

文件編碼組織  
文件編碼交換格式指引 選錄版

TEI Lite: An Introduction to Text Encoding for Interchange

草稿本

路·伯納德  
麥可·斯波伯格－麥昆  
修訂

中央研究院·數位典藏國家型科技計畫後設資料工作組  
研譯



數位典藏國家型科技計畫後設資料工作組

民國九十四年十月



# 目次

目次 .....	I
寫在前面 .....	1
1. 介簡 .....	2
2. 簡易的示範 .....	4
3. TEI 文件的結構 .....	7
4. 文件編碼 .....	8
4.1 文件分項 .....	9
4.2 標頭與結尾 .....	11
4.3 散文、詩及戲劇 .....	11
5. 頁碼與行碼 .....	13
6. 醒目段落的標誌 .....	14
6.1 字型改變等項目 .....	14
6.2 引文與類似的特徵 .....	16
6.3 外國文字或其他特性 .....	17
7. 附註 .....	18
8. 參照與連結 .....	19
8.1 簡單交互參照 .....	19
8.2 延申指標 .....	21
8.3 連結屬性 .....	24
9. 編輯加工 .....	25
10. 刪節、刪除及附加 .....	26
11. 名稱、日期、數字及縮寫 .....	28
11.1. 名稱與關聯字串 .....	28
11.2. 日期與時間 .....	30
11.3. 數字 .....	31
11.4. 縮寫與全稱 .....	31
11.5. 地址 .....	32
12. 列表 .....	32
13. 引用書目 .....	35
14. 表格 .....	36
15. 圖解與圖象 .....	37
16. 詮釋與分析 .....	38
16.1. 拼寫正確的句子 .....	38
16.2. 一般目的的詮釋元素 .....	39
17. 科技文獻 .....	42
17.1. 為科技文件而設的附加元素 .....	42
17.2. 自動產生區段 .....	44
17.3. 產生索引 .....	45
18. 字集、發音符號等 .....	46
19. 正文前與正文後資訊 .....	48
19.1. 正文前資訊 .....	48

19.1.1. 題名頁 .....	48
19.1.2. 正文前資訊 .....	50
19.2. 正文後資訊 .....	51
19.2.1. 正文後資訊的結構區分 .....	51
20. 電子題名頁 .....	52
20.1. 檔案描述 .....	52
20.1.1. 題名敘述 .....	53
20.1.2. 版本敘述 .....	54
20.1.3. 檔案大小敘述 .....	55
20.1.4. 出版敘述 .....	55
20.1.5. 集叢與附註敘述 .....	56
20.1.6. 來源描述 .....	56
20.2. 編碼描述 .....	56
20.2.1. 專案與取樣描述 .....	57
20.2.2. 編輯宣告 .....	57
20.2.3. 標誌、參考及分類宣告 .....	58
20.3. 背景描述 .....	60
20.4. 更新描述 .....	61
附錄 使用元素表 .....	63
附錄 1.1 全域屬性 .....	63
附錄 1.2 TEI 選錄版使用元素 .....	63

## 寫在前面

選錄版的起因，原先只是文件編碼組織的編者們的簡要的示範，說明只要用簡單的 TEI 編碼架構，就可滿足 TEI 社群中九成使用者九成的需求。現在回想起來，我們甚至猜測，多數人可能會誤以為選錄版就是全部的 TEI，或發現選錄版遠甚於需求（為了應付可能的批評，麥可還特別準備了一個「骨架版<sup>1</sup>」）

選錄版奠基於對現行與過去文件編碼經驗的大量觀察，特別是「牛津文件檔案<sup>2</sup>」典藏中實際發生的案例，並輔以個人的經驗。選錄版的內容並不出人意料，似乎早就存在，即便並非標準，至少可以做為全球電子文件與編碼計畫共同的出發點。也許事實就是我們真的作出精簡、易讀、實際合用的東西。

這本手冊理所當然的是為它所描述的 DTD 而編寫，DTD 也依此維護。原文件為 XML 格式<sup>3</sup>，這使得製作各種不同格式的版本，例如 HTML、PDF 等，變得容易，這部分版本可以在「TEI 地下儲藏室<sup>4</sup>」裡找到。

早期的選錄版編寫，包含多個「電子文件中心」自行製作的文獻及授課資料（部分被列在「自學 TEI<sup>5</sup>」的網頁）。

TEI P4，也就是 XML 版的 TEI 指引，是以選錄版的標誌成果作為建置 P4 修正機制的範本。隨著 P4 的出版，這是個微幅修訂現行版本的時機。本文件即為 TEI 節錄版 DTD<sup>6</sup> XML 版的依據。

——路·伯納德于 2002 年 5 月

---

1 barebones version of TEI Lite，<http://www.tei-c.org/Vault/Bare/>。

2 Oxford Text Archive，<http://ota.ahds.ac.uk/>。

3 TEI U5: Encoding for Interchange: an introduction to the TEI，[http://www.tei-c.org/Lite/teiu5\\_en.xml](http://www.tei-c.org/Lite/teiu5_en.xml)。

4 The TEI Vault，<http://www.tei-c.org/Vault/>。

5 Teach yourself TEI，<http://www.tei-c.org/Tutorials/>。

6 <http://www.tei-c.org/Lite/DTD/teixlite.dtd>。

## 1. 介簡

文件編碼格式指引（TEI Guideline，以下簡稱「指引」）的對象，是任一想把資訊轉換成電子形式儲存者。指引強調文字資訊的轉換，但對其他格式的資訊，如圖象及聲音等，也有涉及。建立新的資源或轉換已存在的作品，指引都同等適用。

指引提供使文件特徵明確表現的方法，並以此種方式來輔助電腦軟體在不同的平台上做文件處理。使文件明確的程序稱為標誌或編碼。在電腦中，任何文字的呈現都使用某種標誌的形式，TEI 的形成，部分是因為現有相通且不可省略的編碼架構，很多都是學界設定的，另一部分是因為學術使用電子文件的範圍擴張。

指引描述一個可被擴充到多種不同軟體語言的編碼架構。第一版的 TEI 使用標準通用標誌語言（SGML<sup>7</sup>），最近的版本（TEI P4, 2002），已可使用可擴充標誌語言（XML<sup>8</sup>），未來的版本或許可以擴及其他架構性的語言。這些語言的共通之處，是以元素與屬性來定義文件，並有規則來規範元素與屬性在文件中的用法。TEI 採用 XML 是著重它的完整性與一般性，但 TEI 與其他 XML 的標誌架構並無二致，任何一般目的性的 XML 工具軟體，都可以處理遵循 TEI 的文件。

TEI 之前由計算與人文學會（Association for Computers and the Humanities）、計算語言學會（Association for Computational Linguistics）、及文學與語言學計算學會（Association for Literary and Linguistic Computing）贊助，現在由四所重點大學所主持的獨立的會員性組織維護與發展。目前的經費部分來自美國國家人文研究捐贈基金（U.S. National Endowment for the Humanities）、歐盟委託執行計畫第八指導委員會（Directorate General XIII of the Commission of the European Communities）、安德魯·威廉·美隆基金會（Andrew W. Mellon Foundation）、及加拿大社會科會及人文學研究委員會（Social Science and Humanities Research Council of Canada）。第一版指引於 1994 年 5 月出版，是全球數百位各領域專長的學者投入六年的產物。在接下來的幾年中，指引在數位圖書館的發展、語言工業，甚至是在全球資訊網本身的發展上具漸具影響力。TEI 組織於 2001 年 1 月成立，一年後發行了最新的指引的完整修訂版，修訂成與 XML 完全相容。

TEI 最初工作的總體的目的，是 1987 年 11 月在紐約維薩學院（Vassar College）舉行的預備會議所擬定的封閉條款，這些「帕基普西原則」（Poughkeepsie Principles）在一系列設計的文件中被進一步的闡釋。指引說明這些文件，應該要：

- 滿足表現原文特徵的研究需求。
- 簡單、明瞭、具體。
- 遷就研究者，不需使用特殊的軟體。

---

7 Standard Generalized Markup Language

8 Extensible Markup Language

- 允許文件精確的定義與有效率的處理。
- 提供使用者定義擴張。
- 符合現行以及新出的標準。

學術的世界廣闊而分歧。為了使指引被廣泛的認同，必須確保：

- 文件特徵的核心易於共享。
- 額外特別特徵易於加入（或移除）文件中。
- 同一特徵多重平行編碼的可能性。
- 以最少的規定，讓使用者可自行定義豐富的標誌。
- 提供文件與編碼的充分記錄。

本文獻描述一個自大量元素集中摘取的易用選集，並從設計的目標中推薦範例，稱為 TEI 選錄版。

從完整的 TEI 架構中定義的幾百個元素加以選擇，試圖定義一套有用的「起始集」，歸納出每個使用者都應該知道的元素。對於了解完整的 TEI DTD；以及辨認出在完整的 DTD 中，哪些選用項目對特殊類型的文件是必備者而言，TEI 選錄版的工作經驗是非常寶貴的。

定義本子集的目的，綜述如下：

- 應能包含大多數的 TEI「核心」集，因此它包含了相關的元素以實現所有的文件類型與所有的文件處理工作。
- 在既有經驗能表現的細節層次上（例如，牛津文件檔案的館藏中所展現的）應能勝任合理範圍中的各類型文件。
- 應能運用於製作新文件，如同對已存在的文件進行編碼。
- 應能運用於現有的多種 XML 軟體上。
- 應從完整的 TEI DTD 中衍生，指引所描述的較大結構亦可套用。
- 在符合其他的目標的條件之下，應儘可能短小精簡。

讀者可以自行判斷我們是否成功達成目的。我們的信心源自於實際文件編碼的運用，在文作寫作之時（1995）至少有部分目標已達成。

牛津文件檔案在將原來的文件標誌檔換成 SGML 時採用 TEI 選錄版；維吉尼亞大學與密西根大學的電子文件中心，已採用 TEI 選錄版為館藏編碼。文件編碼組織自己的技術文件——包含本文件，也採用選錄版。

作為一份適當的指導文件，雖然已經嘗試使文件內容完備，但讀者必須注意它並未涵蓋 TEI 編碼架構的所有細節。本文件中所描述的元素，指引中盡皆包含，指引應當做為查詢選錄版以及其他未被選錄元素的權威參考資料。選錄版並假設使用者具備某些 XML 的基本知識。

## 2. 簡易的示範

就以一則小小的例子當開頭，表現當一個略懂標誌的目的為何的人要將散文的一個段落章鍵入電腦時，會發生什麼事，或展現電子文件的潛力。在理想的狀態下，這樣的成果或許要由非常精確的光學掃瞄器來呈現。企圖透過保持換行、記錄頁邊來表現原文件的頁眉標題及頁面的版式等諸如此類的資訊，忠於印刷物的樣貌。鍵盤上未提供的字形（如 faàl 裡的重音節字母或長破折號）也要表現，則試圖模擬這些字的外觀。

### 第三十八章

讀者啊，我嫁給他了。我們舉行了寧靜的婚禮，只有他和我，牧師和執事在場。當我們從教堂回來，我走進宅邸的廚房，瑪麗正在料理晚餐，約翰正在清理刀子，我說：

「瑪麗，我今早已經嫁給羅徹斯特先生了。」管家和她的先生是禮貌週到、鎮定的規矩人，面對他們，任何時間都可以安全的談論一件重大新聞，而不會招致被突如其來的尖叫聲刺破耳朵的危險，接著就被滔滔不絕的冗長訊問弄昏。瑪麗抬起頭，盯著我看，她正在為火裡的烤雞塗油，手上的長柄杓在空中停了大約三分鐘，同一時間內，約翰手中刀子磨光的步驟也暫停了。但是瑪麗又把注意力轉回烤雞，只說：

「是嗎，小姐？嗯，當然囉」

過了一會兒，她繼續說：「我見到你和主人出去，但我不知道你們去教堂結婚。」然後繼續烤雞。約翰呢，當我轉向他時，他正咧著嘴笑。

「我早告訴瑪麗結局了。」他說：「我了解愛德華先生。（約翰是老僕人，主人年幼時就認得他了，所以他通常使用教名）」「我知道愛德華先生會做什麼，而我確定他也不會等太久。就我所知，他做對了。祝你愉快，小姐！」同時有禮貌的拉了一下他前額的頭髮。

「謝謝你，約翰。羅徹斯特先生交代我把這個交給你和瑪麗。」

我把一張五磅的紙鈔放在他手裡。沒等聽到更多的話，我就離開廚房了。不久之後，在通過密室的門時，我聽見對話：

「她或許比那些大小姐更適合他。」接著「就算他不是最美的，但也不醜，脾氣很好。而且情人眼裡出西施，沒有人會否認這點的。」

我立刻寫信到沼澤居和劍橋，說我做了什麼事：完整的說明我

474

簡愛

475

的舉動與原因。黛安娜與瑪麗毫無意見的認同這些手段。黛安娜預告她會給我時間去渡蜜月，然後她會來看我。

「她最好不要等到那時候，簡」當我讀她的信給他聽時，羅徹



斯特先生說：「如果她這麼做，就要很晚才能到，因為我們的蜜月將照耀終生，它的光輝只有到我兩的墓中才會消褪。」

聖·約翰收到個消息的反應，我不知道：在我們通信的過中，他從來沒有回這封信。但是六個月後他寫信給我，不過，也沒有提到羅徹斯特先生的名字或略為提及我的婚姻。他的信友好而平靜，雖然非常嚴肅。自此他維持一貫，但並不密集的書信往來。他希望我快樂，並信賴我並不是那活在世間過著無神生活，只顧俗事的人。

以上的轉錄因下列幾項難處而困擾：

- 頁碼與逐頁標題夾雜在正文中間，造成軟體辨識困難。
- 單引號的引述句與道白句間沒有區別，難以確定哪些是直接敘述。
- 保留連字符號的使用，簡單的搜尋軟體無法搜尋到被切割的字。
- faal 裡的重音節字母與長破折號以特殊的樣式鍵入，但未循任何標準的模式，且僅有在轉錄者在文獻中提及這些樣式時，文件才能被正確的處理。
- 分段僅以首行空白標誌，而且換行符號已經用在行尾做分行，如果用來印刷文字的字體大小改變，重排就成為問題。

以下顯示同樣的段落也可以使用 TEI 指引來編碼。雖然看到其中有許多擴充的編碼方式，但是採用 TEI 至少允許呈現下列的區別：

- 分段標示明確。
- 道白句與引述句得以區分。
- 重音節字母與長破折號使用實體參考。
- 頁面區分以空的元素 <pb> 獨立標誌。
- 為了簡化搜尋及處理，不保留原文的分行，因排版造成的行尾切字無異議重組。若因為重要出版品致使原文分行令人感興趣，雖然此處並未提供，但要保留也很容易。
- 為了便於校讀，每段都重起一行，但取消行首縮排。

