

觀念史研究與數據庫的建立和應用

香港中文大學 劉青峰 教授

DADH2010

建庫三階段與體會

第一階段：嚐試

第二階段：建庫與研究的互動過程

第三階段：研究中尋找更有效的數位工具
「數據庫的基本功能」

DADH2010

「數據庫」的基本功能

1. 快速生成按年代使用次數排序之關鍵詞數量統計圖表
2. 快速給出文獻中所有作者使用、或各篇文獻使用該關鍵詞的次數排序統計表
3. 按每年文字總量的不同，自動給出一個系數求出該詞每年使用的平均值，消除由於各年總字數不等帶來的偏差
4. 可用or或and兩種方式考察兩個關鍵詞的關係。
5. 在研究者做出關鍵詞的不同意義類型以及作者在使用該詞時的主觀評價（0; +; -）的標識後，自動生成按年代對該詞的不同意義及評價之分類統計句庫；
6. 針對中文特有的嵌詞現象，例如，「人民主張」短語中，嵌入了「民主」一詞，並不是研究者找尋的關鍵詞；可以當作「噪音詞」處理，不列入檢索和統計；
7. 其他功能：如筆名管理，文獻瀏覽及原文圖像對照；查找大段相近文字的模糊檢索...等等。

「以含關鍵詞例句為中心的觀念 史研究」及其分析步驟

DADH2010

1. 找尋切合研究目標的關鍵詞

- 與現代漢語中不同，「民主」的古典意義是「民之主」，指皇帝；「主權」也不是指國家主權，而是皇帝的權力，「經濟」與本是儒學中指經世濟民之才能，「個人」在古詩詞中指所愛的人。
- 不能用今天的通用術語去查找相關事件，如要「甲午戰爭」，在1894、95年文獻中是使用「日清戰爭」、「黃海大戰」來指稱這場中日大戰的。

