

本文為109-1學年度國立政治大學宗教研究所[數位人文研究導論]期末報告

## 摘要

本篇文章著重於八字研究中重要的兩本著作：韋千里（1911-1988）的《八字提要》（1946）和徐樂吾（1886-1949）重新校正刊刻古書後再出版的《窮通寶鑑》（1937）。透過數位分析工具：詞頻分析、關鍵詞標記以及相關有利於整理之數位工具，將這兩個文本之間所出現、經常使用的「正負向詞」進行整理與分析，再進一步比對兩文本之間的命例分析風格差異。最後再總結數位分析工具應用於命理文本分析的優勢與限制。

## 文獻回顧

用電腦分析文學文本是近年來嶄新的技術，王汎森認為電腦分析文學文本主要有幾項優勢：典藏上的優勢、速度上的優勢、數量上的優勢、交流上的優勢、教育上的優勢等<sup>[1]</sup> [<https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt1>]。例如可以跨越地域性、館藏限制等，獲取過去研究者可能不易取得的文本文獻資料，例如在日本國會圖書館調閱到明清時代的中文史料等；在檢索調閱彈性變得更靈活迅速，電腦可以短時間處理與掃讀大量文本或文字，擷取特定關鍵字並統計出數量與標定出章節位置，這些都是人類無法短時間超越的能力；這些研究成果除了文字期刊外，還有許多像是3D立體圖像與空間呈現、多媒體影音等，還能將這些數位成果分享、共享至跨越國界的其他研究學者群。

使用數位工具進行人文社會學科之研究，是近年來一大利器，尤其是跨越物理屏障，從其他國家的漢學典藏資料庫迅速擷取並獲得善本書文獻，是筆者感受到最強大、最便利之所在。然而，要透過數位人文工具進行文本分析時就會遭遇到各種挑戰與限制，例如電腦雖然可以迅速擷取並統計文本中出現的關鍵字詞數量，但電腦判讀語句的方式與人類閱讀理解語句的方式完全不同，人類判讀可以透過個人對於語句詞彙的抽象感受（口氣轉折、譬喻、暗示、反諷等）來獲取該與句段落的正向與負向感受，然而電腦無法進行這樣的抽象解讀感受，而是必須要透過「關鍵詞數量」來做出量化統計的數據結果呈現。在隨後文章將逐步討論此問題並嘗試克服。

朗宓謝（Lackner，2017）提到民國初年的命理師面對「五四運動」當時知識份子對於命學、易學為迷信的反對聲浪下，命理師們該如何回應<sup>[2]</sup> [<https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt2>]？韋千里支持並認為「命學應該做為國粹」的論述（pp. 469-471）。韋千里認為，要以「國學知識普及」的態度，將命理知識積極推廣普及給大眾。因此韋千里嘗試整理許多八字命理條例與規則，將過去古書中許多艱澀難解的文字語句和規則，簡化至一般大眾和初學者皆可輕易閱讀理解的程度。根據韋千里上一本著作《韋氏命學講義》（1934）的自序表示：《命學講義》書中規則是對於八字規則進行如同科學性、實驗性的書寫挑戰，建立於前人既有的八字規則上，並刪去許多被韋千里認為是迷信、過時、不規則且艱澀難解的古文詞句，韋千里希望這本書能對渴望學習命學研究的初學者有清楚的指引，因此嘗試用許多清晰簡單的語句來解釋八字規則（頁9）。

## 八字提要

韋千里《韋氏命學講義》的下一本著作即是《八字提要》。此書整理出版之用意，也看得出來是作為一本推廣給大眾的命學講義參考書籍。《八字提要》特色在於將個人的月、日、時分別以月地支、日主天干和時柱干支，從太陽曆的正月（寅月）排至十二月（丑月）；日主天干從甲到癸共十種、加上十二個月、一天共有十二時辰，因此全書共有1440組條目；韋千里在自序中提到：

