

## 後設資料與內容標誌

謝清俊 960621

上一期談到內容的標誌。有讀者問道：「內容的標誌是後設資料嗎？」。也有人說，上一篇文章雖比較了中外標誌的異同，但是越想越弄不清楚。所以本文較詳細的談後設資料和內容標誌，以補上一期未竟之處。

也許是數位化風行以來後設資料太惹人注目，以致有人認為：除了文物數位化的本身之外，所有其他的資料都屬後設資料；也有人引據國外的文章，說後設資料就是「資料的資料<sup>①</sup>」。有了洋人這樣的說法，許多人更認定了：「所有其他的資料都屬後設資料」。其實，這樣的認知是有問題的，它將原本清晰的事理攪得一團模糊。

「資料的資料」這句話，前一個「資料」指文物數位化所呈現的形態，後一個「資料」則指：我們要利用「文物數位化呈現的資料」時，必需搭配的一些背景資料；而這背景資料就稱作「後設資料」。稱「後設」的原因是：它不是位居前線（第一線）的主要資料，僅僅是作後勤（第二線）支援的資料<sup>②</sup>。

說後設資料是「資料的資料」，只是為了闡明後設資料這個概念的性質，並不是將後設資料定義為「資料的資料」；因為，後設資料固然是「資料的資料」，可是並不是所有的「資料的資料」都是後設資料。將後設資料界定為「資料的資料」這種認知，與「不吃豬肉的都是回教徒」犯了同樣的錯誤。

現行的任何後設資料，其表達的方式、訂定的規格，以及欄位（或標籤）的選擇和數目等，都限制了後設資料的範疇。這很明顯表示：不是所有的「資料的資料」都是後設資料。要明白數位化的後設資料，不能把資料二分為資料和「資料的資料」，不能用這樣籠統的概念去理解。

既然說：不是所有的「資料的資料」都是後設資料，那麼目前的後設資料都是些什麼？簡言之，目前的後設資料都是為了某類文物訂定的；比方說，書目資料是一般書籍的後設資料，新聞有新聞的後設資料，玉器、青銅器、畫作、雕刻……等都有各自的後設資料。

後設資料既然是描述「某類」文物的資料，那麼就有它的特徵和它的侷限。首先，它適合敘述文物的共同現象（共相）。目前後設資料中的欄位（或標籤），像是作者、年代、地點、收藏者、檢索詞、摘要……都是一些共相的名目。既是共相的敘述，就無法顧及個別現象（別相）。所以，後設資料充其量只能摘錄文本的一部份，而無法深入觸及文本的內容。了解到這裡，就不難明白「內容的標誌」何指。內容標誌要照顧的正是後設資料無法觸及的——關於文物個別內容描述的這一部份<sup>③</sup>。

其次，一般而言後設資料敘述的多屬事實、屬性這類較客觀可考的資料，不涉及文本內容的理解、感受、比較、批評，以及詮釋等。所以，後設資料是可以由具技術

專業人士查訪、考證的；但是，它不可以作詮釋<sup>④</sup>。比方說，我們可以考證《紅樓夢》的作者是誰，卻不能詮釋《紅樓夢》的作者是誰。

對文本內容的理解（解釋）、感受、比較、批評、詮釋等，正是內容標誌的主要工作。這些工作觸及人文、歷史、社會、美學、哲學等學門的核心問題<sup>⑤</sup>，需要真正了解內容的專業人士為之。

再者，內容標誌，無論作理解（解釋）、感受、比較、批評或詮釋，均觸及一個人文方面最根本的問題——意義（meaning）和了解（understanding）。這是認知科學、語言學、記號學等近幾年來致力研究的重點，也是電腦迄今未能處理的痛處。內容標誌正是為了解決這個困局而設：一個人機合作的構想，由人（專業人士）負責意義和了解的部份，再由機器來處理其餘的工作。所以，內容標誌的重要不言可喻。

後設資料和內容標誌並不相互排擠，它們是兩種類型完全不一樣的工作。若認為：除了文物數位化的本身之外，所有其他的資料都屬後設資料，那麼就犯了不可原諒的大錯——它扼殺了內容標誌生存的空間。換言之，後設資料和內容標誌兩者都是不可缺的，且彼此相輔相成、相得益彰。

目前數位化的工作，對後設資料認識得多，也做得多。相較之下，對內容標誌則幾乎沒有什麼了解，更無實務<sup>⑥</sup>。然而，處此情境，未嘗不是發展內容標誌的黃金時機。時乎，時不再來！願有為者乘此良機大展鴻圖。

至於意義和了解的問題，且留待下期分解。

#### 註：

- ① 「資料的資料」原文為 data about data 或 data of data。國外的論文是有這樣的說法，然而若是較嚴謹的學術論文這麼說，在其前後文中一定會表明這是為了讓讀者了解後設資料的性質，而作如此權宜的說法。換言之，這不是精確的說法
- ② 後設一詞由「meta-」譯來。meta-有兩重意義，一為「形而上的」，一為「背後支撐」的。無論是那個含意，都是指居第二線，而不是居第一線的。
- ③ 內容標誌和後設資料都用同樣的工具——標誌語言來描述。所以稱作「內容標誌」。詳見本專欄上一期（95年6月）〈文章的標誌〉一文。
- ④ 有人將 metadata 譯為「詮釋資料」。這是一個錯譯。
- ⑤ 此所謂「核心」是指這些學門的本體論、認知論（知識論）、方法論，以及研究方法等。例如近代的現象學、釋義學（詮釋學）等。
- ⑥ 關於內容標誌目前還只停留在研究室內，沒有到達工程實務的階段。有興趣的讀者可找一找近幾年的碩士論文，其中有幾本是做內容標誌相關的研究。