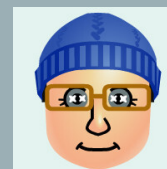




2020/6/26

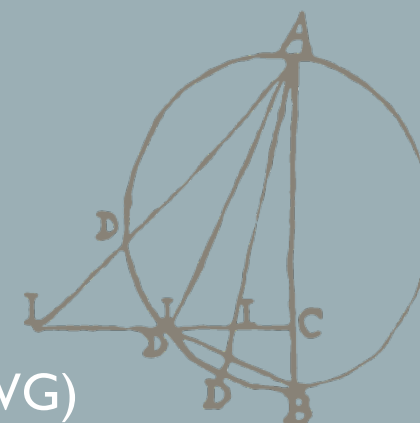
# Exploring HYL's Rare Local Gazetteers with LoGaRT

Shih-Pei Chen, Sean Wang, Calvin Yeh



Max Planck Institute for the History of Science (MPIWG)

Department III “Artefacts, Action and Knowledge”

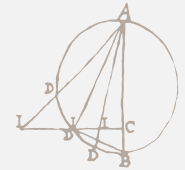




...

Books (18)

[China] : Nei fu



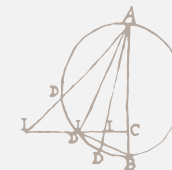
# HYL's Rare Local Gazetteers

- 御製大明一統志序
- 朕惟我
- 太祖高皇帝受
- 天明命混一天下薄海內外悉入
- 版圖蓋自唐虞三代下及漢唐
- 以來一統之盛蔑以加矣顧惟
- 覆載之內古今已然之跡精粗

(seq.

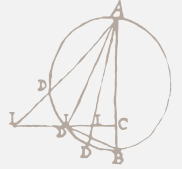






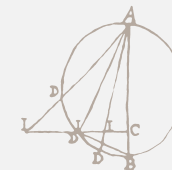
## Today's outline

- Why are we developing LoGaRT?
  - What does LoGaRT do that other LG databases don't
  - Core methodology behind LoGaRT
- Digitizing the HYL local gazetteers (by Sean)
- A glimpse to LoGaRT
  - Walking through its features and a little bit hands-on
  - How LoGaRT helps to explore the HYL LG collection
- Concluding remarks, Q&A



## What are the Local Gazetteers

- Main genre for **recording “local knowledge”** in historical China established or stabilized since Southern Song (12c)
  - Cumulative records about a territorial unit. Mostly compiled by a local government. Published in book format. (Dennis, 2015)
- **Local**: what kind of locality?
  - Mostly administration units – counties, prefectures, provinces, state. Also geographical regions or borderlands.
- **Knowledge**: what kind of information? (topics)
  - Mostly for governing purposes: topography and beyond (eg. institutions, population, taxes, biographies, literature, ... )



# Topics: Extensive and Consistent

順天府志目錄	第一卷地理志	金門圖	畿輔圖	分野
沿革	風俗	山川	疆域	形勝
第二卷營建志	城池	公署	學校	壇社
郵舍	寺觀	創造		
第三卷食貨志	戶口	田賦	徭役	
馬政	經費	物產		
第四卷政事志	歷官	職掌	名宦	
祠典	武備			
第五卷人物志	選舉	功烈	節孝	
鄉賢	隱逸	流寓		
仙釋				
第六卷藝文志	碑刻	題咏		

## Juan 1: **Geography / topography**

Map of Ginmen. Map of Gifu. Field division.

History of establishment. Territory.

Geographical features.

**Local customs. Mountains & rivers.**

Monuments.

## Juan 2: **Infrastructure**

City's infrastructure. **Bureaus. Schools. Altars.**

Postal stations. **Temples.**

## Juan 3: Food and Commodities

**Households. Taxes.** Public services.

Horses. Budgets. **Local products.**

## Juan 4: Politics

**Local officials.** Responsibilities. Famous officials.

Rituals and ceremonies. Military affairs.

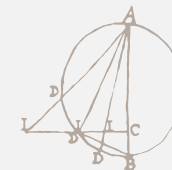
## Juan 5: **People**

Selection. Merits & martyr. Virtue and filial.

Local gentry. Seclusion. Foreigners. Immortals.

## Juan 6: **Literature**

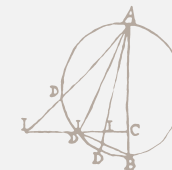
Steles. Poems and Songs.



## Unique features of LGs (I)

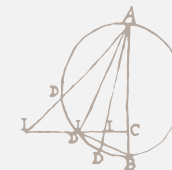
- Many
- Wide geographical coverage: Well-populated and less-well-populated regions in historical China. Several editions for a place.
- Long period: 12c-20c
- A big portion of LGs ever compiled still exist today. A large subset has been digitized as digital texts.
  - 20,000 LGs were compiled in historical China\*
  - About 10,000 distinct titles are extant
  - 5K-7K ( $> 1/4$ ) distinct titles have been digitized as searchable digital texts





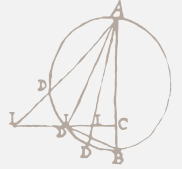
## Unique features of LGs (2)

- Full of data
  - LGs are data collection projects. Each LG is a small database by itself.
  - Extensive and (somewhat) consistent topics, across regions and time periods
  - Structural: topics & lists



## Unique features of LGs (3)

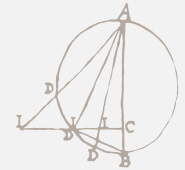
- **Quantity** matters
  - Statistically sufficient to use all the digital LGs (or a big enough subset) to represent the overall patterns of LGs ever existed
  - Running a full text search would give a relatively accurate approximation of the term's overall distribution/pattern in all the historical LGs ever existed
  - Data collected on a certain topic from a big-enough subset would be statistically representative



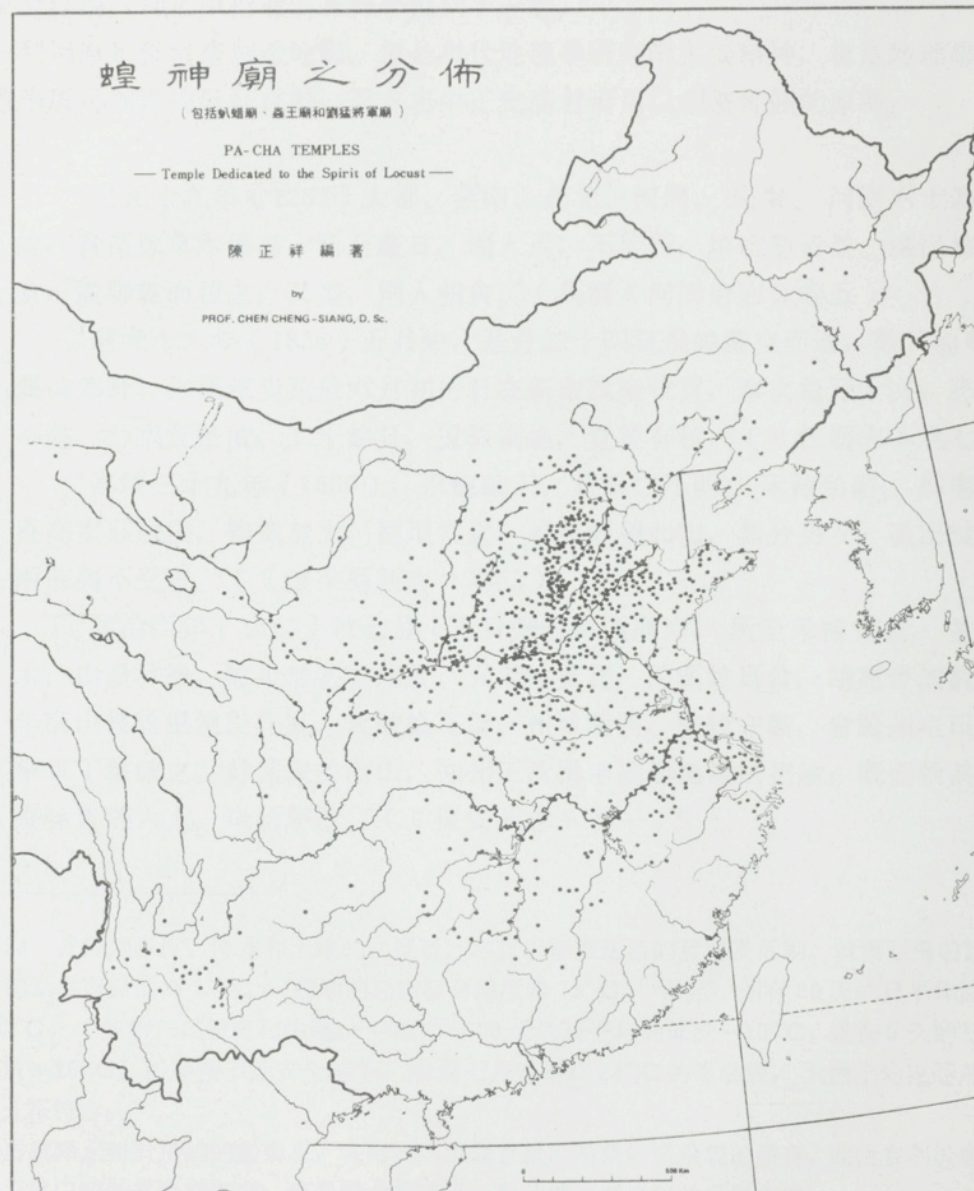
## Core Methodology behind LoGaRT

- Treating the genre as a conceptual “database” for historical inquiry
  - Using the genre to answer questions for the genre / historical China, rather than just using individual LGs
  - To **scale up** research, not necessarily bounded by time, space, and individual efforts
  - LG Working Group: to study the nature of local knowledge and the role of LG in Imperial China
- Only if we have proper digital tools => LoGaRT

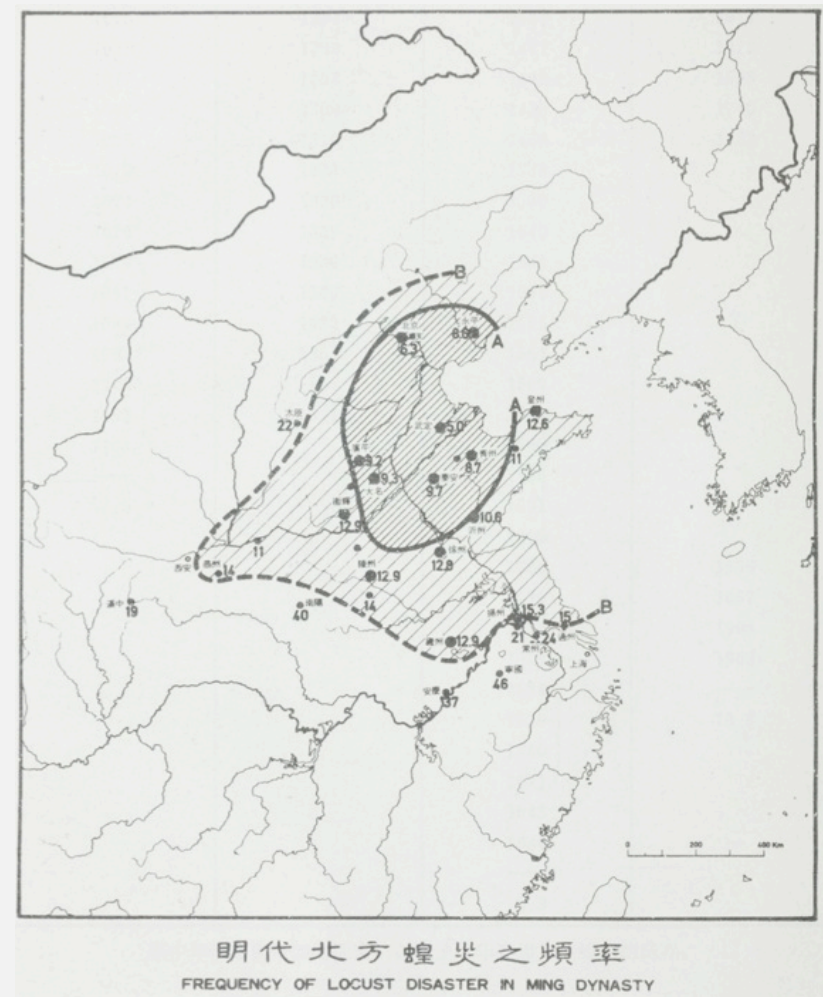
# Example: Maps for Locust Disasters



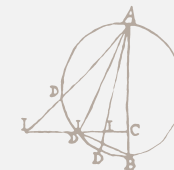
Geographical distribution of locust temples



Frequency of locust disasters in Ming

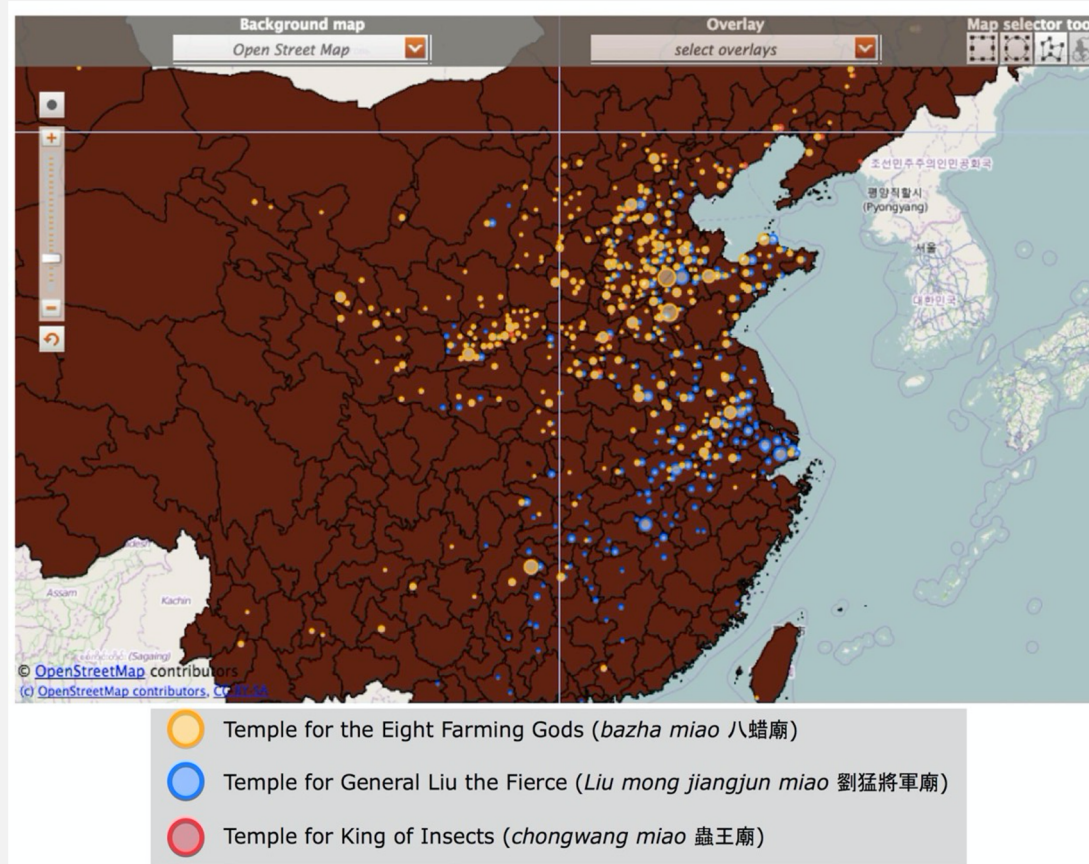


See: Chen Cheng-siang 陳正祥. 1983. *Zhongguo wenhua dili* (Cultural Geography of China). Beijing: Sanlian shu

