# Institutional and Communal Data Repositories: Practices and Tools

## 機構的以及眾有的資料典藏庫: 實務與工具

2015 數位化論文典藏聯盟年會 數位化學位論文與開放取用之發展趨勢

> 2015-11-13 NCCU 政治大學

Tyng-Ruey Chuang 莊庭瑞

IIS, CITI & RCHSS 資訊所、資創中心、人社中心 Academia Sinica 中央研究院

### Institutional Repository

- Self-archiving, with institutional support
  - online, persistent, and accessible
- Creditable sources for finding intellectual works
  - institution members can deposit and access
  - all others can access (and cite)
- Inter-/multi- institutional repositories
  - networks of scholarly communication
  - institution's roles in facilitating the networks
- Interoperable protocols, e.g. OAI-PMH
- Open source software, e.g. DSPACE; Self-hosting

### Communal Repository

- Self-archiving, with communal support
  - online, persistent, and accessible
- Convenient places for storing useful materials
  - community members can deposit, access, and reuse
  - communal norms in the use of the repository
- Often use public licenses to facilitate reuse
- Centralized or distributed networks of repositories
  - federations of catalog services and user activities
- Open source software e.g. GNU MediaGoblin
- Self-hosting and/or relying on third-party services

### Data Repository

- Just for datasets, or anything digital?
  - code, text, image, sound, video, etc.
  - data: bits, formerly known as content
- Metadata, metadata, and metadata
- Discovery and reuse of source materials
- Often place heavy demands on storage, computation, communication, and human resources
- Data provenance and integrity; data citation
- Open source software e.g. CKAN
- Self-hosting and/or relying on third-party services

### Open Repository?

- "Open" is open to different interpretations ...
- Open Content
- Open Data
- Open Source
- Open Collaboration
- Open Government
- Open Science
- "Open to suggestions"

### **Openness**

- Libre
  - can be freely used by people
- Digital
  - can be used by machines and put online
- Raw
  - can be modified and repurposed
- Common format and vocabulary
  - can be exchanged and semantically linked
- Transparent
  - (the process) can be fixed; meta-level

### Public Licenses

- "Some Rights Reserved"; Other rights to use the work are granted to the public with a license.
- The license is worldwide, royaltyfree, non-exclusive, and perpetual.
- The licensor makes clear what is to be licensed, and under what license.
- A licensee is any person who agrees to the license, and uses the work.

### Creative Commons Licenses







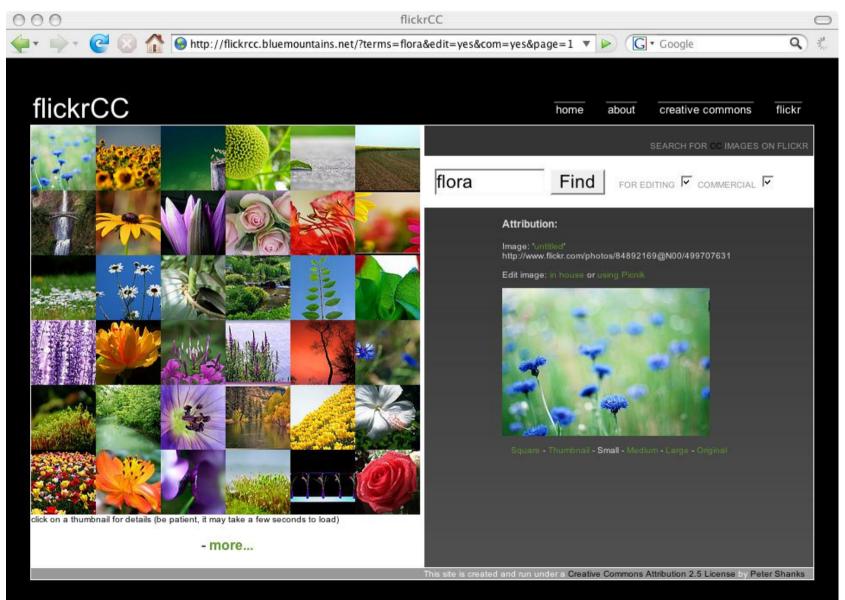






### Search by License

http://flickrcc.net/



### Public Domain Materials

- No longer under copyright
- Never under copyright
- Not copyrightable
- Waived all copyrights
- Other IP analogies (patents etc.)

