鴞科	Strigidae	鴞鴞科	2016-01...	台北市文...	121.586...	24.982227
20	25034	Sturnidae	椋鳥科	Acridoth...	家八哥	2016-01...	金門縣金...	118.413...	24.507908

Grid

Graph

Map

406 records

«

1

-

406

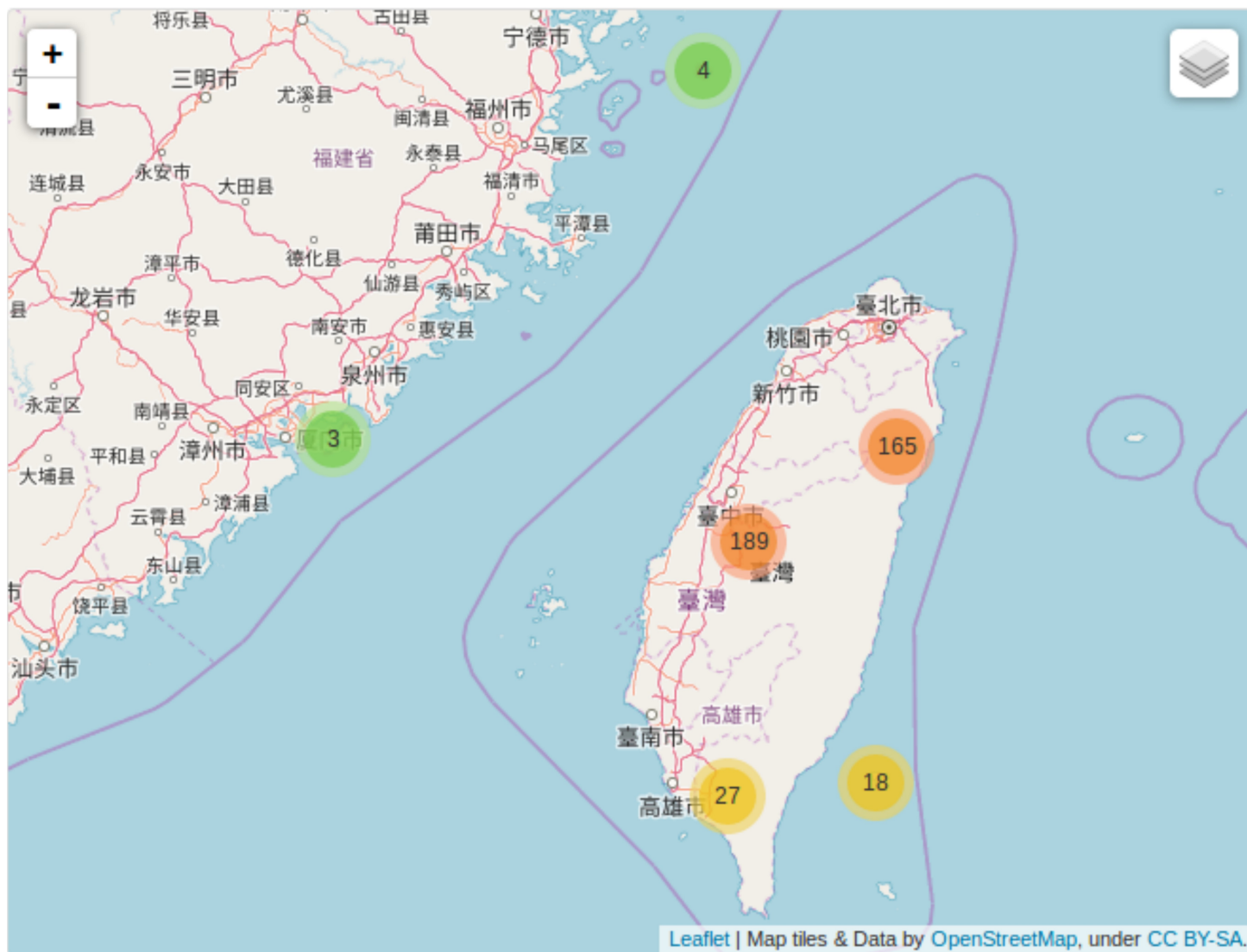
»



Search data ...

