

# 語言學與數位化

謝清俊 911218  
於語言所諮議委員會議

## 壹、數位化的工作內容

### 一、典藏形式的數化

將典藏文物的外觀轉化為多媒體的數位檔案行式。

### 二、背景資料的制式化

各式後設資料 (metadata) 與說明文件等，如版權頁 ...。

### 三、典藏內容的表達與詮注

#### 1. 內容標誌 (markup)

- 注釋：理性的客觀解釋
- 詮釋：含主觀、有立場的說明

『作者不必有此意，讀者何必無此想』—— 清·譚獻語

#### 2. 本文聯接 (intertextuality)

- Hyperlink ...

## 貳、資料整理與語言學

資料的分析、整理、結構化、檢索 ... 無不涉及語言學，而往者只憑常識來規劃、設計，是故舉凡資料的分類、主題辭表、關鍵字集、索引典、乃至於表格中一則欄位的設計（含其語法和語意的規範）等等，都設計的相當膚淺，且欠缺整體性質的宏觀與長遠性質的格局。也正因此，以往的工作僅可稱之為『資料整理』而無法名為『知識表達』。

## 參、數位化與語言學的應用

數位典藏的製作，提供了一個『知識表達』的工作契機，正是語言學的理想舞台；也是語言學走出象牙塔，將理論與實務結合的機會。