- Difficult to draw clear boundaries
- Issues related to orphan works

### Labeling Public Domain Materials

- A matter of (clear) communication
- Public Domain Mark (PD Mark)
  - expression of fact/faith by someone

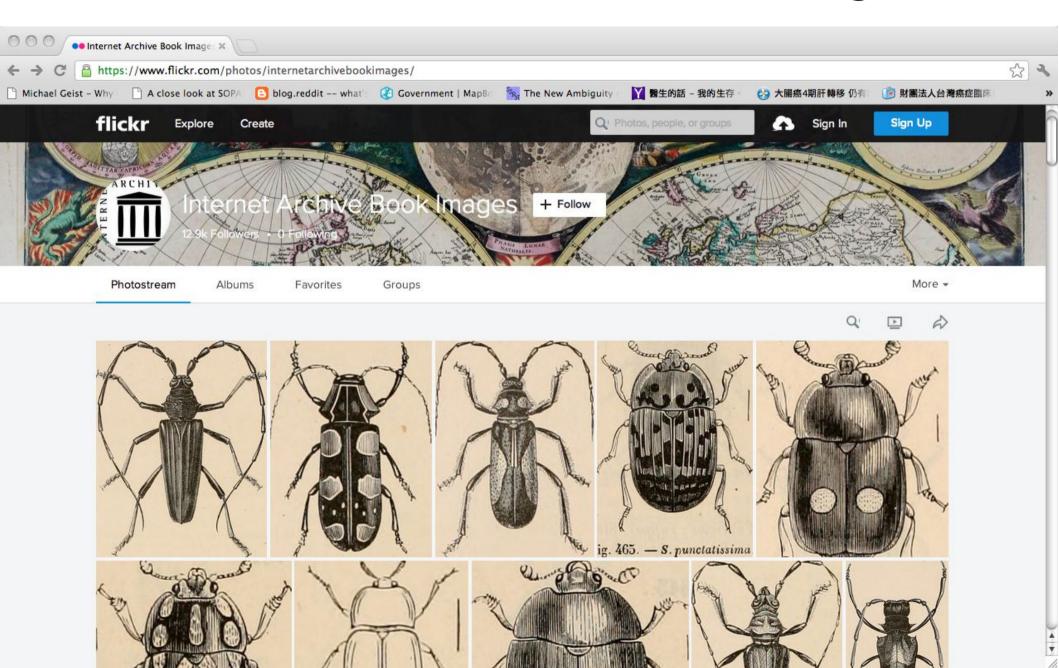


- Public Domain Dedication (CC0)
  - assertion from the rights holder

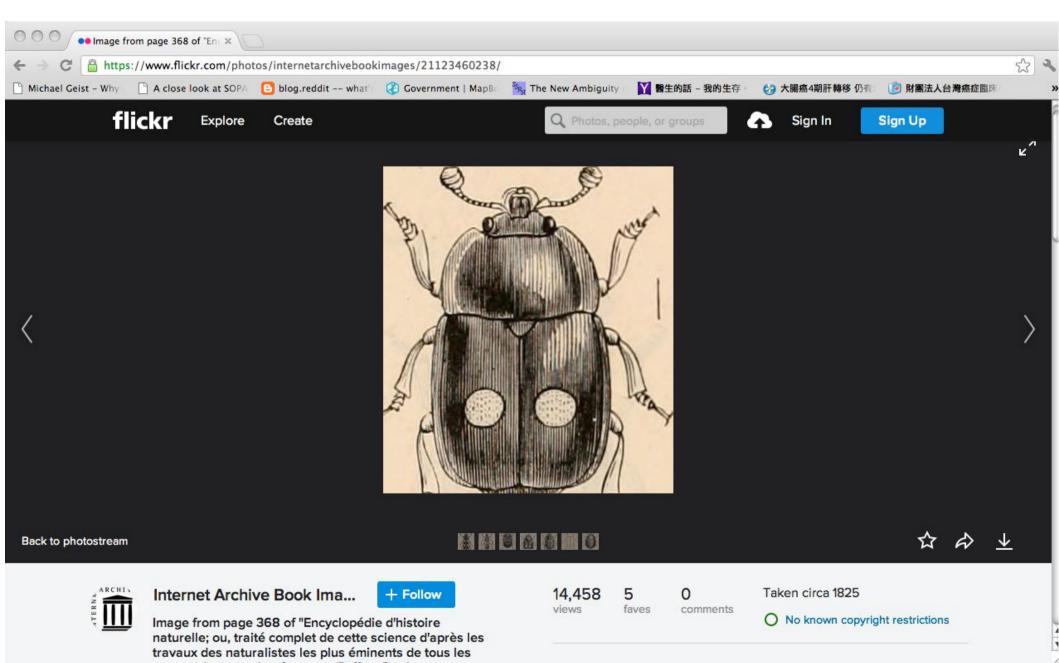




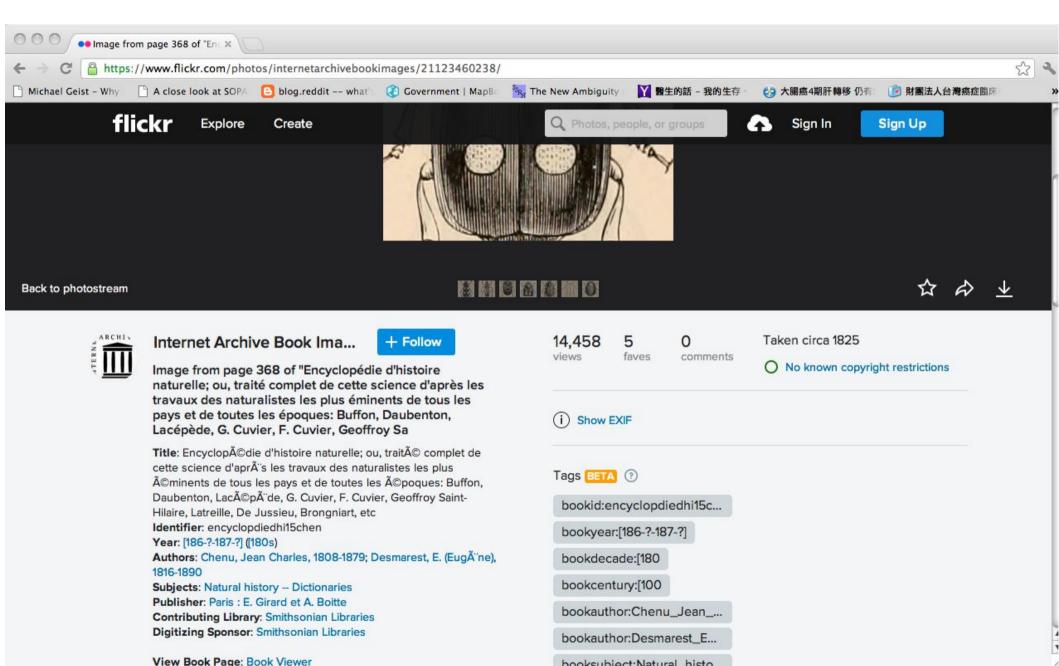
### Internet Archive Book Images



### Internet Archive Book Images



### Searchable Metadata

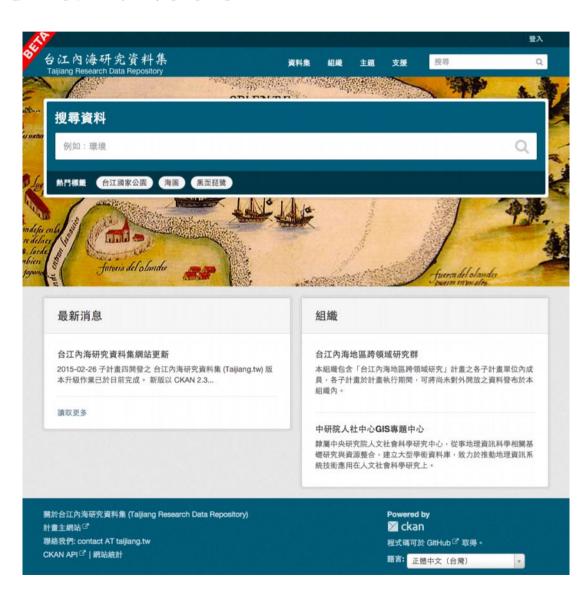


### (Communal) Data Repositories

- Examples taken from projects at Academia Sinica and with external collaborators
- Taijiang Research Data Repository
  - http://taijiang.tw/
- Taiwan Roadkill Observation Network
  - https://roadkill.tw/
- 318 Civil Movement Archive
  - http://public.318.io/
