

第13次中文文獻資源共建共享合作會議

13th Conference on Cooperative Development and
Sharing of Chinese Resources

July 13-14, 2021

跨域經緯·創新翻轉

—圖書館與數位人文的對話



漢學研究中心
Center for Chinese Studies

開放資料服務於漢籍數位人文 研究資源之發展與應用

洪振洲

法鼓文理學院 佛教學系

開放資料 Open Data

開放資料 (Open Data)一詞，近年來經常於科學界與政府機關的發展討論中被提到的重要議題。

其主要目的是希望促進相關各界將具有應用與研究價值的原始資料，開放給社會公眾取用，而不受其使用目的的限制。



哈佛大學神經科學系的 Open Data logo 圖片來源: [JulieBeck](#).

開放資料的優點

- 對於學術研究應用而言，開放資料可以帶來：
 - 增進研究品質與效率、
 - 更快速的產生想法、
 - 促進產生新創應用。



圖片來源: <http://www.pfirereview.com/20141227/>

- 對於十分依賴數位資料的數位人文研究領域來說，資料的取得與累積是本就是研究的關鍵之一。
- 開放資料服務的普及與否，將對於數位人文的研究發展造成重大影響。

研究資料生命週期 (Research Data Cycle)

專案創始階段

- Planning
- Collecting
- Processing and Analysing

專案公開階段

- **Publishing** and **Sharing**

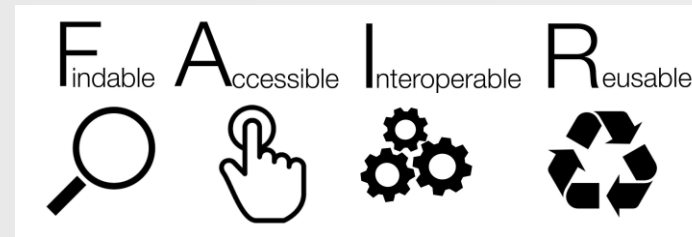
專案維護階段

- Preserving
- **Re-using**



Source: UK Data Service

FAIR 原則



Source: wikipedia

- **F**indable(可查/可查找):

- 後設資料被附予一個**全球唯一且永久不變的識別碼**。
- 資料被豐富的後設資料所描述，並在可搜索的資源中註冊或索引。

- **I**nteroperable(可換/可交換):

- **後設資料**使用正規化、**可取用、可共享**和廣泛適用的語言表達。
- 後設資料使用符合**FAIR**原則的語彙。

- **A**ccessible(可取/可取用):

- 後設資料**使用標準化的通信協議**，以通過其識別碼檢索。
- 即使資料本身不再可取得，後設資料也仍可取用。

- **R**eusable(可再用):