「...，竊以八字云者，年月日時干支，共得八字也，日干為主體，月支為提綱，時辰為歸宿，看命先看此四字，此而知其輕重宜忌，...，爰寫茲『八字提要』，臚舉一千四百四十條，付抒管見...。」（頁一）

韋千里透過這些八字組合關係，再給予條目形式的評論，評論的模式大致上先從出生的月份與日主關係來判斷自身強弱、調候與用神等，再從時柱干支、月支藏干與日主的刑沖合會等組合關係進行判斷，找出格局與喜、忌用神、吉凶等平衡，由此就能看出一個人出生的該組「命盤四字」與命運發展狀況，舉例如下：

「甲日主生寅月甲子時，甲木得祿于寅月，又有時上甲比子印之生扶，強旺可知，惟初春餘寒猶盛，木甫萌孽，得火以暖之則繁榮，木旺見金自可貴，柱中有土斯有財。」（頁8）該八字結構示意如表一：

表一、韋千里《八字提要》之八字結構

八結字構	時柱	日柱	月柱	年柱
十神	比肩		○	○
天干	甲	甲	○	○
地支	子	○	寅	○
藏干十神	癸 正印	○	甲丙戊 比食偏 肩神財	○

甲日主生於正月（寅月）的甲子時，寅月代表初春，而時柱「甲子」中的「子」藏干為「癸水」適合澆灌給日主作為生長養分，月柱地支的「寅」藏干為「甲木、丙火、戊土」其中丙火能溫暖初春的寒冷土地，讓草木得以滋長，「戊土」在甲日主的十神關係又代表「偏財」因此韋千里才會說「有土斯有財」，因此從這樣的評論中可以得出，該組為命運發展良好的「命盤四字關係」。

《八字提要》像是綱要式的整理命盤與分析方法，提供給大眾、初入門或專業命理師作為工具書，可以透過客戶的出生月份、日與時辰，迅速翻書去找出相符合的「四字命盤關係」進行吉凶喜忌等解釋套用，為民眾與命理師省去大量需要推算、推敲判斷的時間、與減少可能計算錯誤的風險，並建立出具有明確條理的邏輯、如同自然科學原則等論述。而且書中用詞遣字盡量以淺白、譬喻等生動描述八字五行關係、季節氣候或其他所影響導致「身強/弱」或干支之間的刑沖合會...等關係之描述，較少直接寫出對於該人格局為「富貴上格」還是「貧賤下格」或只隱晦的談到此命盤的「命運結果」。《八字提要》只寫出八字中的其中四字分析，而不是全盤給出答案，可能是希望保留給命理師與命理學習者對於八字分析的彈性空間，也避免直接給出殘酷的結果，來傷害他人的自尊與信心。也有可能因為此書只有使用到「四字」不足以如「八字」給出完整的資訊，不應該直接給讀者如此嚴厲的用詞，對讀者或客戶容易造成誤解與傷害。

《窮通寶鑑》著書者為徐樂吾（1886-1949）。徐樂吾在自序以及次頁的說明中，對該書進行考究與說明；該書原名為《攔江網》並且未留下原始撰作者年代與姓名，但書中的八字命例多為明代官員，可以理解原書為明代之江湖術士所著書，再由清代的余春台（光緒年間人物）整理出版並命名為《攔江網》。徐樂吾將此書內容進行整理，刪除文中晦澀、字句重疊、闕漏誤字等大幅度的增減刪改等校正工作，並加入許多明清案例之分析，提高八字預測的可靠性和信用度，再將《攔江網》一名改為《窮通寶鑑》。

徐樂吾與韋千里在當時抱持相同的理念，積極將命理技術建立出具有可透過自然科學的邏輯性論述。徐樂吾在《窮通寶鑑》的自序中提到，許多科學性的發明或發現往往是在偶然幾次實驗過程或巧合下完成，人們可能遭遇到一些偶然現象發生，但當時可能認為原因尚未悉知，例如人們發現某些光波可以穿透不透明的固體，這種光波之後被稱為「愛克斯」（X射線，未知數之意）。因此，古人發展出十天干和十二地支的週期循環，卻可能不知其原理運作，因此《窮通寶鑑》將十天干與十二地支月份（日主屬性與季節氣候）整理成大綱，以便查找對於該客戶命盤的相關分析與資訊。

