

## 在项目中使用中国历代人物传记资料库

王宏甦

中国历代人物传记资料库项目经理

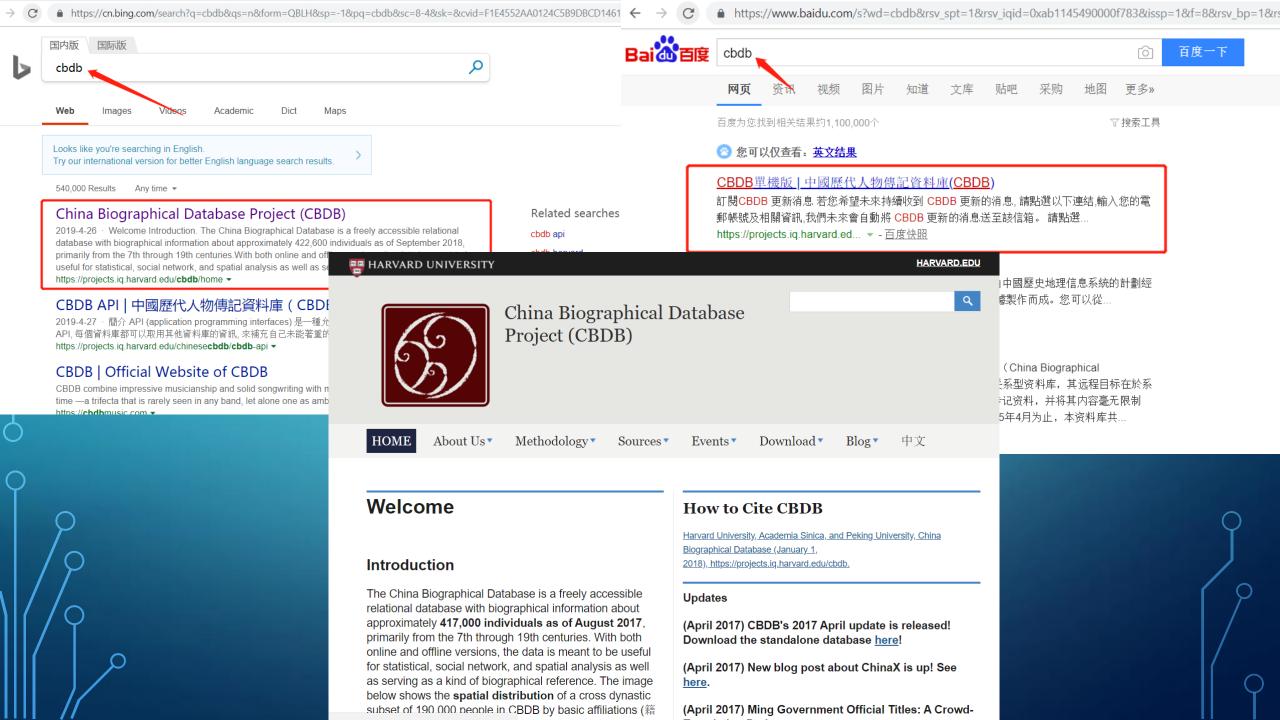
链接: http://t.cn/E9yM5bP 密码: nept

HOME / 關於我們 /

## 核心機構與編輯委員會

#### 核心機構

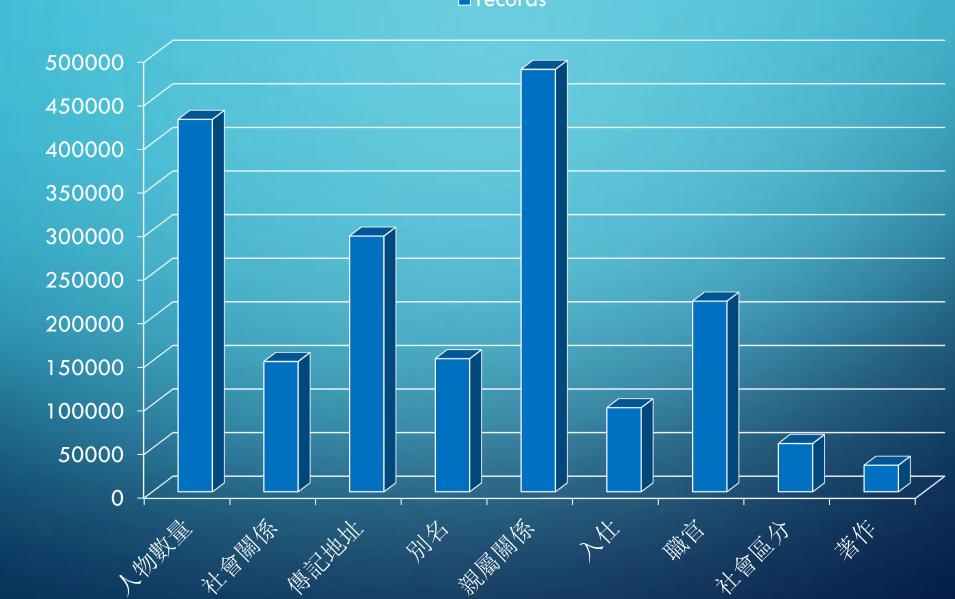
- <u>哈佛大學費正清研究中心 (Fairbank Center for Chinese Studies, Harvard University)</u>
- 中央研究院歷史語言研究所 (Institute of History and Philology of Academia Sinica)
- <u>北京大學中國古代史研究中心 (Center for Research on Ancient Chinese History at Peking University)</u>