- 後設資料具有多重準確且相關的屬性
- **資料符合領域相關的社群標準**

開放資料服務的營運方式

• 集中式 開放資料服務：

- 常見於政府或大型研究組織，所開設的資料匯集與分享中心
- 需要大量資源來建設，並且因為內容資料異質性高，難以呈現單一資料集之特色應用。

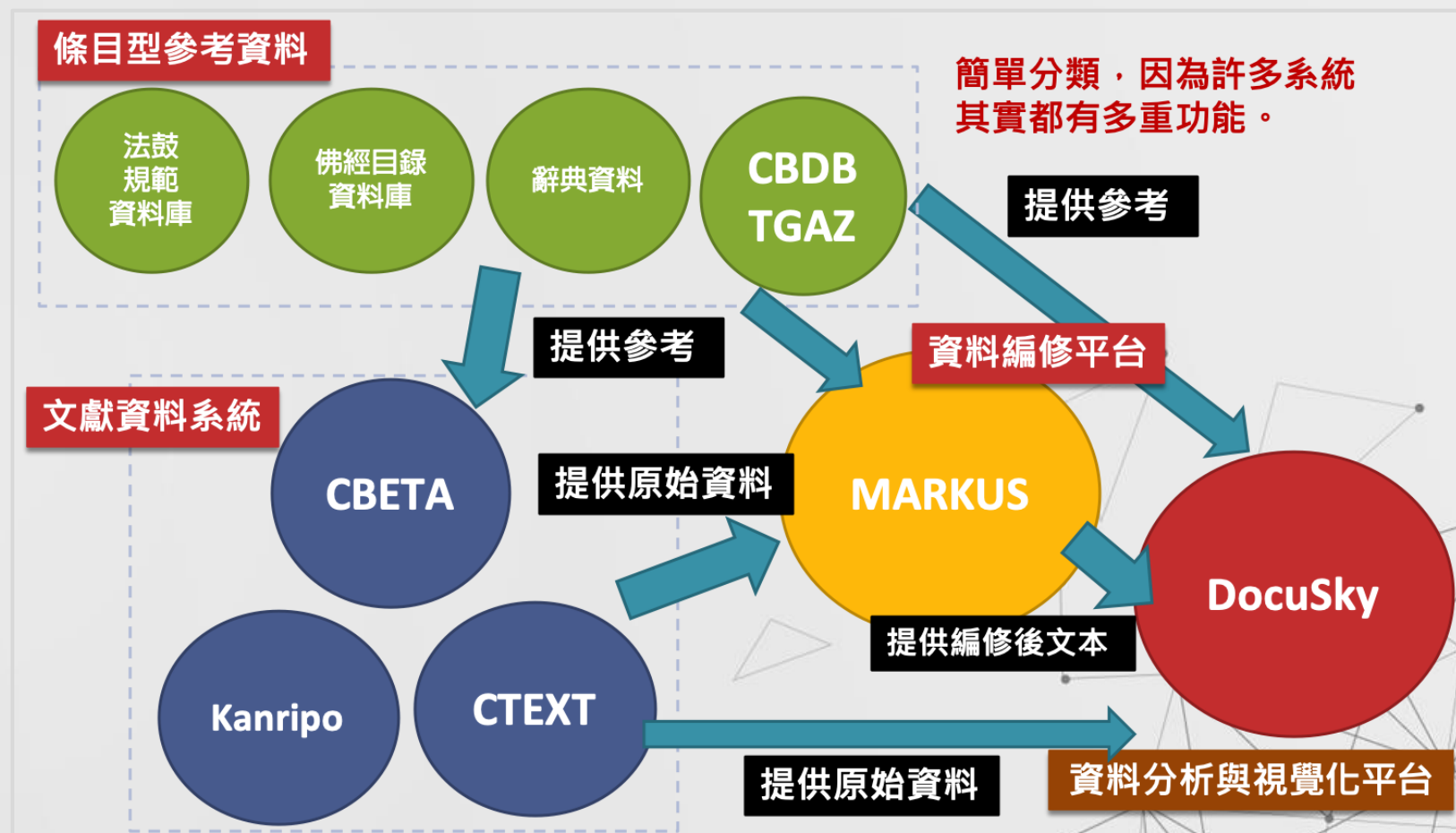
• 分散式 開放資料服務

- 在無合用的集中式開放資料機制時，由各專案自行建置。
- 漢藉數位資源的現狀



漢藉數位資源的開放資料服務概況

- 許多漢藉數位資源，提供開放資料服務。
- 服務包含提供**原始資料的下載**，或提供方便其他系統進行資料交換的開放性**應用程式界面（API）**。
- 有些數位人文工具間，已經進一步利用各自所設置的開放資料服務，建立跨系統連結。



本研究關注的議題

- 本論文的主要目的，便是將幾個知名的數位人文系統所建置的開放資料服務進行整理說明與簡易分析比較。

1. 中國歷代人物資料庫(China Biographical Database Project , 簡稱CBDB)
2. 臺灣歷史人物傳記資料庫 (TBDB)
3. 台灣傳記知識本體資料庫 (TBIO)
4. 法鼓文理學院佛學規範資料庫
5. 中央研究院數位文化中心鏈結開放資料平台
6. CBETA數位研究平台
7. 中國哲學書電子化計劃(Chinese Text Project , 簡稱CText)
8. 漢リポ (Kanseki Repository , 簡稱kanripo)
9. MARKUS半自動標記平台 ,
10. Docusky 數位人文研究平台
11. 中央研究院數位人文研究平台



開放資料的定義

- Open Definition(<https://opendefinition.org/>)網站，對於開放資料的定義：
- 開放資料必須以**開放式授權**的方式公開發行。
- 開放資料必須可由網路**下載**。
- 開放資料必須為**可機讀**格式。
- 開放資料必須為**開放**格式。



Open Definition

DEFINING OPEN IN OPEN DATA, OPEN CONTENT AND OPEN KNOWLEDGE

鏈結開放資料 (Linked Open Data)

- 提姆柏納-李 (Tim Berners-Lee) 在 2006年提出。其最終目的是要產生一個讓電腦能夠直接或間接處理的資料網絡(Web of Data)。
 - 每一個表述的資源必須具有唯一URI (Uniform Resource Identifier , 統一資源識別碼) , 用以表達與存取資料。
 - 資源的存取方式是透過HTTP協定 (Hyper Text Transfer Protocol , 超連結傳輸協定) 與該資源的URI而達成。
 - 資源必須藉由開放標準來提供存取方式, 例如RDF表示方式、SPARQL查詢語言。
 - 當資料發佈於網路上時, 必須提供與網路上其他資料集合的連結。

Open API (Application Programming Interface)

- API就是程式之間的互動方式，可以讓各自獨立的程式於運行的過程中達到相互溝通、協同作業的管道。
- 包含：函式間如何呼叫、如何傳送需求參數、如何解讀回應資料等等。
- 在Open Data Service 中的Open API，意指透過API 提供其他系統方便引用資料的方式



漢籍數位人文資源開放資料服務

- 漢籍數位人文資源，簡單分類為：「名錄型數位資源」、「全文型數位資源」與「純數位研究工具」，等三個類型
- 名錄型數位資源：
 1. 提供人物、地點、組織、家族或重要實體的相關資料。
 2. 此類資源所收錄的物件資料，通常可以看作為一筆具有多個屬性值的資料序列。
 3. 而資源內容除了紀錄各物件的基本屬性資料之外，有時更重視紀錄物件之間的關係。