若同樣是「寅月甲日」的命盤，《窮通寶鑑》會怎麼解釋？《窮通寶鑑》主要是「調和派」風格，「調和」則是希望在特定季節、月份與其他五行的互動關係中，透過特定五行元素來進行平衡或解救。因此在書中第4頁「甲木喜用提要」提到「正月調候為丙癸：調和氣候為要，丙火為主，癸水為佐。」因為書中以八字全盤綜觀，因此五行調和失當的吉凶結果用詞遣字相當嚴厲，例如第6頁說道：「或一派庚辛，主一生勞苦，剋子刑妻。再支會金局，非夭即貧。」因為草木初春時不宜用金（庚辛申酉）制裁，若過度制裁且無法控制，將可能造成命主非貧即夭的嚴重後果。又「如無丙丁，一派壬癸，又無戊己制之，名水泛木浮，死無棺槨。」雖然調候需要「丙癸」但兩者需要並用，若只有「壬癸」代表水的力量過強，將會淹沒草木，又無戊己等土之元素固定草木的話，木頭就會被大水沖走，因此命格也是不佳。因此在《窮通寶鑑》的論述，會強調其五行平衡或調候得宜而成為達官顯貴或富貴豪門，或者是八字五行失調後的貧賤夭寒等悲慘狀況等結果。因此，該書的預設讀者，可以說是主要偏向於給命理師業內人士參考閱讀使用，而不適合給一般大眾閱讀，因為這樣嚴厲的措辭對於一般大眾或初學者容易造成誤解與傷害。

## 文本資料來源

《窮通寶鑑》在 Ctext（Chinese Text Project，中國哲學書電子化計劃 [3] <https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt3>）中有全文資料 [4] <https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt4>，但受限於授權單位無法直接導出檔案，因此要透過複製貼上將文本擷取出來，但該電子全文版本仍有一些字句上的錯誤與闕漏需要補正，例如在電子化版本中的標題為《窮通寶鑑》作者為「餘春台」作者姓氏就是很明顯的錯誤。《八字提要》網路亦有少部分正體中文的電子化文本（絕大部分都是簡體中文化版本，語系轉換後會產生更多字句上的錯誤）但是文本內部仍有少部分的排版問題，與文字缺漏部分；例如「己日主生 亥月 壬申 時 己本濕土，...，申乃壬??????金寒...。」問號部分需要回到實體出版文本中進行校正，得出「申乃壬之長生，水冷金寒...。」（146頁），因此兩文本都需要實體出版書籍，進行交叉比對與文本修正。這或許也是數位人文分析中經常出現的「小問題」。

編號	文件集名稱	文件總數	總長度	總字數	累計單字數	句長範圍	平均句長	句長標準差
1	窮通寶鑑	1	58,991	26,193	1,199	1 ~ 71	4.99	4.19
2	八字提要	1	135,853	112,338	1,994	1 ~ 21	6.33	2.87

表二、「Corpus Grams Tool」分析文獻列表

兩文本的基本元素，先以「DOCUSKY [5] <https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt5>」中的文獻集詞頻分析工具「Corpus Grams Tool [6] <https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt6>」進行分析，分析介紹如下：《窮通寶鑑》總字數約26,193字、八字提要約112,338字（如表二所示）。本次研究分析希

望獲取的「正、負向關鍵詞」嘗試去擷取出命理分析的「結果論斷」相關的字詞。例如「富貴、功名、貧賤」等最終結果，而不是分析過程或八字命理內所使用的「十神、十二長生」等專有名詞上。因此文本分析中需要排除在「十神 [7] [<https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt7>]」與「十二長生表 [8] [<https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt8>]」中所使用的特定使用關鍵詞，並盡量簡化、使用單一文字的擷取，以及注意正負向詞彙的「反用」。簡單來說，若要例如「有錢人」這樣的概念，首先要先排除在八字中「十神」會使用的「正財、偏財」之詞，因此要迴避並重新擷取出關鍵詞，例如採用「富、貴、名利」等。又「短命、促壽、殘疾」的概念需要避開「十二長生表」的「衰、病、死、絕」等字詞，因此需要重新思考關鍵字，例如改用「夭、亡、殘、疾」等。