## CBDB: 427,000 人物的分布时代

时代	数据量
唐 Tang	53607
五代十国 5 Dynasties	1248
宋 Song	50712
辽 Liao	324
金 Jin	530
元 Yuan	22438
明 Ming	185475
清 Qing	85745
民国 Minguo	3875
其他 Other	23093

# CBDB 2019.4 数据种类 records



## 中國歷代人物傳記資料庫

中央研究院歷史語言研究所、哈佛大學、北京大學中國古代史研究中心合作開發 Developed through collaboration among Academia Sinica, Harvard University, and Peking University

#### A Subset of the Data on Sima Guang:

有关司马光的部分数据:

Name	Dates	Offices	Associations
姓名	日期	任官	社会关系
Sima Guang 司马光	1019-1086	(1) 1059 度支勾院 Budget Auditor; (2) 1085 门下侍郎 Executive of the Chancellery; (3) 1086 左仆射兼门下侍郎 Left Executive, Dept of Ministries []	(九佑党); (2) An Dun 安惇 Desires opposed by; (3) Chao Buzhi 晁补之 Sacrificial prayer written by; (4) Chen Jian 陈荐 Sacrificial prayer written for; (5) Chen Min 陈敏 Honored by; (6) Cheng Yi 程颐 Recommended; (7) Ding Du 丁度 Sacrificial prayer written for; (8) Fan Chunli 范纯礼 Patron of; []

#### Reorganizing the Data (2nd Version) Database Normalization

#### 重新编排数据(第二版)数据库三范式

#### People 人物

ID	Name	Dates
1	Sima Guang 司马光	1019- 1086
2	An Dun 安惇	
3	Chao Buzhi 晁补之	
4	Chen Jian(5) 陈荐	
5	Chen Min 陈敏	
6	Cheng Yi 程颐	
7	Ding Du 丁度	
8	Fan Chunli 范纯礼	

#### Office Titles 官名

ID	Office Name
1	度支勾院 Budget Auditor
2	门下侍郎 Executive of the Chancellery
3	左仆射兼门下侍郎 Left Executive, Dept of Ministries

#### Associations 社会关系

ID	<b>Association Type</b>
1	Yuanyou coalition member (元佑党)
2	Desires opposed by
3	Sacrificial prayer written by
4	Sacrificial prayer written for
5	Honored by
6	Recommended
7	Patron of

#### **Postings Data**

#### 任官资料

Person ID	Office ID	Posting Date
1	1	1059
1	2	1085
1	3	1086

#### **Associations Data**

#### 社会关系资料

Assoc Type ID	Person ID	Assoc ID
1	1	-1
2	1	2
3	1	3
4	1	4
5	1	5
6	1	6
4	1	7
7	1	8

## CBDB 用户手册

https://projects.iq.harvard.edu/c bdb/supporting-documents https://projects.iq.harvard.edu/cbdb/supporting-documents



Project (CBDB)