	中國歷代人物 傳記資料庫 (CBDB)	臺灣歷史人物 傳記資料庫 (TBDB)	台灣傳記知識 本體資料庫 (TBIO)	法鼓佛學 規範資料庫	中央研究院數位 文化中心鏈結開 放資料平台
提供原始資料 下載	Yes	No	No	Yes	Yes
下載資料格式	SQL lite, ACCESS	N/A	N/A	XML, SQL, RDF	RDF
更新頻率	約1年1次	N/A	N/A	約每月1次	無資訊
提供鏈結數據 (LOD)	Yes	未來有機會 提供	Yes	Yes	Yes
Open API 數目	1 個	2 個	0	3個	0
資料開放授權	創用CC授權 BY-NC-SA 4.0	未提供說明	未提供說明	創用CC 授權 BY-NC-SA 2. 5	創用 CC 姓名標 示4.0 國際授權 條款授權

中國歷代人物傳記資料庫 (CBDB)

開放資料下載

- CBDB主要以單機離線版資料庫的方式來提供下載使用。
- 其資料庫無限制讀取權限或也無加密，可以輕易匯出所有資料內容。
- 不過其資料庫使用Microsoft Access軟體製成，為非開放格式。因此CBDB另外提供了開放式資料庫SQLite格式的匯出檔案

CBDB單機版

中國歷代人物傳記資料庫 (CBDB) 離線版本記錄了 471,844 位男性與女性人物。以下是Microsoft Access及SQLite 格式的最新版及下載與安裝說明:

Microsoft Access 格式最新版

[CBDB_bc_20210525_build20210531.7z](#)

[客製化下載鏈接](#) (僅供有經驗的用戶下載，如果您不知道該鏈接的具體功能，請不要透過此鏈接下載)

使用 Access 版本前，請務必閱讀[下載與安裝手冊](#)。

在下載時您如果遇到困難，請聯絡 [hongsuwan\(at\)fas.harvard.edu](mailto:hongsuwan(at)fas.harvard.edu)

SQLite 格式最新版

[CBDB_20210525.7z](#)

Mac 版人名詞典最新版

[cbdb dictionary folder.zip](#)

<https://projects.iq.harvard.edu/chinesecbdb/>下載CBDB單機版

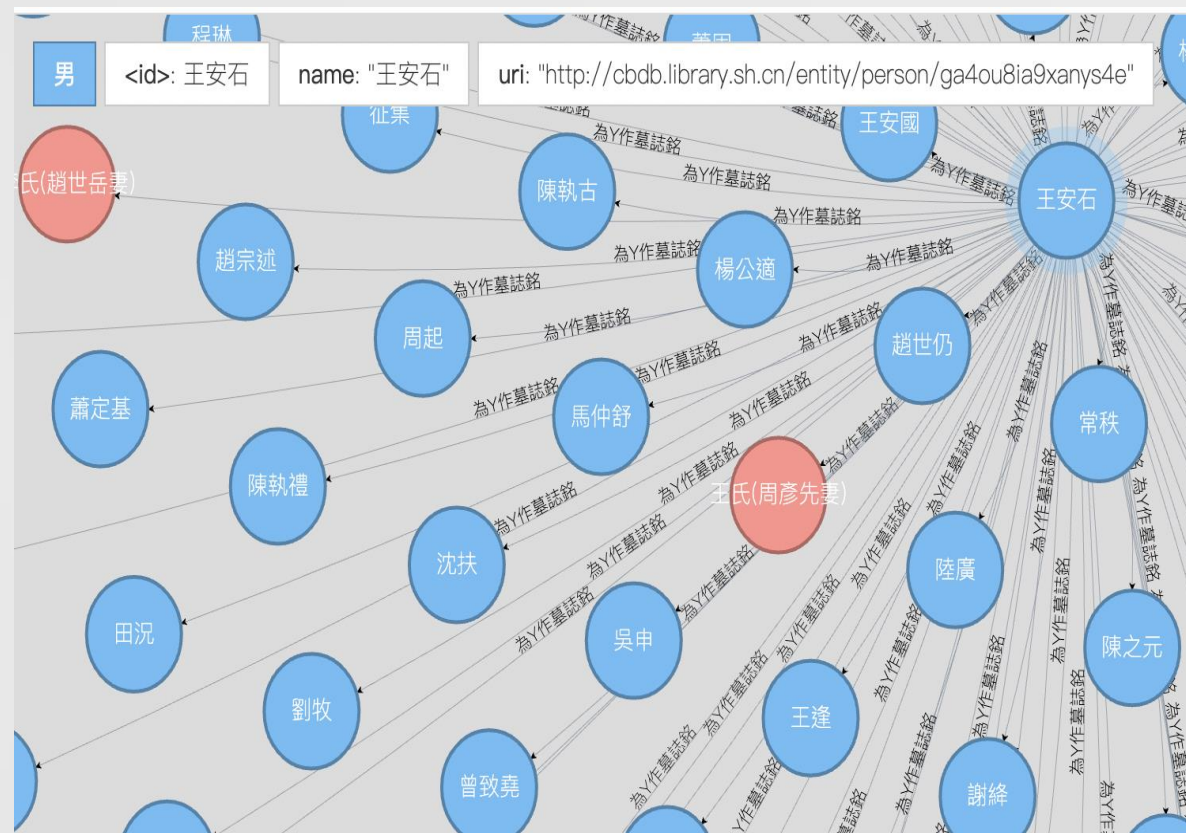
中國歷代人物傳記資料庫 (CBDB)