關鍵字的過濾與定義，先透過「HTML5 Word Cloud」將全文進行詞彙的詞頻整理，《窮通寶鑑》詞頻（字詞頻率在文本重複出現頻率超過2次者）共跑出3679組，再排出在文本中字詞重複出現最高頻率的前20組（以下不列入）正負向詞，如表三：

表三、《窮通寶鑑》的正、負向詞數量

正向詞		負向詞	
頻率數量	字詞	頻率數量	字詞
109	富貴	46	常人
79	科甲	24	僧道
30	功名	23	平常
20	大貴	13	下格（扣除潤下格）
17	主富	14	孤貧
16	大富	14	賤之
14	富中取	13	貧賤
13	富中取貴	12	無用
13	有衣	10	下賤
13	異路功名	10	殘疾
13	科甲定	8	僧道之
13	雙全	8	貧人
12	名利	8	貧苦
12	秀才	7	勞碌
12	衣祿	7	富屋貧人
11	主富貴	6	下賤之
11	主科甲	6	奔波
11	富大貴	6	寒儒
11	恩榮	6	才多身弱
11	科甲定然	6	為僧道

另外再用「Corpus Grams Tool」設定「2 gram」（字元數=2）進行詞頻統計分析，也得到類似結果，由於系統只能顯示到前兩百組高頻率詞組，而負向詞組頻率與數量遠低於正向及一般用詞（在兩百名以外）因此不會被顯示於排行榜中，如表四：

表四、「Corpus Grams Tool」詞頻分析《窮通寶鑑》正、負向詞結果：

No.	Gram	TermFreq
6	富貴	109
12	科甲	79
22	常人	46
51	功名	30
80	僧道	24

85	平常	23
93	異途	22
98	異路	22
112	大貴	20
120	顯達	19
144	衣衿	18
157	主富	17
162	大富	16

可以看出在《窮通寶鑑》使用正向詞的數量遠高於負向詞，若從上表再進行正負向詞的篩選，並且精煉至一個單字的程度已進行關鍵字之擷取與統計，正向詞可以得到「富、貴、功、名、利、科、甲、恩、榮」等；負向詞「孤、貧、賤、殘、疾、僧」等。我們可以發現，為什麼「僧」在窮通寶鑑裡面算是「負向詞」而且頻率還有24組左右？若從《窮通寶鑑》的八字結構來看的話，「僧」大多是出現失衡、無制或缺乏特定五行元素所致，例如：

「三春甲木，或支成金局，多透庚辛，此又不吉，號曰木被金傷，若無丙丁破金，必主殘疾。或支成火局，洩露太過，定主愚懦，常有啣脚災病纏身，終有暗疾。支成木局，得庚為貴，無庚必凶，若非僧道，男主鰥孤。」

「三秋甲木，木性枯槁，金土乘旺，先丁後庚，丁庚兩全，將甲造為畫戟，七月甲堪為戟，非丁火不能造庚，非庚不能造甲，丁庚兩透，科甲定然。庚祿居申，殺印相生，運行金水，身伴明君。或庚透無丁，一富而已，主為人操心太重，不能坐享。或丁透庚藏，亦主青衿小富。或庚多無丁，殘疾之人，若為僧道，災厄可免。」