<分頁 識別號='474' />

<第一層 類型="章" 識別號='38'>

<段落>讀者啊，我嫁給他了。我們舉行了寧靜的婚禮，只有他和我，牧師和執事在場。當我們從教堂回來，我走進宅邸的廚房，瑪麗正在料理晚餐，約翰正在清理刀子，我說：</段落>

<段落><引句>「瑪麗，我今早已經嫁給羅徹斯特先生了。」</引句>管家和她的先生

是禮貌週到、鎮定的規矩人，面對他們，任何時間都可以安全的談論一件重大新聞，而不會招致被突如其來的尖叫聲刺破耳朵的危險，接著被就被滔滔不絕的冗長訊問弄昏。瑪麗抬起頭，盯著我看，她正在為火裡的烤雞塗油，手上的長柄杓在空中停了大約三分鐘，同一時間內，約翰手中刀子磨光的步驟也暫停了。但是瑪麗又把注意力轉回烤雞，只說：</段落>

<段落><引句>「是嗎，小姐？嗯，當然囉」</引句></段落>

<段落>過了一會兒，她繼續說：<引句>「我見到你和主人出去，但我不知道你們去教堂結婚。」</引句>然後繼續烤雞。約翰呢，當我轉向他時，他正咧著嘴笑。<引句>「我早告訴瑪麗結局了。」</引句>他說：<引句>「我了解愛德華先生。」</引句>（約翰是老僕人，主人年幼時就認得他了，所以他通常使用教名）」<引句>「我知道愛德華先生會做什麼，而我確定他也不會等太久。就我所知，他做對了。祝你愉快，小姐！」</引句>同時有禮貌的拉了一下他前額的頭髮。</段落>

<段落><引句>「謝謝你，約翰。羅徹斯特先生交代我把這個交給你和瑪麗。」</引句></段落>

<段落>我把一張五磅的紙鈔放在他手裡。沒等聽到更多的話，我就離開廚房了。不久之後，在通過密室的門時，我聽見對話：</段落>

<段落><引句>「她或許比那些大小姐更適合他。」</引句>接著<引句>「就算他不是最美的，但也不醜，脾氣很好。而且情人眼裡出西施，沒有人會否認這點的。」</引句></段落>

<段落>我立刻寫信到沼澤居和劍橋，說我做了什麼事：完整的說明我<分頁 識別號 = '475' />的舉措與原因。黛安娜與瑪麗毫無意見的認同這些手段。黛安娜預告她會給我時間去渡蜜月，然後她會來看我。</段落>

<段落><引句>「她最好不要等到那時候，簡」</引句>當我讀她的信給他聽時，羅徹斯特先生說：<引句>「如果她這麼做，就要很晚才能到，因為我們的蜜月將照耀終生，它的光輝只有到我兩的墓中才會消褪。」</引句></段落>

<段落>聖·約翰收到個消息的反應，我不知道：在我們通信的過中，他從來沒有回這封信。但是六個月後他寫信給我，不過，也沒有提到羅徹斯特先生的名字或略為提及我的婚姻。他的信友好而平靜，雖然非常嚴肅。自此他維持一貫，但並不密集的書信往來。他希望我快樂，並信賴我並不是那活在世間過著無神生活，只顧俗事的人。</段落>

決定將焦點放在勃朗蒂的文字，而非特定印刷版本上，在編碼的基本議題中只是其中一個面向：關於選擇。只有對編碼者具有重要性的文本特徵才被明確編碼。要在這小小的篇章中想到各種擴充的編碼方式毫無困難。例如：

- 可供使用方言部分的規範格式。
- 可加入對任何段落的詞彙註腳或評註。
- 可加入由此文本連結到其他文本的連結。
- 自背景文件中區辨出各種專有名詞。
- 關於出處與背景的詳細書目資訊可以置於文本前端。

- 可提供篇章的句子、子句、用字等語言分析，每個單位都關連到相應的類碼。
- 文件可切割成敘事或對話的單元。
- 透過潛藏的複雜文件編排或文本與分析間的連結，文本的系統性分析或詮釋可包含在編碼中，或做多種譯本之間的對照。
- 文本的篇章可被連結到其他媒體的聲音或是影象。

TEI 推薦承載此類細節所有的方式，在以下的章節中會敘述。TEI 的整體架構也為可能的做法提供大幅度的空間，以下只提一小部分：

- 各種名詞構件的詳細分析。
- 詳細的後設資訊，提供關於文本來源或主題的索引典式資訊。
- 透過詳細的版本系列資訊，展現關於文本印行歷史或手稿變更的資訊。

本文推薦及其他可用的標誌，請查閱完整的指引。

### 3. TEI 文件的結構

所有遵循 TEI 的文件都包含 (a) TEI 標頭 (以元素<TEI 標頭> (<teiHeader>) 標誌者) 與 (b) 文本的正文轉錄 (以元素<文件> (<text>) 標誌)。

TEI 標頭部分提供相當於印刷出版品題名頁中所提供的資訊。最多包含四個部分：機讀文本的書目描述、編碼方式的敘述、文本的非書目性敘述 (文本背景資訊)、以及修訂史。標頭部分更進一步的敘述詳「20. 電子題名頁」。

TEI 文件可以件 (單一作品)，或合集 (多件單一作品的集合，如選集) 為單位。不論何種狀況，可能出現正文前資訊或正文後資訊。兩者之間為正文，在合集中，或可以由群組組成，每一個群中可包含更多的群組或文本。

以件為單位的作品，編碼的整體的結構如下：

```
<TEI.2>
  <TEI 標頭> [ 標頭部分資訊 ] </TEI 標頭>
  <文本>
    <正文前資訊> [ 正文前資訊 ] </正文前資訊>
    <正文> [ 正文 ] </正文>
    <正文後資訊> [ 正文後資訊 ] </正文後資訊>
  </文本>
</TEI.2>
```

合集可能也有合集的正文前後的資訊，中間會出現一或多個個文本群組，每一個獨立，但可能也各自擁有正文前後資訊。合集編碼使用的整體的結構如下：

```
<TEI.2>
  <TEI 標頭> [ 合集的標題部分資訊 ] </TEI 標頭>
  <文本>
    <正文前資訊> [ 合集的正文前資訊 ] </正文前資訊>
    <群組>
```

```

<文本>
  <正文前資訊> [ 第一份文件的正文前資訊 ] </正文前資訊>
  <正文> [ 第一份文件的正文 ] </正文>
  <正文後資訊> [ 第一份文件的正文後資訊 ] </正文後資訊>
</文本>
<文本>
  <正文前資訊> [ 第二份文件的正文前資訊 ] </正文前資訊>
  <正文> [ 第二份文件的正文 ] </正文>
  <正文後資訊> [ 第二份文件的正文後資訊 ] </正文後資訊>
</文本>
  [ 更多的文件群組 ]
</群組>
  <正文後資訊> [ 合集的正文後資訊 ] </正文後資訊>
</文本>
</TEI.2>

```

也可把 TEI 文件做成合集，每一份文件都有自己獨立的標頭。這種合集稱為 TEI 文集，文集本身也有標頭：

```

<TEI 文集>
  <TEI 標頭> [ 集合的標題部分資訊 ] </TEI 標頭>
  <TEI.2>
    <TEI 標頭> [ 第一份 TEI 文件的標頭部分 ] </TEI 標頭>
    <文本> [ 集合中的第一份 TEI 文件 ] </文本>
  </TEI.2>
  <TEI.2>
    <TEI 標頭> [ 第二份 TEI 文件的標頭部分 ] </TEI 標頭>
    <文本> [ 集合中的第二份 TEI 文件 ] </文本>
  </TEI.2>
</TEI 文集>

```

但是，不能以 TEI 文集再建以一個合集——也就是說，不能將多個元素<TEI 文集> (<teiCorpus>) 合併視為單一物件。這是 TEI 指引現行版本中的限制。

在後續的章節中，主要是討論簡單的文本結構。每一種狀況由幾個相關的 TEI 元素與簡短的定義構成，其後是專屬於該元素的屬性定義。大體而言，都會附帶簡短的範例。

## 4. 文件編碼

如前面章節所述，一個簡單的 TEI 文件在文本的層次下由下列元素組成：

```
<正文前資訊>/<front>
```

標誌正文開始之前的任何前置項目（標頭、書名頁、前言、獻詞等）

```
<群組>/<group>
```

標誌幾篇單一的文本或一組文件。

<正文>/<body>

標誌單一文本的整體部分，不包含正文前及正文後資料。

<正文後資訊>/<back>

標誌隨附於正文之後的附錄等。

正文前置與後置的專用元素，在之後的「19. 正文前與正文後資訊」中敘述。本節只討論標記正文的元素。

## 4.1 文件分項

一篇散文體的文本，其正文可能僅由一系列的段落組成，或說這些段落被集結成章、成節、成附冊等。在第種一情況下，使用<段落> (<p>) 來標記每一個段落。在第二種下， <正文> (<body>) 可以再以元素一系列的<第一層> (<div1>) 或<文字區段> (<div>) 區隔，不論哪一種，都可以再分更深的層次，元素敘述如下：

<段落>/<p>

標記文章的段落。

<文字區段>/<div>

標誌正文前資訊、正文及正文後資訊中文本的分項。

<第一層>/<div1>

標誌正文前資訊、本文及正文後資訊中文本的第一層分項。(如未使用<第零層> (<div0>))。

當需要比<第一層> (<div1>) 還深的結構時，<第一層>可再分出<第二層> (<div2>)，<第二層>又可分出<第三層> (<div3>)，以此類推，可深到<第七層> (<div7>)。如果超過七層，下次兩種方式可擇其一：修改 TEI 的標誌集以接受<第八層> (<div8>)；或採用未編號的元素<文字區段> (<div>)，<文字區段>可分成更小的<文字區段>，且無深度的限制。

每個文字區段都可以使用下列三項屬性：

### 類型/type

指文件的類名。常用的屬性值是書、章節、詩等。其他可能的屬性值包含被視為單元的詩歌類型，如詩節、演講和歌曲。要注意在<文字區段>、<第一層>、<第二層>元素中第一次出現**類型**屬性時，要確保該類型在同一<正文>中必須涵蓋以下各層。也就是說每一種類型要在起始就設定，或時改變屬性值。

### 識別名稱/id

為分項命名指定唯一的識別標記，以便指向「8. 參照與連結」中所討論的交互參照或評注等其他連結。為每一個主要的結構單元提供屬性**識別名稱 (id)**非常有用，識別名稱應以系統化方式命名，如下面例子中，被標誌的文章題名代碼後面再加上章節編號。

### 識別號/n



文字區段的屬性 **n** 記錄好記的短名或是編號，在指定屬性標記上，**識別號 (n)** 比**識別名稱 (id)** 受青睞。如果原文件已有明確的參考或縮寫格式（如書籍/章節/詩節的書目引用格式），屬性**識別號 (n)** 是記錄它們的地方。

屬性**識別名稱 (id)** 和**識別號 (n)** 運用的泛範廣泛，任何 TEI DTD 的元素都能使用，屬於全域屬性。其他 TEI Lite 架構定義的全域屬性，將在「8.3 連結屬性」中討論。

文獻中每一個**識別名稱 (id)** 的屬性值皆為唯一值。確保不重複的簡便方式之一是讓它反映文獻的階層結構。例如，亞當·史密斯的**國富論 (Wealth of Nations)** 的初版共有五冊，每冊又各分章節，某些章節又再分部，則可以此結構定義**識別名稱 (id)** 的屬性值如下：

```
<第一層 識別名稱="WN1" 識別號='I' 類型='書'>
  <第二層 識別名稱="WN101" 識別號='I.1' 類型='章'>
    ... </第二層>
  <第二層 識別名稱="WN102" 識別號='I.2' 類型='章'>
    ... </第二層>
    ...
  <第二層 識別名稱="WN110" 識別號='I.10' 類型='章'>
    <第三層 識別名稱="WN1101" 識別號='I.10.1' 類型="部">
      ... </第三層>
    <第三層 識別名稱="WN1102" 識別號='I.10.2' 類型="部">
      ... </第三層>
  </第二層>
  ...
</第一層>
<第一層 識別名稱="WN2" 識別號='II' 類型='書'>
  ....
</第一層>
...
```

**識別名稱 (id)** 與**識別號 (n)** 中可使用不同的編碼架構，在正規參考架構與作品結構不完全相符時非常有用。例如，某一小說有兩冊，兩冊各分章節，但章節編號連貫而非各冊獨立，則可採用如下架構：

```
<第一層 識別名稱="TS01" 識別號='1' 類型='冊'>
  <第二層 識別名稱="TS011" 識別號='1' 類型='章'>
    ... </第二層>
  <第二層 識別名稱="TS012" 識別號='2'>
    ...</第二層>
</第一層>
<第一層 識別名稱="TS02" 識別號='2' 類型='冊'>
  <第二層 識別名稱="TS021" 識別號='3' 類型='章'>
    ...</第二層>
  <第二層 識別名稱="TS022" 識別號='4'>
```

```
...</第二層>
</第一層>
```

上述作品有兩卷，每卷含兩章。原章節編號是 1 至 4，但記錄識別號 (id) 屬性值時，可允許它們同時被視為編號 1.1, 1.2, 2.1, 2.2。

## 4.2 標頭與結尾

每一個<文字區段>、<第一層>、<第二層>等元素在正文起始處可能有題名或標頭，及（較不常見的）卷末語，如：第一章完。可用下列元素轉錄之：

<標頭>/<head>

可標誌任何的標頭，例如：章節題名，詞彙表或其他表列的表頭。

<卷末>/<trailer>

標誌出現在文本結尾處的結尾辭或尾聲。

其他文本區段的開頭或結尾處可能需要的元素，將後面的「19.1.2 正文前資料」章節中討論。

是否在轉錄時包含標頭或是卷末語，是轉錄者該決定的重大議題。平淡無奇的標頭（如「第一章」）或已由屬性值表示（如<div1 type='Chapter' n=1>）者可省略，除此之外，無法表現的文本都必須被標記。例如，哈代 (Hardy) 的「綠林樹下」 (*Under the Greenwood Tree*) 的卷首可標誌如下：

```
<第一層 識別名稱="UGT1" 識別號='冬' 類型='部'>
<第二層 識別名稱="UGT11" 識別號='1' 類型='章'>
<標頭>梅爾斯鐸克巷</標頭>
<段落>致林中居民樹上的每一種生物...
```