Go »

Filters

☒ Latitude / Longitude fields☐ GeoJSON field

Latitude field

緯度

Longitude field

經度

Update

☒ Auto zoom to features☒ Cluster markers

Filters

Add filter

科名 (中文) ×

蟾蜍科

Update

Grid

Graph

Map

406 records

«

1

-

406

»



Search data ...

Go »

Filters

+

-



經度: 121.065569

物種名 (中文): 黑眶蟾蜍

緯度: 22.736836

地點: 台東縣臺東市

科名 (中文): 蟾蜍科

紀錄時間: 2016-05-11T00:00:00

物種名: *Duttaphrynus melanostictus*

_id: 1925

科名: Bufonidae

id: 27273

☒ Latitude / Longitude fields☐ GeoJSON field

Latitude field

緯度

Longitude field

經度

Update

☒ Auto zoom to features☒ Cluster markers

Filters

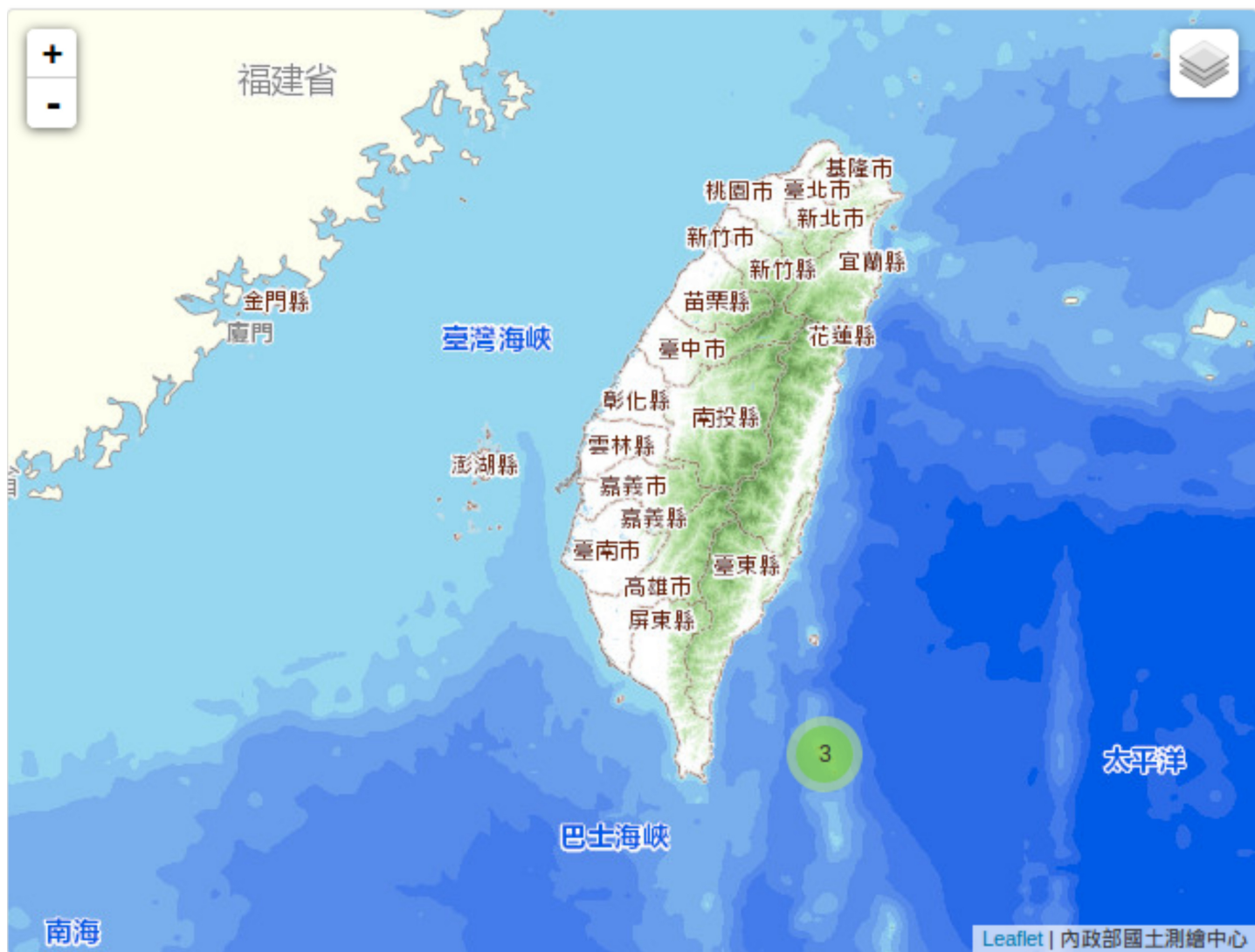
Add filter

科名 (中文) ×

蟾蜍科

Update

Leaflet | Map tiles & Data by OpenStreetMap, under CC BY-SA.



- ☒ Latitude / Longitude fields
☐ GeoJSON field

Latitude field

緯度

Longitude field

經度

Update

- ☒ Auto zoom to features
☒ Cluster markers

Filters

Add filter

科名 (中文) ×

陸寄居蟹科

Update

GridGraphMap

3 records

«1-3»

Search data ...

Go »

Filters

+

-

經度: 121.562714
物種名 (中文): 椰子蟹
緯度: 22.055227
地點: 台東縣蘭嶼鄉
科名 (中文): 陸寄居蟹科
紀錄時間: 2016-04-30T00:00:00
物種名: Birgus latro
_id: 1647
科名: Coenobitidae
id: 26936

✕

Layers

Latitude / Longitude fields

GeoJSON field

Latitude field

緯度

Longitude field

經度

Update

☒ Auto zoom to features

☒ Cluster markers

Filters

Add filter

科名 (中文) ✕

陸寄居蟹科

Update

Leaflet | 內政部國土測繪中心

Grid

Graph

Map

3 records