開放鏈結資料

中國歷代人物傳記資料庫(CBDB)

Query ✕ +

```
1 PREFIX foaf: <http://xmlns.com/foaf/0.1/>
2 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
3 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
4 PREFIX shlib: <http://www.library.sh.cn/ontology/>
5 SELECT ?age
6 FROM <http://lod.library.sh.cn/graph/cbdb>
7 WHERE {
8   ?s shlib:deathAge ?age.
9   filter (?age != '490' && ?age != '0').
10 } order by asc(xsd:int(?age))
```



- 於2018年時，CBDB與上海圖書館合作，建置以鏈結數據服務為基底的服務網站。
- 「CBDB 關聯數據」(<https://cbdb.library.sh.cn>)
- 該鏈結數據服務網站除提供線上查詢CBDB資料的界面外，也提供SPARQL的查詢與知識圖譜查詢界面。

中國歷代人物傳記資料庫 (CBDB)

Open API 服務

- CBDB 提供公開的WEB API 服務。
- 該API可以透過網址呼叫取得相關人物資料內容。
- CBDB所提供的API基本上只包含一個服務網址與兩種應用方式。

CBDB API

CBDB 現在為其他資料庫提供API服務，讓任何資料庫都可以取用CBDB 人物傳記，並於自己的資料庫呈現出來。CBDB API 現在支持兩種查詢方式：

1. 用人物ID查詢 (CBDB ID)
2. 用人名查詢 (漢字或拼音)

如何呼叫CBDB API

例：為王安石創建API Call

以程式送出以下http請求：

1. 用人物ID查詢 (用CBDB ID取得人物傳記)：

<https://cbdb.fas.harvard.edu/cbdbapi/person.php?id=1762> (王安石的ID)

2. 用人名查詢 (漢字或拼音)：

<https://cbdb.fas.harvard.edu/cbdbapi/person.php?name=王安石>

<https://cbdb.fas.harvard.edu/cbdbapi/person.php?name=Wang Anshi>

資料來源：<https://projects.iq.harvard.edu/chinesecbdb/cbdb-api>

法鼓文理學院 佛學規範資料庫 原始資料下載

- 該資料庫提供佛學研究相關之**人物、地點、時間、經典目錄**等基本資料查詢。
- (http://authority.dila.edu.tw/docs/open_content/download.php)
- 人物與地點資料庫的匯出格式為符合TEI格式的**XML檔案** 與 **RDF格式**資料，而時間資料庫則提供**SQL**資料格式匯出資料
- XML/SQL 資料每月更新
RDF 資料 2020/12 日釋出

[Introduction](#) [Entity Relations](#) [Table Schemas](#) [Notes](#) [Download](#)

Person Authority

Current Version **June, 2021**

The zip-archive provided here includes the data of the person authority database in XML/TEI.

 [authority_person.2021-06.zip](#) [7.13mb, 2021-06-01 01:06]

 [authority_person_rdf.2020-12.rdf](#) [58.57mb, 2020-12-04 12:06]

This data is also hosted in a public repository at github. You can find the repository [here](#).

Place Authority

Current Version **June, 2021**

The zip-archive provided here includes those entries of the place authority database that were created by DDBC (c.15.000). The archive is distributed as XML/TEI file. The complete database is accessible through the online interface and API. It contains 38.000 more entries that were provided by the **Center of Geographic Information Science** at Academia Sinica, Taipei.

 [authority_place.2021-06.zip](#) [2.19mb, 2021-06-01 01:07]

 [authority_place_rdf.2020-12.rdf](#) [38.34mb, 2020-12-04 12:06]

法鼓文理學院 佛學規範資料庫

Open API 服務

- 該資料庫提供Web API界面，可藉以取得資料。
- 而該資料庫針對人物、時間、地點等資料庫內容，都分別提供一隻獨立的API。

Query for Person entity

URI format:

<http://authority.dila.edu.tw/webwidget/getAuthorityData.php?type=person&id=<Auth>

Specification:

Argument	Value(s)	Description
type	"person"	Specifies the query type
id	<AuthorityID> <string>	The DDBC Authority ID of the entity. Alternatively, a string may be specified which will serve as a query term.
jsoncallback	<string>	Required by third-party JavaScript libraries (e.g. EXT Js) to perform ajax functions. The value of jsoncallback has no effect on query processing, but will be attached to the beginning of the returned json string.