## 4.3 散文、詩及戲劇

如上文所述，組成文本區段的段落應由<段落> (<p>) 來標誌。例如：

```
<正文>
<段落>我衷心地讚賞波普將軍的輝煌成就及其無價的成果；但你必須知道正規軍中的將帥人才，並不像黑莓一樣隨處可見。</段落>
</正文>
```

下列幾個不同的標誌為韻文和劇本（戲劇、電影等）等具特定結構的文本編碼而設：

<詩行>/<l>

標誌韻文的一行，可能為未完成句。可用的屬性如下：

**分部/part**

用以標示此詩行是否為已完成的韻文，合法的屬性值：

- F：未完成詩行的結尾部分。
- Y：詩行為未完成的韻文
- N：完成或被視為完成的詩行
- I：未完成詩行的開頭部分
- M：未完成詩行的中段。

<詩組>/<lg>

標誌形成一個正式單位的一群韻句，如：詩節、疊句、行韻的段落等。

<講述>/<sp>

標誌劇本中的單一演說，或在散文和詩中以講述方式表現的段落。可用的屬性：

**人物/who**

可以**識別名稱 (id)** 識別該部分的講者。

<講者>/<speaker>

標誌在劇本或文章的片段中一或多個講者的特殊標頭或標籤。

<表演動作>/<stage>

標誌劇本或文章片斷中的任何表演指示。可用的屬性：

**類型/type**

標示各種表演動作指示，建議的屬性值包含進場、退場、佈景、陳述等。

本例中，詩類文本中詩句及詩節被標誌如下：

```
<詩組 識別號="I">
<詩行>我讚頌一個邁向
  不朽的靈魂， </詩行>
<詩行>他的命運，上帝創造，
  但未控制， </詩行>
<詩行>被放在上帝的形像中，始終
  在上帝的律法之前</詩行>
<詩行>結合我們，當時，今後，
  因此我歌頌。</詩行>
<詩行>美好世界到了遲暮之年；</詩行>
<詩行>我畫下早晨新生，經過如日壯年。</詩行>
<詩行>那金色的迦勒底，銀色的波斯臂膀揮動，</詩行>
<詩行>希臘的銅，或羅馬的鐵，俱在一軀之中；</詩行>
<詩行>耗盡精力慢慢形成磚與石的塞特柱，</詩行>
<詩行>（神聖的文件被接受）開始向無人吶喊，</詩行>
</詩組>
```

注意，元素<詩行>（<l>）標誌的詩的行，並非排版的行：上例中，前幾行的原編排方式，在此一編碼中因此未被明確表現，而且可能被遺落。如果真的想要標誌排版的行，可使用「5. 頁碼與行碼」中敘述的元素<分行>（<lb>）。

有時，特別是在劇本中，詩行會被兩個講者表演而割裂。最簡單的標誌方式是用**分部 (part)** 屬性來表示詩行是破碎而未完成的，如下例：



```

<第一層 類型='幕' 識別號='I'><標頭>第一幕</標頭>
<第二層 類型='景' 識別號='1'><標頭>第一景</標頭>
<表演動作 樣式="斜體">兩個衛兵巴納度與法蘭西斯柯入場，在某個門旁</表演動作>
<講述><講者>巴恩<詩行 分部="Y">誰在那裡？</詩行></講者></講述>
<講述><講者>法蘭<詩行>喂，回答我。站住，你不要躲了。</詩行></講者></講述>
<講述><講者>巴恩<詩行 分部="i">吾皇萬歲！</詩行></講者></講述>
<講述><講者>法蘭<詩行 分部="m">巴納度？</詩行></講者></講述>
<講述><講者>巴恩<詩行 分部="f">他啊。</詩行></講者></講述>
<講述><講者>法蘭<詩行>你最好小心別為在當班時找來麻煩</詩行></講者></講述>

```

同樣的機制可應用在由兩個不同講者輪流表演的詩節上：

```

<講述><講者>第一個聲音</講者>
<詩組 類型="詩節" 分部="I">
<詩行>但是什麼讓船跑那麼快</詩行>
<詩行>沒有浪或風？</詩行>
</詩組>
<講述><講者>第二個聲音</講者>
<詩組 分部="F">
<詩行>風被截斷在前。</詩行>
<詩行>並緊跟在後。</詩行>
</詩組>

```

本例子顯示將散文中的對話像劇本一樣的標誌。同時顯示使用屬性人物 (**who**) 記載代碼以識別與對話段落相關的說話者：

```

<講者 人物="歐彼"><講者>令人敬畏的歐彼米安博士</講者>
<段落>我不認為我只在為一條難登大雅之堂的魚命名。</段落>
<講者 人物="葛生"><講者>葛雷爾先生</講者>
<段落>鯛魚，博士，鯛魚沒什麼好談的。</段落>
<講者 人物="歐彼"><講者>令人敬畏的歐彼米安博士</講者>
<段落>恰恰相反，先先，我認為它有一堆可談。首先...</段落>
<段落>...魚，葛雷爾小姐——我可以花一整個小時和你談魚，但是現在我得節制...</段落>
</講述>

```

## 5. 頁碼與行碼

分頁和分行可以下列元素來標誌。

```
<分頁></pb>
```

以標準參考系統標記文本一頁與下頁間的邊界。

```
<分行></lb>
```

標誌某些編輯或版本上（印刷時）的新起行。

上述兩個元素標誌文本裡的一個點，而非文本其中一段。標誌時，全域性屬性**識別號 (n)** 可記載頁碼或行碼。此外，這兩個元素共用以下屬性：

### 版本/ed

標誌被標誌的某一點所屬版本或修訂本的頁數。

有編頁碼的原件在數位化時，即便只為了讓日後的校對更容易，記錄頁碼通常也有相當的作用。同樣的理由也適用於標誌行碼。來源文件為印刷品時，行末的連字符號的處理也需要考慮。

若標誌的頁碼中有多個版本，使用屬性**版本 (ed)** 來指明所標誌的版本，若需要時可應儘可能多標記。如下例，可指出某一段落在不同版本中出現的頁數。

```
<段落>我立刻寫信到沼澤居和劍橋，說我做了什麼事：完整的說明我<分頁 識別號='475' />的舉措與原因。黛安娜與<分頁 版本="版本 1" 識別號="475" />瑪麗毫無意見的認同這些手段。黛安娜預告她<分頁 版本="版本 2" 識別號="485" />會給我時間去渡蜜月，然後她會來看我。</段落>
```

<分頁>及<分行>兩個屬性為通用元素<分界>（<milestone>）的特例，<分界>（<milestone>）可標誌文本中的參考點。TEI Lite 也包含通用的<分界>元素，而且沒有特例的限制，可標誌任何類型的參考點：例如，分欄或不會被用在別處的分段類型等。這個元素的說明及可用屬性如下：

### <分界>/<milestone>

標明標準參考系統的變更以標誌文本分段的界限。可用屬性包括：

#### 版本/ed

記錄使用分界的版本或修訂本。

#### 單元/unit

記錄分界類型的變更。

屬性**版本 (ed)** 及**單元 (unit)** 所指涉的版本及單位，命名格式可以自由選擇，但必須記錄在 TEI 標頭裡。

元素<分界>可取代<行碼>或<頁碼>，<行碼>或<頁碼>可視為一組，但不可相互混淆。

## 6. 醒目段落的標誌

### 6.1 字型改變等項目

醒目的單字或段落，是看起來與其他文字不同者，一般而言是透過改變字型、書寫體，或墨色來吸引讀者的注意力。

全域屬性樣式 (**rend**) 可隨附於任何元素，任何需要特別標誌的醒目細節都可以使用。例如，一個粗體的標頭可用「標頭 樣式="粗體"」標誌，為斜體則「標頭 樣式="斜體"」。

文本中字型或色彩的改變，未必都能、也未必都需要標記。在這種狀況下，元素 `<醒目標示>` 可用來標誌一連串醒目字，且不需要宣告它的狀態。

但若遇到醒目部分，不論其功用為何，都可用元素 `<醒目標示>` (`<hi>`) 來標誌。  
`<醒目標示>/<hi>`

標誌與背景文字顯著不同的單字或段落，不論其目的。

下例中，只標誌副題名與人名使用的字型變化，未標誌其類型。

`<段落><醒目標示 樣式="哥德體">與本契約其他見證</醒目標示>`  
`它宣稱<醒目標示 樣式="斜體">華爾特·仙迪</醒目標示>，零售商，上述婚約的`  
`受益者...</段落>`

此外，醒目字使用的確實原因可被識別，還有其他更精確的元素。

`<強調>/<emph>`

標誌語言或修詞效果的重音或強調的字或片語。

`<外國語言>/<foreign>`

定義與背景文字不同語言的字或片語。

`<指涉>/<mentioned>`

標誌被提到的不用字或片語。

`<專有名詞>/<term>`

標誌被視為技術性詞彙的單字、複合字或設計圖象。

`<題名>/<title>`

標誌作品的題名，不論是是文章、書籍、期刊或叢書，其他題名或是副題名也包含在內。可使用的屬性包含：

#### 層級/level

標明此為文章、書籍、學報、期刊、叢書或其他未出版物。合法屬性值：

*m*：專論的題名（書籍、合集或其他明顯為出版品者，多冊作品中的各單冊也包含在內）。

*s*：（叢書題名）

*j*：（學刊題名）

*u*：未出版物（包含未商業發行的論文及學位論文）。

*a*：分析題名（合集中所出版的文章、詩或其他類型文件）。

#### 類型/type

將題名分成某些簡易的類型。參考值如：簡稱 (*abbreviated*)、主要題名 (*main*)、副題名及部分題名 (*subordinate*) 及共同題名 (*parallel*，作品的又稱，通常是另一種語言)。

在文本間出現的某些特徵（明確的引文或行間註），有些使用醒目標示，有些使用引號。以上兩種情形，可使用元素<引句>（<q>）和<註解>（<gloss>）（下節會討論到）。若要記錄採用的醒目方式，可使用全域屬性樣式（rend）。

本節中所定義元素的例子，請看下列句子：

一方面，與尼布龍根之歌相關的是十二世紀法國新興的騎士文學——<外國語言>romans d'antiquité</外國語言>（中世紀的騎士文學）；克萊第安·德·托耶爾所作的克里特與特洛伊浪漫文集，以及海因利希·馮·菲爾德克、哈特曼·馮·奧、以及弗爾富蘭·馮·艾仙巴赫所作的德文改編本。

要轉錄醒目字的特性，句子可標誌如下：

<段落>一方面，與<題名>尼布龍根之歌</題名>相關的是十二世紀法國新興的騎士文學——<外國語言>romans d'antiquité</外國語言>（中世紀的騎士文學）；克萊第安·德·托耶爾所作的克里特與特洛伊浪漫文集...<段落>

只描述原文的外觀，可標誌如下：

<段落>一方面，與<醒目標示 樣式="斜體">尼布龍根之歌</醒目標示>相關的是十二世紀法國新興的騎士文學——<醒目標示 樣式="斜體">romans d'antiquité</醒目標示>（中世紀的騎士文學）；克萊第安·德·托耶爾所作的克里特與特洛伊浪漫文集...<段落>

## 6.2 引文與類似的特徵

就像改變字形，引號也常用在標示一些不同的文本特徵，其中最常用在引文。比起簡單的標誌引號在文本中出現，可能的話，建議使用下列元素標誌引號隱含的特徵：

<引句>/<q>

標誌引文或顯然為引文者——標誌引述其他人言語或思考的表現（不論是否實際引用）。在述敘中，是演員或講者的文字；在字典中，<引句>可用在標誌真實或編製的使用範例。屬性包括：

### 類型/type

可用在標明引文內容屬言語或思考，亦或其他更精確的特性。簡單的屬性值包括：

言語 *spoken*（表示口說的言語，通常以引號標記）

思考 *thought*（表現思考，例如內心獨白）

### 人物/who

識別口說言語部分的講者。

<指涉>/<mentioned>

標誌被提到的非用字或片語。

<另稱>/<soCalled>

標誌作者或編寫者標示為放棄責任的個字或詞，如：使用顯目的引號或斜體者。

<註解>/<gloss>

標誌一個字或詞，為其他字或片語提供夾註或定義。可用屬性：

#### 目標/target

識別相關的字或片語。

以下為引文的簡例：

```
<段落>字典的製作者幾乎不可能忘記約翰生博士把字典的編纂者描述成<引句>無  
害的苦工。</引句></段落>
```

記錄引文印刷的方式（如：夾在文章中，或拉出獨立展現，使用引文區塊），可使用屬性樣式（**rend**）。也可以用來指示使用引號的類別。

口說言語被編排打斷者，可以簡單地先結束引文，在打斷部分結束後再重標，如下例：

```
<段落><引句>誰打敗你了？</引句>;他終於開口;<引句>你閉嘴，該死，我要  
幹掉你。</引句>他邊說，黑暗中，亮晃晃的戰斧開始在我周身遊走。</段落>
```

如果把兩個元素<引句>結合重製成單一的言語以表達完整的意念是重要的，欲，可使用「8.3. 連結屬性」敘述的連接屬性下一元素（**next**）和上一元素（**prev**）。

引文或可伴隨來源或講者的參考資料，不论文本中是否提供來源，都可用屬性人物（**who**），如下例：

```
<q 人物="威爾森">斯柏汀，他下樓進到辦公室，就在八週前的今天，手上握著正  
是這份文件，他說:<引句 人物="斯柏汀">上帝保佑，威爾森先生，我是紅髮。  
</引句></引句>
```

上例同時顯示如何在引文中嵌入其他引文：一個講者（威爾森）引述另一個講者（斯柏汀）。

電子文件的建置者必須決定：引號是否以標誌取代，或保留引號再另加標誌。如果文本上的引號被移除，可用屬性樣式（**rend**）來記錄原文件上的樣式。

如同醒目標示，用這種方式轉錄文本中引號的功能時，不可能完全轉錄也無法盡然合意。遇到這種狀況，標誌「<醒目標示 樣式=引句>（<hi rend=quoted>）」或可用來標誌未宣告其狀態的引文。

## 6.3 外國文字或其他特性

非文本的主要語言的字或片語，可就兩種方式擇一來標誌。若字或片語有某種原因已被標誌，標誌的元素應以全域屬性語言（**lang**）的值，標明使用的語言。若無適用的元素時，可使用元素<外國語言>（<foreign>），同樣要使用屬性語言（**lang**），如下例：



```
<段落>約翰真有手腕（<外國語言 語言="fra">savoir-faire</外國語言>）。</段落>
```

```
<段落>你讀過三便士歌劇（<題名 語言="deu">Die Dreigroschenoper</題名>）嗎？</段落>
```

```
<段落><指涉 語言="fra">Savoir-faire</指涉>法文的手腕</段落>
```

```
<段落>法院簽署執行令（<專有名詞 語言="lat">mandamus</專有名詞>）。</段落>
```