- An exemplar public domain depository
  - http://publicdomain.tw/

### Taijiang Research Data Repository 台江內海研究資料集

- Part of interdisciplinary project
- Datasets, documents, maps, etc.
- XLS, PDF, Shapefile, etc.
- Some control vocabularies
- Use CKAN
  - Comprehensive Knowledge
     Archive Network
  - http://ckan.org/



### (Customized) Features

### Spatiotemporal Search 時空間搜尋機制



### Spatial Metadata Helper 空間欄位填寫輔助



#### Preview 資料預覽



### More Features

- Batch upload
- Usage statistics





- 特定格式欄位 (欄位文字以藍色表示):參考下頁特定格式欄位說明
- 特定選項欄位 (欄位文字以紅色表示):選項請參考範例 excel 檔之「清單選項」頁籤
- 接受多重值之欄位(主題關鍵字、空間範圍關鍵字、使用史料、群組、標籤、參考來源, 後兩者可自由填寫),請以「:」做區隔

# Taiwan Roadkill Observation Network 台灣野生動物路死觀察網



















Miseru Wang | 尋找朋友 | 首頁



四處爬爬走(路殺社, Reptile Road Mortality)

成員

活動

相片

檔案

✓ 通知

+ 建立社圏



搜尋人、地點和事物









留個言吧……

置頂贴文



#### 林德恩

各位夥伴大家好:

由於「台灣野生動物路死觀察網」資料彙整進入最後階段 目前電腦程式已順利自動擷取下6成的資料 但還有4成資料是在「相簿」中,無法被程式讀取到 為了盡快將這4成資料擷取到資料庫裡……

更多

划閣

3,536 位成員

❷ 公開社團

社團四大宗旨:

1.改善路死--藉由大量的資料蒐集,分析路 死嚴重的路段、季節及種類,改善道路設計或 增加廊道、圍籬等設施,以減少野生動物因國 內道路開發與車流量日漸增加而造成的直接死 亡(roadk ····· 更多

3,536 位成員 (56 位新成員) 以電子 郵件激請

+加成員到社團









#### 主撰單

#### ▼ 關於路級社

- ▶ 社團宗旨
- 。 管理員介紹
- 。創用CC授權
- 。 最新消息
- 。 加入我們
- ▼ 路級資料庫 o 物種分類樹
  - 。物種圖庫
  - · 圖片審核總覽
  - 。 路殺觀測紀錄
  - o 路級分布地圖

#### ▼ 學習教室

- 。 如何記錄路死動物
- 。 照片上傳驗書社剛說明
- 。 路死動物拍照說明
- o 定位教學
- 。 路死標本寄送SOP
- 全台同步大調查說明
- 生物小百科

#### ▼ 下載

- · App
- o 記錄表格下載
- 。路殺社第1屆年會
- 保育類野生動物名錄
- 下載路死紀錄

#### miseru

- 0 首百
- 建立路死紀錄
- 。我的路死紀錄
- o 使用者列表
- o ADMIN內容管理
- 。(管理者)我的內容
- 。 文章内容管理
- ▼ 建立内容
- o Panel
- 。 文章頁面 o 最新消息
- o 路死紀錄
- 。 資料匯入管理

#### 首頁

#### **国庫總管**

前往分類圖庫 動物圖庫 地圖瀏覽



521357714617045

▶ 作者: 吳銘
(cc) BY

663605453660080

2013-10-07 04:18:35

▶ 作者: 陳秀晴 (C) ALL RIGHTS



▶ 作者: 劉珊珊 (C) ALL RIGHTS

2013-10-07 13:46:04 2013-10-07 13:20:11 2013-10-07 04:28:15

663603120326980

▶ 作者: 陳秀晴

(C) ALL RIGHTS

▶ 作者: 蘇哲弘

(C) ALL RIGHTS



663608226993136 ▶ 作者: 陳秀晴 (C) ALL RIGHTS













499196486843671 ▶ 作者: 謝中和 (cc) BY-NC-SA

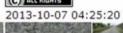


662815877072371 662815113739114 ▶ 作者: 陳秀晴



663607350326557 ▶ 作者: 陳秀晴

(C) ALL RIGHTS



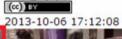


521017964651020 521017287984421 ▶ 作者: 吳銘 ▶ 作者: 吳銘 (cc) BY





▶ 作者: 陳秀晴





▶ 作者: 莊克敏 (cc) BY-NC-SA

013-10-06 01:55:36





521016774651139

▶ 作者: 吳銘

(cc) BY

663605733660052

▶ 作者: 陳秀晴

(C) ALL RIGHTS

▶ 作者: AWei H... (C) ALL RIGHTS 2013-10-05 12:30:24



662814860405806 520160304736786 ▶ 作者: 吳銘



4918961631743 ▶ 作者: Ya-wan...



521014844651332

▶ 作者: 吳銘

(cc) BY

662816190405673 ▶ 作者: 陳秀晴



▶ 作者: 陳秀晴





499195176843802

▶ 作者: 謝中和

(cc) BY-NC-SA

▶ 作者: 陳秀晴

### You Gotta Have Apps!







Privacy Policy Public Licensing Anonymous Posting







#### 台灣野生動物路死觀察網

#### Taiwan Roadkill Observation Network

關於路殺社 參與指南

消息與公告

保育文章

主題行動

路殺社成果

社團四大宗旨 ●

#### 改善路死

藉由大量的資料蒐集,分析路死 嚴重的路段、好發季節及受威脅 種類,改善道路設計或增加廊 道、圍籬等設施,以減少野生動 物因國內道路開發與車流量日漸 增加而造成的直接死亡 (roadkill) •

#### 環境教育

藉由民眾參與「路死動物」的記 錄活動,推廣並提升國人生態保 育觀念。

#### 全民科學

推廣公民參與科學調查的概念 (Citizen science),由全民一同 發掘生活中的環境議題,進而提 出想法、試驗設計、資料蒐集、 分析討論和合作解決。

#### 珍惜生命

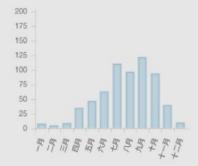
將路死動物個體製成標本典藏, 無償提供學術研究、教學或展示 之用,減少因學術研究所需而造 成的活體犧牲。



#### Bungarus multicinctus multicinctus (雨傘節)

#### 641筆路殺紀錄!





# 318 Civil Movement Archive 318 公民運動文物紀錄典藏庫

- Historians at Academia Sinica arranged to acquire artifacts in the main chamber when students decided to evacuate from the legislature.
- Thousands of art works and notes have been digitized. Are these orphan works?
- More videos and photos were acquired. CC0?