«

1

- 3

»



Search data ...

Go »

Filters

+

-

經度: 121.510102

物種名 (中文): 短掌陸寄居蟹

緯度: 22.070254

地點: 台東縣蘭嶼鄉

科名 (中文): 陸寄居蟹科

紀錄時間: 2016-05-02T00:00:00

物種名: Coenobita brevipennis

_id: 1795

科名: Coenobitidae

id: 27120

☒ Latitude / Longitude fields☐ GeoJSON field

Latitude field

緯度

Longitude field

經度

Update

☒ Auto zoom to features☒ Cluster markers

Filters

Add filter

科名 (中文) ×

陸寄居蟹科

Update



Leaflet | 內政部國土測繪中心

Openness Benefits Research

- Help disseminate research findings.
- Help reproduce and re-purpose research results.
- Help encourage research collaborations.

Openness

- Libre
 - can be used by people
- Digital
 - can be used by machines and put online
- Raw
 - can be modified and re-purposed
- Common (format & vocabulary)
 - can be exchanged and interlinked
- Transparent
 - (the process) can be fixed; meta-level

FAIR Data

<https://www.force11.org/group/fairgroup/fairprinciples>

- **Findable:** Have sufficiently rich metadata and a unique and persistent identifier.
- **Accessible:** Retrievable by humans and machines through a standard protocol; open and free by default; authentication and authorization where necessary.
- **Interoperable:** Metadata use a “formal, accessible, shared, and broadly applicable language for knowledge representation”.
- **Reusable:** Metadata provide rich and accurate information; clear usage license; detailed provenance.