如上例所示，若有其他更明確的元素，如<題名>(<title>)、<指涉>(<mentioned>)或<專有名詞>(<term>)可用，外文字不應以元素<外國語言>標誌。全域屬性語言(<lang>)，可附加在任何元素中，表示與背景文字不同的語言。

## 7. 附註

所有的註釋，不論印成註腳、卷末註、旁註或其他形式，通通使用同一元素：

```
<附註>/<note>
```

包含附註或註釋，屬性包括：

### 類型/type

描述附註的形式。

### 責任類型/resp

標明對附註負責任者：如作者、編者或譯者等。屬性值可以是作者、編者等，或是加註者個人的縮寫。

### 地點/place

標明附註在來源文件中的位置。參考值包括：

行中註 (inline)：夾在正文間

行間註 (interlinear)：夾在兩行間

卷左旁註 (left)、卷右旁註 (right)：在頁扉左側或右側

註腳 (foot) 及卷末註 (end)：在章節或整卷末端

### 目標/target

標明附註所註的點，或附註所註段落的起點。

### 目標終點/targetEnd

如果附註不是嵌在正文中，標誌附註所註段落的終點。

### 錨點/anchored

標明附註在文本中確實的參考位置。

可能的話，註文應該插在文本中由它附註或第一次出現的地方。但是像旁註就辦不到，它無法確實定位。為了簡化，把旁註放置在相關段落或是其他元素的前面還算差強人意。附註也可以獨立置於文本之外（像書籍中的卷末註），並使用屬性目標（**target**）連結到相關的部分。

如果需要，屬性**識別號(n)**可用來記錄附註的編號或識別碼。屬性**責任類型(resp)**可逐一區別作者註與編者注。此外，TEI 標頭中需說明附註的種類。

```
<段落>合集是各種不同實體或物件的群體
<附註 地點="地腳" 識別號=1>
以下解釋為何使用不尋常的詞
<指涉>合集 (collection) </指涉>
取代了預期中的
<指涉>集合 (set) </指涉>。
這是對應在數學寫作中使用<指涉>集叢 (aggregate) </指涉>，而在較早的邏輯寫作中使用<指涉>類別 (class) </指涉>的概念。
</附註>
本元素 ...</段落>
```

```
<詩組 識別名稱="RAM609">
<附註 地點="頁邊">詛咒終於失效</附註>
<詩行>咒語已解除：再一次</詩行>
<詩行>我目睹海之青綠</詩行>
<詩行>再行遠眺，幾無所見</詩行>
<詩行>除已見者外</詩行>
</詩組>
```

## 8. 參照與連結

在同一份 SGML 文件中，要從文本中的一點明確地參照或連結到另一點，可以使用「8.1. 簡單交互參照」一節中所敘述的元素來編碼。參照或連結到其他 SGML 文件或非 SGML 文件中的某部分時，可以使用「8.2. 延申指標」一節中所敘述的 *TEI 延申指標 (TEI extended pointers)* 來編碼。不明確的連結（如兩個平行文本或在文本裡被中斷部分的結合），可用「8.3. 連結屬性」一節中敘述的連結屬性來編碼。

### 8.1 簡單交互參照

單一文件中的交互參照要從一點連到另一點時，可在下列二個元素之中擇一編碼：

```
<參照>/<ref>
```

指向本文件中的另一個位置，即一或多個可識別的元素，可能為額外的修正文字或意見。

## <指標>/<ptr>

指向本文件中的另一個位置，即一或多個可識別的元素。

以上元素共用下列屬性：

### 目標/target

以一或多個 SGML 的識別碼標示指標的目的地。

### 類型/type

使用任何簡明的類型為指標做適當的區分。

### 目標類型/targType

標示指標指向元素的類型。

### 建立日期/crDate

標示指標建立的時間。

### 責任類型/resp

標示指標的建立者。

<指標> (<ptr>) 與<參照> (<ref>) 兩個元素的不同之處在於：<指標> (<ptr>) 是一個空的元素，只的簡單的標誌指標連結的起點，但是<參照> (<ref>) 最好包含一點文字——最典型的就交互參照的本身的文本。元素<指標>將可用在非文字的交互參照，例如符號、圖示、或用一個按鈕連到一個電子文件。<指標>在文件製作系統中也很有用，格式化工具可從交互參照中產生正確的動詞形式。

以下兩例的形式在邏輯上是相同的(假設文獻某處交互參照的正確動詞語法已用元素<指標>表現)：

```
特別要看<參照 目標="SEC12">34 頁的 12 節</參照>。  
特別要看<指標 目標="SEC12"/>。
```

目標 (**target**) 的屬性值必須是同一文件中已使用元素的識別代碼。這表示，被連結的章節或段落必需要具有識別代碼，因為必需被某一種元素標誌。下例中，交互參照連結到元素<第一層> (<div1>)：

```
...  
特別要看<指標 目標="SEC12"/>。  
...  
<第一層 識別名稱="SEC12"><標頭>相關辨識碼...  
...
```

因為屬性**識別名稱 (id)** 具有全域性，文獻中的任何元素都可因此被連結。下例中，一個段落因被賦予識別名稱所以可被連結：

```
...  
這是在討論<參照 目標="pspec">連結的段落</參照>  
...  
<段落 識別名稱="pspec">連結可以用在任何類別的元素上...
```

屬性**目標類型 (targType)** 可標示出被連結的元素具特定類型，如下例：



```
...
這是在討論<參照 目標="dspec" 目標類型='第一層 第二層'>連結的段落</參照>
```

上例中，若識別名稱為「**dspec**」的元素不是<第一層> (<div1>) 或<第二層> (<div2>)，則此參考無效。注意，在只執行 SGML 或 XML 的分析器時，這種額外的查核無法查到，因為這類分析器只能檢查以「**dspec**」識別的元素是否存在。

屬性**類型 (type)** 可使用任何簡明的類型為指標做適當的區分。屬性**責任類型 (resp)** 可用來表示連結的建立者、建立機構；屬性**建立日期 (crDate)** 可表示建立日期，如下例：

```
...
這是在討論<參照 類型="xref" 責任類型="auto" 建立日期="950521" 目標="dspec" 目標類型='第一層 第二層'>連結的段落</參照>
```

這些屬性最適同用在內含大量具各種目的與義意的指標的超文本系統。

有些時候，交互參照連結的目的無法對應到任何文本特徵，且多數未被任一種元素標誌。如果想連結的目標只是同一文件中的一點，最簡單的方式是用<錨點> (<anchor>) 在適當地點插入標誌。如果連結的標的是未標誌的文字串，可用元素<分割段落> (<seg>) 標誌。這兩個元素的敘述如下：

<錨點>/<anchor>

標記文件中被連結的位置或特定点。

<分割段落>/<seg>

標示文件中被連結的文字區塊或片段。可用屬性包括：

**類型/type**

為連結段落分類。

下例中（設有此例），元素<參照> (<ref>) 用來表示要以某種方式連結到文本中的其他部分。第一例中連結到某一點，第二例中連結到一串文字：

```
回到<參照 目標="ABCD">我打瞌睡的那點上</參照>，我注意到有<參照 目標="EFGH">三個字</參照>被前一個讀者用紅筆圈起來
```

這種編碼方式要求元素具備明確的識別碼（例中的 **ABCD** 與 **EFGH**）必須在同一文件中的其他地方出現。假設含這些識別代碼的元素並不存在，可使用元素<錨點>和<分割段落>：

```
....<錨點 類型="書籤" 識別名稱='ABCD' /> ....
....<分割段落 類型="目標" 識別名稱='EFGH'> ... </分割段落> ...
```

屬性**類型 (type)**（如上例所示）可為文本中不具特定目的的元素作不同目的的區分。其他的使用方法，在下面的「8.3. 連結屬性」一節中討論。

## 8.2 延申指標

元素<指標> (<ptr>) 及<參照> (<ref>)，只能用在同一文件的交互參照或連結。只能參照到文獻中明確標示的元素。本節討論的元素不限於這種用法。

<外部指標>/<xptr>

定義連向本文件或外部文件中另一位置的指標。

<外部參照>/<xref>

定義連向本文件或外部文件中另一位置的指標，可能為額外的修正文字或意見。

除了在「8.1. 簡單交互參照」一節裡已討論指標屬性，這兩個元素共用下列額外的屬性，用來標明屬性目標 (**target**) 的交互參照或連結的地點：

#### 出現位置/doc

標明可發現被連結地點的文件，內定值為現用文件。

#### 起點/from

標明以 TEI 延申指標語法表現的指標目的地的起點，內定值為屬性**出現位置 (doc)** 的所記錄的整個文件。

#### 終點/to

標明以 TEI 延申指標語法表現的指標目的地的終點，只能在已有屬性**起點 (from)** 時使用。

內定值為屬性**出現位置 (doc)** 的所記錄的整個文件。

用來表現 TEI 延申指標所指的目的的完整敘述語言，超出本文獻的範圍，用來表現的完整敘述語言，此處只列舉部分常用的特性，更進一步的細節可查閱完整的指引。

只需為屬性**出現位置 (doc)** 提供實體名稱做為屬性值，元素<xptr> (或<xref>) 即可指向某些其他文件的整體，如下例：

```
見<外部參照 出現位置="P3">TEI 指引，各處</外部參照>
```

上例假設某些系統或是公共版權物件被宣告的名稱是 P3。文獻被分析器處理時，此項宣告已被強迫包含在 DTD 內，此種方法是針對相關軟體的使用（如「15. 圖解與圖像」中所討論的）。

屬性**起點 (from)** 用來標誌任何文件中以屬性**出現位置 (doc)** 標誌的某個位置。這種標誌使用一種名為 TEI 延申指標語法 (TEI extended pointer syntax) 的特殊語言，此處只提供其中的部分細節。在此一語言中，地點被定義成一連串的步驟，每一步驟試別文件中的某一部份，通常是指每位置是受前一步驟所識別。例如，欲指向第二章第二段第三句，第一個步驟是指向第二章，第二個步驟是指向第二段，最後一個步驟是指向第三句。步驟可使用自己的文獻樹，如以父層、子層、*同前一層*等，或更寬鬆的以文本樣式、單詞或是字的角來定義。也可以用外來（非 SGML）的標誌，或是座標系統的觀點來標明圖象中的位置。

屬性**起點 (from)** 與**終點 (to)** 使用同種標誌法。每一指標記指向目標文件的部分內容，延申的指標為一個整體，始於**起點 (from)** 終於**終點 (to)**。

在位置的路徑中，第一步驟通常會標明目的文件中某個元素的識別碼，如下例：

```
<外部指標 出現位置="P3" 起點='id (SA)'/>
```

這個元素在 P3 的中被稱為「SA」。如果需要更好的目的敘述，可再使用其他的元素。下列的關鍵詞可用來敘述本元素和其他元素的之間關係。

選擇文件實體 P3 中被識別為 SA 的任何元素。如果需要更細膩的目標，可進一步使用其他的步驟。以下的關鍵詞有助於以相互關係做為選擇元素屬性的依據。

#### 下層元素/child

被此元素包含的元素。

#### 上層元素/ancestor

直接或間接包含此元素的元素。

#### 上一元素/previous

此元素在同一層中前一個出現的元素。

#### 下一元素/next

此元素在同一層中下一個出現的元素。

#### 已出現元素/preceding

此元素在不同一層中前一個出現的元素。

#### 尚未出現元素/following

此元素在不同一層中下一個出現的元素。

每一個關鍵詞對應到一個特定的元素（子元素、父元素、同元素中的前一段等），為標明指向的元素，這些關鍵詞後面可以選擇用括弧來說明。

- 正負數，說明指標中可能使用的元素數目（+1 表示所指向的目標就在第一個元素中，-1 表示目標在上一個元素中），使用 *all* 的時候，表示指向所有的元素。
- 通用識別，指向符合類型的元素，或指向任何同類型的元素的起點。
- 一組屬性名稱和屬性值，表示被選擇的元素必須擁有特定的屬性名稱與屬性值。

續上例，以下的參考指標將選擇第三個<段落>元素中，直接以 SA 識別任何的元素：

```
<外部指標 出現位置="P3" 起點='id (SA) child (3 p)'/>
```

類似的情況，假設實體 P3 確實為 TEI 指引的 XML 格式，則指向出版品（假設該出版品存在）中的章節 14.2.2 的延申參考指標規範的語法定義為：

```
完整的細節，見  
<參照 出現位置="P3" 起點='id (SA) child (2 div 2) child (2 div 3)'/>TEI  
延伸指標語法定義</ref>
```

通常，交互參照的範圍完全由屬性起點（**form**）定義。但在某些文件中，可以簡便地以起始和終止的範圍來定義。如前文所述，屬性終點（**to**）即為此目的，如下例：

```
<外部指標 出現位置="P1" 起點='id (xyz)' 終點='id (abc)'/>
```

這個延申指標的目的是文件 P1 中以 XYZ 為起點，以 ABC 為終點所標誌包夾任何元素，兩者之間任何元素都被包含，與結構無關。但若結尾的 ABC 在起始的 XYZ 之前，則為錯誤的指標。

非常複雜的格式可以用這種語法簡單的建立。例如，以下的參考指標將指向最近的<標頭>中語言（**lang**）屬性值為 LAT 者，而此元素出現在以 SA 識別的元素之前：

```
<外部指標 出現位置="P3" 起點='id (SA) preceding (1 head lang lat)'/>
```

如果屬性出現位置（**doc**）沒有屬性值，則假設指向目前的文獻。因此，以下的參考指標在語意上是相同的，都指向目前使用文獻中的 X1 段落。

```
<指標 目標="X1" />
<外部指標 起點='id (X1)'/>
```

TEI 延申指標語法定義比較晚近的 XLink 規格要早，XLink 對 TEI 有一些延申。最近進行的工作在調和這兩個語言的規格。

### 8.3. 連結屬性

以下特殊用途的連結屬性是為 TEI Lite DTD 裡列舉的所有元素定義的：

#### 分析/ana

連結元素與它的分析。

#### 對應元素/corresp

連結元素與一或多個與之相對應的元素。

#### 下一元素/next

元素同屬一個集叢時，連結到下一個元素。

#### 上一元素/prev

元素同屬一個集叢時，連結到上一個元素。

屬性分析（**ana**）標誌同一文件中分析或詮譯的所在處，進一步的說明在「16. 詮釋與分析」。例如「John loves Nancy」的語言學分析可被標誌如下：

```
<分割段落 類型="句" 分析="SVO">
  <分割段落 類型="字" 分析="NP1">約翰</分割段落>
  <分割段落 類型="字" 分析="VVI">愛</分割段落>
  <分割段落 類型="字" 分析="NP1">南西</分割段落>
</分割段落>
```