HOME

About Us▼

Methodology ▼

Sources •

**Events** 

Downlo

#### **DOWNLOAD**

Download CBDB Standalone Database

**Supporting Documents** 

Presentations and Papers

Download Supporting Software

CBDBRegexMachine

Data Sets

CBDB API

Report an Error

HOME / DOWNLOAD /

#### **Supporting Documents**

**Supporting Resources** 

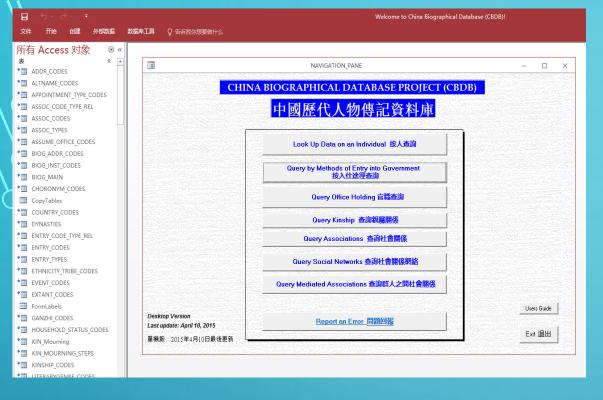
**CBDB Videos** 

The CBDB User's Guide, February 28 2015

The latest User Guide. You can find instructions for using the stadocument.

The CBDB User's Guide [Chinese], 2017

Translated by CBDB editors in CBDB.



## 离线查询系统 在线查询、在线录入系统

#### 中國歷代人物傳記資料庫

中央研究院歷史語言研究所、哈佛大學、北京大學中國古代史研究中心合作開發 Developed through colleboration among Academia Sinica, Harvard Liniversity, and Peking University

人名查詢 地名查詢 官名查詢 關鍵詞查詢 進階查詢 製壓/人際關係查詢 检索結果輸出 社會關係分析 操作說明 名詞說明





人名查詢可以根據姓名、指數年、朝代、地址等人物的基礎生平資訊,從 CBDB 中篩選出相應的人物。預設狀態下,查詢範圍包含 CBDB 中的所有人物,但您也可以通過調整左邊任一查詢欄位中的參數來進一步篩選查詢结果。

請進一步閱讀【傑作說明】中的"分類查詢"

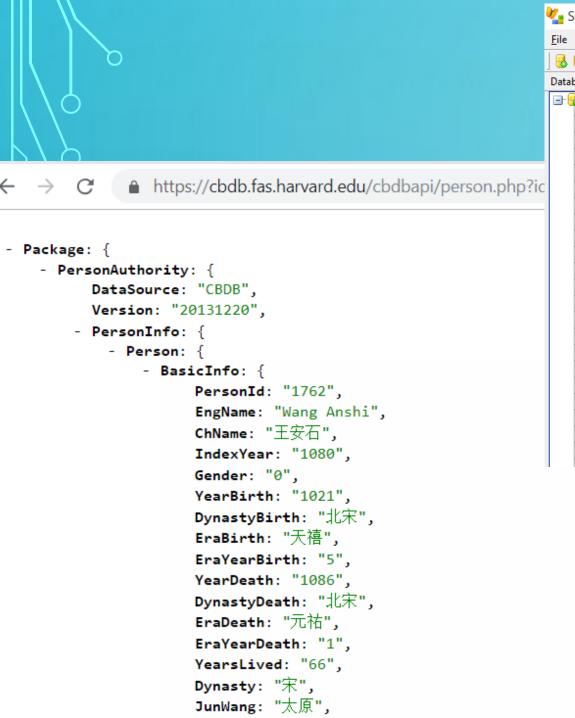
#### 範例

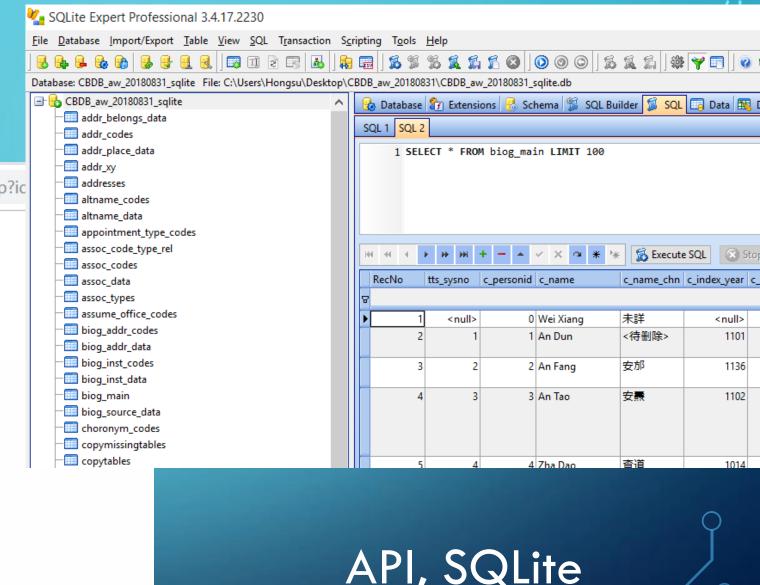
利用人物姓名中的關鍵字來查找人物。比如說,在**【姓名】欄**輸入 Wang Anshi (拼音)或王安石 (漢字),都將產生一筆查詢結果(截止至2011年二月份為止):十一世紀的王安石(Wang Anshi)。