<http://authority.dila.edu.tw/docs/services/>

法鼓文理學院 佛學規範資料庫 鏈結資料服務

- 「DILA 鏈結開放資料平台」的**開放鏈結數據服務**
- 該系統主要功能：資料搜尋、資料集下載與SPARQL查詢與資料視覺化。

■ 查詢範例

1. 查詢常名為"金總持"的人物相關資料

2. 查詢別名為"三藏"的人物相關資料

3. 查詢僧人鳩摩羅什的別名

4. 查詢人名規範資料ID為：A001204的中文人名與相關說明

5. 查詢玄奘法師的所有弟子

SPARQL QUERY 語意輸入框

namespace

PREFIX dila: < http://purl.dila.edu.tw/resource/>

PREFIX skos: < http://www.w3.org/2004/02/skos/core#>

PREFIX bdo: < http://purl.bdrc.io/ontology/core/>

PREFIX rdf: < http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX rdfs: < http://www.w3.org/2000/01/rdf-schema#>

SELECT ?subject ?predicate ?object

WHERE {

?subject ?predicate ?object.

?subject skos:prefLabel "金總持"@zh-Hant.}

1) 檢索有中文常名欄位(skos:prefLabel)資訊的資料，其資料值為"金總持"

2) 列出符合上述檢索條件者的所有相關資訊

reset

submit

中央研究院數位文化中心鏈結開放資料平台

<https://data.ascdc.tw/>

中央研究院數位文化中心 鏈結開放資料平台

檢索本鏈結開放資料平台已收錄並轉置為LOD的**120,359**筆條目，資料類型包含圖片、影像、聲音、文字，類別則涵蓋生物, 人類學, 宗教, 藝術, 影音, 歷史……等等。

請輸入物件名稱或關鍵字

全部類別

search

- 該資料庫已經完成**13個資料集**，總和超過**60萬筆資料**的鏈結開放資料的資料轉置工作。內容含括生物學、人類學、宗教、藝術、影音、歷史等面向，十分精彩多元。
- 所有資料庫內容，皆可以完整打包下載，其**下載格式**也包含JSON、RDF、Turtle、N-Triples 等各種鏈結資料相關格式下載。建立了**開放性的SPARQL Endpoint**，並提供許多範例來引導使用者撰寫SPARQL查詢。

「名錄型數位資源」之資料開放服務特性 小結

- 幾乎每個資源都(預計)提供以**鏈結資料格式**為基礎的資料開放服務。
- 名錄型資料庫的內容多半以提供人物、時間、地點、組織、書目、實體物件等實體的概念性資料為主，其資源內容的不僅相當結構化，也十分著重於**實體之間關係**。
- 而另一方面，我們也發現到一些專案項目以**十分積極的方式在經營資料開放服務**，如**CBDB與佛學規範資料庫**，兩者皆幾乎提供了上述所有資料分享的可能性。或許這樣的策略，是導致我們現在能在許多數位工具中，都能看見這些資料作為重要輔助內容的原因。

「全文型數位資源」開放資料服務

- 所謂的全文型漢籍數位資源，泛指內容主要由漢文典籍全文資料所組成的資源。
- 全文資料通常具有較為複雜的階層式結構。例如：
 - 一篇文章可能向上匯總成為書冊，而書冊向上匯總成為叢書。
 - 一篇文章的內容包含有多層的章節，而章節之下還可分為段落，段落內包含多個句子等等。
- 這樣的差別導致全文型與名錄型數位資源在資料分享的策略上，有著相當不同的選擇。



CBETA Online

漢リポ Kanseki Repository

百諸
家子

中國哲學書電子化計劃

	中國哲學書電子化計劃 (Chinese Text Project)	Kanripō (Kanseki Repository)	CBETA 數位人文研究平台
提供完整原始資料載	Yes	Yes	Yes
開放資料格式	純文字	純文字	TEI/XML
開放資料取得方式	透過網站功能下載	放置於GitHub中	放置於GitHub中
更新頻率	無特定說明 隨時可能更新	批次性更新 約每年更新一次	約三個月會更新一次
提供鍊結數據下載	No	No	No
開放式資料存取API 數目	約12個	無	約30個
資料開放授權	自訂授權，完全開放給 個人或學術使用	CC BY-SA 4.0	CC BY-NC-SA 3.0