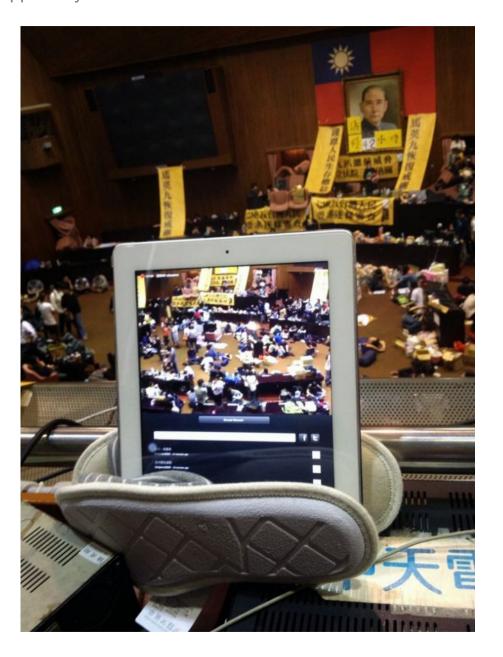
 Goals: Collecting and preserving the artifacts not only for the purpose of future research, but also for deepening collective memories about this contemporary/historical event.



"Occupy Taiwan Legislature by VOA (1)" by Voice of America https://commons.wikimedia.org/wiki/File:Occupy\_Taiwan\_Legislature\_by\_VOA\_(1).jpg

#### 專訪 LIVE 阿北熱血轉播 | 即時新聞 | 20140328 | 蘋果日報 http://www.appledaily.com.tw/realtimenews/article/new/20140328/368726/





"Occupy Taiwan Legislature by VOA (8)" by Voice of America https://commons.wikimedia.org/wiki/File:Occupy\_Taiwan\_Legislature\_by\_VOA\_(8).jpg



Photo by Lai Chi-Lin.

CC BY-NC-ND 3.0 TW http://creativecommons.org/licenses/by-nc-nd/3.0/tw/



Photo by Lai Chi-Lin.

CC BY-NC-ND 3.0 TW http://creativecommons.org/licenses/by-nc-nd/3.0/tw/

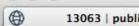


Photo by Lai Chi-Lin.

CC BY-NC-ND 3.0 TW http://creativecommons.org/licenses/by-nc-nd/3.0/tw/

### March 18, 2015





000

13063 | public.318.io

+

















#### 13063

識別號: 13063 主要形式: 平面

形式分類:海報 印刷

內容分類: 繪畫

產製日期: 產製地點:

測量: 599x427mm

描述:

兩張海報組成。左邊畫面為一個披著中華民國國旗的 人,身邊有多朵太陽花。右邊畫面為兩隻手正在握手, 周圍有許多太陽花,標題為「謝謝在前線的你們台灣加

油」

內容:

材質: 合成紙

姓名標示值:

< 上一筆 温湯難

(共7389筆・本頁顯示第1筆)



14017 | public.318.io

+ |













♠ > 找 TERM:CC-BY > 14017



License: CC BY 4.0

Attribution: BOOKSHOW

Status: Publicly Licensed

#### 14017

識別號: 14017

主要形式:數位媒體

形式分類:影片

內容分類: 剪輯實錄

所屬事件: 318占領立法院

產製日期: 03/18/2014

產製地點: 立法院院區周圍

測量: 0:08:08

描述:

318占領立法院過程紀實「反黑箱服貿、守護民主之

夜」

內容:

釋出條款: CC BY 4.0 @

姓名標示值: BOOKSHOW

權利狀態:已採用公眾授權 ①

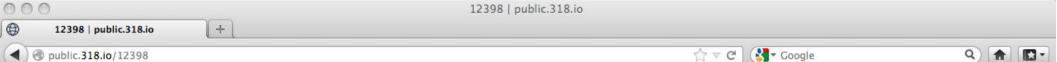
温測算 | 工一業

下一筆 >

(共145筆,本頁顯示第2筆)

### Identification of Works

- The public can search the archive to identify works of their own
- Self certification of authorship
- Change in the copyright status of a work
  - Unknown Copyrights Status
  - Copyrights Claimed and Reserved
  - Publicly Licensed (Creative Commons Licensed)
  - Dedicated to the Public Domain
- Example



關於我們 About 藏品指認 Identification 首頁 Home 使用規範 Terms of Use 登入 Login

↑ > ALL > 12398



#### 12398

識別號: 12398 主要形式: 平面

形式分類:海報 黏貼 內容分類:太陽餅

產製日期: 產製地點:

測量: 210x296mm

描述:太陽堂太陽餅一塊

內容:

材質:紙 塑膠 姓名標示值:

權利狀態: 權利狀態

If this is your work, click here to identify it.