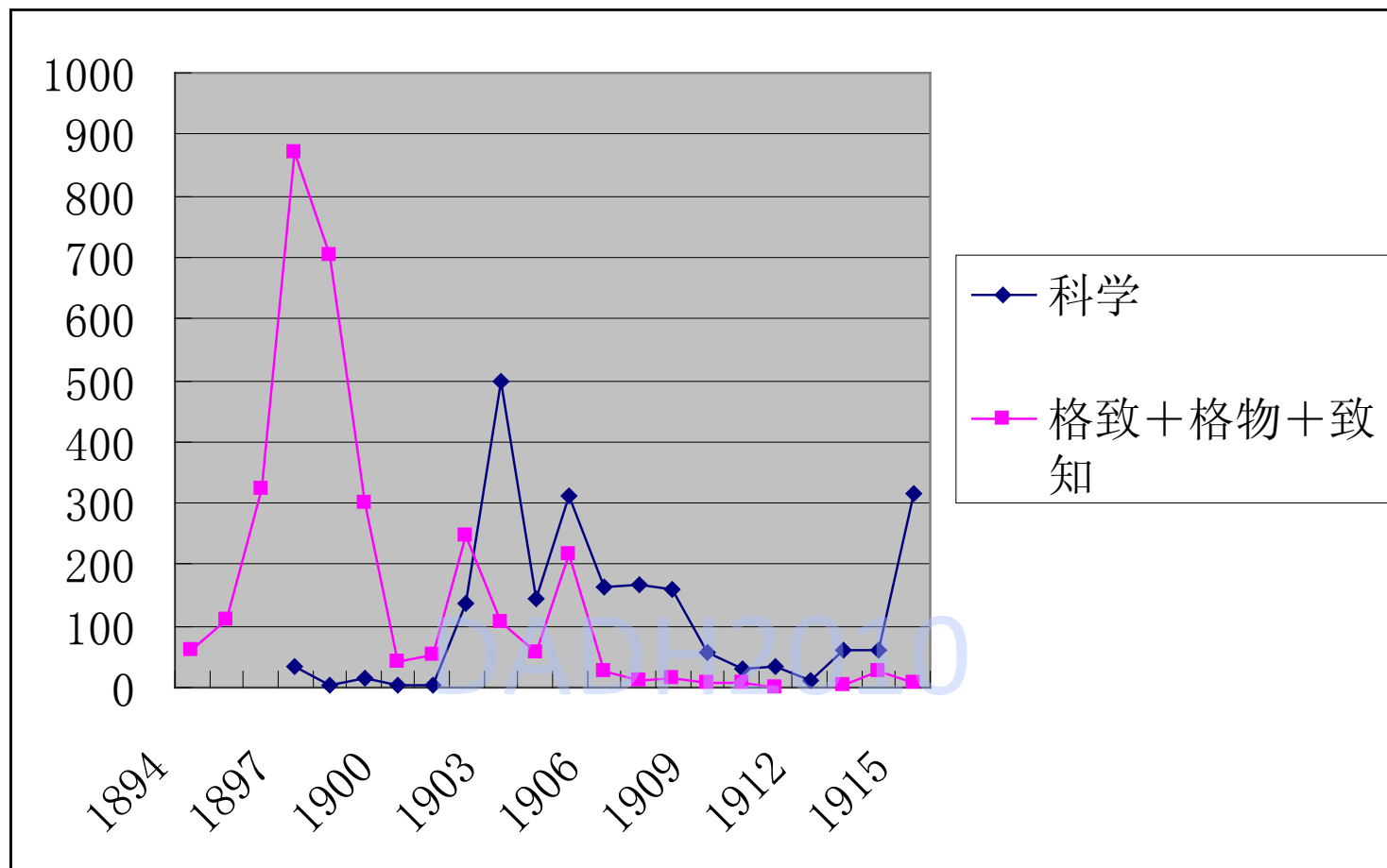
DADH2010

2. 選擇了關鍵詞之後，觀察該詞在歷史上使用時的宏觀走勢

- 以「科學」取代「格致」爲例，看宏觀統計對研究者的啓發。
- 「科學」的古典用法中是科舉學校的簡稱。如「自科學之興，世之爲士者往往困於一日之程文，甚至於老死而或不遇。」〔宋〕陳亮〈送叔祖主筠州高要簿序〉。

DADH2010

甚麼年代發生了「科學」取代「格致」呢？



3. 建立含對關鍵詞意義類型分析的句庫

以「民主」一詞爲例，四種意義類型。

1. 第一種，古典意義，由「民之主」來定義，指皇帝、帝王。「如，民主，天子也。」蔡邕註班固〈典引〉中「肇命民主，五德初始」
2. 第二種，現代意義，與古典意義相反，是「民作主」。
3. 第三種，是指實行君權有限制的政治制度的國家(民主之國)、法律(民主法)、組織(民主黨)等。最早，1864年丁韜良主譯《萬國公法》介紹西方有一類「君權有所限制」的國家，任何法令的實施必先由「部院會議，議定後，其君方能施行；民主之國，多由長老院，同議同准。」
4. 第四種，指民選國家最高領袖。如，「北美洲民主林肯，與南美連兵數年。」(唐才常，〈交涉學第四〉)「華盛頓拿破侖，民主中之傑出者也」(王仁俊，〈實學平議〉)

與中國君主世襲相反，由選舉產生。林樂知，《萬國公報》1874年12月〈選舉民主〉一文中說「美國民主，曰伯理璽天德，自華盛頓爲始已百年矣。例以四年換舉」

「民主」四種意義類型(1864_1915)

分類	君主			民爲主或人民統治			和世襲君主制相反的政治制度			民選領袖			合計
	+	0	-	+	0	-	+	0	-	+	0	-	
1864								18					18
1867					1			3					4
1873								2			1		3
1875					5			28					33
1876					1			5					6
1877					1			6	1				8
1878							2	11	1		1		15
1879							5	3					8
1880								9			1		10
1881								10					10
1882								21					21
1883								10					10
1884								3					3
1885								23	2				25
1886		1			1			22					24
1887					2			22		1			25
1888								9					9
1889					6			2					8
1890								16					16
1891								7			2		9
1892		4			1			20					25
1893								48					48
1894								14			1		15
1895								117			39		156

分類	君主			民爲主或人民統治			和世襲君主制相反的政治制度			民選領袖			合計		
	+	0	-	+	0	-	+	0	-	+	0	-			
1896						12		1	11	72	3		7	106	
1897			5					64	29	155	7	2	13	275	
1898								66	20	84	22		9	201	
1899			2					12	60	5			31	110	
1900						6			2	30	11		1	50	
1901						30			15	87			1	133	
1902						1	15		6	211			3	236	
1903						52		10	34	153			8	257	
1904							10	11	9	411			1	442	
1905						16	8	20		77	1		1	22	145
1906						58	52	54		114	1		2		281
1907						13	19	3		49	2		1		87
1908						7	12			40					59
1909									22		132				154
1910						8		3		78	1		1		91
1911								31		9	2		9		51
1912								10		43					53
1913								2		186			1		189
1914										39				5	44
1915						10	24	1		157				33	225

數據庫方法對人文研究的重要功能

1. 可檢驗性
2. 開放性

DADH2010

客觀性

- 相對於以往思想史的文本分析，不同研究者在判定句中關鍵詞意義時，能取得相一致看法。
- 例[句：「適值美國更換民主之時，美商以公舉總統兩人。」中「民主」是何意義？《皇朝蓄艾文編》卷三十三「稅則一」(西人葛顯禮，〈光緒二十二年通商各口華洋貿易情形總論〉)，

DADH2010

一、簡單檢驗：

- 佐藤慎一曾指出，20世紀初「『專制』這一詞語本身在十年前也不存在。『專制』和其他許多專門用語一樣，都是20世紀初從日本引進來的語言。」(《近代中國的知識份子與文明》)查數據庫，1893年鄭觀應：「論者謂變法之易，莫如專制政治」《盛世危言》，已在現代意義上使用「專制」一詞，時間顯然早於佐藤慎一說的二十世紀初。

DADH2010

二、考察某一著名、已成定論的論斷是否成立

- 例一：「科學、民主是新文化運動的兩面大旗」。
- 例二：「十月革命一聲炮響，給中國送來了馬克思主義」。

DADH2010

期待

1. 期待人文研究者能更密切地與電腦工程師合作。
2. 期待不同學科、不同專業學者共同使用數據庫。
3. 期待中、日、韓三國各自建立本國的思想史研究數據庫，並在此基礎上整合為更大型的數據庫。

DADH2010

感謝聆聽！

DADH2010