以GitHub 分享資料

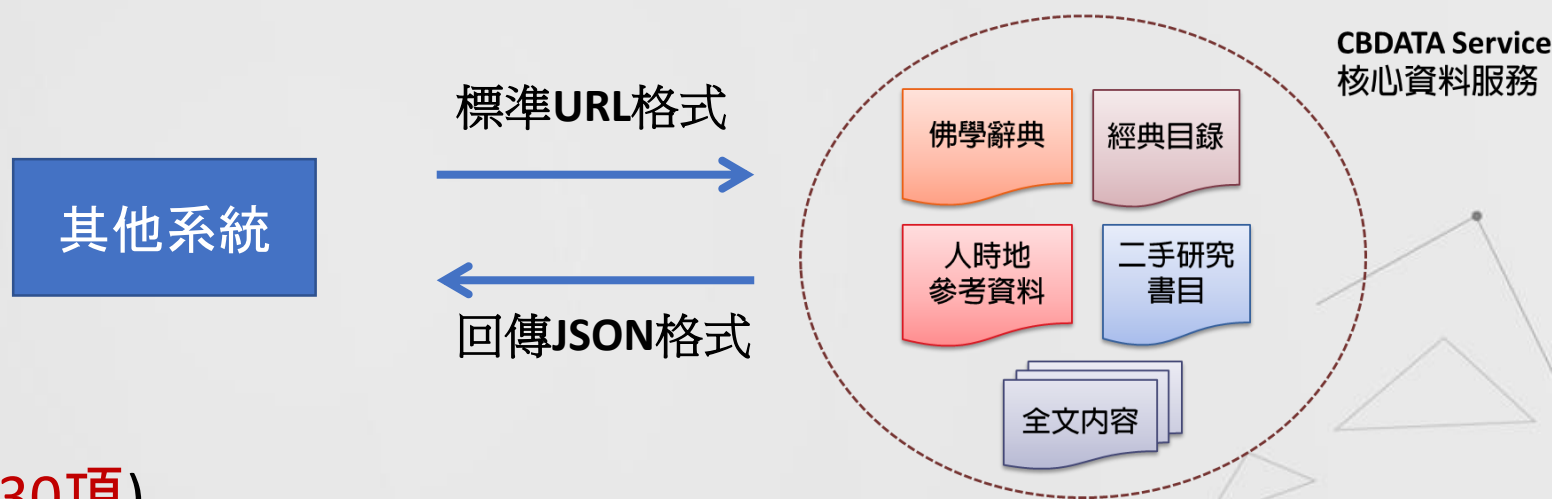
CBETA GitHub Repo

- CBETA 完整公開XML原始資料，使用者可以透過以下方式取得完整資料：
 - 由CBETA光碟內複製
 - 於GitHub網頁公開下載(TEI 格式資料)
 - <https://github.com/cbeta-git/xml-p5a>
 - 還有HTML格式與純文字格式資料，都可以下載到
- 資料採用 Creative Commons License 釋出



CBETA 數位平台 Open API

- 定義了一套完整的API 開放給所有人使用。
- 說明網址：<http://CBdata.dila.edu.tw>



□ API 內容 (約30項)

- ▣ 部類目錄、經目搜尋、以作譯者、時間條件搜尋目錄
- ▣ 取得某部典籍的Metadata、典籍目次、內容。
- ▣ 全文搜尋結果。

CBETA 數位平台 Open API

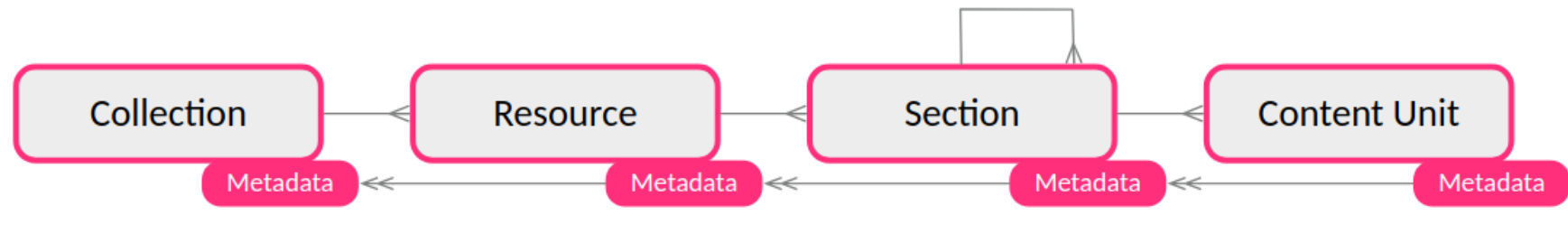
RISE and SHINE API 標準

- CBETA 的開放 API 中，也實作了文獻資料交換的國際標準 RISE and SHINE API (<https://rise.mpiwg-berlin.mpg.de/>)



Implementing SHINE API

API Data Structure



Collection	Text	Fascicles	Content
Taishō Tripiṭaka Shinsan Zokuzōkyō Chinese translations of the Pali Canon	T0001 長阿含經 T0002 七佛經 T0003 毘婆尸佛經	T0001 長阿含經 卷1 T0001 長阿含經 卷2 T0001 長阿含經 卷3	則導之以契經；演 幽微，則辨之以法 相。然則三藏之作 也，本於殊應，會 之有宗，則異途同 趣矣。禁律，律藏 也，

「全文型數位資源」之資料開放服務特性

- 由於全文資料結構複雜、內容多、使用需求集中於取用文字內容。使此類型的資源項目較少提供鏈結資料格式的內容。
- 全文型數位資源，較傾向於提供完整原始資料開放下載，並考慮如何有讓資料能被有效取用。而其中一種常用的方式，便是透過GitHub服務公開原始資料。
- 另外，全文數位資源，也十分傾向於提供資料存取服務的API。
- 相較於名錄型數位資源，全文型數位資源的Open API，通常會有數量較為龐大的API函式。

「純數位研究工具」開放資料服務

- 定位為純粹輔助人文研究的數位工具。
- 這類的數位工具並不預先提供預先處理好的資料，而是讓使用者上傳數位文件或抓取由系統中表列的開放資料，然後進行後續分析運用。
- 因此這些數位分析工具，通常扮演的是「資料消費者」的角色。透過了解這些工具如何應用上述「資料提供者」的開放資料服務