#### 指認與授權

若您是本藏品著作人,請點選上區

認授權作業,想了解更多,請參考相關説明

引用網址: http://public.318.io/12398

温測戰

〈上一筆 下一筆 〉 (共7366筆,本頁顯示第1筆)



## Declaration and License Agreement; To Be Signed and Send Back

### 318 公民運動文物紀錄典藏庫

### 藏品著作權利聲明及同意書

#### 1. 聲明及同意事項

- 1. 僅以下列表格之場寫、列印、觀筆簽名及書面交寄之流程,向 318 公民運動文物紀錄典議及(以下稱 典藏座) 聲明收錄藏品之權利狀態,以利典藏廉於藏品展示時爲本人標註通當之顧名聲明。此外,與藏 庫並得就本聲明書內之指定藏品,依本人所選定之公眾授權及宣告方式,釋出高解析度之數位重製 品,提供公眾使用。
- 11. 透過表格資訊之場寫,代表本人明確知悉並同意,若以己名或冒名聲明非本人創作之藏品及其權利, 頻自負法律責任,且不得轉和關責任以任何理由歸請於與藏庫或其他和信此權利聲明資訊之善意第三
- 111. 本人周蠡典藏崖於收受藏品著作權利聲明及同意書後,可對本人聲明擁有著作權利之藏品,進行下列
  - 1. 就本人選定為對公眾授權之藏品,與藏庫得依資源負載以及風險評估之規劃,自主決定是否以高 解析度之數位重製品為該藏品進行釋出。
  - 2. 典藏廉僅對本聲明權利流程進行步驟控管,並不對聲明資訊作實質客驗。當個別藏品有不同權利 主體進行複式聲明時,典藏庫或轉以揭露併存的方式處理;倘若個別聲明人對其主張藏品權利款 態有釐清之需求,與藏庫當徵詢相關人並獲其聯絡資訊披露之同意後,橋接雙方、多方聲明人自 行對權利狀態進行協商處理。

#### 11. 著作品資訊及其權利聲明與授權行使

本人聲明以下藏品爲本人所創作,擁有著作權利並同意行使以下授權及宣告方式。對採行公眾授權方式或公 思領越首獻宣告的創作,本人並同意與藏庫得釋出高解析度之數位重製品,提供公思使用 (\* 必維資訊):

- 1. 藏品識別號: 12398
- 11. 聲明人本名\*: 莊庭瑞
- iii, 自訂姓名標示方式(於展示該藏品時標示使用): trc
- iv. 對該聲明藏品所行使的授權方式\*: 創用 CC 「姓名標示 3.0 台灣」授權條款及其援版本 (CC BY 3.0 TW and later)
- v. 身分證字號後四碼: 0000
- vi. 連絡電話: +886 2-2788-3799
- vii. 電子郵件\*: trc@iis.sinica.edu.tw
- viii, 通讯地址\*: Institute of Information Science, Academia Sinica, Nangang District 115, Tainei, Taiwan

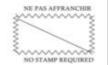
除姓名標示外,以上個人資料僅為身分識別及供與藏庫洽簽署人確認資訊之用,不對外揭露。

### 111. 答章及具结

借以下列與肇署名或萬章,表達本聲明及同意書所繼密之資料係為盈實。

立聲明及同意書人:_			_ (親筆簽名/蓋章	
日期:公元	年_	月	E	





REPLY PAID/PÉPONSE PAYÉE TAIWAN, ROC

Institute of History and Philology, Academia Sinica

中央研究院 歷史語言研究所 收

(歷史事件巨量資料的收集與研究計劃)

No.130, Sec. 2, Academia Rd., Nangang Dist., Taipei City 11529, Taiwan (R.O.C)