如果只知道人物姓名中的一部份的話,您可以通過調整時間範圍來篩選查詢結果。比如說,<mark>维字在</mark>人名中很常見(截止至2011年二月份共有463例記錄),但當我們將查詢的時間範圍限定在唐朝時,查詢結果就會被縮減到5例,在**【朝代】**一欄的選單中選擇**唐**,則王維(Wang Wei) 將出現在查詢結果中的第五位。

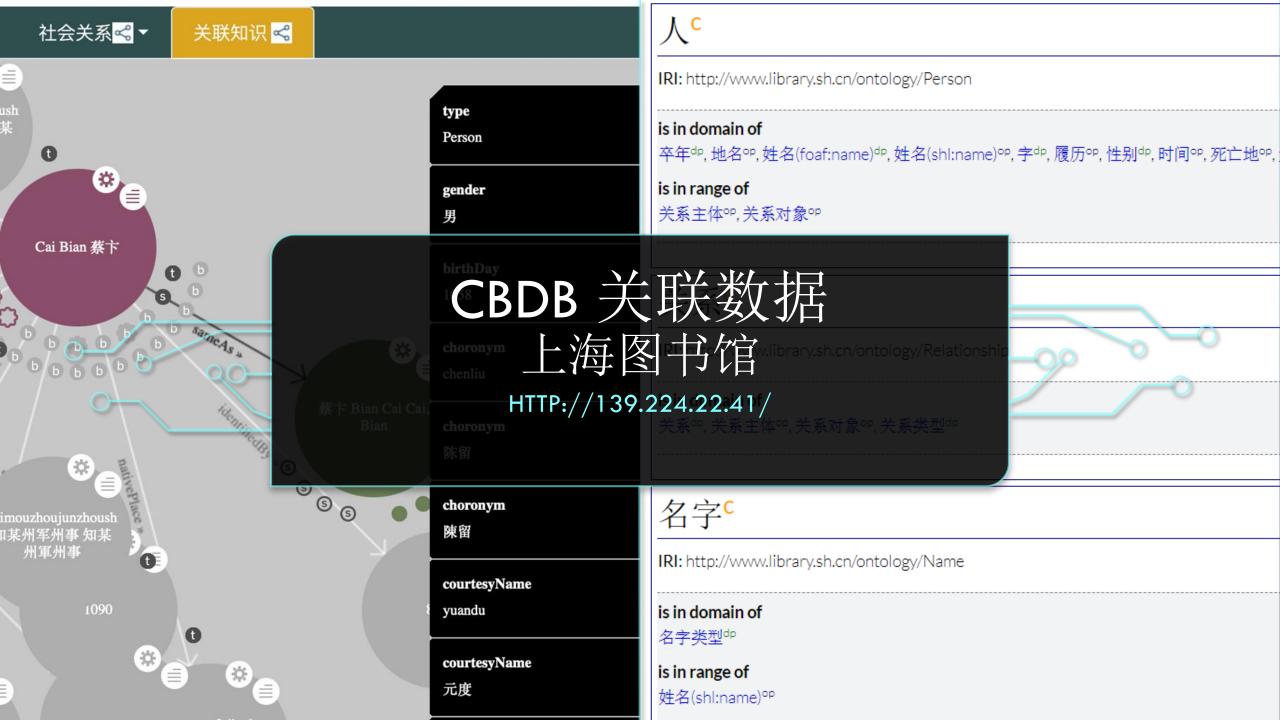
另一種常用的情况是,透過輸入人物姓氏來篩選出一批相同姓氏的人物。例如,在【姓名】欄中輸入<mark>獨孤</mark>兩字,將會產生54例結果。我們可以對這54例結果進行進一步篩選,使其只顯示活躍在北魏(5例)或者活躍在唐朝(32例)的人物。