「四月丙火，或壬癸俱無，愚頑之輩，火炎無制，僧道之流。」

若從中國社會結構的預設理想條件（妻、財、子、祿、壽）來看的話，「僧道」等於放棄伴侶、錢財與傳宗接代等這些俗世追求，然而有時成為僧道反而可以避免災禍，因為不再追求俗世的價值標準和世間規則，若在俗世標準來看「孤殘」終身無伴侶或無勞動能力是有所缺憾的不足，但在僧道中，無伴侶屬於正常而且是必要的規則；殘疾之人也能在僧道中得到棲身之處，因此僧道雖然是等於放棄俗世規則的追求，但也讓這些不完美的人們有棲身之處。

那麼在《窮通寶鑑》所使用的關鍵字詞，同樣也適用於《八字提要》嗎？可能會有哪些問題？分析結果發現許多在《窮通寶鑑》常出現的關鍵字，在《八字提要》是相當少見、甚至完全不會使用到。而且《八字提要》進行詞頻統計，就跑出16596組字詞，其中統計字詞出現頻率，從最高頻率到重複20次以上的字詞，就已經有804組，而這800多組如何篩出正負向詞？一一檢閱相當耗時，因此嘗試以《窮通寶鑑》既有關鍵字，用「CTRL + F」搜尋關鍵詞數量並進行比對，對應整理列如表四。

《八字提要》的文字策略與使用與《窮通寶鑑》不太相同，例如「名」大多指的是「名稱為...」例如「名為截脚」、「名謂食神生財」、「名為籐蘿繫甲」<sup>[9]</sup> <https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt9>」而只有少數作「功名」用；另外「功」也是「功用」之意而非「功名」；「孤」並不是單純指「孤獨」還有五行中的「勢單力薄」例如「孤陽失輔」等。因此《八字提要》的用詞遣字策略與《窮通寶鑑》略有不同。《窮通寶鑑》對於該對象的真實狀況，與該對象的命盤分析結果，有相當明確的論斷（官員或科舉仕人）。然而《八字提要》僅是韋千里的對於八字內的其中四字，進行內部細節的技術討論，並且以隱喻、譬喻的方式進行描述，尤其強調五行間的結構平衡關係，並由此做出的推算整理；這些整理不見得有明確的命例對應與論斷結果。而且《八字提要》不重視最終個人命運的發展結果。

表五、《八字提要》正負向詞使用頻率（紅色字體為《窮通寶鑑》前三高詞頻詞）

正向詞		負向詞	
頻率數量	字詞	頻率數量	字詞

155	貴	48	孤
97	佳	46	枯
87	富	40	榮
87	秀	19	無用
83	功	18	貧
75	名	6	夭
65	美	4	凶
40	榮	4	鈍
38	中和	4	劣
22	富貴	3	賤
9	大富	3	疾
9	上乘	3	下格 (扣除潤下格)
6	名利	3	下命
4	錦	2	平常
3	大貴	2	亡
3	壽	1	貧賤
0	功名	1	僧
0	科甲	0	常人

若用「Corpus Grams Tool」進行「2 gram」(字元數=2)的詞頻分析，發現正、負向字詞，完全沒有在前兩百組排行榜內，而「1 gram」(字元數=1)的前兩百排行榜如同表五的排序一樣，但只能顯示到頻率數量91之前的單字，也就只有「貴(155)、佳(97)」兩字而已。

### 《窮通寶鑑》與《八字提要》相同命例對照

那麼再進一步問，《窮通寶鑑》裡面的好、壞命格，在《八字提要》也會有相同、或類似、甚至是相矛盾的解釋嗎？舉《窮通寶鑑》幾例嘗試分析：「丙寅、甲午、丁丑、乙巳。甲透庚得所，富貴極品。己丑相合，庚金暗藏也，亦以甲木引丁為用。」(116頁，如表六)

在《八字提要》提到相同命格者：「丁日主生 午月 乙巳 時 巳午氣成南方，時乙洩之於火，木燥火炎，宜以潤澤為先，仲夏丁火，用神不離壬癸之水，今柱中以火烈為病，故見水最宜，如有辰丑濕土之洩火尤妙。」而此八字的確有丑土(日柱地支)底下又有藏干癸水，因此這樣的八字的確可以算佳，但真的有到「極品」的程度嗎？韋並沒有給出絕對的論斷答案。