http://public.318.io 318公民連動文物記錄典網 3. 請用【部報】封司·寧到郵筒投鑑。 元的同縣邀出,「



不顧閩語, 次至銀小®

【容簽群五】龍,各簽人書即豐立 ◎

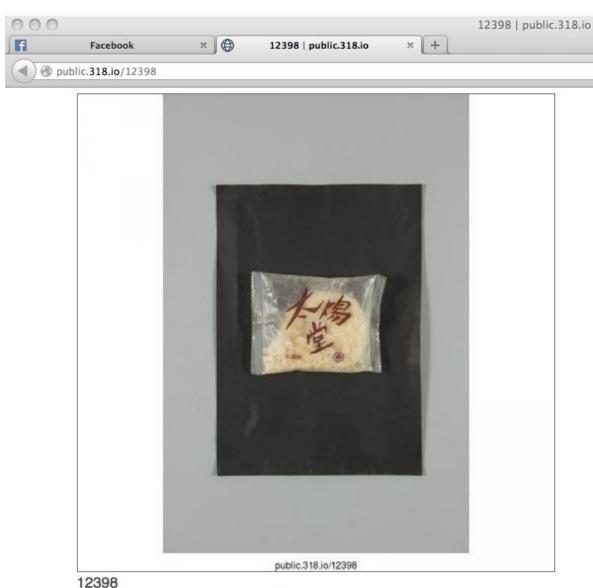
學戏咨匯

《据篇》《题》及张《初述明期目篇》



Taiwan Citizens Movement

**値重另公81**5



12398

識別號: 12398 主要形式: 平面

形式分類:海報 黏貼

↑ ▼ C Google

內容分類:太陽餅

產製日期: 產製地點:

測量: 210x296mm

描述: 太陽堂太陽餅一塊

內容:

材質:紙 塑膠

釋出條款: CC BY 3.0 TW 🗗

姓名標示值: trc

權利狀態:已採用公眾授權 🗊

## 指認與授權 (已被指認)

若您是本藏品著作

記授權作業→ License: CC BY 3.0 TW

引用網址: Attribution: trc

Status: Publicly Licensed

Q 🛖 🖸 -

▲下載典藏檔案

瀏覽 < 上─≌

下一筆 🗦

(共7389筆,本頁顯示第1筆)





Q 🛖 🖸 -

## 首頁

## 登入

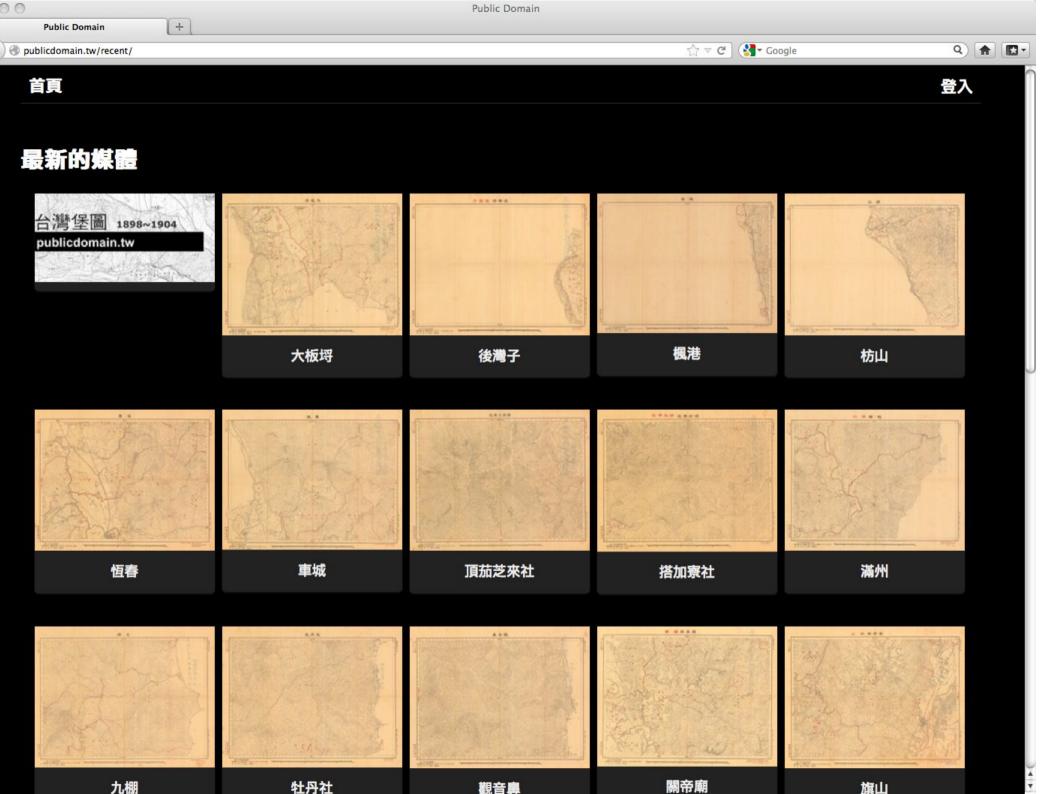


■ 本人繪製的「台灣堡圖」始於西元1898年,測繪時以劉銘傳的清丈區域為準,並加上 現代三角與高程測量,精度大為提高,是日後日人統治台灣的重要基礎資料,由於此 圖的繪製是以清代堡圖為準,故而稱之為「台灣堡圖」。(引用來源 作者:廖泫銘 提供單