### API, SQLite



## 开源录入系统

### 中央研究院数位文化中心、北京邮电大学

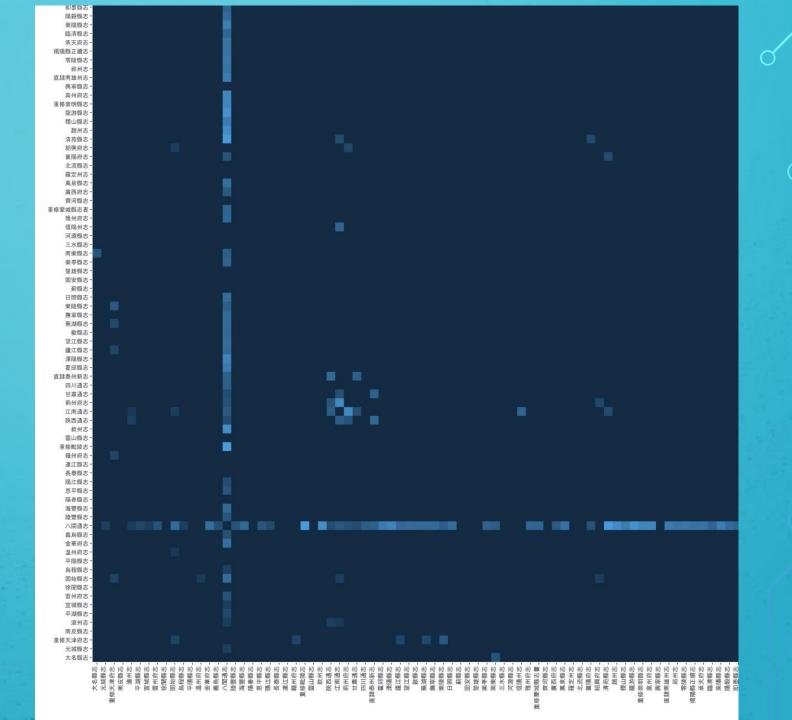
https://github.com/cbdb-project

develop	Fr	om: 8/10/2017 V To: 11/22/	2018 ~	)
Graph	Actions	Message	Author	Date
•	•	允许修改别名	funqc	10/24/2.
♦	<b>⊕</b> ♣	update	funqc	10/7/20.
<b> </b>	•	update	funqc	8/20/20.
<b>+</b>	•	Merge branch 'develop' of github	funqc	7/19/20.
	•	typo fixed	Hongsu	7/17/20.
	•	修改了部分tts依赖的模块	funqc	7/19/20.
<b>+</b>	<b>(i) (i)</b>	Merge pull request #4 from cbdb	funqc	7/15/20.
<b> </b>	<b>•</b>	Merge branch 'develop' into to5.5	funqc	7/15/20.
<b> </b>	<b>•</b>	增加出处	funqc	6/5/201.
<b>†</b>	•	update	funqc	6/5/201.
	<b>•</b>	Merge pull request #3 from cbdb	funqc	7/15/20.
	<b>•</b>	update	funqc	7/15/20.
	•	reroll to Qunchao's router version	Hongsu	6/6/201.
	•	reset default router for biog	Hongsu	6/6/201.

中國歷代人物傳記資料	<b>=</b>		
MAIN NAVIGATION	Basicinformation 编辑人物基本信息		
❷ 个人基本信息	386 ELY CLASS   1810		
	人名查询		
❷ 最近编辑列表			
CODES			查询人物
<b>S</b> Codes	Search		
❷ 地址编码表	共查询到422622条记录		
❷ 别名编码表	c_personid	c_name_chn	c_name
。 たみ米利伯可主	0	未詳	Wei Xiang
∞ 任命类型编码表	1	安惇	An Dun
❷ 著作编码表	2	安邡	An Fang
❷ 行政單位等級编码表	3	安燾	An Tao
♥ 11以半世寸級無時化	4	査道	Zha Dao