以上標誌顯示了文件隱含的部分，元素以 **SVO**、**NP** 及 **VV1** 為識別代碼，這些特別的代碼解釋了元素的義意。請注意使用元素<分割段落> (<seg>) 標誌分析的特殊部件時，是使用屬性**類型 (type)** 來區分。

屬性**對應元素 (corresp)** 標誌提供簡單的方式，可表現同一文本中兩個元素之間形式上的對應關係。例如，在多語的文本裡，可用來連結對應的翻譯，如下例：

```
<分割段落 語言="FRA" 識別名稱="FR1" 對應元素="EN1">Jean aime
Nancy</分割段落>
<分割段落 語言="ENG" 識別名稱="EN1" 對應元素="FR1">John loves
Nancy</分割段落>
```

同一機制可用在不同的目的。下例中，「該節目」對應到「雪萊」、「該公司」對應到「國家廣播公司」：

```
<段落><題名 識別名稱="雪萊">雪萊</題名>，於週五首播不過一個月，已經不在
<名稱> 識別名稱="國家廣播公司">國家廣播公司</名稱>的新節目表上，雖然
<分割段落 識別名稱="公司" 對應元素="國家廣播公司">該公司</分割段落>表示
<分割段落 識別名稱="節目" 對應元素="雪萊">該節目</分割段落>仍在考量
之中。</段落>
```

屬性**下一元素 (next)** 與**上一元素 (prev)** 提供簡便的方式，將不連續的元素連結在一起，如下例：

```
<引句 識別名稱="Q1a" 下一元素="Q1b">誰打敗你了？</引句>
他終於開口<引句 識別名稱="Q1b" 上一元素="Q1a">你給我閉嘴，該死，我要
殺了你。</引句>他一邊說，發亮的戰斧在黑暗中開始在我身邊揮舞。
```

## 9. 編輯加工

電子文件編碼的程序與編輯手稿或其他印刷出版品的程序之間，有許多相同之處。在這兩者中，一個認真的編輯或許想要記錄來源文件的原始狀態，以及任何編輯訂正或其他對文本的改變。本章與下一章討論的元素，就提供了幾種符合這些需求的便利工具。

下列成對的元素可用來標誌編輯相信原件被誤植的編輯訂正：

<更正>/<corr>

標誌對文本中明顯錯誤的更正。可用的屬性包括：

**原文/sic**

為文件中的明顯錯誤提供原始格式

**責任類型/resp**

標誌可為<更正>元素所建議的內容負責的編輯或轉錄者。

**正確度/cert**

標誌<更正>元素內容來源的可信度。



<原文照錄>/<sic>

標誌雖有明顯錯誤卻仍照原件所錄者。可用的屬性包括：

**更正/corr**

標誌對文本中明顯錯誤的更正。

**責任類型/resp**

標誌可為此更正建議內容負責的編輯或轉錄者。

**正確度/cert**

標誌更正來源的可信程度。

下列成對的元素或可用來規格化，指為了文本一致化或現代化的原因所做的編輯改變：

<原始格式>/<orig>

標誌文件材料原始的格式，調整後的格式可放入屬性值中。屬性包括：

**採用格式/reg**

記錄文本的一般（標準）格式。

**責任類型/resp**

識別能為此字或詞規格化負責的個人。

<採用格式>/<reg>

標誌文件材料因某種原因調整或是規格化後的格式。屬性包括：

**原始格式/orig**

標誌原件中未被規格化文本的格式。

**責任類型/resp**

識別能為此字或詞規格化負責的個人。

例如，文件的原文為

```
... for his nose was as sharp as a pen and a' table of green  
feelds
```

Gifford 指出原文包含錯誤(1) table 被誤植為 babbled，以及(2)應為 he 和 fields 的非標準拼字 a'和 feelds。則 Gifford 的推測可以被標誌如下：

```
... for his nose was as sharp as a pen and <採用格式 原始格式="a'">he  
</採用格式> <更正 原文="table" 版本="Gifford">babbl'd</更正> of  
green <採用格式 原始格式="feelds">fields</採用格式>
```

## 10. 刪節、刪除及附加

除了字與詞的更正或規格化，編輯者與轉錄者也可以補遺、刪節，或抄入來源文件中被刪除或劃掉的部分。此外，還有一些特別難以轉錄的資料，因為它們很難標示在頁面上。下列的元素可以用來記錄這類現象：

<加入文字>/<add>

標誌由作者、代筆者、註解者或更正者在文本中加入的字母、字或詞。屬性包括：

### 出現地點/place

如果附加的部分是寫在原件中，標誌附加文本被書寫的位置。可參考的屬性值包含：文中 (*inline*)、行上方 (*supralinear*)、行下方 (*infralinear*)、左 (左頁扉)、右 (右頁扉)、天頭 (*top*)、地腳 (*bottom*) 等。

### <省略資料>/<gap>

標示資料在轉錄時被省略的部分，可能依 TEI 標頭中敘述的編輯理由刪去，或為抽樣練習的部分，或因為是非法或是限閱的資料。屬性包括：

### 節文描述/desc

描述被省略的文字。

### 責任類型/resp

標示可為決定不提供任何原文轉錄而改使用<省略資料> (<gap>) 標記負責的編輯者、轉錄者或是編碼者。

### <刪除>/<del>

標記由作者、代筆者、附註者或更正者在原件中刪除的字母、文字或段落，標記為刪除者，或指明為多餘的或偽造的部分。屬性包括：

### 類型/type

以任何方便的類型為刪除部分的類型分類。

### 狀態/status

可用來標示誤刪的內容，例如，多刪或少刪的文本。

### 經手者/hand

標明經手此一刪除者。

### <不明>/<unclear>

標誌因來源文件中非法或是限制公開的內容，而不可被確實轉錄的字、詞或段落。屬性包括：

### 未錄原因/reason

標明資料難以轉錄的原因。

### 責任類型/resp

標明可為<不明>元素所標誌字母、文字或段落負責的個人。

這些元素可用來記錄編輯者、轉錄者，或由 (手稿資料中的) 作者或代筆者做的變更。例如：若電子文獻的來源文件中寫著：

以下的元素是供供編輯做簡單更動之用。

內含令人想要更正的明顯錯誤，但是，同時也記錄多餘部分刪除的動作，因此：

以下的元素是供<刪除 經手者="LB">供</刪除>編輯做簡單更動之用。

屬性經手者 (**hand**) 的屬性值 LB，表示 LB 更正了重複的「供」字。

若來源文件中寫著：

以下的元素供編輯做簡單更動之用。

(例如, 「是」無意中被落掉了) 則可能被更正為:

以下的元素<加入文字 經手者="LB">是</加入文字>供<刪除 經手者="LB">供</刪除>編輯簡單更動之用。

屬性經手者 (**hand**) 的屬性值 LB, 表示 LB 更正了重複的「供」字。

這些元素不限由編者變更的部分, 它們也可以用來記錄手稿上的作者塗改。在手稿中, 作者原先寫「它真煩人啊, 好的跟屁蟲」, 然後劃掉「煩」並插入「纏」, 則可以標誌如下:

它真<刪除 經手者="DHL" 類型="取代">煩</刪除><加入文字 經手者="DHL" 地點="填入">纏</加入文字>人啊, 好的跟屁蟲

類似的, 元素<不明> (<unclear>) 與 <省略資料> (<gap>) 可以並用, 以標明非法資料的省略, 下例中顯示使用<加入文字> (<add>) 一併做校訂。

一百二十名令人滿意的常備兵加入我<不明><省略資料 未錄原因="難解"/></不明>馬上就對我有很大幫助<加入文字 經手者="編輯">在?</加入文字>對抗威明頓的大業上。

元素<刪除> (<del>) 標誌的資料是轉錄到電子文獻後被標誌為刪除的部分; 然而, <省略資料> (<gap>) 標誌電子文獻中被刪除資料的位置, 不論其是否非法。例如, 一組語料可能因其使用的外文引文太長而被省略:

<段落>...<日期>1611 年 12 月</日期>在<名稱> 類型="地點">柯汀屋</名稱>封地帳冊的表列範例如下所示。顯示賣蜂蜜的現金收入...</段落>  
<引句><省略資料 節文描述="引自帳冊" 未錄原因="丹麥文"/></引句>  
<段落>帳目總體結構的描述又再次...</段落>

其他語料 (特別是在掃描器廣為流傳之前建構者) 有系統的省略圖形與算式者:

<段落>在你螢幕的最下面, 模式行的後方是<專有名詞>小暫存區</專有名詞>。這個區域是 Emacs 回應你輸入指令、指定 Emacs 尋找的檔名, 以及蒐尋與取代的值等<省略資料 節文描述="圖解 Emacs 的螢幕" 未錄原因="圖像"/></段落>

## 11. 名稱、日期、數字及縮寫

TEI 架構中定義了標誌「具資料性」特性的元素, 它們可能出現在任何文件的任何位置。這些特性在某些原則之下特別令人感興趣: 它們與文本之外的物件相關, 例如: 人名、地名、數字及日期。在許多自然語言處理的應用中它們造成特別的問題, 因為在文件中它們會以各種不同方式的表現。本章節所描述的元素, 是為了使這些特質更為明確, 以降低包含這些特性的文本在處理時的複雜度。

### 11.1. 名稱與關聯字串



**關聯字串**，是一個關係到某人、某地、某物的詞彙。可使用以下兩個元素來標誌這類的字串：

<關聯字串>/<rs>

標誌一般性質的名稱或是關聯字串。屬性包括：

#### 類型/type

標示關聯字串所指的物件。屬性值可為人物、地點、船舶、元素等。

<名稱>/<name>

標誌一個正式的名稱或名詞片語。屬性包括：

#### 類型/type

標示由詞彙名命的物件的類型。

屬性**類型 (type)**用來分辨名稱屬於(例如)人名、地名、組織名或是其他可能者：

<引句>我親愛的<關聯字串 類型="人物">班奈特先生</關聯字串>,</引句>有一天他的妻子跟他說,<引句>你聽說了嗎<關聯字串 類型="地點">尼特菲爾德園</關聯字串>終於租出去了? </引句>

這是<關聯字串 類型="機構">債務辦公室</關聯字串>的原則之一,絕不,任何客戶都一樣,給簡單明瞭的答案,<關聯字串 類型="人物">巴納可先生</關聯字串>說,<引句>或許吧。</引句>

下例顯示,元素<關聯字串>(<rs>)可作為任何與人、地及其他種類的參考,且不必為正式的名稱或是片語。

<引句>我親愛的<關聯字串 類型="人物">班奈特先生</關聯字串>,</引句>有一天<關聯字串 類型="人物">他的妻子</關聯字串>跟他說...

相對於元素<關聯字串>(<rs>),元素<名稱>(<name>)為關聯字串中的特殊狀況,僅供正式名稱使用。可視為與元素<關聯字串>(<rs>)同義,或是在關聯字串包含正式名稱的混用狀況中,在元素<關聯字串>(<rs>)內層使用。

簡單地將某些東西標誌成名稱,通常並不足自動處理以全名為參考檢索詞的人名。出現在文本中的人名,或許拼字不全、僅含部分或是曖昧模糊。甚至人名的前置詞如「van」或「de la」,頭銜的來源可能因語言及國家的不同,未被納入人名的參考格式中。

下列屬性可以在上述兩個或是其他類似的元素中使用,來協助克服類似的困擾：

#### 其他識別/key

為已命名過的物件提供另一組識別代碼,如資料庫中的資料鍵。

#### 採用格式/reg

為使用名稱提供標準或規範的格式。

屬性**其他識別** (**key**) 相當有用，可整合四散在整個文獻中所有指向同一人或同一地的參考點：

```
<引句>我親愛的<關聯字串 類型="人物" 其他識別="BENM1">班奈特先生</關聯字串>，</引句>有一天<關聯字串 類型="人物" 其他識別="BENM2">他的妻子</關聯字串>跟他說，<引句>你聽說了嗎<關聯字串 類型="地點" 其他識別="NETP1">尼特菲爾德園</關聯字串>終於租出去了？ </引句>
```

這種用法與屬性**採用格式** (**reg**) 有所區別，屬性**採用格式** (**reg**) 提供關聯字串的標準格式的義意，如例下所示：

```
<名稱> 類型="人物" 其他識別="WADLM1" 採用格式="德·拉·馬瑞，華爾特·華爾特·德·拉·馬瑞</名稱> 1837年生於<名稱> 其他識別="KT1" 類型="國家">肯特</名稱>的<名稱> 其他識別="Ch1" 類型="地點">查爾頓</名稱>。
```

使用名稱與日期的附加標誌集，可以標誌更多正式名稱組件的細節。

## 11.2. 日期與時間

標誌更多日期與時間細節的標誌包含：

<日期>/<date>

標誌任何格式的日期，屬性包括：

### 曆法系統/calendar

標示日期使用的曆法系統

### 標準格式值/value

為日期賦與某種標準格式值，通常使用 yyyy-mm-dd。

<時間>/<time>

標誌定義一天之內時間的詞彙。屬性包括：

### 標準格式值/value

為時間賦與某種標準格式值。

屬性**標準格式值** (**value**) 的值可使用經認證的格式，如 ISO 8601，來規格化日期或時間的標準格式。不完整的日期或時間（如 1990、1990 年 9 月、12 日），通常可以簡單的略去未使用到的部分，不明確的日期與時間（例如：早秋、十點到十二點之間）可以日期或時間的間距的來表現。如果日期的結束或是時間的間距難以確認（例如，1230 年前的某一時間、萬聖節後幾日），可使用屬性**確定** (**exact**) 來說明。

範例：

```
<日期 標準格式值="1980-02-21">1980年2月21日</日期>
<日期 標準格式值="1990">1990</日期>
<日期 標準格式值="1990-09">1990年9月</日期>
```

```
於<日期 標準格式值="1977-06-12">（史丹佛）大學第八十六週年、美國建國
```