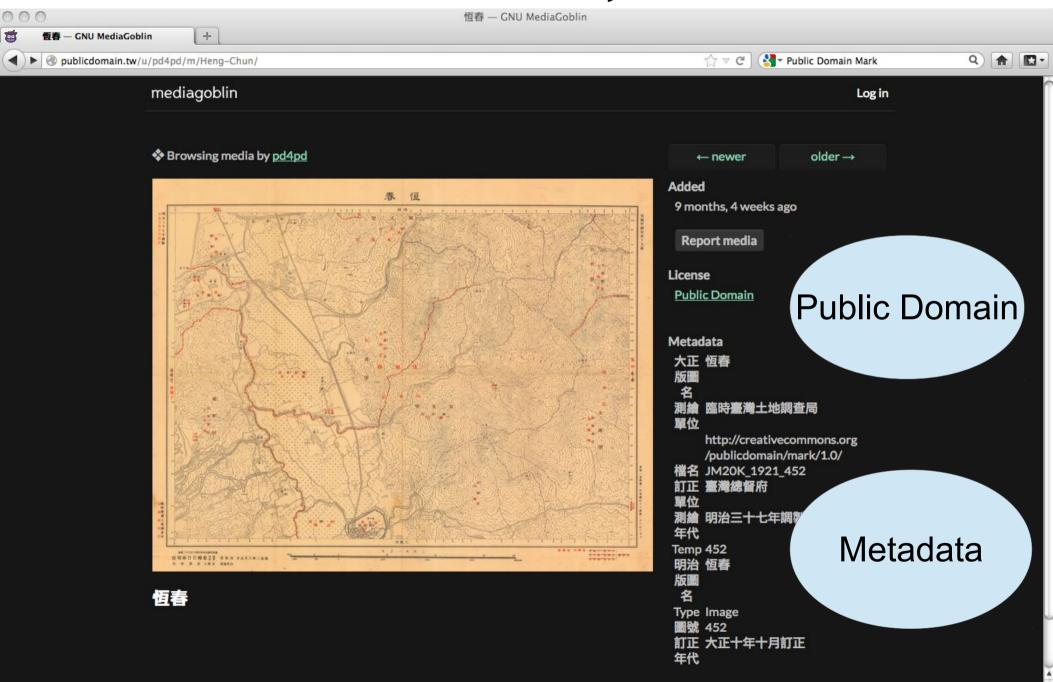
位:中央研究院數位典藏計畫文章出處:中央研究院數位典藏資源

網, http://digiarch.sinica.edu.tw/content.jsp?option\_id=2621&index\_info\_id=6924)

## 查看所有的媒體



# Public Domain, Labeled



## "Practices and Depositories for the Public Domain" (PD4PD)



tools

communities





# Learning from Wikipedia

## WikipediA

## **English**

The Free Encyclopedia 4 000 000+ articles

## **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

## 日本語

フリー百科事典 815 000+ 記事

## Deutsch

Die freie Enzyklopädie 1 430 000+ Artikel

## **Français**

L'encyclopédie libre 1 270 000+ articles

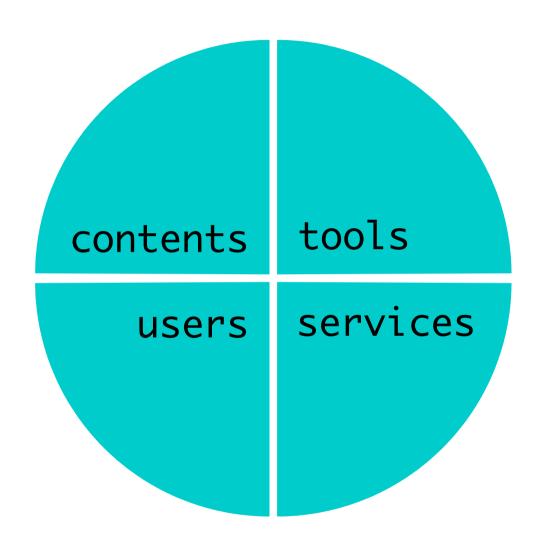
## Polski

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 · بستبر · Căutare · كما · Hl'adat' · Søg · Serĉu · Претрага · Paieška · Poišči · Cari · יפוש · Търсене · Iздеу · Bilatu · Suk · Bilnga · Traži · खोर्जे





[ MediaWiki

contents

tools

users

services





Thank You!

謝謝!

# Acknowledgements

- Anonymous contributors
- Academia Sinica
  - Academia Sinica Digital Center
  - Institute of History and Philology
  - Institute of Information Science
  - Research Center for Biodiversity
  - Research Center for Humanities and Social Sciences
  - Research Center for IT Innovation
- Digital Memory, Asia
- Word Gleaner Ltd.
- Taiwan Endemic Species Research Institute
  - ... and many more