

數位化的文章

謝清俊 950328

最近，「火星文」出現在考題中的事件，引了一陣陣議論。觀這件事的始末，不由得令人覺得又好氣、又好笑。最後只有長嘆一聲掩報收場。無獨有偶，余光中教授和教育部長對文言文教學比例的爭議，則顯露：即使是知識份子，面臨文字政策時，見解也南轅北轍，莫衷一是。本來，像這樣的大事，應該辯個明白，說個分明。這樣，我們才能找到正確的未來方向。可是，最後卻掩旗息鼓，不了了之。留下的是一頭霧水的老百姓，以及「只要官喜歡，有什麼不可以？」的「好官」^①。

上述的兩個例子，不約而同的說明：時下文章的生態已今非昔比。本文也來湊熱鬧，從數位化的角度來談談文章的生態。人人都知道，數位化改變了溝通的生態。溝通的生態既然改變了，擔任溝通中介的文字紀錄，其生態也必然風行草偃。觀察文章的生態變化，是件既有趣、又有意義的事。

談到數位時代文章的變遷，首先聯想到的可能是網際網路上「輕薄短小」的文章；也許是迎合青少年口味「圖多字少，膚淺花俏」的圖文夾雜；也可能是有關閱讀習慣和認知行為變遷的研究報告。關於這些，已有許多研究和專文報導，本文不談這些，僅從數位資訊情境的生態，來看文章生態的改變。

首先說明一個事實：文章經數位化存在電腦中時，只存文章的信息是絕對不夠的，必需把一些有關背景的情境資料也存起來，並與文章作適當的連接^②。所以，我們習以為常的文章，到了電腦裡就必需含蓋文章情境的某些信息，並且要和文章構成一個整體。這些文章情境的信息，是依文章應用和管理文章的需要而決定的。比方說，作者、寫作日期，出版處，發行人……等皆是文章情境的信息。談到這裡，引發了一個文章範疇的問題：數位化的文章，該不該包含傳統文章外的情境信息？這就對文章的界定產生了疑惑，使文章的概念變得複雜了。

做數位化的人都知道，文物數位化時必需搭配指定的後設資料，否則就等於沒有數位化。後設資料就是情境信息的一種。不太了解數位化的人常常以為後設資料只有一種。這就錯了！後設資料的內容是依應用的目的而異，一件數位文物的後設資料可以有許多種。例如，記者撰寫的新聞稿，對記者而言，有一種後設資料；對報社來說，同一則新聞有編輯用的、管理用的，甚至於是與其他通訊社交流用的各種後設資料。這些後設資料之間，會有些重複，但也有獨特之處。所以，應用時會要求獨特者能彼此互通，重複者需彼此一致。再者，人們通常要求電腦做的事情越來越多，而且人世間的事情常有變化，所以後設資料不會是固定不變的，它會與時遷移，需要花很多力氣更新、維護和保養。一般人不容易想像，後設資料竟有這麼複雜。

後設資料既然如此複雜，就不是電腦常用的欄位結構可以處理得了的。所以，描述後設資料現在都用後設語言（meta-language），如 HTML、XML，也只有語言才有能力描述後設資料的種種規格和後設資料之間的相容關係，以符合應用的需求。

我們用紙筆寫文章，已經歷了二千多年，早就熟知文章就只是文章，不須顧及情境（後設資料）的信息。然而，數位化以來，越來越將文章與其情境信息結為一體的趨勢。撇開數位化工作不談，例如，以前出書不必提供書目資料，現在則必需提供。又如，現在投稿亦需附檢索詞、摘要、作者的信息等。這都是將文章與情境信息合為一體的作法。這做法，也可視為是為讀者、管理者和工作者設想，有了這些情境信息使大家都方便。然而，對作者而言，文章就不只是文章本身了，還需提供必需的情境信息。換言之，只有文章本身是不夠的，必需和情境信息一起打包，文章才算完成。

情境信息有許多種，小到可以僅僅是作者的識別信息，也可以是寫作時作者的意圖、身心狀態等；大則可能包括文章寫作時的社會情境、政治情境、文化情境等。這些都需要用一種通用的語言—即後設語言描述。因為這些情境信息必需不分國家、種族，甚至於不分電腦機種都要能夠處理，所以，後設語言一定是一種通用的人工語言（artificial language），需要電腦會處理。

後設語言不僅僅可以描述情境信息，文章內容的注疏、註釋，以及文章之間彼此的參照，甚至於文章內容與實物之間的聯繫關係等，也都可以用後設語言描述，如下表所示。從下表知，文章的結構在電腦中產生了根本的改變，它變成以自然語言和後設語言相輔表達的雙重結構：以自然語言寫文章本身，以後設語言描述數位化的文章與外界的各種關係。

表：文章數位化後信息之表達

數化之文章	使用之表現系統
文章本身	自然語言
情境描述	後設語言
參照聯繫	
內容詮注	

有些學者認為，後設語言越來越重要：往後的年輕人除了母語外，其次最重要的不是外語，而是後設語言。因為它是超越國界、不分種族、文化，能跨越時空、讓你充份表達思想、意念的人類共同工具^①。

註：

① 這兩件事的過程，都是典型的負面教材，事情的本身已不堪，事件過程更是雪上加

霜。若是勇於面對事實、真正關心我們的未來，就應該辯個明白來決定我們的方向、掌握未來。這兩件事的過程，難免「不求甚解」、「粉飾太平」之譏，更脫不了被人詬病百餘年的「差不多先生」模式。

- ② 此謂連接即指在電腦中構成某種資料結構，以利往後的運用。
- ③ 人們會越來越依賴電腦才能做事，發表文章也逃不出這個魔咒。