二百零一年，耶穌基督一千九百七十七年，六月的第二十一日</日期>授證。

<詩行>特別是在距午夜零時還有九個鐘點</詩行>  
<詩行>與<時間 標準格式值="15:00">下午的三點</時間></詩行>

### 11.3. 數字

數字可以正體或位數表現（二十一、xxi、21），且這些數字的表現方式與使用的語言有關（例如：英文的 5th，希臘文寫成 5；英文的 123,456.78，法文寫成 123.456,78）。在自然語言或機器翻譯的應用上，將文件中的數字與其他較「字彙的」部分區分開來相當有用。在其他應用上，用標準數記法記錄數值的能力很重要。元素 <數字>（<num>）提供的用法：

<數字>/<num>

標誌數字，不論其書寫形式。屬性包括：

#### 類型/type

標示數值的類型，建議的屬性值包括：分數（*fraction*）、序數（*ordinal*，例如，第二十一）、百分比（*percentage*）、基數（*cardinal*，實數，例如 21、21.5 等）。

#### 標準格式值/value

以應用的標準格式說明數字的值。

範例

```
<數字標準格式值="33">xxxiii</數字>
<數字 類型="基數" 標準格式值="21">二十一</數字>
<數字 類型="百分比" 標準格式值="10">百分之十</數字>
<數字 類型="百分比" 標準格式值="10">10%</數字>
<數字 類型="序數" 標準格式值="5">第五</數字>
```

### 11.4. 縮寫與全稱

就像名稱、日期及數字，縮寫也可以轉錄它的狀態或是全稱，縮寫可以不處理，或是以下列的元素來標誌：

<縮寫>/<abbr>

標誌任何形態的縮寫。屬性包括：

#### 全稱/expand

標示縮寫的全稱

#### 類型/type

允許編碼者以方便的類別為縮寫分類。簡單的屬性值包含：合併、刪節、圖略、及頭縮字。屬性類型（**type**）的值也可以使用標頭、圖形、組織等，描述物件所指的本質。

元素<縮寫> (<abbr>) 是分辨半字彙的項目的義意時很有用的方式，例如頭縮字或專業術語：

上述的討論可歸結如下：<縮寫>CC</縮寫>的識別由口徑定義，它激發每一個元素<縮寫>GSP</縮寫>的頻譜。每一個<縮寫>3GL</縮寫>或 <縮寫>4GL</縮寫>製造者的目前的標記訂在<縮寫>OOP</縮寫>的擴大部分。

屬性類型 (type) 也可以功能來區分縮寫字的類型，屬性全稱 (expan) 可以用來說明全稱：

<名稱><縮寫 類型="頭銜" 全稱="Doctor">Dr.</縮寫><縮寫 類型="簡寫" 全稱="Marilyn">M.</縮寫>Deegan</名稱>is the Director of the<縮寫 全稱="Computers in Teaching Initiative" 類型="頭縮字">CTI</縮寫> Centre for Textual Studies.

本元素在轉錄經常出現縮寫字的手稿資料時特別有用。

## 11.5. 地址

元素<地址> (<address>) 用來標誌各種郵寄的地址。內含一或多個<地址行> (<addrLine>) 元素，以標示每一行地址。

<地址>/<address>

標誌郵寄或其他種類的地址，例如，出版商地址、機構地址、個人地址。

<地址行>/<addrLine>

標誌地址中的一行。

以下為簡單的範例：

<地址>  
<地址行>Computer Center (M/C 135)</地址行>  
<地址行>1940 W. Taylor, Room 124</地址行>  
<地址行>Chicago, IL 60612-7352</地址行>  
<地址行>U.S.A.</地址行>  
</地址>

地址中的個別部分，可以更進一步以之前所討論的元素<名稱>( 章節「11.1. 名稱與關聯字串」) 來區分。

<地址>  
<地址行>Computer Center (M/C 135)</地址行>  
<地址行>1940 W. Taylor, Room 124</地址行>  
<地址行><名稱> 類型="城市">Chicago</名稱>, IL 60612-7352</地址行>  
<地址行><名稱> 類型="國家">USA</名稱></地址行>  
</地址>

## 12. 列表

任何形式的表列表，皆可以元素<列表> (<list>) 標誌。列表是一串排序、未排序，或字彙條列的文本項目。每個項目可以用一個標籤來處理（在字彙表中，標籤即為被定義的詞目）：

<列表>/<list>

標誌被組織成表的項目排列。屬性包括：

#### 類型/type

描述列表的形式。建議的屬性值包括：序數式 (*ordered*)、點列式 (*bulleted*，未以數字或文字編碼，而以點列符號分別條列者)、字彙表 (*gloss*，由一組技術字彙組成的表列，每一個字彙以元素<標籤> (<label>) 標誌，並伴以元素<項目> (<item>) 標誌的解釋或定義)、及簡單表列 (*simple*，各項目未編碼或符號分點的表列)

<項目>/<item>

標誌列表中的一個部件。

<標籤>/<label>

在列表中，標誌與某一項目相對應的標籤；在字彙表中，標誌被定義的詞彙。

列表中各別項目以元素<項目> (<item>) 標記。第一個<項目> (<item>) 可以選擇以元素<標頭> (<head>) 標誌成列表的標題。列表的編號可被省略（若可被重建），在每一項目中使用屬性**識別號** (**n**) 來表示，或（在極少數情況下）把它視為內容而以元素<標籤> (<label>) 標記。下列的例子都指同一表列：

```
<列表>
<標頭>小表列</標頭>
<項目>表列第一項</項目>
<項目>表列第二項</項目>
<項目>表列第三項</項目>
</列表>
```

```
<列表>
<標頭>小表列</標頭>
<項目 識別號="1">表列第一項</項目>
<項目 識別號="2">表列第二項</項目>
<項目 識別號="3">表列第三項</項目>
</列表>
```

```
<列表>
<標頭>小表列</標頭>
<標籤>1</標籤><項目>表列第一項</項目>
<標籤>2</標籤><項目>表列第二項</項目>
<標籤>3</標籤><項目>表列第三項</項目>
</列表>
```

上述的樣式在同一列表中不可混用。



簡單的兩欄式表單，可以詞彙表 (*glossary list*) 的方式來處理，標誌成「<list type="gloss">」。詞彙表中，一個項目由一個詞目 (*term*) 及一個解釋 (*gloss*) 組成，各以元素<標籤>及元素<項目>標記。與這兩個元素相對應的元素為<專有名詞> (<term>) 與<註解> (<gloss>)，這兩個元素可出現在文本中的任何位置。

```
<列表 類型="詞彙表">
<標頭>詞彙</標頭>
<標籤 語言="enm">nu</標籤>      <項目>now</項目>
<標籤 語言="enm">lhude</標籤>   <項目>loudly</項目>
<標籤 語言="enm">bloweth</標籤> <項目>blooms</項目>
<標籤 語言="enm">med</標籤>     <項目>meadow</項目>
<標籤 語言="enm">wude</標籤>    <項目>wood</項目>
<標籤 語言="enm">awe</標籤>     <項目>ewe</項目>
<標籤 語言="enm">lhouth</標籤>  <項目>lows</項目>
<標籤 語言="enm">sterteth</標籤> <項目>bounds, frisks</項目>
<標籤 語言="enm">verteth</標籤> <項目 語言="lat">pedit</項目>
<標籤 語言="enm">murie</標籤>   <項目>merrily</項目>
<標籤 語言="enm">swik</標籤>    <項目>cease</項目>
<標籤 語言="enm">naver</標籤>   <項目>never</項目>
</列表>
```

內部結構複雜的列表項目，可傾向視之為表單 (*table*)，此一特殊目的的標誌，定義見後面的章節「14. 表單」。

不論何種列表，其中的項目當然可以應需要層層包夾。下例中，一個詞彙表有兩個項目，每一個項目中都各自包含一個簡表：

```
<列表 類型="字彙表"><標籤>劣者</標籤>
<項目><列表 類型="簡表">
  <項目>我被棄置在於恐怖的無人島，放棄所有獲救的希望。</項目>
  <項目>我被孤立與隔絕，整個世界因此變得悲慘。</項目>
  <項目>我自人群中隔離——一個隱居者、一個被逐出人類社會者。</項目>
</列表> <!-- 第一組內嵌表列結束--></項目>
<標籤>優者</標籤>
<項目><列表 類型="簡表">
  <項目>但我還活著，不像其他船上的伙伴，我沒有淹死。</項目>
  <項目>但我也是僅有的一名，在所有船員中，得以免除一死...</項目>
  <項目>但我沒有捱餓，不會在無法提供食物的不毛之地上等死...</項目>
</列表><!-- 第二組內嵌表列結束--></項目>
</列表><!-- 字彙表結束-->
```

列表未必要以條列方式呈現。例如：

```
在這些互不相干的記錄上，寫著動物分成<列表 樣式="接續"><項目 識別號="a">屬於皇帝的，<項目 識別號="b">被薰香防腐的，<項目 識別號="c">受過訓練的，<項目 識別號="d">乳豬，<項目 識別號="e">美人魚，<項目 識別號="f">傳說動物，<項目 識別號="g">流浪狗，<項目 識別號="h">這些是包含在
```

分類中的，<項目 識別號="i">如果發瘋會顫抖的，<項目 識別號="j">無法計數的，<項目 識別號="k">以極細的駝毛刷畫成的，<項目 識別號="l">其他類，<項目 識別號="m">剛打破花瓶的，<項目 識別號="n">遠看像蒼蠅的。</列表>

書目的列表應以元素<書目表列> (<listBibl>) 標誌，下一節章節會說明。

## 13. 引用書目

區別引用書目在被轉錄文本中出現的位置，對研究通常有所助益，即使是只為了使此引用書目在文本印行時使用正確的格式。元素<引用書目> (<bibl>) 即為此目的而設：

<引用書目>/<bibl>

標誌結構鬆散的引用書目，可選擇是否明確標誌其中的構件。

若需要區分書目參考資料的構件，可適當的使用下列的元素。將這些需要特殊格式的部分（例如文章、書籍及學刊的題名）標誌出來，通常相當有用。其他的元素可因應特殊情況加入類似的細節。

<作者>/<author>

在書目參照中，標誌作品個人或團體作者的名字，或任何書目中的原始責任敘述。

<引用範圍>/<biblScope>

定義書目參照的範圍，例如：頁碼表列，或大型作品中被命名的部分。

<日期>/<date>

標誌任何格式的日期。

<編者>/<editor>

文件的第二責任類型敘述，例如編輯、編譯、翻譯等角色的個人、機構或組織（或其他類似者）的名稱。屬性包括：

### 角色/role

標明責任類型的性質。簡單的屬性包括：翻譯者、編譯者、插畫者等，預設值為編輯者。

<印行資訊>/<imprint>

集合文件出版或發行的相關資訊。

<出版者>/<publisher>

提供關於負責文件出版或發行的組織名稱。

<出版地>/<pubPlace>

標誌文件出版地的名稱。

<集叢>/<series>

標誌書籍或其他類型文件被集叢收錄的資訊。

<題名>/<title>

標誌作品的題名，不論是是文章、書籍、期刊或叢書，其他題名或是副題名也包含在內。屬性包括：

**類型/type**

對題名進行某種方式的分類，例如主要或附屬等類型。

**層級/level**

標示題名的書目層級 (*level*) 或分類。合法的屬性值詳見章節「6.1. 字型改變」。

例如，以下的編輯註可被轉錄如下：

他在 1445 年為華威克夏的國會的一員，卒於 1470 年 3 月 14 日。(據基特瑞吉，哈佛研究 五，摺頁第 88)

他在 1445 年為華威克夏的國會的一員，卒於 1470 年 3 月 14 日卒。(據<引用書目><作者>基特瑞吉</作者>，<題名>，哈佛研究</題名> <引用範圍>五，摺頁第 88</引用範圍></引用書目>)。

引用書目的表列，可以使用元素<書目表列> (<listBibl>)，它可以包含一系列的<引用書目>元素。

## 14. 表格

在任何文字處理系統中，表格代表巨大的挑戰。但是最起碼的簡單表格出現在大量的文本中，甚至出現在這個簡化過的 TEI 標籤集中。表格的標誌是必須的。下列的元素即在應因此需求：

<表格>/<table>

標誌以欄與列組成的表格形式呈現的文本。屬性包括：

**列數/rows**

標示表格的列數。

**欄數/cols**

標示表格每一列的欄數。

<列>/<row>

標誌表格的一列。屬性包括：

**角色/role**

標示此一系列中各格子所載資訊的種類。建議的屬性值包含做為標籤或是描述性資訊的標籤 (*labels*)，與填入實際資料的資料 (*data*)。

<格>/<cell>

標誌表格中的一格。屬性包括：

**角色/role**

標示此一格子所載資訊的種類。建議的屬性值包含做為標籤或是描述性資訊的標籤 (*labels*)，與填入實際資料的資料 (*data*)。

**欄數/cols**

標示此一格子所佔的欄數。

**列數/rows**



標示此一格子所佔的列數。

例如，狄福（Defoe）在《瘟疫年紀事》中使用如下的死亡人數表來顯示疫情的昇降：

```
<段落>它確實來得迅速，因為在同一週相鄰教區的葬禮數如下所列：
<表格 列="5" 欄="4">
<列 角色="資料">
<格 角色="標籤">聖李奧納多，岸溝區</格>
    <格>64</格> <格>84</格> <格>119</格></列>
<格 角色="標籤">聖伯多菲，主教門區</row>
    <格>65</格> <格>105</格> <格>116</格></列>
<格 角色="標籤">聖吉利斯，克里波門區</列>
    <格>213</格> <格>421</格> <格>554</格></列>
</表格>
<段落>...這些房舍的關閉，起先是以非常殘酷及非基督教的方式處理的，受到禁閉
的可憐人們吟唱痛苦的哀歌...</段落>
```

## 15. 圖解與圖象

文獻並不需要完全由文字組成。許多由行列排成的文本常包含圖表或插圖，圖象與文字糾結無法分解的文獻就更不用說了，電子資源中的圖文可互為補充。

編碼者可以使用本節所描述的元素，簡單地記錄圖象出現在文本中，可能的話附上簡短的內容敘述。同樣的元素也可以用在把數位版的圖象嵌入電子文獻中。

<圖象>/<figure>

標記文獻中圖像插入的位置，屬性包括：

**實體名稱/entity**

已被定義過的系統實體名稱，標記欲插入的數位版圖象。

<圖象描述>/<figDesc>

標誌圖象的出現或內容的文字敘述，只記錄而不顯示圖形時使用。

任何伴隨圖形出現的文字資訊，例如標題或圖說，可包含在該元素<圖象>（<figure>）內，使用元素<標頭>（<head>）或多個元素<段落>（<p>），也可以標誌任何出現在圖形本身的文字。強烈推薦補充影象的文字敘述，做為元素<圖象描述>（<figDesc>）的內容，以便應用程式無法顯示圖形時使用，同時向無法看到圖像的讀者描述文獻的內容。（這類的文字通常不視之為正式文獻的一部分。）