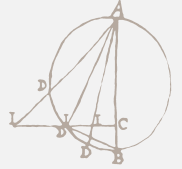


# Model Applicable to Many Questions

- With the digital databases on LGs, we can repeat what he did. But how easy / fast?
- Recreating Chen's method with digital technologies
  - Repeat Chen's work in mins
  - Geo-distribution almost the same
  - Everyone can ask different questions – Page Text Search in LoGaRT
  - Results can be analysed in different ways

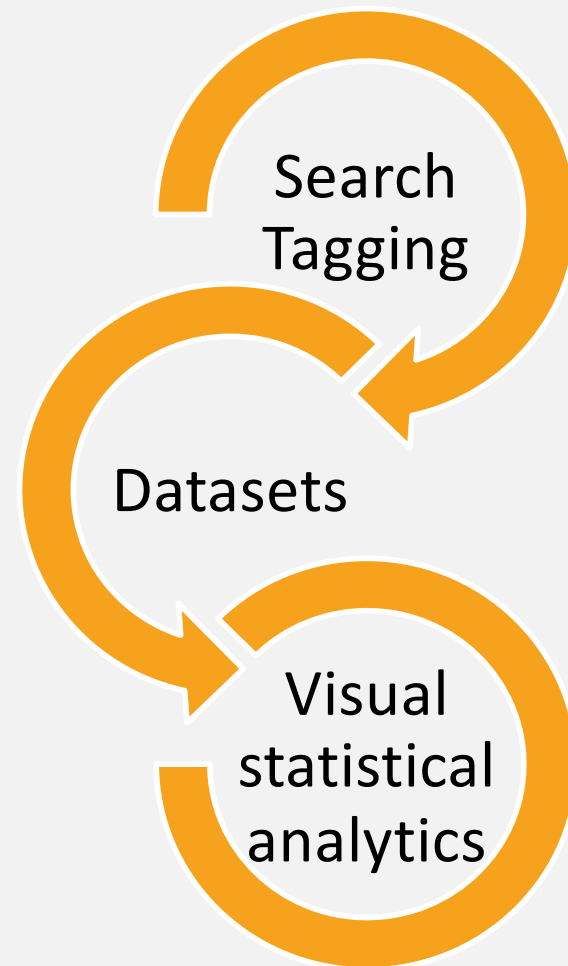


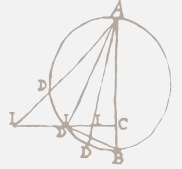




## Practically what you can do with LoGaRT

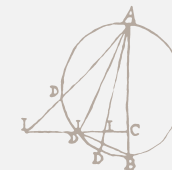
- Recreate “models” that can take input from scholars by digital technologies
- I) Data collection & analyses
  - Data collection (via full text search)
  - Obtain datasets
  - Visualize the dataset with visual analytic methods (eg. maps), and identify patterns
  - => Collective analysis





## Practically what you can do with LoGaRT

- 2) Visual / statistic analytics: to quickly grasp **overviews of a collection / search results / datasets**
  - Help us to understand the composition
  - Help to verify / examine our assumptions toward the genre
- 3) **Quick switch between close / distant reading**, help to quickly adjust one's assumption and dig into the contents through multi-perspectives



## Practically what you can do with LoGaRT

- 4) Takes advantage of the **structures** embedded in LGs: chapters/sections, lists
  - Section search: cut through the database and get sections on particular topics
  - Text-tagging: convert lists into computer-operable data tables



# The Full-text Production Process for HYL Local Gazetteers

Sean Wang

Project Manager, MPIWG

# Multi-step process for full-text production

## 1. Title selection

2019 HYL typing list (master, sorted by priority) ☆ 📁 🔄

File Edit View Insert Format Data Tools Add-ons Help Accessibility Last edit was on June 11, 2019

100% \$ % .0 .00 123 Arial 11 B I S A

Worksheet

	A	C	D	E	F	G	
1	Worksheet	Is Dup.	Scanned?	Call #	8-digit Hollis #	書名	
51	[2] "1065 与爱如生元数据不重复计算结果"		Yes	T 3035.27 021.7	08094479	齊雲山志	<a href="http://id.lib">http://id.lib</a>
52	[2] "1065 与爱如生元数据不重复计算结果"		Yes	T 3035.29 7012	08094491	明州阿育王山志	<a href="http://id.lib">http://id.lib</a>
53	[2] "1065 与爱如生元数据不重复计算结果"		Yes	TNC 3027 4478d	08103464	大明一統志	<a href="http://id.lib">http://id.lib</a>
54	[2] "1065 与爱如生元数据不重复计算结果"		Yes	TNC 3035.29 867.7	08104997	普陀山志	<a href="http://id.lib">http://id.lib</a>
55	[2] "1065 与爱如生元数据不重复计算结果"		Yes	TNC 3035.31 145.75	08105035	武夷志略	<a href="http://id.lib">http://id.lib</a>
56	[2] "1065 与爱如生元数据不重复计算结果"		Yes	TNC 3079.1 6903	08106838	冰嶺紀程	<a href="http://id.lib">http://id.lib</a>
57	[2] "1065 与爱如生元数据不重复计算结果"		Yes	T 3544 4193	08988967	太平省通志	<a href="http://id.lib">http://id.lib</a>
58	[3] "49-1 出版商选目 (T) "		Yes	T 3185 7332.81	07465015	〔康熙〕隨州志四卷	<a href="http://id.lib">http://id.lib</a>
59	[3] "49-1 出版商选目 (T) "		Yes	T 3200 3627.83	07467070	〔乾隆〕潛山縣志二十四卷首一卷	<a href="http://id.lib">http://id.lib</a>
60	[3] "49-1 出版商选目 (T) "		Yes	T 3210 2773.82	07472253	〔康熙？雍正？〕山陰縣志三十八卷	<a href="http://id.lib">http://id.lib</a>
61	[3] "49-1 出版商选目 (T) "		Yes	T 3210 3530.833	07473482	〔乾隆〕寧志餘聞八卷	<a href="http://id.lib">http://id.lib</a>
62	[3] "49-1 出版商选目 (T) "	Yes_A	Yes	T 3150 1428.81	07482541	〔康熙〕平順縣志十卷	<a href="http://id.lib">http://id.lib</a>
63	[3] "49-1 出版商选目 (T) "	Yes_A	Yes	T 3150 2532.81	07486999	〔康熙〕絳州志四卷	<a href="http://id.lib">http://id.lib</a>
64	[3] "49-1 出版商选目 (T) "		Yes	T 3150 4386.82	07487091	〔乾隆四年〕太谷縣志八卷	<a href="http://id.lib">http://id.lib</a>
65	[3] "49-1 出版商选目 (T) "		Yes	T 3150 7422.83a	07488845	〔乾隆五年〕陵川縣志二十八卷	<a href="http://id.lib">http://id.lib</a>
66	[3] "49-1 出版商选目 (T) "		Yes	T 3145 2369.80	07493620	〔順治〕息縣志十卷	<a href="http://id.lib">http://id.lib</a>
67	[3] "49-1 出版商选目 (T) "		Yes	T 3133 0814.81	08026931	〔康熙〕廣平府志：二〇卷	<a href="http://id.lib">http://id.lib</a>
68	[3] "49-1 出版商选目 (T) "		Yes	T 3134 3930.83	08032949	〔乾隆〕直隸深州總志二十卷首一卷	<a href="http://id.lib">http://id.lib</a>
69	[3] "49-1 出版商选目 (T) "		Yes	TNC 3271 8202.74	08105120	〔萬曆〕朔方新志五卷	<a href="http://id.lib">http://id.lib</a>
70	[5] 初選十五種		Yes	T 3159 5232.83	07453876	〔乾隆〕重修肅州新志三十卷	<a href="http://id.lib">http://id.lib</a>
71	[5] 初選十五種		Yes	T 3128 82	08026777	〔雍正〕續補通志一百二十卷	<a href="http://id.lib">http://id.lib</a>



# Multi-step process for full-text production


1. Title selection
2. Develop production specifications
  1. For the main body texts
  2. For tables of contents

Main body:

<https://github.com/SeanyWang/desp-ecs/blob/master/LGspecs.md>

TOC:

<https://github.com/SeanyWang/desp-ecs/blob/master/TOCspecs.md>



```
1 <page seq="x">
2   <ls>
3     衛河道
4     小西門城角起至三皇廟南扇止 長一百二十九步係左位河道
5     [...]
6     杜家墳起至下八里堂止 長一里係德州河道
7   </ls>
8   <pb />
9   以上截斷分管本州爭管河道一十五里零五十四步北接吳橋地方
10  <img>
11    □
12    <char no="1" />
13    駟 桑園 第□ 墩台 □□ □□屯 第三屯 □□ □□ □□
14  </img>
15 </page>
16 <page seq="y">
17   <img>
18     第□屯 □□ □□ □□ 新□ □家□ □□
19   </img>
20   <pb />
21   以上截斷分管本州爭管河道一十五里零 | 西步北接□橋縣地方
22   <img>
23     良店駟 桑園鎮 第六屯 德州 第三屯 八□堂 □米□ 第一屯 浮橋口 老君堂 白□ 四
24   </img>
25 </page>
```

# Multi-step process for full-text production

1. Title selection
2. Develop production specifications
  1. For the main body texts
  2. For tables of contents
3. Main production

Some statistics:

- 1 year of production (2018-19), followed by 6 months of QC
- >99% accuracy rate
- 122 million characters typed
- 411 titles of local gazetteers, from 1,520 image viewer catalog records
- \$744,000 (funded by the Max Planck Society and the Chiang Ching-kuo Foundation)

Project page: <http://bit.ly/MPIWG-HYL-LG>

# Multi-step process for full-text production

1. Title selection
2. Develop production specifications
  1. For the main body texts
  2. For tables of contents
3. Main production
4. Post-production processing

Project page: <http://bit.ly/MPIWG-HYL-LG>



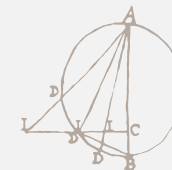


# A Glimpse to LoGaRT

Walking through its features

Hands-on

How LoGaRT helps to explore the HYL LG collection

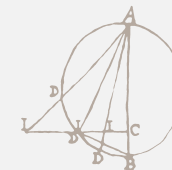


## Before we start...

- First public release of LoGaRT (rather than in-house)
  - Technical infrastructure might need adjustment to support large number of concurrent users
- Watch me demo & short breaks for trying-it-out. Slides will be shared as reference points.
- Questions can be raised during demo **in the Chat window**
- Online support (via Chat):
  - **Calvin Yeh**: “Chief” developer for LoGaRT
  - **Sean Wang**: Project manager



# Technical Components (User Interfaces) of LoGaRT



## Table of Contents (TOC) Editor

book id	book name	level1	level2	period	pages	version	editor	date
16751	順天府志	北京		高曆	1210	2	fxcal	2015-08-13 15:00:59
Show editing records: <a href="#">Show all</a> <a href="#">Close</a>								
1	順天府志序				1 - 10	level 1	<a href="#">delete</a>	<a href="#">+</a>
2	順天府志序				11 - 25	level 1	<a href="#">delete</a>	<a href="#">+</a>
3	順天府志序				26 - 39	level 1	<a href="#">delete</a>	<a href="#">+</a>
4	順天府志目錄				40 - 42	level 1	<a href="#">delete</a>	<a href="#">+</a>
5	第一卷地理志				43 - 145	level 1	<a href="#">delete</a>	<a href="#">+</a>
6	金門圖				40 - 47	level 2	<a href="#">delete</a>	<a href="#">+</a>
7	畿輔圖				48 - 56	level 2	<a href="#">delete</a>	<a href="#">+</a>

## Portal (main UI)

Book Name	Level	Dynasty	Adm.	Years	Edition	Volume	Author	Source	Total Page	TOC Date	Action
沙州府志	33001	甘肅	肅州	1714	1714	新修志	不分卷	(清) 孫承澤	中國方志網	52	<a href="#">edit</a> <a href="#">delete</a>
西安府志	33007	陝西	西安	1077	1077	新修志	不分卷	(清) 孫承澤	中國方志網	629	<a href="#">edit</a> <a href="#">delete</a>
順天府志	34248	浙江	杭州府	1171	1171	清乾隆七年	3卷	(清) 周亮工	中國方志網	170	<a href="#">edit</a> <a href="#">delete</a>
四庫全書	10772	浙江	杭州府	1173	1854	清乾隆七年	12卷	(清) 孫承澤	中國方志網	403	<a href="#">edit</a> <a href="#">delete</a>
廣州府志	10723	浙江	杭州府	1187	1187	清乾隆七年	3卷	(清) 孫承澤	中國方志網	200	<a href="#">edit</a> <a href="#">delete</a>

## Text-tagging Interface

Welcome to Extraction Interface. [Info] book name: 晉安州志, period: 高曆, section id: 5214572, section name: 書籍, branch id: 1441907, file id: 5214572

【88】

- <晉安州學> <書籍> <領降書籍> 易經大全一部
- <晉安州學> <書籍> <領降書籍> 書經一部
- <晉安州學> <書籍> <領降書籍> 詩經一部
- <晉安州學> <書籍> <領降書籍> 春秋一部
- <晉安州學> <書籍> <領降書籍> 禮記大全一部
- <晉安州學> <書籍> <領降書籍> 四書大全一部
- <晉安州學> <書籍> <領降書籍> 性理大全一部
- <晉安州學> <書籍> <領降書籍> 五倫書一部
- <晉安州學> <書籍> <領降書籍> 李順事實一部
- <晉安州學> <書籍> <領降書籍> 為晉陽編一部
- <晉安州學> <書籍> <領降書籍> 佛曲一部
- <晉安州學> <書籍> <領降書籍> 大明一統志一部 已上俱係領降書籍
- <晉安州學> <書籍> <領降書籍> 提學毛公科發下 五倫書一部
- <晉安州學> <書籍> <領降書籍> 提學毛公科發下 通鑑綱目一部
- <晉安州學> <書籍> <領降書籍> 提學毛公科發下 大學衍義一部
- <晉安州學> <書籍> <領降書籍> 提學毛公科發下 韓柳文各一部