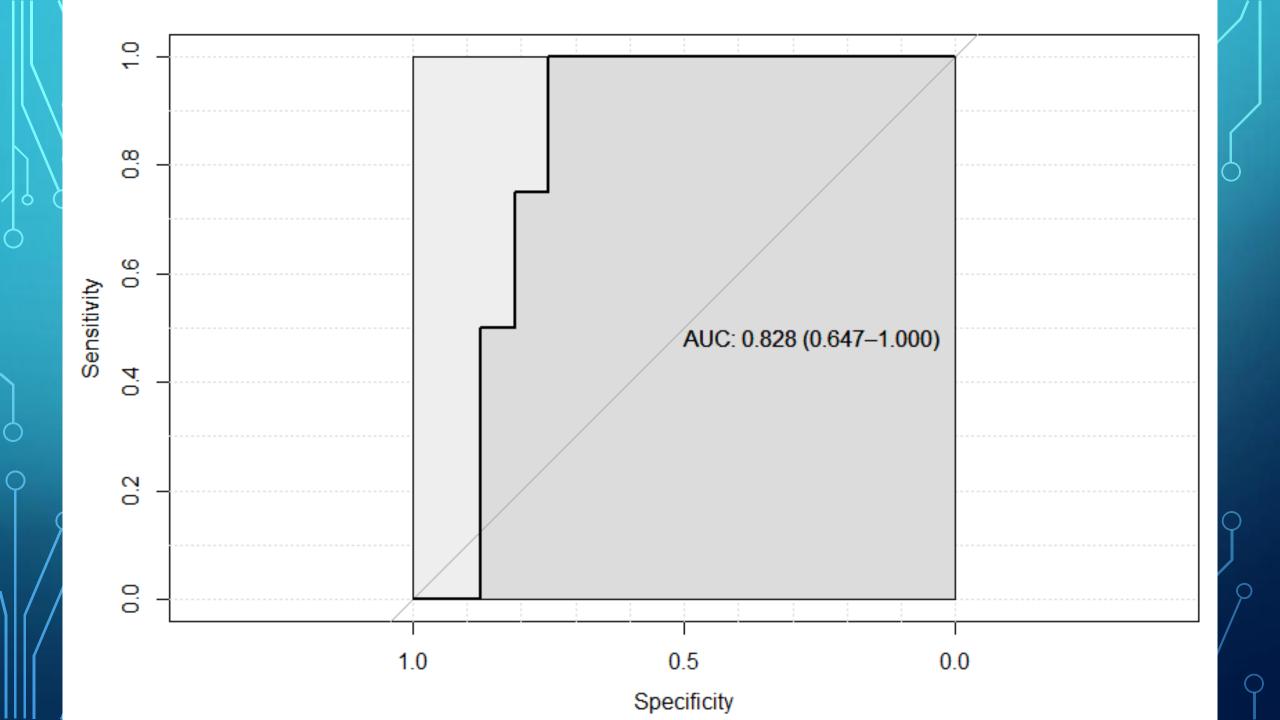
方法论与项目示例
METHODOLOGIES AND SAMPLES

## PROSOPOGRAPHY 群体传记学



张三千@哈佛统计系

		overlap	count.adj	count.adj	percentage	percentage		Y: overlap
Book1	Book2	7	.1	_	overlap	overlap		N:separate
							86.6449511	
平湖縣志	平湖縣志	266	307	2688	86.64495114	9.895833333	4	·Υ
							85.0806451	
元城縣志	大名縣志	422	496	1766	85.08064516	23.89580974	6	Υ
宣城縣志	宣城縣志	462	583	550	79.24528302	. 84	84	·Y
	重修天津府						83.6363636	
滄州志	志	506	605	7037	83.63636364	7.190564161	4	Y
							80.0366300	
鳥程縣志	烏程縣志	437	688	546	63.51744186	80.03663004	4	Υ
							79.9212598	THE PARTY
固始縣志	固始縣志	203	254	449	79.92125984	45.21158129	4	Υ
雷州府志	徐聞縣志	297	2019	432	14.7102526	68.75	68.75	Υ
							66.8539325	
重修天津府志	南皮縣志	357	7037	534	5.073184596	66.85393258	8	Υ
溫州府志	平陽縣志	485	3578	811	13.55505869	59.8027127	59.8027127	Υ
義烏縣志	金華府志	450	866	1771	51.9630485	25.40937324	51.9630485	Υ
							34.1059602	
海豐縣志	陸豐縣志	103	632	302	16.29746835	34.10596026	6	Υ
							27.5362318	
長泰縣志	八閩通志	133	483	14762	27.53623188	0.900961929	8	Υ
							26.6666666	
八閩通志	漳平縣志	24	14762	90	0.162579596	26.66666667	7	Υ
							21.2121212	<b>7</b>
八閩通志	內黃縣志	28	14762	132	0.189676196	21.21212121	1	N