這些元素最簡單的使用方式是標記圖象的位置，例如：

```
<分頁 識別號="412" />
<圖象></圖象>
<分頁 識別號="413" />
```

(注意，結束標記不可省略，即便元素不包含內容)更常用的是，一個圖象至少有一個識別題名，且應該以元素<標頭> (<head>) 標誌。標誌影象的簡述也很方便，如下例：

```
<圖象>
  <標頭>費茲威格先生的舞會</標頭>
  <圖象描述>克魯威克珊克之夜，費茲威格先生帶頭狂歡。</圖象描述>
</圖象>
```

可提供相關的數位版圖象的時候，把它嵌在文獻的適當位置顯然更好。圖象元素，例如照片，原則上與文獻所包含的文本是分開儲存的實體(檔案)，而且使用不同的標註(儲存格式)。TEI Lite 的 DTD 在 SGML 下支援使用 CGM、PNG、TIFF、GIF 或 JPEG 標準編碼的圖象，相對應的檔案標註為 cgm、png、tiff、gif 及 jpeg。

不論圖形編碼採用何種格式，都可以同樣的方式嵌入文獻中。首先要宣告一個特殊類型的實體來指定實體的名稱、一個外部的識別碼(例如檔案名稱)，以及使用的標註。例如，假設費茲威格先生的舞會的數位圖象被存在一個名為「fezzi.tff」的 TIFF 檔中，就需要實體宣告如下：

```
<!ENTITY fezziPic SYSTEM "fezzi.tff" NDATA tiff>
```

所有這一類的宣告，必須在文獻本身製作前就先行處理，處理的方式超出本文獻的範圍，但是在「XML 入門<sup>9</sup>」及多種 SGML 與 XML 的介紹文件中都有討論。

在上述宣告的實行之後，要將數位化的影象嵌在文獻的正確位置，只需要在元素<圖象> (<figure>) 補上屬性**實體名稱** (entity) 的屬性值：

```
<圖象 實體="fezziPic">
  <標頭>費茲威格先生的舞會</標頭>
  <圖象描述>克魯威克珊克之夜，費茲威格先生帶頭狂歡。</圖象描述></圖象>
```

## 16. 詮釋與分析

有一種說法：所有的標記都是詮釋或分析的一種形式。但是要以任何通用的方式確實分辨「客觀」與「主觀」的資訊確實有困難，而且甚至於不可能的。相較於客觀判斷，主觀判斷基本上被認為較易引發的爭議仍然是事實。許多學者傾向只有當這類詮釋可能提醒讀者：相較於其他的標誌，該部分被視為開放討論，才會使用。本章節描述 TEI 架構中幾個符合這種需求的元素。

### 16.1. 拼寫正確的句子

<sup>9</sup> <http://www.tei-c.org/Guidelines2/gentleintro.html>

詮釋的基本上涵括的範圍橫跨整個文本，與其他結構單元無特殊的關係。一個要被用在詮釋的有用序言，因此被分割成分散與可識別的單元，每一個單元帶著一個標籤，做為一種「正規的參考資料」。為了便於使用，這些單元不可相互重疊，亦不可互相包含。各單元可以簡單地使用下列元素來表現：

<成句單位>/<s>

標誌文獻中的成句單位 (*s-unit*)。目的在建立一個適用於文本整體的簡單標準參考架構。屬性包括：

#### 類型/type

為單元分類 (例如：敘述句、疑問句等。)

如名稱所建議，元素<成句單位> (<s>) 最常用 (至少在在語言學的應用上) 來標誌拼寫正確的句子 (*orthographic sentences*)，也就是說，由標點符號等拼寫正確的特徵來定義的單元。例如，之前討論選自《筋愛的》章節可被切割為如下的成句單位：

```
<分頁 識別號="474" />
<第一層 類型="章" 識別號="38">
<段落><成句單位 識別號="001">讀者啊，我嫁給他了。</成句單位>
<成句單位 識別號="002">我們舉行了寧靜的婚禮，</成句單位>
<成句單位 識別號="003">只有他和我，牧師和執事在場。</成句單位>
<成句單位 識別號="004">當我們從教堂回來，我走進宅邸的廚房，瑪麗正在料理晚餐，約翰正在清理刀子，我說：</成句單位>
<段落><引句><成句單位 識別號="005">「瑪麗，我今早已經嫁給羅徹斯特先生了。」</成句單位></引句> ...
```

注意元素<s>不可以使用巢狀結構：一個<成句單位>的開始，表示上一個<成句單位>已經結束。當成句單位以上述的方法標誌，建議將整個文本悉數逐一標誌，如此被分析的文本中的每一個字，都只被唯一的<成句單位>標誌，可用識別代碼為它指定唯一的參考值。如果使用的識別代碼在文獻中為唯一值，則上述例子中的屬性識別號 (*n*)，改為屬性識別名稱 (*id*) 會更好。

## 16.2. 一般目的的詮釋元素

一般目的的分割元素<分割段落> (<seg>)，是用來識別未能以其他類型元素標誌的交互參照與超連結的，前面已經介紹過 (見「8. 參照連結」)。<分割段落>識別文本中某些片語層次的部分，編碼者可為這些片語指定一個使用者設定的類型 (*type*)，最好同時指定唯一的識別碼。如此則可標記已出版的 TEI 指引中未提供的文本特徵。

例如，指引中並未提供元素<道白句> (<apostrophe>) 來標誌在文學作品中講述者直接向讀者 (或聽者) 講述的部分。其中一種處理的方式，是將它們視為元素<引句> (<q>) 的情境，並以相對應的人物 (*who*) 屬性值作區別。另一種較為簡單，實際上也較普遍的解決方式，則是使用元素<分割段落> (<seg>)，如下所示：

```
<第一層 類型="章" 識別號="38">
<段落><分割段落 類型="道白句">讀者啊，我嫁給他了。</分割段落>
我們舉行了寧靜的婚禮，...
```

元素<分割段落>中的屬性**類型 (type)**可使用任何屬性值，所以可用來記錄任何片語層次的現象。最好可以把這些屬性的值及其特性記錄在標頭裡。

元素<分割段落>具某一類型（不同於元素<成句單位> (<s>) 只是表面的形似）者，可以在相同或其他的類型的<分割段落>中以巢狀結構疊套。這種方式可表現相當複雜的結構，部分例子在之前的章節「8.3 連結屬性」中已提供。但是，因為必須遵守元素正確地疊套的條件，元素間不得互相疊壓，使它無法應付無視於文獻階層的武斷分割的一般詮釋需求。詮釋的本身也只能以單一的屬性值表現**類型 (type)**。

但元素<詮釋> (<interp>) 並不套用上述的限制，它提供強而有力的特性，可以用相對直接的方式為複雜的詮釋資訊編碼。

<詮釋>/<interp>

提供詮釋性的註解，可以連結到一段文字。屬性包括：

**標準格式值/value**

識別註釋的特別現象。

**責任類型/resp**

標明對註釋內容負責任者。

**類型/type**

標明文章中註釋現象的種類。參考值包括圖象 (*image*)、特性 (*character*)、主題 (*theme*)、暗示 (*allusion*)、或是已被實際識別的特殊類型名稱。

**實際位置/inst**

指向由元素標誌的分析或詮釋的實際內容。

<詮釋群組>/<interpGrp>

將<詮釋>集合在一起。

本元素允許編碼者具體說明詮釋的類別，以及詮釋相關類別的特定例子，當使用元素<分割段落>時，只能簡單地說某部分是道白句，而使用元素<詮釋>時，則可說這是一個大類（敘述角色的）中的實例（道白句部分）。

除此之外，<詮釋>是一個空的元素，必須連結到對應的段落，不論是藉由章節「8.3.連結屬性」討論過的屬性**分析 (ana)**，或是藉由元素所有的屬性**實際位置 (inst)**。這意味著任何分析類型都可以被表現，不需遵循文獻的階層。為便於集合特定類型的分析，元素<詮釋群組>應運而生。

例如，假設使用者想要標誌文本的各種面向，如主旨、主題素材、敘述的角色及敘事中各場景的位置。以取自《劬愛》中的篇章為例作不同的配置，可以連結到道白、誇飾與隱喻的修辭特性；教堂、僕役、烹飪、郵務及蜜月等參考主題素材；或教堂、廚房及不特定地點（客廳？）的場景地點。

這些詮釋可以放在元素<文件內容> (<text>) 的任何地方，將所有的詮釋放在同一處（例如，文前或文後資訊的獨立區段）似乎是個好辦法，如下例所示：

```
<正文後資訊>
<第一層 類型="詮釋">
<段落><詮釋 識別名稱="性質—道白" 責任類型="LB, MSM"
    類型="講述性質" 標準格式值="道白"/>
<詮釋 識別名稱="性質—誇飾" 責任類型="LB, MSM"
    類型="講述性質" 標準格式值="誇飾"/>
<!-- ... -->
<詮釋 識別名稱="場景—教堂" 責任類型="LB, MSM"
    類型="場景" 標準格式值="教堂"/>
<!-- ... -->
<詮釋 識別名稱="相關—教堂" resp="LB, MSM"
    類型="相關" 標準格式值="教堂"/>
<詮釋 識別名稱="相關—神職人員" resp="LB, MSM"
    類型="相關" 標準格式值="神職人員"/>
<!-- ... -->
</段落></第一層>
```

此處明顯重覆的編碼，可以使用元素<詮釋群組> (<interpGrp>) 集合所有的共用相同屬性值的元素<詮釋> (<interp>)，達到相當程度地減省，如下所示：

```
<正文後資訊>
<第一層 類型="詮釋">
<段落>
<詮釋群組 類型="講述性質" 責任類型="LB, MSM">
<詮釋 識別名稱="性質—道白" 標準格式值="道白"/>
<詮釋 識別名稱="性質—誇飾" 標準格式值="誇飾"/>
<詮釋 識別名稱="性質—隱喻" 標準格式值="隱喻"/>
<!-- ... -->
</詮釋群組>
<詮釋群組 類型="場景" 責任類型="LB, MSM">
<詮釋 識別名稱="場景—教堂" 標準格式值="教堂"/>
<詮釋 識別名稱="場景—廚房" 標準格式值="廚房"/>
<詮釋 識別名稱="場景—未設" 標準格式值="未設"/>
<!-- ... -->
</詮釋群組>
<詮釋群組 類型="相關" 責任類型="LB, MSM">
<詮釋 識別名稱="相關—教堂" 標準格式值="教堂"/>
<詮釋 識別名稱="相關—神職人員" 標準格式值="神職人員"/>
<詮釋 識別名稱="相關—烹煮" 標準格式值="烹煮"/>
<!-- ... -->
</詮釋群組>
</段落></第一層>
```



這些詮釋元素一旦被定義，就可與文本的部分連結，應用方式有兩種，可僅採其一，或是二者並用。屬性分析（ana）可在任何適用的元素內使用：

```
<第一層 類型="章" 識別號="38">
<p 識別名稱="P38.1" 分析="場景—教堂 場景—廚房"></段落>
<成句單位 識別名稱="P38.1.1" 分析="性質—道白">讀者啊，我嫁給他了。</成句單位>
...
```

注意此例因段落中有兩個場景（教堂內及廚房內），兩者都有識別代碼。

另一種方式是，使用屬性實際位置（inst），元素<詮釋>（<interp>）即可指向文本中所有相關的部分：

```
<詮釋 識別名稱="性質—道白" 類型="講述性質" 責任類型="LB, MSM" 標準格式值="道白" 實際位置="P38.1.1"/>
<!-- ... -->
<詮釋 識別名稱="場景—教堂" 類型="事件—場景" 標準格式值="教堂" 實際位置="P38.1" 責任類型="LB, MSM"/>
<詮釋 識別名稱="場景—廚房" 類型="事件—場景" 標準格式值="廚房" 實際位置="P38.1" 責任類型="LB, MSM"/>
<!-- ... -->
```

元素<詮釋>（<interp>）不限制任何特定類型分析，上列的文學分析不過是可能用法之一，使用者同樣也可善用<詮釋>（<interp>）來標誌局部演說的語言分析。例如，章節「8.3 連結屬性」中的例句所假設的語言分析，可以標誌如下：

```
<詮釋 識別名稱="NP1" 類型="pos" 標準格式值="名詞片語，單數"/>
<詮釋 識別名稱="VV1" 類型="pos" 標準格式值="語尾變化動詞，單數現在式"/>
...
```

## 17. 科技文獻

雖然本文獻的焦點是用 TEI 架構對已存在的「未電子化」文獻編碼，但是同一架構也可以用在新的文獻的編碼。在新文獻（例如本文獻）的預製上，XML 有多處可資推薦：文獻結構可清楚的呈現，而且同一電子文本可依不同目的重覆使用——例如，提供線上超文件或可瀏覽的版本，以及一般媒體使用的規格化排版本。

為簡化起見，TEI Lite 收錄了一點附加元素為主 TEI DTD 作擴充，用來標誌一般科技文獻，以及特殊的 XML 相關文獻特徵。

### 17.1. 為科技文件而設的附加元素

下列元素或可被用來標誌科技文獻的特徵：



<範例>/<eg>

標誌討論中的技術主題所舉的一則短例，例如：程式的片斷或是 SGML 編碼的樣本。

<程式碼>/<code>

標誌使用正規語言（通常是程式語言）的程式片斷。

<識別詞>/<ident>

標誌某種識別名稱，例如，XML 元素或屬性的名稱及其異稱。

<類名識別>/<gi>

標誌識別名稱的特殊類型：XML 的通用識別或是元素名稱。

<關鍵術語>/<kw>

標誌某種正規程式語言中的關鍵詞。

<公式>/<formula>

標誌數學或化學公式，可選用部分非 XML 的標記法。屬性包括：

#### 標記法/notation

標示用來表示公式本體的標記法。預設值是 *tex*，意為行公式使用 TeX 計算機表示系統呈現。

下例顯示這些元素如何為福傳（Fortran）程式語言的課程簡介中的段落編碼：

<段落>用下列程式來介紹一種語言是傳統：

<範例>

```
CHAR*12 GRTG
GRTG = 'HELLO WORLD'
PRINT *, GRTG
END
```

</範例></段落>

<段落>這個簡單的例子首先在<code>CHAR\*12 GRTG</code>這一行中以宣告變數<ident>GRTG</ident>，它識別了<ident>GRTG</ident>是由 12 個位元組組成，類型為<kw>CHAR</kw>（文字）。這個變數的值隨後被指定為<指涉>HELLO WORLD</指涉>。接著是一個<kw>PRINT</kw>（列印）指令及一個<kw>END</kw>（結束）指令。