Smart Reader

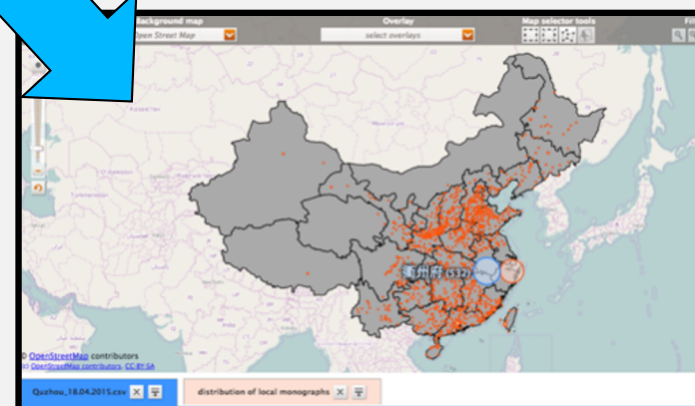
Tag only for this range:  to

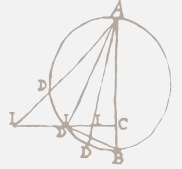
☐ Default ☐ With line break ☐ Allow space within block

[Compose](#) [Load](#) [Save](#) [Clear](#)

[Save text](#)

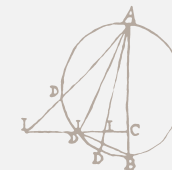
## LGMap





## Grouping technical features

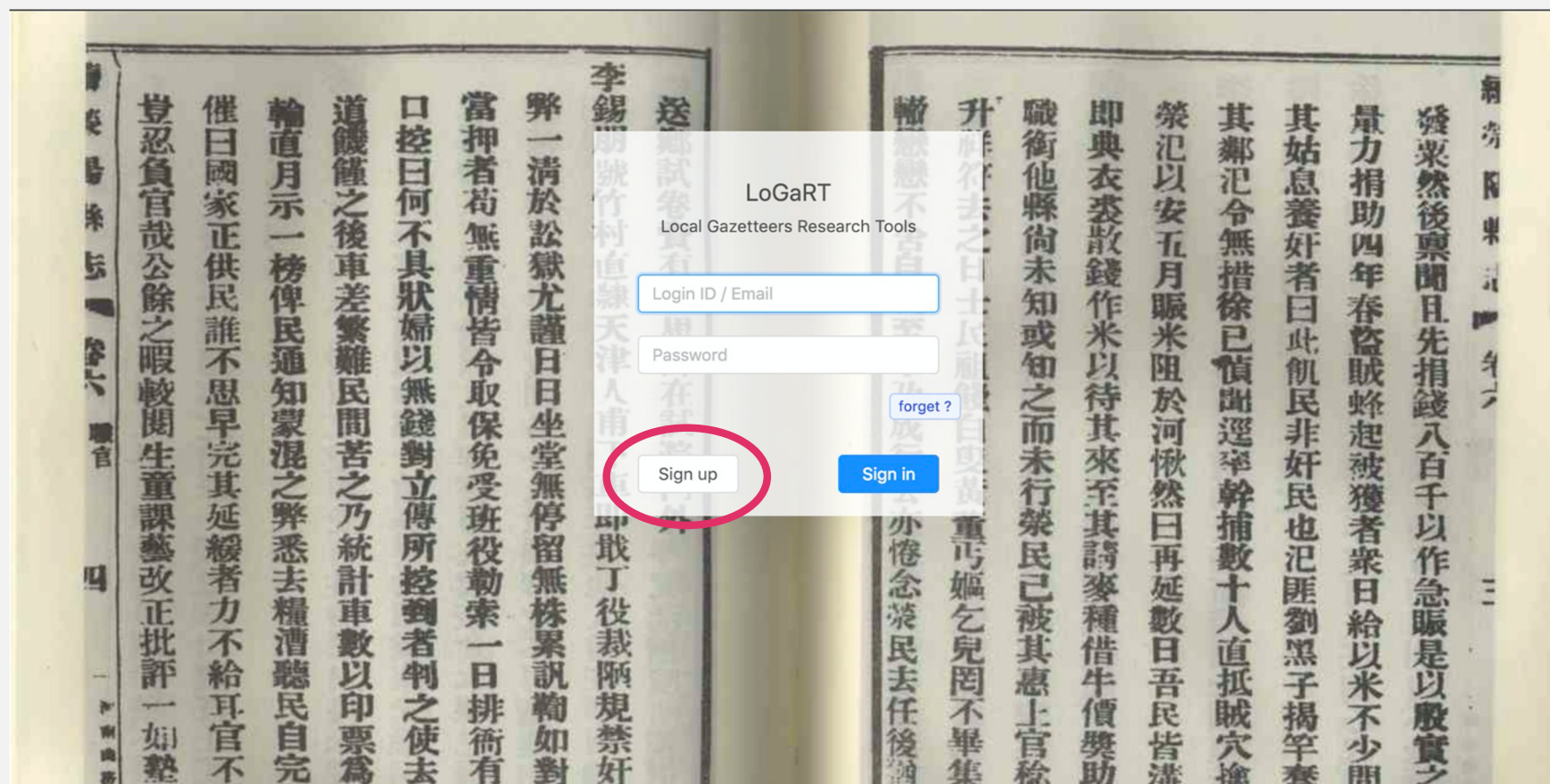
- Search and obtain overview
  - Search at three levels: Books, Sections, Pages (top menu modules)
  - Statistics view of search result
  - Visualize search result in LGMap / CHMap
- Text-tagging for data collection
  - Top menu modules: Topics, Tasks
- Work with visual materials (*tu* 圖) in LGs
  - Top menu modules: PWI (Pages with Images)
  - TU in LG working group (2018-)



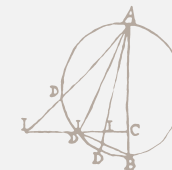
# Sign up for an Account

Visit: <https://logart.mpiwg-berlin.mpg.de>

Or, Google “logart mpiwg” then “Click here to access LoGaRT”







# Fill in Your Information

## LoGaRT

Local Gazetteers Research Tools

\* Login ID:

Login ID is required

\* Password:

Password is required

\* Retype Password:

\* Given Name:

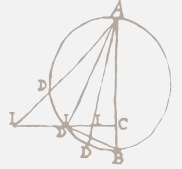
\* Family Name:

\* Email:

Affiliation:

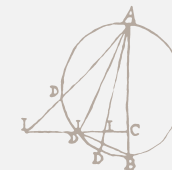
Position:

Sign up



## About page (After log-in)

- Design philosophy. The kind of research scenarios that LoGaRT supports.
- An overview to the technical components
- Sources
- Collaborators
- How to cite us
- Contact



# Top to Bottom Left to Right

## Top menu (black bg): main modules

LoGaRT

Books

Sections

Page Texts

PWI

Topics

Tasks

Compare

About

Gail Chen

Book Name: 一統 ×

Level 1: Please input level 1

Dynasty: 明 × 清 ×

Admin Type: Please select Admin T...

Level 2: Please input level 2

Period: Please input period

Source: Please select Source

Clear All

Search

Data

Statistics

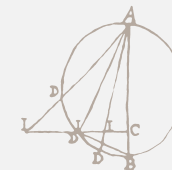
Total: 410 book(s)

< 1 2 3 4 5 ... 17 > 25 / page Columns

Book Name	Level...	Dyna...	Admi...	Years...	Edition	Volume	Author	Source	Total Page	TOC Done	Action
天下一統志 H00F4	北京 北京	明 天順	一統	1461 1735		90卷	李賢等奉勅修.	哈佛燕京圖 書館藏珍稀 方志	2260	<input checked="" type="checkbox"/>	
大明一統志 H00F5	北京 北京	明 天順	一統	1505 1505		90卷	李賢等奉勅修.	哈佛燕京圖 書館藏珍稀 方志	2821	<input checked="" type="checkbox"/>	
嘉善縣志 H016D	浙江 嘉興	明 正德	縣	1517 1517	明正德丁丑 [1517]修.	6卷	倪璣等纂修.	哈佛燕京圖 書館藏珍稀 方志	163	<input checked="" type="checkbox"/>	

To CSV

To LG-Map



LoGaRT

Books

Sections

Page Texts

PWI

Topics

Tasks

Compare

About

Gail Chen

# Filters

Book Name: 一統 ×

Level 1: Please input level 1

Dynasty: 明 × 清 ×

Admin Type: Please select Admin T...

Level 2: Please input level 2

Period: Please input period

Source: Please select Source

X Clear All

Search

Data

Statistics

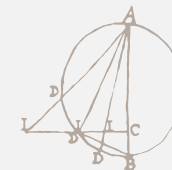
Total: 410 book(s)

< 1 2 3 4 5 ... 17 > 25 / page Columns

Book Name	Level...	Dyna...	Admi...	Years...	Edition	Volume	Author	Source	Total Page	TOC Done	Action
天下一統志 H00F4	北京 北京	明 天順	一統	1461 1735		90卷	李賢等奉勅修.	哈佛燕京圖書館藏珍稀方志	2260	<input checked="" type="checkbox"/>	
大明一統志 H00F5	北京 北京	明 天順	一統	1505 1505		90卷	李賢等奉勅修.	哈佛燕京圖書館藏珍稀方志	2821	<input checked="" type="checkbox"/>	
嘉善縣志 H016D	浙江 嘉興	明 正德	縣	1517 1517	明正德丁丑 [1517]修.	6卷	倪璣等纂修.	哈佛燕京圖書館藏珍稀方志	163	<input checked="" type="checkbox"/>	

To CSV

To LG-Map



LoGaRT

Books

Sections

Page Texts

PWI

Topics

Tasks

Compare

About

Gail Chen

Book Name: 一統 ×

Level 1: Please input level 1

Dynasty: 明 × 清 ×

Admin Type: Please select Admin T...

Level 2: Please input level 2

Period: Please input period

Source: Please select Source

X Clear All

Search

Data

Statistics

Total: 410 book(s)

Search result display

< 1 2 3 4 5 ... 17 >

25 / page

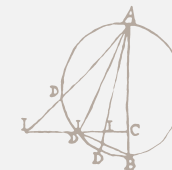
Columns

Book Name	Level...	Dyna...	Admi...	Years...	Edition	Volume	Author	Source	Total Page	TOC Done	Action
天下一統志 H00F4	北京 北京	明 天順	一統	1461 1735		90卷	李賢等奉勅修.	哈佛燕京圖書館藏珍稀方志	2260	<input checked="" type="checkbox"/>	
大明一統志 H00F5	北京 北京	明 天順	一統	1505 1505		90卷	李賢等奉勅修.	哈佛燕京圖書館藏珍稀方志	2821	<input checked="" type="checkbox"/>	
嘉善縣志 H016D	浙江 嘉興	明 正德	縣	1517 1517	明正德丁丑 [1517]修.	6卷	倪璣等纂修.	哈佛燕京圖書館藏珍稀方志	163	<input checked="" type="checkbox"/>	

To CSV

To LG-Map





LoGaRT

Books

Sections

Page Texts

PWI

Topics

Tasks

Compare

About

Gail Chen

Book Name: 一統 ×

Level 1: Please input level 1

Dynasty: 明 × 清 ×

Admin Type: Please select Admin T...

Level 2: Please input level 2

Period: Please input period

Source: Please select Source

X Clear All







Search

Data

Statistics

Total: 410 book(s)

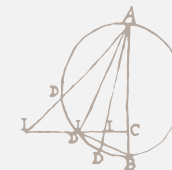
< 1 2 3 4 5 ... 17 > 25 / page Columns

Book Name	Level...	Dyna...	Admi...	Years...	Edition	Volume	Author	Source	Total Page	TOC Done	Action
天下一統志 H00F4	北京 北京	明 天順	一統	1461 1735		90卷	李賢等奉勅修.	哈佛燕京圖書館藏珍稀方志	2260	<input checked="" type="checkbox"/>	 
大明一統志 H00F5	北京 北京	明 天順	一統	1505 1505		90卷	李賢等奉勅修.	哈佛燕京圖書館藏珍稀方志	2821	<input checked="" type="checkbox"/>	 
嘉善縣志 H016D	浙江 嘉興	明 正德	縣	1517 1517	明正德丁丑 [1517]修.	6卷	倪璣等纂修.	哈佛燕京圖書館藏珍稀方志	163	<input checked="" type="checkbox"/>	 

To CSV

To LG-Map

Action buttons for each record



LoGaRT

Books

Sections

Page Texts

PWI

Topics

Tasks

Compare

About

Gail Chen

Book Name: 一統 ×

Level 1: Please input level 1

Dynasty: 明 × 清 ×

Admin Type: Please select Admin T...