</> MARKUS



</> MARKUS

- 在MARKUS系統中，有兩個運用到其他數位資源所提供的開放資料服務。
- 首先，第一部份是其內建的地名、人名、時間、官職等資料庫的清單內容。
- 而其取用的方式，便是透過數位資源的提供的開放完整原始資料檔案，成為MARKUS比對清單的一部分。

automated markup options

- ☒ paragraph division (default one line of blank space)
- ☐ full name (last name + first name, 姓名) dynasties selection
- ☐ partial name (alternative names, 別名)
- ☒ date (日期)
- ☒ place name (地名) ☐ including Taiwan historical place name (Alpha)
- ☒ official title (官名)
- ☐ enable muzhiming (墓誌銘) experimental
- ☐ person authority database for Buddhist Studies (法鼓佛學人名規範資料)
- ☐ place authority database for Buddhist Studies (法鼓佛學地名規範資料)
- ☐ glossaries for Buddhist Studies (法鼓佛學詞彙)
- ☐ Korean historical figures (韓國歷代人物)
- ☐ Bureaucratic offices of the Joseon dynasty (朝鮮時期官職名)
- ☐ Korean place names (韓國歷代地名)
- ☐ Korean books (韓國歷代書名)

</> MARKUS

- 與其他全文系統所提供的資料 API 相**界**接。
- 提供使用者**抓取由其他開放資料來源的全文資料內容**，進行資料標記。
- MARKUS 透過 **SHINE API** 規範，平台中，連接 **CBETA 資料與 Kanripo** 的資料。
- MARKUS **應用 CTEXT API**，抓取所需的 CTEXT 資料。

MARKUS open access text (漢リポ Kanseki Repository), powered by RISE and SHINE

Search text title
虎

1 2 3

上清太 上元始 耀光金 虎鳳文 章寶經 FETCH	九還七 返龍虎 金丹析 理真訣 FETCH	古文龍 虎上經 註 FETCH	古文龍 虎經註 疏 FETCH	大還丹 金虎白 龍論 FETCH	洞真太 上神虎 玉經 FETCH	洞真太 上神虎 隱文 FETCH
--	---------------------------------------	------------------------------	------------------------------	-------------------------------	-------------------------------	-------------------------------

MARKUS open access text (Ctext)

Search text title
Search book title

1 2 3 4 5 6 7 8 9

莊子 ctp:zhuangzi 續古逸叢書 FIND SIMILAR CTEXT MARKUS +	淮南子 ctp:huainanzi 四部叢刊初編，景上海涵芬樓藏景鈔北宋本 FIND SIMILAR CTEXT MARKUS +
--	--

- 資料由使用者上傳。可以是純文字檔案或由MARKUS 檔案匯入之外。
- Docusky也連接多種開放資料來源連接：
 - CBETA API，連接CBETA 文獻資料。
 - 以及透過CTEXT API取得CTEXT資料，
 - 針對Kanripo的資料格式爬梳，提供Kanripo的資料匯入功能。



The screenshot displays the DocuSky interface for connecting external data sources. It features a list of sources with expand/collapse icons (+/-) and a detailed view for the '中國哲學書電子化計劃' (Chinese Philosophy Electronic Text Project).

Source	Action
CBETA	+
CTEXT	-
開發者 德龍 教授	
中國哲學書電子化計劃是一個線上開放電子圖書館，為中外學者提供中國歷代傳世文獻，力圖超越印刷媒體限制，通過電子科技探索新方式與古代文獻進行溝通。收藏的文本已超過三萬部著作，並有五十億字之多，故為歷代中文文獻資料庫最大者。	
使用說明 進入工具	
Kanripo	+
WikiSource	NEW +
CBDB	+

- 另一個Docusky 使用開放資料的方式，則是發生在系統所提供的閱讀界面中。
- Docusky 透過DILA 規範資料庫與CBDB資料庫的API，動態取得文章內標記實體的參考資訊。

1. 釋玄奘

釋玄奘。本名禪。姓陳氏。漢太丘仲弓後也。子孫徙於河南。故又為洛州緱氏。

士。父

素出

之。日

閑缺。

懷。尚

擇。僧

專門受

之。行

理又高

蘊結胸

講肆多

攝論

率多昏

釋玄奘

authorityID	A000294
name	玄奘
class	譯經
gender	1
bornDateBegin	+0602-02-01
bornDateEnd	+0603-02-18
diedDateBegin	+0664-03-10
diedDateEnd	+0664-03-10
	俗姓陳，大慈恩寺僧人，唯識宗創始人，世稱唐三藏。幼隨兄長捷法師住淨土寺，習佛五年。大業八年（612），洛陽度僧破格以沙彌入僧籍。武德五年（622）在成都受具足戒。貞觀年間西

。識者以為

都淨土寺。以

爾卓然梗正不

為兒戲。可謂

槃攝論輪馳林

。時年十五！

道基化開井絡

毘曇論。一聞

披弘唱極繁。

古罕類也。基

益。相與和

住義。中表鍾

眉構。如斯甚

- 可以存取多種開放文本資源
- 關係資料以Linked Open Data 方式紀錄

文本查詢 ▾

上傳與匯入 ▾

上傳.txt格式文本

上傳平台格式文本

上傳MARKUS匯出檔

上傳Excel目錄文本

上傳Excel目錄文本(多選模式)