A Work In Progress

- User manual (Chinese and English languages)
- Interoperable base map services when previewing spatial datasets
- Outreach
 - to researchers who have been curating datasets
 - to students, by being part of teaching in classes
- Research Data Management Workshop
 - currently planned in 2018 Winter

搜尋資料

例如：環境



熱門標籤

台江國家公園

海圖

黑面琵鷺

最新消息

系統更新 / System Upgrade v6.0 (2017/11)

十一月 3, 2017

[English Below]

跨領域區域研究資料集 版本升級為 v6.0。

同時更新 CKAN 核心版本至 2.6.4。(depositar-io/ckan@85bb3fe)

更新內容如下：

- ...

系統更新 / System Upgrade (2017/9)

九月 5, 2017

[English Below]

跨領域區域研究資料集 版本升級作業已於日前完成。

組織

測試區 / Sandbox

僅供測試用途。 For testing purposes only.

嘉南大圳研究群 / Chia-nan Project

本組織包含「嘉南平原水文、農業空間及聚落生活空間之百年變遷與連動關係之研究」之各子計畫單位內成員，各子計畫於計畫執行期間，可將尚未對外開放之資料發布於本組織內。

中研院人社中心GIS專題中心

隸屬中央研究院人文社會科學研究中心，從事地理資訊科學相關基礎研究與資源整合，建立大型學術資料庫，致力於推動地理資訊系統技術應用在人文社會科學研究上。

File Edit View History Bookmarks Tools Help

組織 - 跨領域區域研究資料集 x +



https://data.depositar.io/organization

80%



Search



登入

跨領域區域研究資料集

資料集

組織

主題

支援

搜尋



家 / 組織

❗ 組織是什麼？

您可以使用 CKAN 組織來建立、管理與發佈多個資料集。在組織中，每位使用者可擁有不同角色，而得以建立、編輯與發佈資料集。

搜尋組織



找到4個組織

排序依照: 根據名稱遞增排序

as-rchss-gis

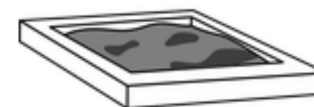
中研院人社中心GIS專題中心

隸屬中央研究院人文社會科學研究中心，從事地理資訊科學相關基礎研究與資源整合，建立大型學術資料庫，致力於推動地理資訊系統技術應用在人文社會科學研究上。

Chia-Nan

嘉南大圳研究群 / Chia-nan Project

本組織包含「嘉南平原水文、農業空間及聚落生活空間之百年變遷與連動關係之研究」之各子計畫單位內成員，各子計畫於計畫執行期間，可將尚未對外開放之資料發布於本組織內。



測試區 / Sandbox

僅供測試用途。 For testing purposes only.



台江內海地區跨領域研究群 / Taijiang Project

本組織包含「台江內海地區跨領域研究」計畫之各子計畫單

https://data.depositar.io/organization/sandbox

Free Software for Serving Open Source/Data/Media



ckan



GitLab

WIKIPEDIA

English

The Free Encyclopedia
4 000 000+ articles

日本語

フリー百科事典
815 000+ 記事

Español

La enciclopedia libre
900 000+ artículos

Русский

Свободная энциклопедия
875 000+ статей

Italiano

L'enciclopedia libera
939 000+ voci

Português

A enciclopédia livre
742 000+ artigos

Deutsch

Die freie Enzyklopädie
1 430 000+ Artikel

Français

L'encyclopédie libre
1 270 000+ articles

Polski

Wolna encyklopedia
908 000+ haseł



中文

自由的百科全書
500 000+ 條目

Search · Suchen · Rechercher · Zoeken · Ricerca · Szukaj · **Buscar** · Поиск · 検索 ·
Busca · Sök · 搜尋 · Tìm kiếm · Пошук · Cerca · Søk · Haku · Hledání · Keresés · 찾기 · Cari ·
Ara · جستجو · Căutare · بحث · H'adat' · Søg · Serçu · Претрага · Paieška · Poišči · Cari · חיפוש ·
Търсене · Izdey · Bilatu · Suk · Bilnga · Traži · खोजें



contents 内容

工具 tools

users 用戶

服務
services



Thank you!

<http://data.depositar.io/>

<http://demo.depositar.io/>

Please contact us at:

data.contact@depositar.io