Level 2: Please input level 2

Period: Please input period

Source: Please select Source

X Clear All







Search

Data

Statistics

Total: 410 book(s)

< 1 2 3 4 5 ... 17 > 25 / page Columns

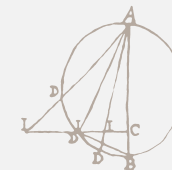
Book Name	Level...	Dyna...	Admi...	Years...	Edition	Volume	Author	Source	Total Page	TOC Done	Action
天下一統志 H00F4	北京 北京	明 天順	一統	1461 1735		90卷	李賢等奉勅修.	哈佛燕京圖 書館藏珍稀 方志	2260	✓	 
大明一統志 H00F5	北京 北京	明 天順	一統	1505 1505		90卷	李賢等奉勅修.	哈佛燕京圖 書館藏珍稀 方志	2821	✓	 
嘉善縣志 H016D	浙江 嘉興	明 正德	縣	1517 1517	明正德丁丑 [1517]修.	6卷	倪璣等纂修.	哈佛燕京圖 書館藏珍稀 方志	163	✓	 

To CSV

To LG-Map

Action buttons for the whole search result

# Buttons used in several modules



Expand/collapse filter area

LoGaRT Books Sections Page Texts PWI Topics Tasks Compare About Gail Chen

Book Name: 一統 × Level 1: Please input level 1 Level 2: Please input level 2 Dynasty: 明 × 清 × Admin Type: Please select Admin T... Period: Please input period Source: Please select Source

X means clear X Clear All Search

Data Statistics Total: 410 book(s)

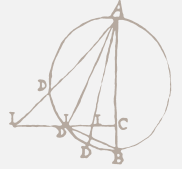
Switch view (data or statistics)

Book Name	Level...	Dyna...	Admi...	Years...	Edition	Volume	Author	Source	Total Page	TOC D
天下一統志 H00F4	北京 北京	明 天順	一統	1461 1735		90卷	李賢等奉勅修.	哈佛燕京圖書館藏珍稀方志	2260	
大明一統志 H00F5	北京 北京	明 天順	一統	1505 1505		90卷	李賢等奉勅修.	哈佛燕京圖書館藏珍稀方志	2821	<input checked="" type="checkbox"/>
嘉善縣志 H016D	浙江 嘉興	明 正德	縣	1517 1517	明正德丁丑 [1517]修.	6卷	倪瓚等纂修.	哈佛燕京圖書館藏珍稀方志	163	<input checked="" type="checkbox"/>

Columns

To CSV To LG-Map

Display/hide columns

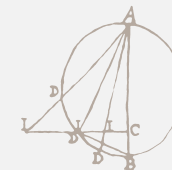


## Main Modules

- Books: Lists all the books linked in LoGaRT
- Sections: Search for sections
- Page Texts: Search in page texts (full-text search)
- Topics: Themes for tagging
- Tasks: Retrieve tagged texts
- PWI (Pages with images): lists all identified images

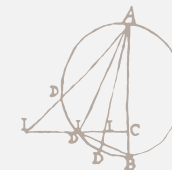


# Demostration of LoGaRT




# Books

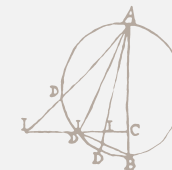
- Lists all the titles currently linked in LoGaRT
  - Can be filtered
  - Each record in the table is a title / a local gazetteer
- Display mode: Data tables or statistical overview
- Metadata fields
  - Book Name (book title), identifier, Level 1 / Level 2 (modern provinces and subordinated admin units that a LG belongs), Dynasty / Reign period of publication, Admin type (administrative units: county, prefecture, etc.), Book year / Edition year (publication year in AD and year of edition) , Edition, Volume, Author / compilers, Source (from which database/collection), Total page (total number of pages/images), Last editor / Modified (for TOC), Comments (problem reports)



# Books

- Action buttons for each record 
  - TOC Editor: View / edit a LG's Table of Contents
  - Show record in source website (Hollis)
  - Post comment: (internal use only!) for reporting problems
- Data / Statistics views
  - Data tables: listing titles
  - Statistics: showing the composition of a dataset (e.g. the whole HYL set) along several dimensions (represented by metadata fields)





# Table of Contents (TOC) Editor

Book ID	Book Name	Level 1	Level 2	Period	Total Page	Version No	Last Editor	Created Date
H016D	嘉善縣志	浙江	嘉興	正德	163	1	calvinyeh	2020-02-04 13:13:35

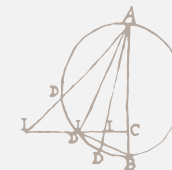
En	🕒	📄	<input checked="" type="checkbox"/> hide deleted sections	<a href="#">View</a>	<a href="#">Edit</a>	<a href="#">Save</a>
----	---	---	---	----------------------	----------------------	----------------------

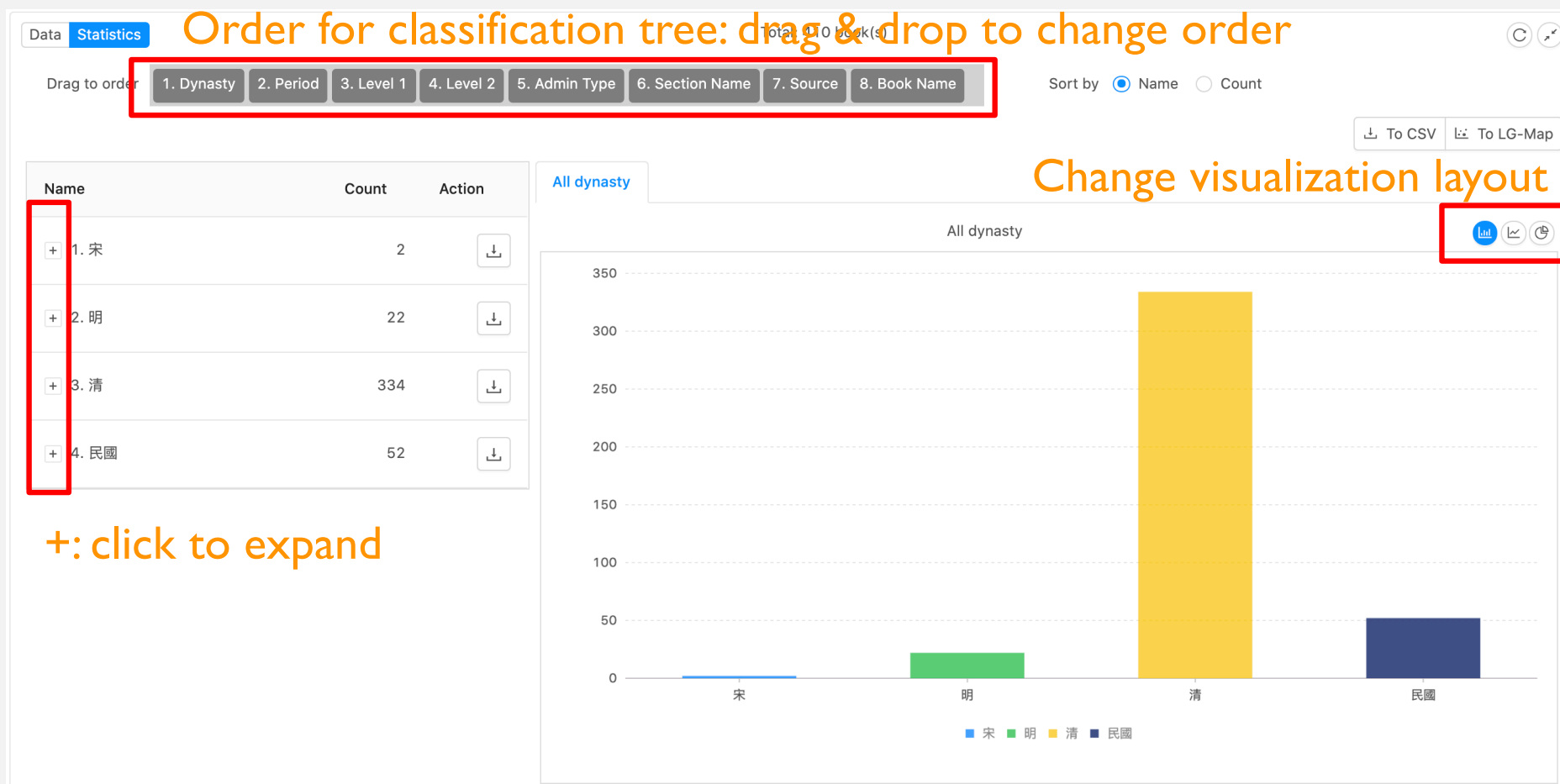
Seq No.	Section Name	Short Section Name	Start Page	End Page	Level
1	嘉善縣志序		4	- 6	1
2	嘉善縣志目錄		6	- 8	1
3	嘉善縣境圖		8	- 9	1
4	嘉善縣志卷一		9	- 30	1
5	建置		9	- 10	2
6	疆域		10	- 11	2
7	山川		11	- 17	2
8	市鎮 卷附		17	- 19	2
9	鄉都 村附		19	- 22	2

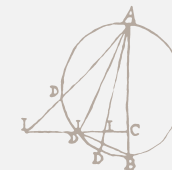
Each record includes:

- Section heading
- Begin & end page numbers
- Level in hierarchy: up to 3
- Show modification history
- Display scanned images

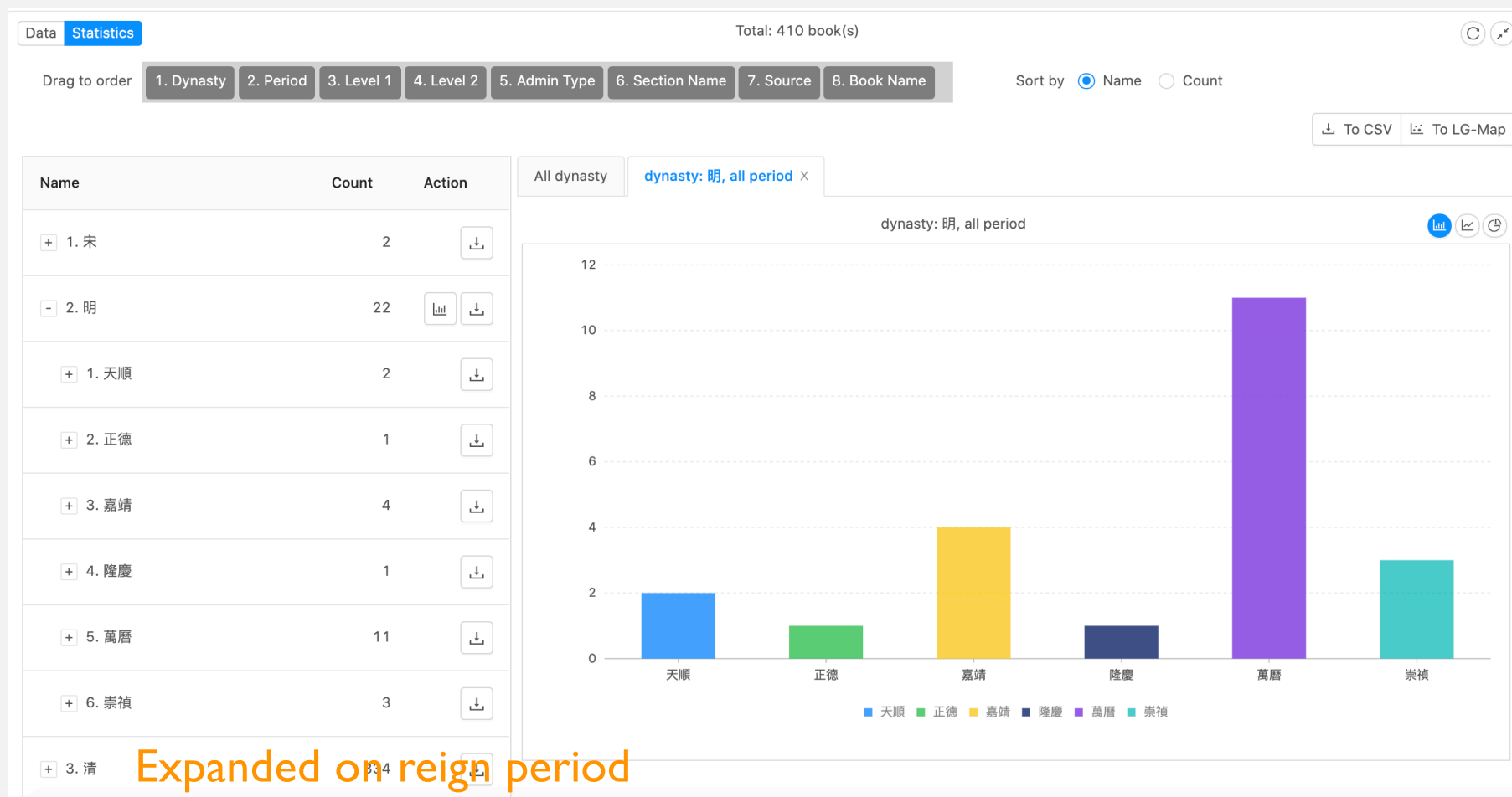


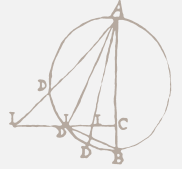
# Statistical view of the HYL set





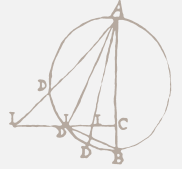
# Statistical view of the HYL set





## Books

- Action buttons for the entire search result
  - To CSV: Download result as a Comma Seperated Value file (a text file, in which each record is a row and each column is separated by comma). This allows the search result to be imported to other software such as spreadsheet & GIS
  - To LGMap: Bring the search result to LGMap

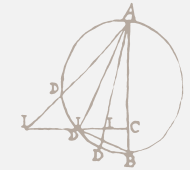


## LGMap: Visualizing a dataset (search result)

- Four display modules, inter-linked
  - Geographical map, data table, a timeline, pie charts
- Can overlay with Chinese historical maps
  - From CHGIS: 1820 provincial, prefectural boundaries and names

LGMap is customized from the open source software PLATIN, Place and Time Navigator: <http://platin.mpiwg-berlin.mpg.de/>.

Icons to call up: data  
management, pie charts,  
timeline, about



Overlay: to overlay with existing map layers

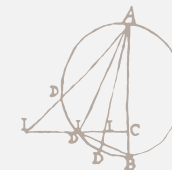
**Data table and manipulation**

search all fields / filter by columns

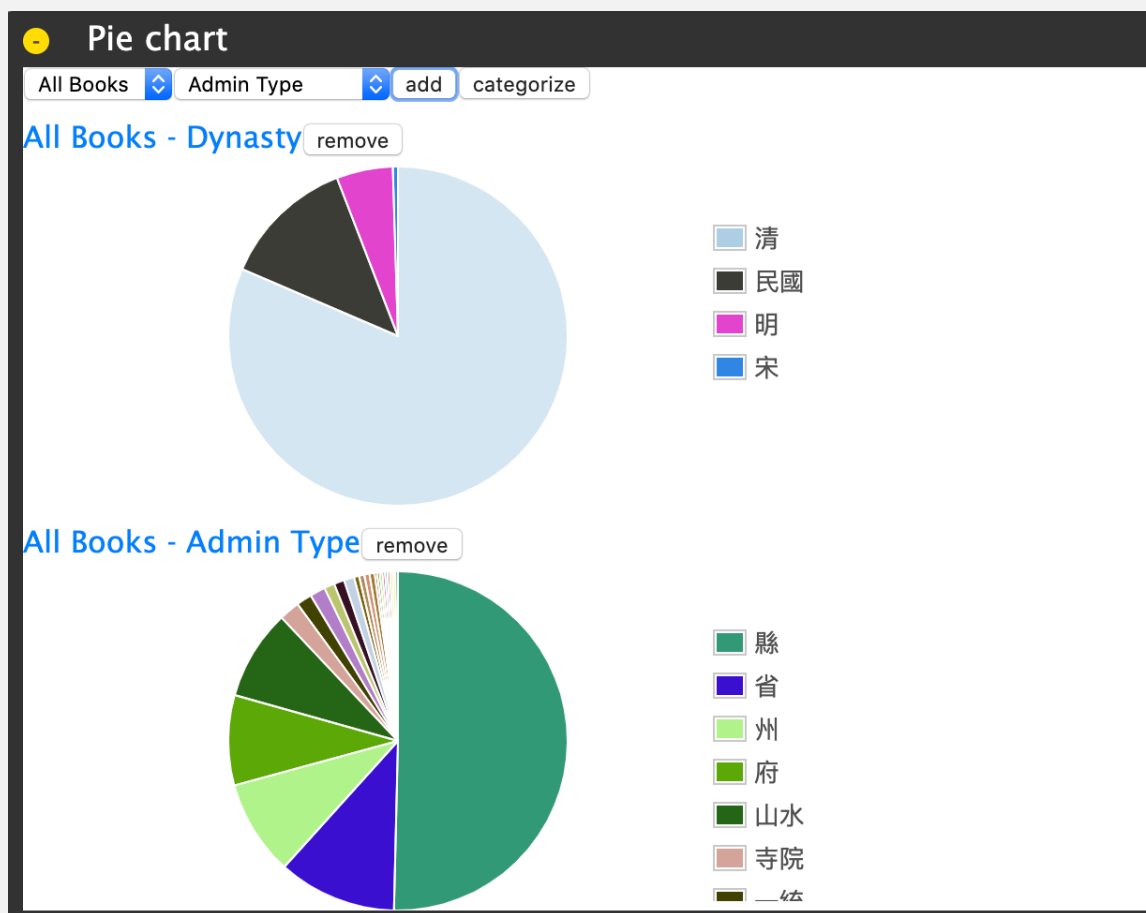
1-10 of 410 Results

10 1 Page 1 of 41

Book ID	Book Name	Admin Type	Level1	Level2	place	Dynasty	Period	Edition	Volume	Author	Source	Book Year	Edition Year	TimeSpan:begin	TimeSpan:end	description
H00F4	天下一統志	一統	北京	北京	順天府	明	天順		90卷	李賢等奉勅修.	哈佛燕京圖書館藏珍稀方志	1461	1735	1461	1461	90卷 / 李賢等奉勅修. /
H00F5	大明一統志	一統	北京	北京	順天府	明	天順		90卷	李賢等奉勅修.	哈佛燕京圖書館藏珍稀方志	1505	1505	1505	1505	90卷 / 李賢等奉勅修.