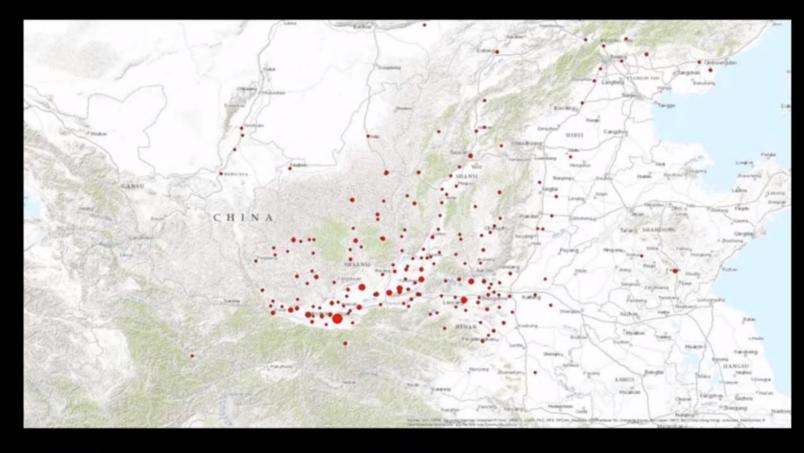


## GIS: 唐代的两京走廊

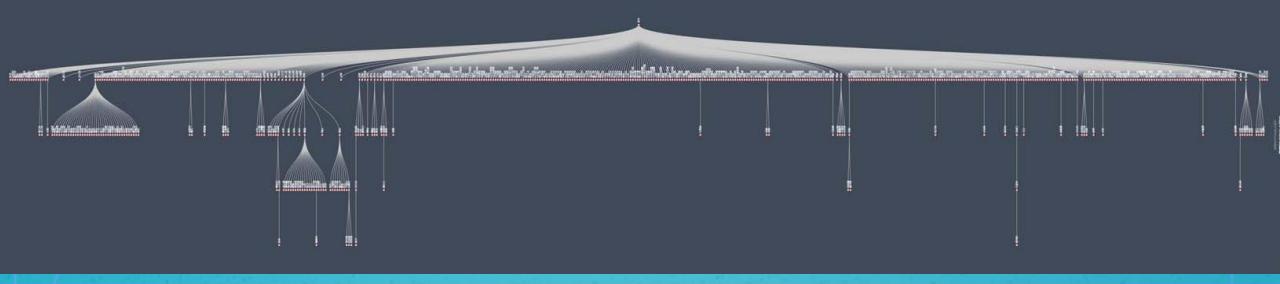
文欣 @普林斯顿东亚系

https://www.youtube.com/watch?v=S 3afqSFvNYw&list=PLGgLzyv7BMga2 4KUn dHULCYKY Te2umr&index=3