表六、《窮通寶鑑》116頁案例之八字結構

八結 字構	時 柱	日 柱	月 柱	年 柱
十 神	偏 印		正 印	劫 財
天 干	乙	丁	甲	丙
地 支	巳	丑	午	寅
藏 干 十 神	丙戊庚 劫傷正 財官財	己辛癸 食偏七 神財殺	丁己 比食 肩神	甲丙戊 正劫傷 印財官

假若用《窮通寶鑑》所提到的「貧、賤」等案例，能否反推至《八字提要》得到相同解釋？舉其中一個案例：「辛卯、辛丑、丁卯、甲辰。柱無庚丙，乙木寒濕，至乙運身死。才印交差，又不能用壬癸通其氣，貧困之造也。」（第126頁，如表七）因為無丙丁火助己暖身，乙木無法生長，因此認為此命造為貧困之人。然而在《八字提要》卻寫道：「丁日主生丑月甲辰時時甲引丁為喜，辰丑洩火為忌，去病莫如得寅，則可去其寒濕之土，而生垂絕之火，如柱中疊逢木印，祇要金財得所，反成有用之材，土多可作食傷生財格論。」

表七、《窮通寶鑑》126頁案例之八字結構

八結字構	時柱	日柱	月柱	年柱
十神	正印		偏財	偏財
天干	甲	丁	辛	辛
地支	辰	卯	丑	卯
藏干十神	乙癸戊 偏七傷 印殺官	乙 偏印	己辛癸 食偏七 神財殺	乙 偏印

先決條件可能在於「辰丑洩火為忌」在月、時柱中已經分別出現「丑、辰」之忌神，那會導致什麼樣的負面結果？韋千里並沒有繼續講下去；而且《八字提要》後面卻又補充一句「土多可作食傷生財格論」丑土和辰土難道不算嗎？似乎有些矛盾，雖然丑、辰底下藏干有「食神」和「傷官」的確作為生庚辛財之用（土生金）形成「食傷生財」，然而日柱底下為卯，藏干乙木偏印，這裡的財卻又會造成自身的「財印互剋」（金木互剋）但在八字中最重要的規則「身弱不可任財官」，日主若無其他得以扶助自身的情況下，「財多身弱」自身能量洩出太多而無補充，以及無法控制的剋洩將反致貧困災厄，就像此命例的貧困情況。

在《八字提要》唯一提到的「僧道」命格，與《窮通寶鑑》是否相同？《八字提要》內文如下：「癸日主生酉月辛酉時癸水生仲秋，月建在酉，又逢辛酉時元，梟神太重，其太過之病，與上條庚申時相等，金多水濁，以金為病，以火為藥，所以干支有財，為第一義，庶藥來病去，如見木而無火剋金，無濟於事，若滿盤金水或土，則非僧即道，最多不駟僧之流。」

而在《窮通寶鑑》只提到類似的案例：「辛酉、丙申、癸酉、辛酉。火無力，又被辛合，身旺無依，貧僧。丙火被合去，身旺而無洩，即不為僧，亦無子嗣，氣勢寒冷，貧困之象也。」（265頁）雖然此處的案例生於申月，而非酉月；但這段可能是兩本書中提到相同、最為類似的命格特質。


為什麼《八字提要》有時會與《窮通寶鑑》下的真實情況案例不太相符合？因為《八字提要》是經由韋千里自行推算整理、從「甲日主生、寅月、甲子時」一直推算到最後一條「癸日主生、丑月、癸亥時」類似數學在理想模型狀態下，用公式一一推導所得出的演算結果，輔以古書已經給出的方法結果（例如『籐蘿繫甲』等）進行補充或修正。但書中並沒有給出真實案例進行佐證。反過來說，《窮通寶鑑》雖然有給出命例與實際可能的情況，但並沒有將各種不同時間、月份下所造成的各種可能結果推算出來，只以書中給出的案例或特殊組合關係之分析結果。因此大多數在《八字提要》所提到的特定好、壞案例，在《窮通寶鑑》並沒有其相對應的參照組可供比對；而《窮通寶鑑》的特定論斷案例，在《八字提要》可能也只會給出一個模糊、曖昧的解釋。