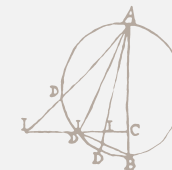


# Pie charts

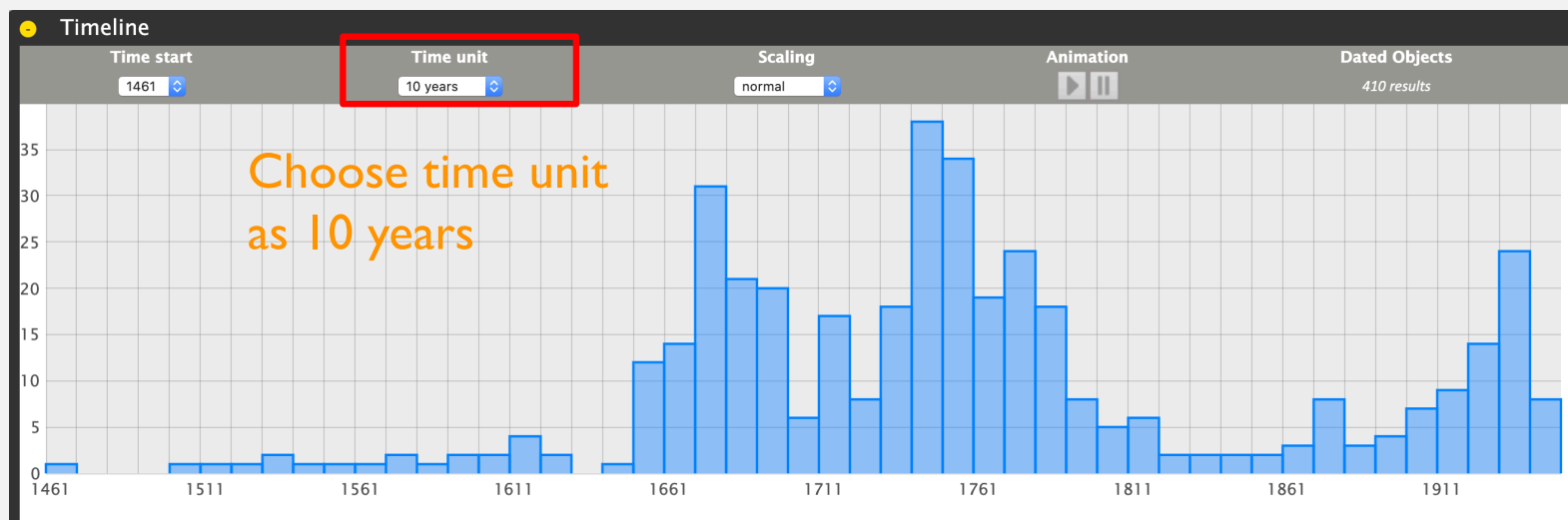
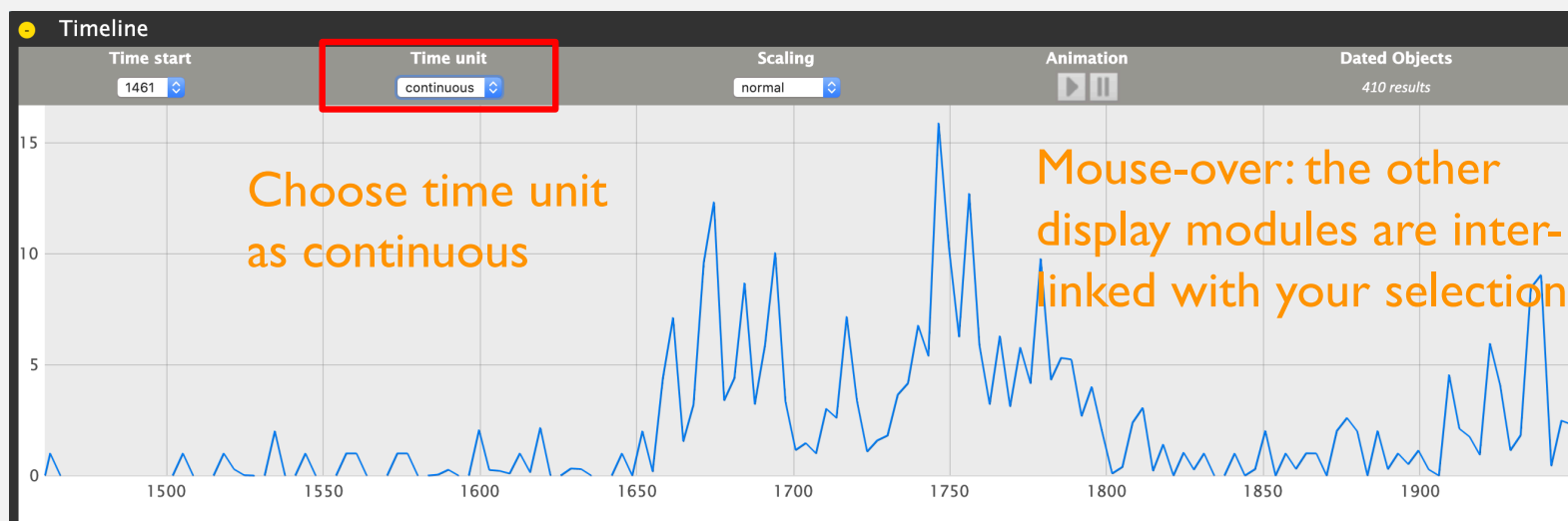


- Select a metadata field to create a pie chart (categorical analysis)
- Multiple pie charts
- Mouse-over: the other display modules are inter-linked with your selection



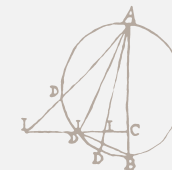


# Timeline



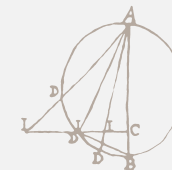


Break I (3 mins)  
Try it out  
Q&A



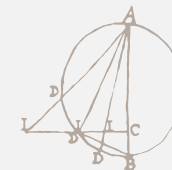
## Sections

- Search for sections
  - Each record in search result is a **section**!
  - Search by section headings: allow multiple keywords. Eg., type in “物産, 土産” (including ,) to find sections on local products
  - Can be filtered by Book titles, Level 1/2, Dynasty, Reign period, administration level (county, prefecture, province, ...)
- Metadata fields:
  - Section Level (in TOC hierarchy), Pages (begin & end), Section Name (heading)
  - Task: whether a tagging task has been created for a section



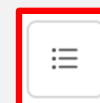
## How to input multiple keywords in filters

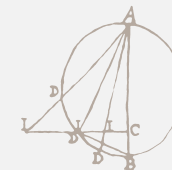
- Common across all modules
- For example, in Book Name, Section Name, Level I, Period
- Keywords have to be in **Traditional Chinese!**
- Type *keyword I* then **click comma**. *Keyword I* should become a label now. Then, type another keyword.



## Sections

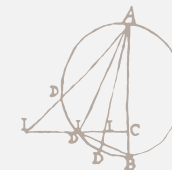
- Action buttons for each record
  - Load text / New task: 载入全文 / 新建标记工作
  - Edit TOC
- Action buttons for search result
  - To CSV
  - To LGMap
  - Save query in LoGaRT (can be called up later)





## How to compare two datasets in LGMap

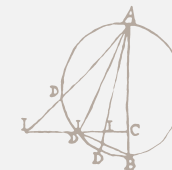
- Method 1: aggregate (search) then divide (in LGMap)
  - Use any of the four modules to filter data
  - A filtered dataset can be shown on a separate tab with a different color
  - Helps to compare two or more sub-datasets
  - Can always call up the whole collection as baseline
  - Can import datasets produced in LoGaRT (via downloading CSV and then uploading to LGMap)
- Method 2: save different searches as several CSVs. Then load them in LGMap.



## How to interpret visualization

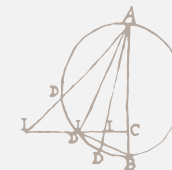
- Need to first understand your “baseline”
  - The distribution / composition of the whole HYL set
- Always compare a search result distribution with the baseline





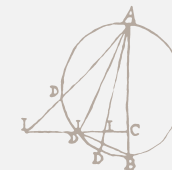
## Page Texts

- Search for pages by keywords (combined with filters)
  - Each record in the result is a **page!**
  - **Total number: number of pages instead of number of the keyword's occurrence**
- Multiple keywords
  - Space between keywords means “OR search”: a page is returned when one or more of the keywords appears on that page. Useful when finding something with different literal expressions
    - Eg. “桑災 殺桑 害桑” are all expressions for damages of mulberry trees .”  
八蜡廟 蝗神廟 劉猛將軍廟” are names that can refer to locust temples.
  - Add + or – before a keyword: to identify “must-have” and excluded keywords
    - +: keywords with + in front must appear on a page for it to be returned
    - -: pages with this keyword are excluded and are not returned




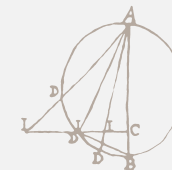
## Page texts

- Metadata fields :
  - Page (number), Section Name (which the page appears in), Content (full text of the page)
  - Select: by default all the records are selected (check box). Use these check boxes to remove records from search result (effected only for Saved queries)
  - Task: whether a tagging task has been created



## Page texts

- Action buttons for each record
  - View text: load text of the section (that includes this page)
- Action buttons for search result
  - To CSV
  - To LGMMap
  - Save query 
  - Create a tagging task from search result



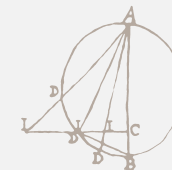
## Try-it-out examples

- Click on Books
  - See the how the HYL collection of LGs is composed of in terms of periods, geographical regions, admin types. This would illustrate the “baseline” of this collection in your mind for understanding and interpreting search results later on.
- Click on Sections
  - Use it for identifying sections with similar contents across LGs. Eg., Search for “物產, 土產”.
  - Or, use it for getting chronological/geographical overview of a section. Eg., “凡例”, “災”.
- Page Texts:
  - Use it to find the pages where certain keywords appear. Eg., “八蜡廟 八腊廟 劉猛將軍廟 蝗神廟 蟲王廟” (an OR search).



Break II (3 mins)  
Try it out  
Q&A

Use examples in the slides



## PWI: Pages with images

- Lists all the identified pages with images in the collection
- While LGs is mainly a textual genre, they may also include visual materials such as maps, city layout, star maps.
- By PWI, we pull all the images together out of their original positions, and hope that would help scholars to treat all the images of LGs as a database.
- See: **TU in LG** working group page @MPIWVG
  - Kenneth Hammond, Shellen Wu, Anne Gerritsen, Jiajing Zhang
  - **Upcoming conference (2021 spring) CFP**

# PWI

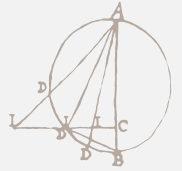


Image Source: All valid images  
or My collection

Keywords that appear on page

Expand  
filter area

LoGART Books Sections Page Texts PWI Topics

Image Source: 1. All valid images input search keywords, eq : +辣椒 -五皇廟 八蜡廟

Data Statistics Keywords : ; Filters: imageSource=ValidImage ; Found: 4,156 page(s)

Table Carousel All

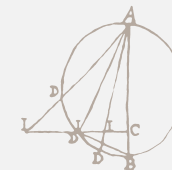
< 1 2 3 4 5 ... 167 > 25 / page

H00F5\_33 H00F5\_32

地 內 師 京 圖 之 統 一 明 大

皇明一統之盛冠乎古今者庶萬世有足徵云

To CSV To LG-Map To CH-Map Save



# PWI

- Display modes under Data: Carousel mode (show images in a row, default) or Table mode.
- In Carousel mode, scroll left and right to see images
- 4 buttons above each image: further actions



From left to right:



Information about this image



(Tags) Current tags applied to this image

- Image identifier: Book ID & page number

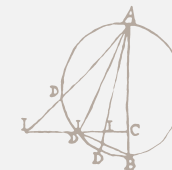


(Pen) Enlarge and edit this image



(Star) Add this image to My Collection





# Image enlarged view

Next page

Previous page

Display consecutive pages

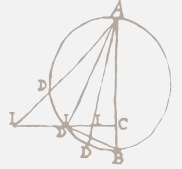


(thumb) Valid / invalid image

☆Collect

Zoom in / out,  
rotate,  
reload image

Add personal note



# PWI

- For each image:
  - One can collect images of interests (by clicking Collect)
  - One can add personal notes (visible to self) and tags (keywords) for categorization (visible to all)
  - Pre-defined tags:
    - Form tags: map, diagram, photograph
    - Content tags: agriculture, building or building complex, flora and fauna, human figure, landscape, military, musical, object, ritual
  - The Images in the HYL collection are not yet given tags. We encourage you to tag them!

- Filter with keywords that appear on page (they are typed in; see our digitization spec: <https://github.com/SeanyWang/despecs/blob/master/LGspecs.md>)
- Filter with metadata: all the metadata fields mentioned before, plus filter by Note (personal note)
  - Please ignore: filter by Tag (categories of images; not yet tagged) & by Text inside (a field for internal production; please disregard).

LoGaRT

Books

Sections

Page Texts

PWI

Topics

Tasks

Compare

About

Gail Chen

Image Source: 1. All valid images

八蜡廟 八腊廟 劉猛將軍廟 蝗神廟 蟲王廟

Book Name: Please input book name

Level 1: Please input level 1

Dynasty: Please select Dynasty

Admin Type: Please select Admin ...