匯入中研院開放文本(由史語所、台史所、近史所提供)

匯入漢籍授權版部分開放文本(限中研院IP)

匯入CText文本(限中研院IP)

匯入Kanripo文本

匯入開放的文本

	Subject	Predicate	Object
1	賈寶玉	知己	林黛玉
2	賈政	父子	賈寶玉
3	王夫人	母子	賈寶玉
4	賈寶玉	夫妻	薛寶釵
5	賈政	父女	賈元春
6	王夫人	母女	賈元春
7	賈赦	父女	賈迎春
8	賈政	父女	賈探春
9	趙姨娘	母女	賈探春
10	賈珍	兄妹	賈惜春
11	賈璉	夫妻	王熙鳳
12			

小結

- 在上述三個數位研究平台，都相當積極的**致力於運用**各系統的**開放資料**或利用**Open API**來進行系統間的連接。
- 這也讓我們更清楚的看見，在實際的數位人文系統應用中，**開放資料**確實於跨系統的整合協作的實現，扮演相當重要的角色。
- 我們發現，上述三個數位平台，雖然也匯入各資源的開放原始資料的行為，但**似乎更趨向於利用Open API進行系統間的溝通**。



結語

- 總結來說，我們歸納出下面幾個趨勢與未來值得努力的方向。
- **開放資料服務已經是現在進行式**。開放資料服務對於漢籍數位人文資源而言，已不僅只是一個口號，而已經是許多數位人文資源實際進行的工作項目。
- **似乎知名度越高，越為人所使用的資源，其資料開放程度便越加完整**。究竟是因為採取各種開放策略，讓資源容易被使用，而推升了資源的知名度；還是因為資源知名度提昇、吸引了更多使用者後，也連帶產生更多不同需求，而導致需要製作更多開放資料的服務來滿足各種需求，其實際原因尚待釐清。

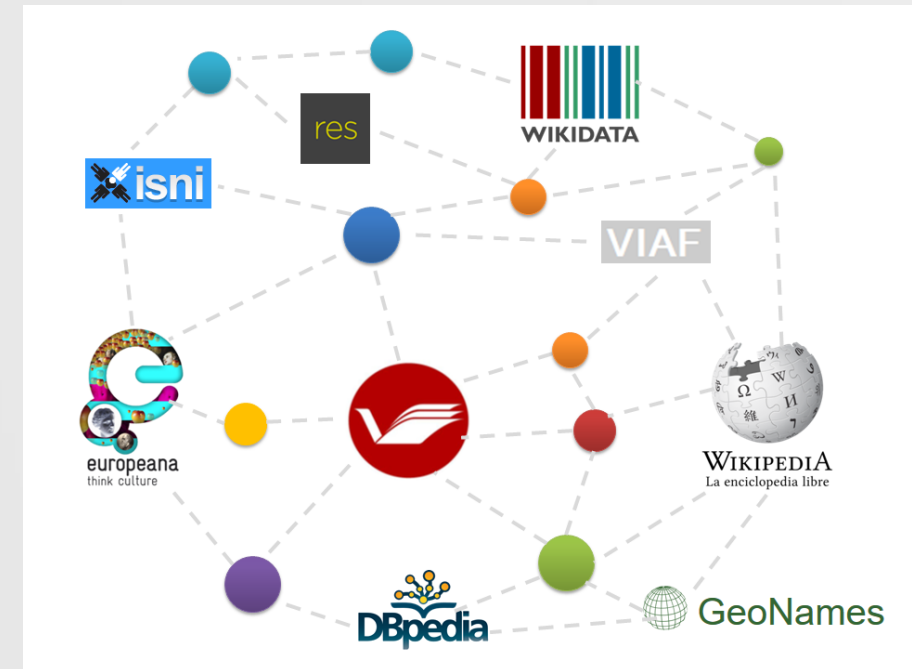
需建立共同的全文資料交換API的標準

- 數位平台十分熱衷於對於透過各資源的Open API取得所需的資料。但目前各數位資源所提供的API仍具有相當高的異質性。可能導致後續整合上與維護上的負擔。
- 因此為了促進工具與資源間能夠更方便的達到資料交換與流程整合的目的，應朝向開發全文資料交換的標準API。
- 一個較為接近此方向的實際成果，
前述所提及的 RISE & SHINE API。



鏈結資料的應用仍然十分缺乏

- 根據本研究結果：
- **名錄型資料庫**似乎都相當熱衷於提供以**鏈結資料**為基礎製作之開放資料服務。
- 作為資料消費者的數位人文平台卻似乎興趣缺缺。僅有中研院數位人文平台的**關係資料**採用**類鏈結資料概念**。
- 目前尚未出現真正深入運用鏈結開放資料的應用。



數位人文的資源整合仍待努力

- 本文所提及的具有資料開放服務的數位人文資源，實際上還是只佔人文相關數位資料庫建設成果的一小部份。
- 許多國內重點單位過往所建立的數位資源，多半欠缺在資料開放方面的考量。
- 若能有更多相關單位，願意投入為資料庫製作開放資料的服務，我們應該可以期待我國的數位人文發展能快速的提昇到另一個水平。