## 反思「數位人文分析」的功能與角色

經過一系列的分析之後，來反思一下「數位人文分析工具」在本篇文章的特色與限制。本篇文章採用詞頻分析工具「HTML5 Word Cloud」、關鍵詞標記工具「MARKUS」<sup>[10]</sup>  
[\[https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt10\]](https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt10)、線上試算表格工具「Google Sheets」<sup>[11]</sup>  
[\[https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt11\]](https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt11)、以及這篇文章就是用google線上文件工具編寫完成。雖然我們已經生活在數位化相當普遍、且習以為常的電腦世界，但要如何用手邊的各種工具去完成分析與研究工作？我想至今仍有許多研究者（包括我自己在內）也是相當頭痛的、用著熟悉且有限的手邊工具，去想盡辦法、痛苦的完成自己所想打造的成果。

在這篇文章研究之前，原本想要嘗試進行數位人文研究的材料，是希望分析「四字斷終身」<sup>[12]</sup>  
[\[https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt12\]](https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt12) 這個技術；但掃描建檔進入OCR文字辨識<sup>[13]</sup>  
[\[https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt13\]](https://hiroaki78450.blogspot.com/2021/01/blog-post.html#ftnt13) 的時候，卻遭遇到相當麻煩的問題：原本的書本文字編排模式是直書、右至左的書寫方式，然而OCR辨識工具時而正常的中文辨識、時而以英文書寫模式（左至右的橫向書寫）進行辨識解讀，文件如果有十首詩，大概會有四首得到如此的分析結果（如表八）。嚴格來說只有60%以下的成功辨識率，因為還有亂碼、甚至整首詩完全沒辨識讀出的情況；而整本書共有1440首詩，若要不斷修正錯誤辨識文字亂碼、橫直書等各種錯誤，相當曠日廢時而且進度痛苦且緩慢，最後只能忍痛暫時放棄此研究分析工作。

表八、OCR辨讀常見問題

掃描文件圖檔	OCR正常辨讀 (理想狀態)	OCR錯誤辨讀 (實際狀況)
 <p><a href="https://1.bp.blogspot.com/-DCN2A6z2Q8A/X_yKe4xNIYI/AAAAABcwA/1AGgyVgLECMJxniHAJnq3foWJEoO-RpNgCLcBGAsYHQ/s374/image1.png">[https://1.bp.blogspot.com/-DCN2A6z2Q8A/X_yKe4xNIYI/AAAAABcwA/1AGgyVgLECMJxniHAJnq3foWJEoO-RpNgCLcBGAsYHQ/s374/image1.png]</a></p>	<p>春來芳草依然綠，雪裡圖花見孤木；          親生兄弟有如無，只因生來命孤獨；          楚樹吳花景色鮮，燕飛鶯舞日和天；          狂風驟雨來軒檻，月暗山深聞杜鵑。</p>	<p>狂楚親春          風樹生來          驟吳兄芳          雨花弟草          來景有依          軒色如然          檻鮮無綠          月燕只雪          暗飛因裡          山鶯生圖          深舞來花          聞日命見          杜和孤孤          鵑天獨木</p>

進行《八字提要》和《窮通寶鑑》之分析過程反而沒有遭遇到較大的困難與挑戰，由於八字技術與書寫方式，是相當具有規律性、條理性的編排整理，因此採用EXCEL或MARKUS進行標記與排序等，是相當簡單並且迅速完成的工作。然而在數位文本的取得仍然是相當辛苦的工作，除了直接購買書籍、或從圖書館借閱（學校對於命理相關書籍購入量相對來說仍相當不足）進行文本掃描和OCR辨識之外，從網路上





輸入您的留言...



發表留言的身分：

lfguog@gmail.c ▼

登出

發布

預覽

通知我