Section Name: Please input section name

Level 2: Please input level 2

Period: Please input period

Source: Please select Source

Tag: Please input tag name

Note:

Text Inside: Please select text status

X Clear All

Search

Focus on the input or type keywords inside, then top 50 tags will be shown.

Data

Statistics

Keywords: ; Filters: imageSource=ValidImage section\_name=分野,星野 ; Found: 127 page(s)

Table

Carousel

All

H00F9\_23

H00F9\_22

H00F9\_20

H00F9\_19

1

2

3

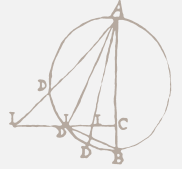
4

5

6

>

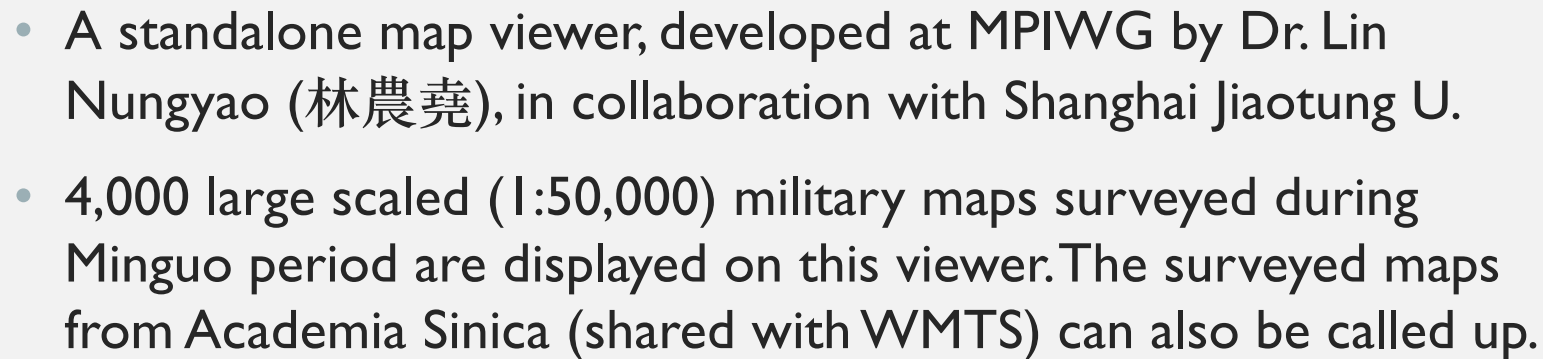
25 / page



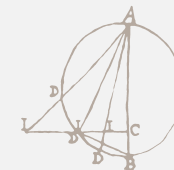
## PWI: Action buttons for search result

- To CSV: save the metadata of resulting images as a CSV text file. Can be brought to other software.
- To LGMap: visualize the result on LGMap
- To CH-Map: display images on CH-Map, where one can call up Republican military maps of China
- Save: save the query / filter in LoGaRT so that it can be retrieved later, by clicking

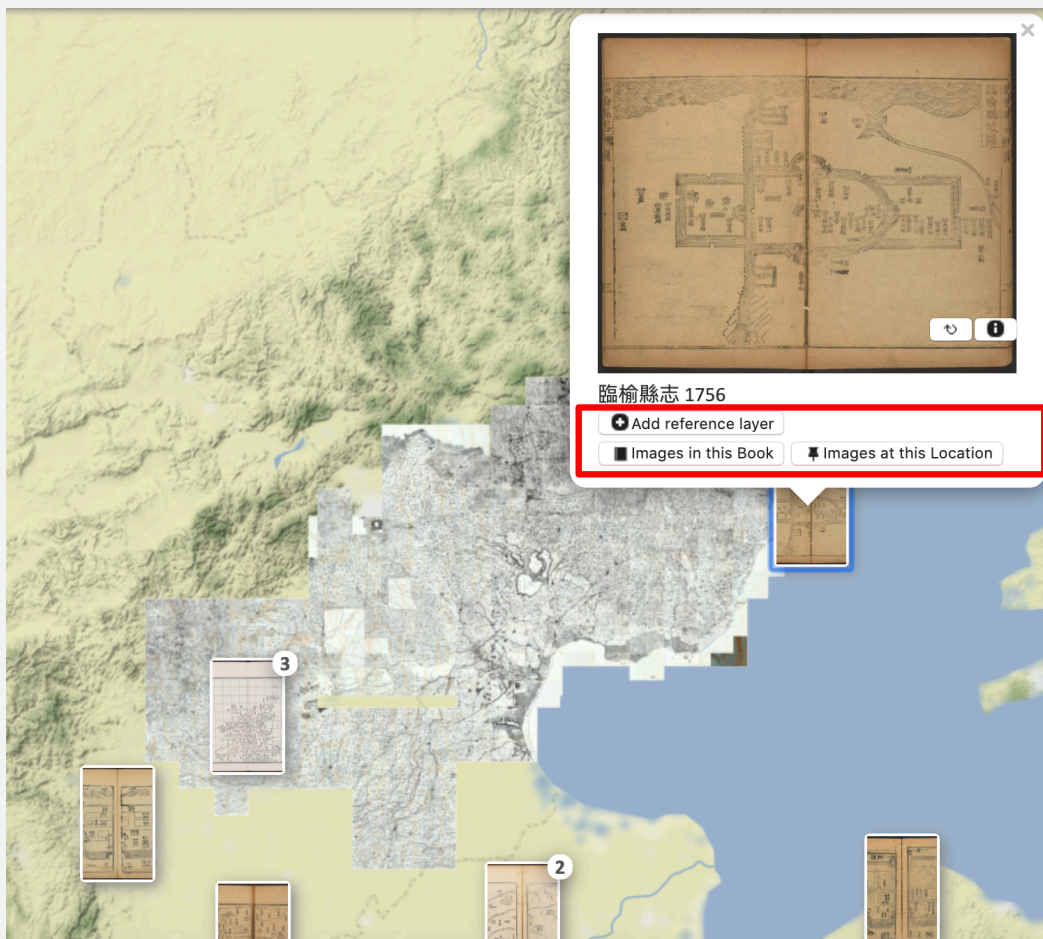
<https://chmap.mpiwg-berlin.mpg.de/>



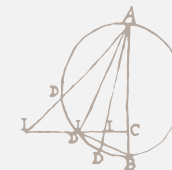




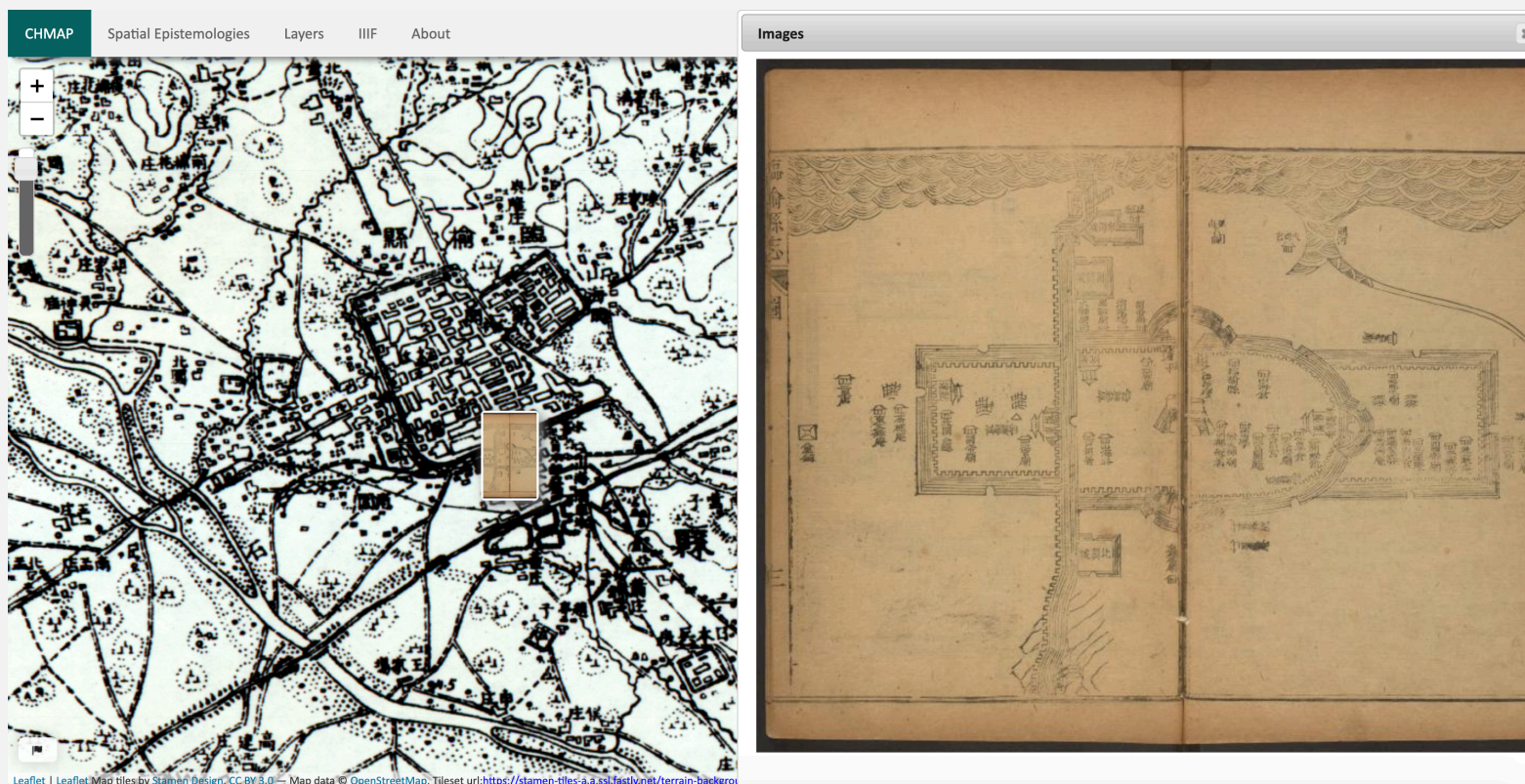
## PWI: Visualizing LG images on CHMap



- PWI search result can be viewed on CHMap by one click
- Zoom in/out by scrolling mouse cursor
- Clicking on any image thumbnail brings up further action buttons:
  - **Add reference layer:** to display the Republican survey maps which this image's location falls within
  - **Images in this book:** show all the images from the same LG (and that appear in this search result)
  - **Images at this location:** show all images at this location (and that appear in this search result)



# Displaying different spatial epistemologies



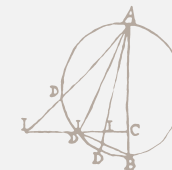
See: Lin, N.Y., Chen, S. P., Wang, S., & Yeh, C. (2020). Displaying spatial epistemologies on web GIS: using visual materials from the Chinese local gazetteers as an example. *International Journal of Humanities and Arts Computing*, 14(1-2), 81-97.



Break III (1 mins)  
Try it out  
Q&A

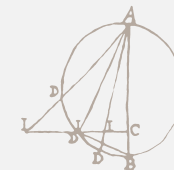
Use examples in the slides





## Text tagging

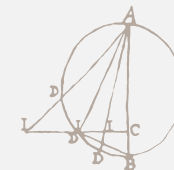
- Text to data
  - By tagging, to “teach” computers which part of the text mean something; to teach computers the hidden structure of a text.
- When do you need it?
  - To convert lists given in certain sections into a spreadsheet (table): Eg. lists of local products, officials ...
  - To convert full-text search result into a spreadsheet, esp. when each record contain consistent types of information.



## Example: Page text from the section of “local products” 物產

椰樹似欒櫚而高大外堅中虛斫其間有類子作穗生木端葉下有鬚木紫黑色有文理可以制器  
竹木其狀類竹可為杖漳州志有桃枝竹疑即此也吉兆藤亦名烏理藤紫色其根為降真香金剛藤根赤色俗呼菠蘿香藤緣木而生其氣清香又有南藤土人或取以釀酒王孫藤柔韌可為纜  
**竹之屬**戴凱之竹譜竹有生日謂五月十三日移植最宜亦謂竹醉將枯則生花或云劈竹多用辰日斷**猫竹**方言謂之麻竹大者徑七八寸土人或取其葉以裹角黍笋夏生**江南竹**龐大而堅直笋冬生**淡竹**肉薄節間有粉南人燒以取瀝者葉亦可入藥**慈竹**任昉述異記云南中有子母竹生不離本今慈竹是也王勃云宗生族茂天長地久萬抵爭盤千株競紉**斑竹**永福縣鶴洋多產之差及湘江者**紫竹**小而色純紫**苦竹**  
千株競紉**斑竹**永福縣鶴洋多產之差及湘江者**紫竹**小而色純紫**苦竹**  
入藥**慈竹**任昉述異記云南中有子母竹生不離本今慈竹是也王勃云宗生族茂天長地久萬抵爭盤千株競紉**斑竹**永福縣鶴洋多產之差及湘江者**紫竹**小而色純紫**苦竹**  
亦可入藥**慈竹**任昉述異記云南中有子母竹生不離本今慈竹是也王勃云宗生族茂天長地久萬抵爭盤千株競紉**斑竹**永福縣鶴洋多產之差及湘江者**紫竹**小而色純紫**苦竹**  
萬抵爭盤千株競紉**斑竹**永福縣鶴洋多產之差及湘江者**紫竹**小而色純紫**苦竹**  
千株競紉**斑竹**永福縣鶴洋多產之差及湘江者**紫竹**小而色純紫**苦竹**

椰樹似欒櫚而高大外堅中虛斫其間有類子作穗生木端葉下有鬚木紫黑色有文理可以制器竹木其狀類竹可為杖漳州志有桃枝竹疑即此也吉兆藤亦名烏理藤紫色其根為降真香金剛藤根赤色俗呼菠蘿香藤緣木而生其氣清香又有南藤土人或取以釀酒王孫藤柔韌可為纜**竹之屬**戴凱之竹譜竹有生日謂五月十三日移植最宜亦謂竹醉將枯則生花或云劈竹多用辰日斷**猫竹**方言謂之麻竹大者徑七八寸土人或取其葉以裹角黍笋夏生**江南竹**龐大而堅直笋冬生**淡竹**肉薄節間有粉南人燒以取瀝者葉亦可入藥**慈竹**任昉述異記云南中有子母竹生不離本今慈竹是也王勃云宗生族茂天長地久萬抵爭盤千株競紉**斑竹**永福縣鶴洋多產之差及湘江者**紫竹**小而色純紫**苦竹**



# Example: Water conservancies 水利

(Credit: Che Qun 车群)

Page no: 718

Page no: 715

## 岳州府志卷之十四

岳州府知府黃凝道重修

原任衡陽縣知縣謝仲坑編纂

### 水利志

或謂岳陽之水卜其害未卜其利夫害也而防之室廬以聚畦塋以修利之轉移孰大於是況涸流而陸歲爭此土於汪洋所云卜害在水亦止仍其故道哉然則善都水者亦惟以故道予水而毋盡水之利則水若其性民利溥存高下溝隄無施而不可爾作水利志