#### From Capitals to "Capital Corridor"



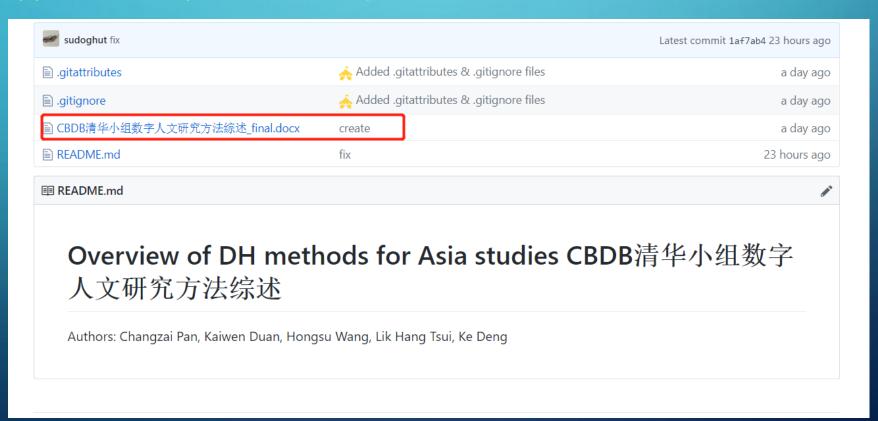


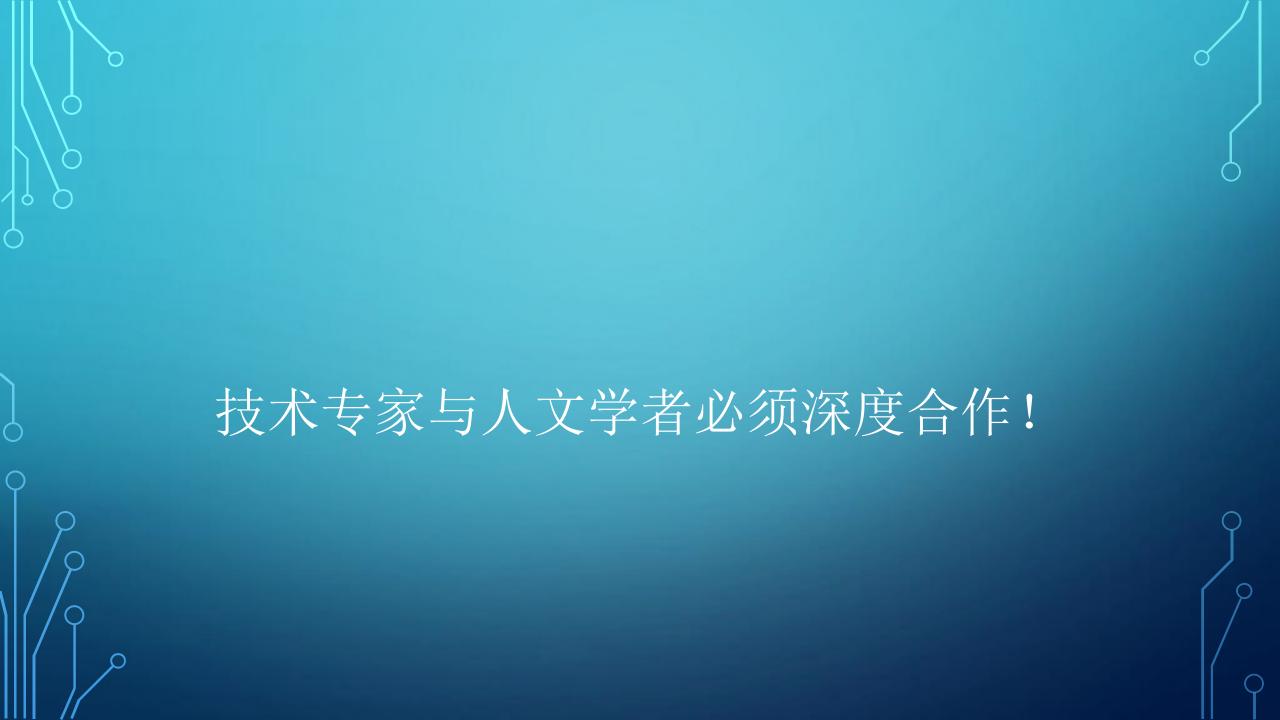


邱勇@北京大学信息管理系数字图书馆实验室http://dh.kvlab.org/cbdb\_vis/part2\_dizinet.html

## CBDB 清华小组数字人文研究方法综述

https://github.com/cbdb-project/overview-of-DH-methods-for-Asian-studies





## 技术专家所忽视的: 唐代人物迁徙

- 地点(橙色表示死亡人数>出生人数,蓝色反之)
- 🚃 迁徙路线 (生地->卒地)

A network framework of cultural history (2014/8)
Ohttp://science.sciencemag.org/content/345/6196/558



## 人文学者所忽视的:父子、爷孙籍贯的传递性

∠ c_personi	BIOG MAIN	ADDR COD	ES.c_r ▼	c_kin_code 🕶	c_kinrel_chn	Ψ.	c kin id 🔻	BIOG MAIN	→ ADDR COD	ES_1.( +
	3 安燾	開封		180	子		3000	安扶	開封	
	3 安燾	開封		180	子		3001	安日華	開封	
	4 查道	如皋		180	子		13314	查拱之	如皋	
	4 查道	如皋		180	子		119821	查循之	如皋	
	4 查道	如皋		180	子		16338	查塾	如皋	
	4 查道	如皋		180	子		13317	<b>查應辰</b>	如皋	
	4 查道	如皋		180	子		13315	查慶之	如皋	
	4 查道	如皋		180	子		13316	查沖之	如皋	
	. <del>*</del> \ <del>*</del> \	1			7			A +++		,

Peter Bol, Michael Fuller, CBDB清华小组

父子籍贯的传递几率是 92.4%,爷孙籍贯的传递几率是 66.9%

# 多统计学语言/模型: 概率与似然 STATISTICAL LANGUAGE/MODEL: POSSIBILITIES AND LIKELIHOOD

- <Grandfather, Grandson>
- =  $\Sigma$  P(<Grandfather, Son>| $\rho$ )
- = 66.9%

- <Grandfather, Grandson>
- =  $\Sigma P(\langle Father, Son \rangle | \rho) P(\langle Father, Son \rangle | \rho)$
- $= \Sigma P(\langle Father, Son \rangle | \rho)^2$
- $792.4\%^2 = 85.3\%$

## THANKS!

HONGSUWANG@FAS.HARVARD.EDU