水利志

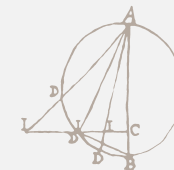
水利

### 巴陵縣水利

偃虹隄在岳陽門西距金鷄之右宋慶曆間知州

### 軍勝宗諒築

土石兼用長一千丈高三丈厚三丈二尺殺其上厚二丈有奇凡爲工一萬五千五百初荆潭黔蜀往來湖中者苦無所泊常多覆溺宗諒乃議築隄上之轉運使聞於朝決之三司報可舉役隄成舟之往來皆泊隄下行旅頓之而城垣亦有所捍禦不爲患有歐陽修記後



# Example: Water conservancies 水利

(Credit: Che Qun 车群)

Page no: 720

Page no: 719

白荆隄在縣東一十五里一名紫宋築明成化八年知府吳節增築嘉靖十年知府蕭晚重修四十年邑紳方鈍增修

永濟隄在縣北一十五里明成化二十年知府李鏡築上自城北演武亭起下至城陵磯止長四十二丈廣二丈高七八尺樹柳二萬株有石橋二堯石為閘於二橋之北廣五丈高一丈二尺長一丈五尺架木梁以通車馬建亭列室以為官屬迎候之所凡為工二十有七萬費金三千餘兩民賴安居永無水患墾田至數百頃行旅舍舟而陸亦無覆溺百年誦之今名國朝順治十年知縣王李公隄有李東陽記

篤慶康熙五年知縣李价重修

舊江村隄在縣西二十里明築國朝順治十二

護城隄在西門城下明隆慶初知府李時漸築址久失工費浩繁事不果行

九龍隄在城北附郭明洪武初築郡城既修北鑿以衛城隄即河之岸今汴河墾田此隄資為禦水

固城垸隄在江北與監利接界明初築嘉靖三十九年隄潰隆慶元年知府李時漸知縣李之珍重修

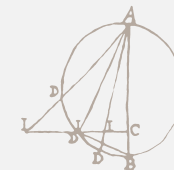
南津隄在縣南明弘治間知府張金築



Page no: 721

75

年增築舊隄止二百五十丈至是監利潰隄乃增築巴隄其土名竹荊河何家壩瓦子灣魯  
家埠四處新舊共築一千一十二丈六尺康熙五十五年水決巡撫李  
發甲奏發帑銀檄委岳常道裴衝度監工知縣  
王國英承修時湖南諸隄共發銀六萬二千兩有奇巴邑新隄分發銀二千六百兩有  
雍正五年水決復奉發帑銀重修銀數與康熙五十五年同  
陂凡一十七曰楓梓陂古江陂大塘陂河風陂燕  
竹陂瘦蛇陂黃沙陂白家陂馬家陂龍西陂龍過  
陂龍潭陂檀樹陂官陂蓮花陂金陂櫻桃陂  
塘凡二十曰橫塘君山塘麻塘李婆塘段家塘金  
林塘寺塘黃塘萬工塘西婆塘道塘清塘石塘長

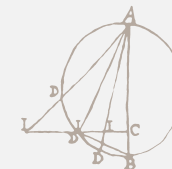


## How the section text looks like (Section search “水利”)

【718】○正派為二十二流歸武漢故時無江患自穴口盡○○塞惟調弦虎渡二口在江之南江水泛濫調弦分○○江流之四湖南近江州縣悉皆受害按調弦古名○○揚口即今華容河元以前穴口盡開水無漲決故○○以開揚口為利迨明初穴口盡塞江水並注調弦○○故又有以塞調弦為利者然開固漫而東塞亦障○○而西終非上策夫穴口不可猝復則惟嚴謹隄防○○且禁沿江枝河毋許塞流報墾貪目前小利而增○○他日浸溺之憂此不獨為岳言之而治○○必總其成無分一郡而可偏治之水也○○巴陵縣水利○○偃虹隄在岳陽門西距金鷄之右宋慶曆間知州○○軍勝宗諒築土石兼用長一千丈高三丈厚三丈○二尺殺其上厚二丈有奇凡為工一○○萬五千五百初荆潭黔蜀往來湖中者苦無所泊○○常多覆溺宗諒乃議築隄上之轉運使聞於朝決○○之三司報可舉役隄成舟之往來皆泊隄下行旅○○賴之而城垣亦有所捍禦不為患有歐陽修記後○

【719】○廢守土者屢議興修皆以故○○址久失工費浩繁事不果行○○護城隄在西門城下明隆慶初知府李時漸築土○隄○○自岳陽樓而南凡二百六十丈有奇初分巡副使○○施篤臣倡議創築所費不貲當事者皆難之時漸○○奉行唯謹卒○○以有成今廢○○九龍隄在城北附郭明洪武初築○郡城既修北鑿○汴河週千餘丈○○以衛城隄即河之岸今汴○○河墾田此隄資為禦水○○固城坑隄在江北與監利接界明初築嘉靖三十○○九年隄潰隆慶元年知府李時漸知縣李之珍重○○修○○南津隄在縣南明口治間知府張金築○

【720】○○白荆隄在縣東一十五里○一名紫○荆隄○宋築明成化八○○年知府吳節增築嘉靖十年知府蕭晚重修



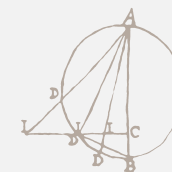
## Text after tagging

【722】湖塘千工塘黃泥塘李公塘石城塘百家塘沙塘平江縣水利隄無陂凡三十一曰芭蕉陂。  
在二都長三十丈濶二十丈。  
烟州陂在三。  
都即堰陂長五十丈。  
謝家陂在三都長二十五丈有奇。  
源澤陂在五都即鼃潭陂長一百丈。  
肥田大陂在五都長二百丈唐宋古陂明嘉靖間都御史顧璘檄築國朝康熙二十七年知縣鹿廷瑄修乾隆十年知縣楊富崙勸諭鄉民重修。  
柞樹陂在五都清。  
江陂在五都。  
觜頭陂在五都。  
冷水陂即邱家陂在下五都。  
杖子陂在七都長三十三丈。  
塘陂在八都。

Color coded tags: Name of water conservancy, location, 长length, width, height



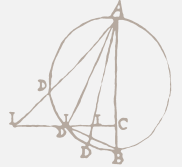
## Export as a spreadsheet (table)



name	部	Address	Longitude	Latitude	Descriptio	頁數	水利	地點	描述	種類	毀修	別名	長度	寬度	高度
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	718	偃虹隄	在岳陽門西距金鷄	隄	宋慶曆間知州車	長一千丈	厚三丈二	高三丈		
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	719	護城隄	在西門城下;自岳陽	隄	明隆慶初知府李時	二百六十丈有奇				
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	719	九龍隄	在城北附郭;北盤汴	隄	明洪武初築郡城既	週千餘丈				
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	719	固城垸隄	在江北與監利接界	隄	明初築;嘉靖三十九年	隄潰;隆慶;元年知府李時漸				
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	719	南津隄	在縣南	隄	明 治間知府張金築					
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	白荊隄	在縣東一十五里	隄	宋築明成 紫荊隄					
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	永濟隄	在縣北一十五里;上	隄	明成化二十年知府	長四十丈 廣二丈;廣高七八尺				
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	李公隄		隄	國朝順治十年知縣王篤慶康熙五年知縣李煥重修					
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	舊江村隄	在縣西二十里;竹莊	隄	明築國朝 巴隄	新舊共築一千一十二丈六尺				
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	楓梓陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	古江陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	大塘陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	河風陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	燕竹陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	瘦蛇陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	黃沙陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	白家陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	馬家陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	龍西陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	龍過陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	龍潭陂		陂						
岳州府志	水利志	岳州府	113.0977	29.37126	岳州府志	720	檀樹陂		陂						

See: Che Qun. Environmental Changes and Water Conservancy on the Yangtze River. (tbp)

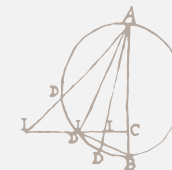




## The (semi-automatic) process of text-tagging

1. Load text to the Tagging Interface via:
  - Section search
  - Page text search
2. Create a Topic for tagging (or choose from shared ones): via the “Topic” module
3. Create Tags associated with a Topic: on the Tagging Interface
4. Tag manually or automatically with SmartRegex
5. Preview current tagging result (as a HTML table), or export as a file (CSV, XML, HTML)

Iterative from steps 3-5



# Load text via Section Search













Search for “壇廟”, and click on Load text for the record from “乾隆南寧府志”, p309-327.

Sometimes having scanned images is useful when tagging: click Edit TOC to call up.

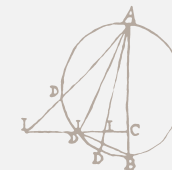
LoGaRT Books Sections Page Texts PWI Topics Tasks Compare About Gail Chen En

Book Name: Please input book name Level 1: Please input level 1 Dynasty: Please select dynasty Admin Type: Please select admin ty...  
Section Name: 壇廟 x Level 2: Please input level 2 Period: Please input period Source: Please select source  
X Clear All Search

Data Statistics Filters section\_name=壇廟 ; Found 79 section(s)

Book Name	Level...	Dyna...	Sec. lvl.	Page No.	Section Name	Task	Action
H0129	保定	雍正	3	285 - 294	壇廟	No	 
深澤縣志 H0131	河北 石家莊	清 雍正	3	88 - 91	壇廟 生祠附	No	 
陵川縣志 H00C7	山西 晉東地區	清 乾隆	3	108 - 108	壇廟	No	 
南寧府志 H008D	廣西 南寧	清 乾隆	2	309 - 327	壇廟 祀田附		  
南寧府志 H008D	廣西 南寧	清 乾隆	2	327 - 329	壇廟	No	 

To CSV To LG-Map Save



# Load text via Page text search

Note: a limit of 400 records for tagging is applied to avoid crashing the Tagging Interface

LoGaRT

Books

Sections

Page Texts

PWI

Topics

Tasks

Compare

About

Gail Chen

En

蝗神廟 蟲王廟

🔍

🗑️

Data

Statistics

Keywords: 蝗神廟 蟲王廟; Found: 25 page(s)

<

1

>

25 / page

Columns

Book Name	Level...	Dyna...	Section Name	Page No	Content	Select	Task	Action
清豐縣志 H015C	河南 濮陽	清 康熙	清豐縣志卷之四 祀典	104 📄	<List_Start> 八蜡廟 舊在寶固村知縣王寵從置南關外知縣李汝寬重修 邑厲壇 在北郭外祭以清明中元孟冬朔 以上俱於康熙十 江濟廣源王廟 在縣南十五里相傳神即楚三閭大夫旁有井 鄉李奎翰林范瑟諸靈應碑記 國朝康熙十四年因禱雨靈應 康王廟 在縣東南二十五里上和村世傳宋康王質于金間道 土地祠 在儀門外東 三官廟 在南門內小街  關王廟 襄城內俱有 龍王廟 在東門外 真武廟 在西關外金堤上 東嶽天齊廟 在縣東二里李家庄 蟲王廟 縣西南二十五里豆固村 顯聖靈源公廟 祥符元年宋真宗幸廟的奠以頓 令□□令又 泉源祠 寰宇記澶州頓 縣東北三十五里有泉源祠詩云泉 確山廟 縣東南三十里神當頭村 <List_End>	<input checked="" type="checkbox"/>	No	👁️
畿輔通志 H0129	河北 保定	清 雍正	畿輔通志卷之四十 九 祠祀 順天府	1813 📄	祀河神康熙三十四年知州張朝琮建 張相公廟 一在通州境內舊志神本越人爲水神明萬歷間有 盛 河神廟 盧溝河至永清宋家莊康熙三十七年五月河工告成 聖祖賜名永定河建河神廟	<input checked="" type="checkbox"/>	No	👁️

↓ Refine With Regex

📄 To CSV

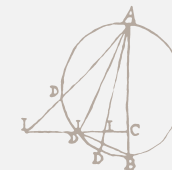
🗺️ To LG-Map

Create Tagging Task

Save

# Tagging interface

Load text for the “壇廟” section from “乾隆南寧府志”



book name: 南寧府志, period: 乾隆, section name: 壇廟 祀田附

## Tools

【309】

<List\_Start>

灋州在城南隋大將劉方始開此路置鎮守尋廢唐貞觀十二年李宏節遣首領寧師京循故道開置以達交趾因州界有灋水故名後廢。

南寧府志卷十三終。

。

南寧府志卷十四。

輿地志。

壇廟。祀田附。

宣化縣。附郭。

社稷壇在城東二里明嘉靖二十五年知府郭楠建久圯。

國朝雍正十年知府張漢知縣陳旭奉旨遵式改建。

先農壇在東門外楊梅嶺。

國朝雍正五年署宣化縣親寧州知州徐德秩遵式

<List\_Start>

【310】

<List\_Start>

建置藉田四畝九分祭器農具俱全。

風雲雷雨山川城隍壇在社稷壇後明嘉靖二十五年知府郭楠建。

## Text area

Topic : Basic Demo 1 (基本示範第一...

Main tag: 壇廟

▼ Edit, Tag, Save text

Tagging

Editing

Undo changes

Smart Regex©

壇廟

Load

Compose

Clear

☒ Without line break

☐ With a preceding line break

☐ With a succeeding line break

Apply

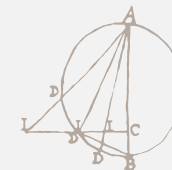
Preview

☐ embedded

Create task

> Replace By Regex

> Export



# Topic and tags

Topic : Basic Demo 1 (基本示範第一...



Main tag: 壇廟



Create / choose a Topic for tagging

Create tags

▼ Edit, Tag, Save text

Tagging Editing Undo changes

Smart Regex©

壇廟

Load Compose Clear

☒ Without line break  
☐ With a preceding line break  
☐ With a succeeding line break

Apply

Preview Create task

☐ embedded

> Replace By Regex

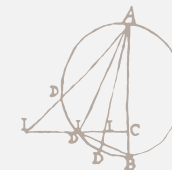
> Export

Tags of Basic Demo 1 (基本示範第一組, jī běn shì fàn)

Name	Tag	Color	Actions
壇廟	basic_demo	rgb(255, 0, 174)	
壽命	age	rgb(241, 196, 15)	
食物	food	rgb(39, 174, 96)	
性別	gender	rgb(6, 82, 221)	
長	basic_demo_length	rgb(142, 68, 173)	

A **tag** is a marker telling the meaning of a segment of text. Eg., names of temple, built date, built by whom, ...





## Automatic tagging with SmartRegex

- Useful when the text segments you wish to tag have literal patterns
  - Eg. 社稷壇、先農壇、關帝廟、城隍廟...
  - Observed pattern: **two characters** followed by **壇** or **廟**
- SmartRegex: a tool that helps to compile regular expressions in a more intuitive way
  - Choose the Topic of “Basic Demo 1”
  - Under Smart Regex, click “Load”.  
Choose “Miao” from existing Regex.  
Click “Compose”.

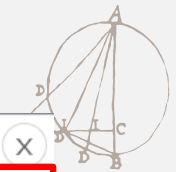
Smart Regex©

壇廟

Load Compose Clear

☒ Without line break  
☐ With a preceding line break  
☐ With a succeeding line break

Apply



Smart Regex© composing editor

+ Add

🗑

⬇ Load

壇廟

✎

Components of a Regex

Miao

▶ trial apply

Actions ▾

Tagged 154 entities!

\* Name: 壇廟

+ Word list

⬇ Load

[^○>]{1,4}[廟壇]

?

\* Tag as: 壇廟

▶ trial apply

Actions ▾

Tagged 129 entities!

【309】

<List\_Start>

灋州在城南隋大將劉方始開此路置鎮守尋廢唐貞觀十二年李宏節遣首領寧師京循故道開置以達交趾因州界有灋水故名後廢。

南寧府志卷十三終。

。

南寧府志卷十四。

輿地志。

壇廟。祀田附。

宣化縣。附郭。

社稷壇在城東二里明嘉靖二十五年知府郭楠建久。

國朝雍正十年知府張漢知縣陳旭奉旨遵式改建。

先農壇在東門外楊梅嶺。

國朝雍正五年署宣化縣親寧州知州徐德秩遵式

<List\_Start>

【310】

<List\_Start>

建置耜田四畝九分祭器農具俱全。

Only one component in this case.

Its regex is: `[^○>]{1,4}[廟壇]`

It means : Any character **except** ○ and > **appears consecutively for 1-4 times**, followed by a character of 廟 or 壇.

Click on Trial apply to see result.





### Smart Regex© composing editor

+ Add
🗑️
⬇️ Load

\* Name: 
📖 Word list
⬇️ Load

壇廟

⬅️ [^O>]{1,4}[廟壇]

⬇️

Miao
▶️ trial apply
Actions ▾

\* Tag as: 
▶️ trial apply
Actions ▾

Tagged 154 entities!
Tagged 129 entities!

壇廟。祀田附。  
 宣化縣。附郭。  
社稷壇在城東二里明嘉靖二十五年知府郭楠建久圯。  
 國朝雍正十年知府張漢知縣陳旭奉旨遵式改建。  
先農壇在東門外楊梅嶺。  
 國朝雍正五年署宣化縣親寧州知州徐德秩遵式  
 <List\_Start>  
 【310】  
 <List\_Start>  
 建置藉田四畝九分祭器農具俱全。  
 風雲雷雨山川城隍壇在社稷壇後明嘉靖二十五年知府郭楠建。  
 國朝雍正十年知府張漢知縣陳旭遵式改建。  
厲壇在城北二里。  
關帝廟在府治西城隍廟大門內宋建明正德十六年有司概以淫祠毀之嘉靖八年都御史姚鏌命知府蔣山卿重建。  
 國朝雍正三年奉旨追封三代公爵遵。  
 。  
 欽定封號設主後殿奉祀又一廟在城外鎮北橋。  
城隍廟在城內府治西街廟基高聳爲五花嶺之一峯明洪武初建宏治十五年知府劉芳修嘉靖二十八年知府王貞吉復修後燬。

Close the composing editor

Immediately see the result of trial apply:  
helps “debugging” and refine the Regex

壇廟○祀田附○

宣化縣○附郭○

社稷壇在城東二里明嘉靖二十五年

國朝雍正十年知府張漢知縣陳旭奉

先農壇在東門外楊梅嶺○

國朝雍正五年署宣化縣親寧州知州

<List\_Start>

【310】

<List\_Start>

建置藉田四畝九分祭器農具俱全○

風雲雷雨山川城隍壇在社稷壇後明嘉靖二十五年知府郭楠建○

國朝雍正十年知府張漢知縣陳旭遵式改建○

厲壇在城北二里○

關帝廟在府治西街廟基高聳爲五花嶺之一峯明洪武初建宏治十五年知府劉芳修嘉靖二十八年知府王貞吉復修後燬○

國朝雍正三年奉旨追封三代公爵遵○

○

欽定封號設主後殿奉祀又一廟在城外鎮北橋○

城隍廟在城內府治西街廟基高聳爲五花嶺之一峯明洪武初建宏治十五年知府劉芳修嘉靖二十八年知府王貞吉復修後燬○

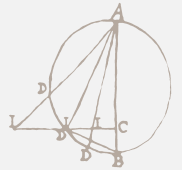
國朝康熙十二年知府韓章通判顧鼎植重建有碑記五十七年知府黃之孝修郡紳李雲燦記○祀田附一卜東鄉陳誠戶屯糧米二石計租一十六坵塘一口坐落朝天張村每年收租穀七十二斗○一下東鄉朱子柏蔣義夏人三戶屯糧米二十二石五斗計田八十坵坐落奇羅村每年收租穀四十三石○一下東鄉蕭繼先戶屯糧米九斗計田二十六坵坐落三义村每年收租穀七石○一上東鄉何七戶屯糧米三石五斗三升一合九勺計田一十六坵坐落站墟陳村

Step 2: Go through the text to check whether everything tagged is correct, and everything should be tagged is tagged. If not, manually tag/remove them.

Step 1: Apply

Step 3: Create task

Load the other Smart Regex named “where”.  
It contains two components: 在 and 哪裡.



Click Apply and see how the text changes color.

南寧府志卷十四。

輿地志。

壇廟。祀田附。

宣化縣。附郭。

社稷壇在城東二里明嘉靖二十五年知府郭楠建久圯。

國朝雍正十年知府張漢知縣陳旭奉旨遵式改建。

先農壇在東門外楊梅嶺。

國朝雍正五年署宣化縣親寧州知州徐德秩遵式

<List\_Start>

【310】

<List\_Start>

建置藉田四畝九分祭器農具俱全。

風雲雷雨山川城隍壇在社稷壇後明嘉靖二十五年知府郭楠建。

國朝雍正十年知府張漢知縣陳旭遵式改建。

厲壇在城北二里。

關帝廟在府治西城隍廟大門內宋建明正德十六年有司概以淫祠毀之嘉靖八年都御史姚鏌命知府蔣山卿重建。

國朝雍正三年奉旨追封三代公爵遵。

。

欽定封號設主後殿奉祀又一廟在城外鎮北橋。

城隍廟在城內府治西街廟基高聳爲五花嶺之一峯明洪武初建宏治十五年知府劉芳修嘉靖二十八年知府王貞吉

復修後燬。

Topic : Basic Demo 1 (基本示範第一...

Main tag: 壇廟

▼ Edit, Tag, Save text

Tagging Editing

Undo changes

Smart Regex©

在 哪裡

Load

Compose

Clear

☒ Without line break

☐ With a preceding line break

☐ With a succeeding line break

Apply

Preview

☐ embedded

Create task

> Replace By Regex

> Export

社稷壇在城東二里明嘉靖二十五年知府郭楠建久圯。

國朝雍正十年知府張漢知縣陳旭奉旨遵式改建。

先農壇在東門外楊梅嶺。

國朝雍正五年署宣化縣親寧州知州徐德秩遵式

<List\_Start>

【310】

<List\_Start>

建置藉田四畝九分祭器農具俱全。

Go through the text and manually correct missing and wrong ones.

風雲雷雨山川城隍壇在社稷壇後明嘉靖二十五年知府郭楠建。

國朝雍正十年知府張漢知縣陳旭遵式改建。

厲壇在城北二里。

關帝廟在府治西城隍廟大門內宋建明正德十六年有司概以淫祠毀之嘉靖八年都御史姚鏌命知府蔣山卿重建。

國朝雍正三年奉旨追封三代公爵遵。

Click Preview

Topic : Basic Demo 1 (基本示範第一...

Main tag: 壇廟

▼ Edit, Tag, Save text

Tagging

Editing

Undo changes

18

Smart Regex©

在 哪裡

Load

Compose

Clear

☒ Without line break

☐ With a preceding line break

☐ With a succeeding line break

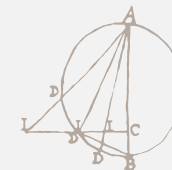
Apply

Preview

Create task

☐ embedded

Count	Book ID	Book Name	Section ID	Section Name	Book Year	Edition Year	Place Name	Longitude	Latitude	Page No.	壇廟	性別
1	H008D	南寧府志	37481114739796	壇廟 祀田附	1741	1742	南寧府	108.312073	22.806145	309	社稷壇	城東二里
2	H008D	南寧府志	37481114739796	壇廟 祀田附	1741	1742	南寧府	108.312073	22.806145	309	先農壇	東門外楊梅嶺
3	H008D	南寧府志	374811147397	Preview the table in a separate browser window.						310	風雲雷雨山川城隍壇	社稷壇後
4	H008D	南寧府志	37481114739796	壇廟 祀田附	1741	1742	南寧府	108.312073	22.806145	310	厲壇	城北二里
5	H008D	南寧府志	37481114739796	壇廟 祀田附	1741	1742	南寧府	108.312073	22.806145	310	關帝廟	府治西城隍廟大門內
6	H008D	南寧府志	37481114739796	壇廟 祀田附	1741	1742	南寧府	108.312073	22.806145	310		城外鎮北橋
7	H008D	南寧府志	37481114739796	壇廟 祀田附	1741	1742	南寧府	108.312073	22.806145	310	城隍廟	城內府治西街廟基高
8	H008D	南寧府志	37481114739796	壇廟 祀田附	1741	1742	南寧府	108.312073	22.806145	311	旗纛廟	倉西門外三里
9	H008D	南寧府志	37481114739796	壇廟 祀田附	1741	1742	南寧府	108.312073	22.806145	311	三界廟	鎮江門內一新廟;倉西門外面大江與
10	H008D	南寧府志	37481114739796	壇廟 祀田附	1741	1742	南寧府	108.312073	22.806145	311	東嶽廟	北門內明洪武十

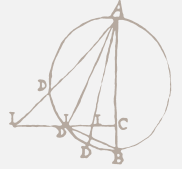


## Example on tagging text search result

- Schaefer, Chen, Che. JCH special issue
  - Used 13 keywords (OR) to search for mulberry disaster records. Read all the records and verified half of them.
  - Tagged dates, related disasters, expression, etc.
  - Add geo-coordinates, visualized on map and by timeline.
  - Half of the mulberry disasters took place in Yuan. Does it imply that Yuan had more natural disasters? No!

### 【538】

嘉泰十五年宋大名忠義彭義斌復京東州縣  
金興定二年廢東明縣為通安堡  
正大元年東明王鶚進士及第是年曹州渰廢乃割東明隸○○曹州  
○九年罷通安堡為縣改名儀封仍割蘭陽六鄉為蘭○○陽縣○元世宗  
至元元年置白馬東明濬州等綱  
○二十二年秋河水○○害田  
○二十五年河北徙東明曹濮等處被其害  
元貞二年水○六月蝗○八日旱蝗○  
大德元年大名路旱  
○三年罷大名路黃河故道田所輸租  
○○五年四月蟲食桑○○[duplication]  
六月旱○  
七月復○水  
至大元年五月○○蟲食桑○

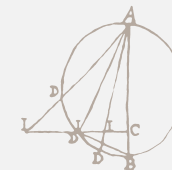


## LoGaRT's tagging is designed for sharing

- Topics, tags, and regexes are **shared among all users**
- They can be **shared among different texts**. Regexes can be modified to fit slightly different expressions (without saving).
- All tagged texts can be found via the Tasks module (top menu). Tasks by default can only be seen by the creator, but can also be shared with other users (by adding **contributors** under Tasks->Manage Task).
- LoGaRT keeps **versions** of a task (click Save text). Older versions can be retrieved (via Manage Task).
- Exported spreadsheets from the same Topic can be merged by oneself, and can be brought to other software for analyses (Excel, GIS...).



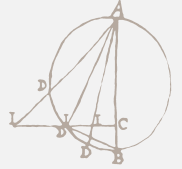
## Concluding remarks



## Summary

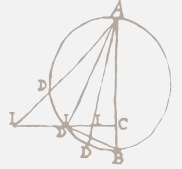
- Philosophy: Treating the whole genre as a conceptual database to ask questions at large scales
  - Using digital local gazetteers as a proxy; collection large enough?
- Data collection across time and space. Scale up research scope.
  - From text to data
  - Taking advantage of the structures in this genre
  - Via search or text-tagging
- LoGaRT provides alternative perspectives via visual / statistical tools; helps to discover global patterns and phenomena.
  - Inspires more questions
  - Data driven, empirical?





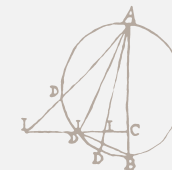
## Our observations

- Through the quick turn-around time between search and results, scholars can quickly test their hypotheses toward the genre (or a sub-collection) and adjust their understanding
  - Quick switch between distant and close reading; examine phenomenon and abnormalities
- Through text-tagging it helps scholars to pay more attentions to the structure behind the texts when reading
- Interpretation of data statistics/visuals: can be misleading. Rely on solid background knowledge and cautious analyses.



## Recent research results derived from LoGaRT

- Journal of Chinese History, special issue on Digital Sinology
  - A utility paper on LoGaRT, featuring the TU 圖 Group
    - Ken Hammond, Anne Gerittsen, Shellen Wu, Zhang Jiajing
  - A research paper "What is Local Knowledge"
    - Dagmar Schäfer, Che Qun, Chen Shih-Pei
- Joe Dennis, The Role of Donations in Building Local School Book Collections in the Ming Dynasty, Ming Qing Yanjiu 24 (2020)
  - Based on the dataset of school library holdings in LGs, 30K records



## Call for collaborations!

- LoGaRT is a suite of software, and can link to different collections of LGs
  - We hope database / content owners can consider linking their LGs with LoGaRT (in a secure way that protects contents)
- Research collaboration:
  - Use Open LoGaRT
  - Apply for visiting scholars / postdocs at MPIWG (to use the larger set)
  - Collaboration projects
    - HYL, CBDB (collecting posting data), Renmin U.

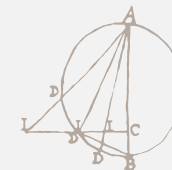


# Thank you!

Contacts:

Shih-Pei Chen <[schen@mpiwg-berlin.mpg.de](mailto:schen@mpiwg-berlin.mpg.de)>

Sean Wang <[swang@mpiwg-berlin.mpg.de](mailto:swang@mpiwg-berlin.mpg.de)>



## Existing digital lg

- Commercial databases
  - Erudition: 4,000 titles, all full-text searchable
  - Diaolong: 4,131 titles (partially with full texts)
  - EastView: 7,000 titles (partially with full texts)
- Public institutions
  - National Library of China: ~4,500 titles. 1/3 with full texts
  - Shanghai Library
  - Ctext, Kanripo: hundreds