在上面的文本中，一個格式的運用，可以適切地架構成格式範例（例如，保留空行，或是使用醒目的字體）。相同的，<識別詞>（<ident>）與<關鍵術語>（<kw>）之類標記的使用，對建構有用的索引大有幫助。

應使用元素<公式>（<formula>）將文本中的數學或化學公式標誌成特殊項目。因為公式通常包含一般文本不會出現的大量特殊印刷排版特徵，通常必須以特殊的標記法呈現公式的本體。所用的記符法應以屬性標記法（**notation**）標明，如下例所示：

```
<公式 標記法="tex">
  \ (E = mc^{2}) \
</公式>
```

Tex 並非 TEI Lite DTD 裡已定義的標記法，因此必須在 DTD 的子集裡由記符法宣告定義。

當 XML 編碼即為討論主題的科技文獻本身就以 XML 編碼時，就會出現特殊的問題。在這類的文獻中，最基本的部分顯然就是從標誌文獻的內容以及高度相似的終止標記中，清楚地區辨範例出現的標誌。簡單的解決方案之一是使用預先定義的實體參考「lt」，來表示以一個「<」字元標記 XML 標誌範例的起始。較常用的解決方案是將每一例範例的整體切分出來視之為內含資料，當檢查器檢查文件時，這部分不會被當做 XML 標誌來掃描。這是以被稱為字元資料 (CDATA) 標誌區段的特殊 XML 結構來包覆記錄，如下例所示：

```
<段落>一個列表應編碼如下：
<範例><![ CDATA [
  <列表>
  <項目>表列第一項</項目>
  <項目>第二項</項目>
  </列表>
]]>
</範例>
<類名識別>列表</類名識別>元素由一系列<類名識別>項目</類名識別>組成。
```

上例中使用的元素<列表> (<list>)，將不會被視為正文的正規部分，因為它被嵌在標誌的區段中 (以特殊的標誌「<![CDATA[」宣告起始並以「]]>」結束)。

也要注意元素<類名識別>/<gi>在標誌參照文本內容中的元素名稱 (或通用識別) 的使用。

## 17.2. 自動產生區段

大多數的現代文獻製作系統都有自動產生完整的目次或是索引的能力。TEI Lite 架構提供一個元素來標誌這類自動產生區段的位置。

<自動產生區段>/<divGen>

標示由文字處理應用程式自動產生的文字區段的位置。屬性包括：

### 類型/type

標示產出的文本為何種類型 (例如，索引、目次等)。參考值包括：索引 (*index*，自動產生並插入在這個點上的索引)、目次 (*toc*)、圖形表列 (*figlist*)、表單表列 (*tablist*)。

元素<自動產生區段> (<divGen>) 可依其章節的合理性插在任何位置，如下例所示：

```
<正文前資訊>
<題名頁> ... </題名頁>
```

```

<自動產生區段 類型="目次"/>
<文字區段 類型="前言"><標頭>前言</標頭> ... </文字區段>
</正文前資訊>
<正文> ... </正文>
<正文後資訊>
<第一層><標頭>附錄</標頭> ... </第一層>
<自動產生區段 類型="索引" 識別號="索引"/>
</正文後資訊>

```

上例同時展示使用屬性來區辨不同類型的自動產生章節：首先為目次，其次是索引。

對既有的索引或目次進行編碼時（相對於自動產生者），應使用章節「12. 列表」所討論的元素<列表>（<list>）。

### 17.3. 產生索引

從標誌良好的文獻中產生目次，對自動化文書處理器來說通常不是問題，但是，製作高品質的索引，通常需要更仔細的標誌。雖然抽取所有出現過的<專有名詞>（<term>）或<名稱>（<name>）元素對索引來說是一個好的出發點，但是，簡單地以某種特殊方法製造一個被標誌部分的總表是不夠的。

TEI DTD 提供一個特殊功能的<索引>（<index>）標記，可用來標誌文獻中應被索引的部分，及索引產生的方式。

<索引>/<index>

標誌一個因某種目的而被索引的位置。屬性包括：

#### 主要款目/level1

索引的主要款目。

#### 第二層複分/level2

索引的第二層複分，需要時可採用。

#### 第三層複分/level3

索引的第三層複分，需要時可採用。

#### 第四層複分/level4

索引的第四層複分，需要時可採用。

#### 索引/index

標示索引款目屬於哪一個索引（若有多個索引）。

例如，本章節的第二段可標誌如下：

```

...
TEI 選錄版也提供具特殊目的<類名識別>索引</類名識別>標記
<索引 主要款目="索引"/>
<索引 主要款目="索引（標記）" 第二層複分="用來產生索引"/>

```

## 可能被用在...

元素<索引>也可以提供詮釋或分析資訊的形式。例如，在「奧維德」(Ovid)的研究中，要做文體比較研究，可能需要記錄詩所參照的所有不同角色。下例兩行取自變形記(*Metamorphoses*)，這類研究會記錄詩人參照朱彼特(Jupiter，同宙斯(deus)、同祂 se、及懺悔(confiteor)的主詞〔由第 227 號的語尾變化而來〕)，參照朱彼特化成公牛(公牛形象的化身(imago tauri fallacies)與載(teneo)的主詞)等等。<sup>10</sup>

```
<詩行 識別號="3.001"> iamque deus posita fallacis imagine tauri
<詩行 識別號="3.002"> se confessus erat Dictaeaque rura tenebat</
詩行>
```

(譯註：詩的大義為宙斯化身公牛的形象，載歐羅芭到了狄克特之岸)

這些需求可適用章節「7. 附註」中所討論到的元素<附註> (<note>)，或章節「16. 詮釋與分析」中所討論的元素。此處展示如何使用元素<索引> (<index>) 來滿足此一需求。

假設物件會產生一或多個索引：一個是神名索引(稱為 *dn*)，另一個是專有名詞索引(稱為 *on*)，第三個是代名詞索引(稱為 *pr*)，餘以此類推。可行的記錄方法之一如下所示：

```
<l 識別號="3.001"> iamque deus posita fallacis imagine tauri
  <索引 索引="dn" 主要款目="朱彼特" 第二層複分="宙斯" />
  <索引 索引="on" 主要款目="朱彼特(白色公牛)"
    第二層複分="公牛形象的化身" /></詩行>
<l 識別號="3.002"> se confessus erat Dictaeaque rura tenebat
  <索引 索引="pr" 主要款目="朱彼特" 第二層複分="祂" />
  <索引 索引="v" 主要款目="朱彼特" 第二層複分="懺悔(v227)" />
  <索引 索引="mons" 主要款目="狄克特" 第二層複分="狄克特之岸" />
  <索引 索引="regio" 主要款目="克里特島" 第二層複分="狄克特之岸" />
  <索引 索引="v" 主要款目="朱彼特(白色公牛)"
    第二層複分="載(v9)" /></詩行>
```

上例中每一個元素<索引> (<index>)，都會在相對位置產生索引款目，屬性**主要款目**(*level1*)的值为主標目，屬性**第二層複分**(*level2*)的值为第二關鍵字，所包含引用字的形式為主格。實際的參考資料取自元素<索引> (<index>) 中出現的脈絡，例如，上例中元素<詩行> (<l>) 所包含的識別代碼。

## 18. 字集、發音符號等

隨著 XML 與 XML 採用 Unicode 做為所有文件的必備字集，原先與全球語言及書寫系統分歧相關的大部分問題已大幅減少。特別是對以歐語標準格式的工作，幾乎

10 apply-templates select="text//note[@place='foot']" mode="printnotes"/>

不需要特別的動作：任何 XML 編輯器都允許直接輸入重音節字母或其他「非 ASCII 碼」字元，而且不論是採 Unicode 或採參考字元碼來編碼，這些字元應設法存成一個可在不同系統間直接轉換的最終檔。

然而，顧及其他舊系統的相容性，TEI Lite DTD 宣告了一些最廣泛使用的字碼集，讓這些字形可以字元助記法輸入與儲存。

在合法的 TEI 檔案中，使用者可使用自己的字碼。如果使用者想要並提供這些字碼的宣告，可將這些字碼對應到相應的 Unicode。標準編碼（使用長編碼）擁有的優點是清晰；預設的字元對任何說英文者而言是當然是明確的，他們認識已被編碼的字元，通常甚至沒有來源表列。但許多表現重音節字母的舊架構並非如此。

當使用者需要的字元並不在公用字集中，使用者可能想用 ISO 公用字集的編碼規則產生一個字碼，如下面所述者：

### 連音字/digraphs

表示連音字（兩個字母發一個音節）的字碼，由一串綁在一起的字母表現。如果需要大寫格式，兩個字母都要大寫（記得，這種狀況在字碼中通常很明顯），例如：aelig (æ), AElig (Æ) szlig (ß)。

### 發音與音節符號/diacritics and accents

用來表示多數西歐語言中含音節字母的編碼，在含音節標記的字母後面附加下列字串中的一種，大小寫均可。

#### 曲音符號/umlaut

曲音以 uml 表示，例如：auml (ä), Auml (Ä), euml (ë), iuml (sic: ï), ouml (ö), Ouml (Ö), uuml (ü), Uuml (Ü)。

#### 重音節符號/acute

急促音或是重音節以 acute 表示，例如：aacute (á), eacute (é), Eacute (É), iacute (í), oacute (ó), uacute (ú)。

#### 抑音節符號/grave

抑音節以 grave 表示，例如：agrave (à), egrave (è), igrave (î), ograve (ò), ugrave (ù)。

#### 曲折符號/circumflex

曲折音符號以 circ 表示，例如：acirc (â), ecirc (ê), Ecirc (Ê), icirc (î), ocirc (ô), ucirc (û)。

#### 鼻音符號/tilde

鼻音符號以 tilde 表示，例如：atilde (ã), Atilde (Ã), ntilde (ñ), Ntilde (Ñ), otilde (õ), Otilde (Ö)。

#### 子音 consonants

以下是西歐語言遇到特殊子音時推薦的字碼：ccedil (ç), Ccedil (Ç), eth (þ, 小寫的 eth 或盎格魯撒克遜或冰島語的劃線 d), ETH (大寫的 eth), thorn (小寫的 thorn), THORN (大寫的 thorn), szlig (德文中的 s-z 的連字或 *esszet*、ß)。

#### 標點符號/punctuation marks



以下推薦某些常見標點符號使用的字碼：ldquo（左雙引號括弧，在文字上方像 66 者）、rdquo（右雙引號括弧，在文字上方像 99 者）、mdash（一個內夾的連字符號）、hellip（水平省略符號，三個相近的點號）、rsquo（右單引號，在文字上方像 9 者）。

## 19. 正文前與正文後資訊

### 19.1. 正文前資訊

為了各種目的，特別是在早期的文本中，正文前面的資料如題名頁、長篇序文等，可額外提供非常有用的語言或是社會資訊。TEI P3 提供一組推薦元素來區別正文前資訊中最常遭遇的文本元素，本章為其擇要。

#### 19.1.1. 題名頁

題名頁的開頭應以元素<題名頁>（<titlePage>）標誌。題名頁上包含的所有的文本都應轉換，並以下列適當的元素標誌：

<題名頁>/<titlePage>

標誌文本的題名頁，出現在正文前或正文後的資訊之中。

<文獻題名>/<docTitle>

標誌題名頁上所載的文獻題名，包含構成題名的所有組件。必須再標誌成不同的<題名部件>（<titlePart>）。

<題名部件>/<titlePart>

依題名頁上所示，標誌作品題名的部分或區段，也可以標誌題名頁上不屬文獻標題、著作身分等無目的性的片段。屬性包括：

#### 類型/type

標明題名附屬區段的特性。建議的屬性值包括：主要題名（*main*）、副題名（*sub*）、敘述題名（*desc*，包含在題名中的敘述性作品解說）及其他題名（*alt*）。

<題名頁責任敘述>/<byline>

標誌在題名頁、作品開頭或結尾處原始責任敘述。

<文獻作者>/<docAuthor>

標誌在題名頁所載的文獻作者名字（通常，用但並不一定要在元素<題名頁責任敘述>（<byline>）中使用）。

<文獻日期>/<docDate>

標誌題名頁所載的文獻日期。

<文獻版本>/<docEdition>

標誌文獻的題名頁所載的編輯敘述。

<文獻印行資訊>/<docImprint>



標誌在題名頁下方（通常如此）所載的印行敘述（出版地、出版日期、出版者名稱）。

<卷首引句>/<epigraph>

標誌出現在章節開端或題名頁上具名或不具名的引文。

字樣的差別應以前面章節已敘述的屬性樣式（rend）標誌。題名中裝飾性字母所佔的空間與大小等細節描述，指引尚未提供。語言的變更應視需求適切使用屬性語言（lang）或是元素<外國語言>（<foreign>）標誌。名稱，不論出現在何處，應像其他部分一樣使用元素<名稱>（<name>）標誌。

題名頁範例兩則如下：

```
<題名頁 樣式="羅馬體">
  <文獻題名><題名部件 類型="主要">
    復樂園<醒目標示>第四冊</醒目標示>中的一首詩。
  </題名部件>
  <題名部件>
    附加<題名>力士參孫</題名>。
  </題名部件>
</文獻題名>
<題名頁責任敘述>作者<文獻作者>約翰·米爾頓</文獻作者></題名頁責任敘述>
<文獻印行資訊><名稱>倫敦</名稱>，
  印行者<名稱>約翰·米爾頓</名稱>
  <名稱>約翰·史塔基</名稱>
  近<名稱>教堂區</名稱>
  <名稱>佛利特街</名稱>
  的<名稱>主座教堂</名稱>
</文獻印行資訊>
<文獻日期>1671</文獻日期>
</題名頁>
```

```
<題名頁>
  <文獻題名><題名部件 類型="主要">
    諾曼人的征服以後的英格蘭皇后生活；</題名部件>
  <題名部件 類型="副標">她們的宮廷軼聞
</題名部件></文獻題名>
  <題名部件>官方記錄與其他可靠的私密文獻首度公開</題名部件>
  <文獻版本>新版，附更正與更新</文獻版本>
  <題名頁責任敘述>作者<文獻作者>阿涅斯·史崔克蘭</文獻作者></題名頁責任敘述>
  <卷首引句>
    <引句>舊時英雄角色堆疊的古老寶藏，由我開啟。</引句>
  <引用書目>包蒙</引用書目>
```

```

</卷首引句>
<文獻印行資訊>費城：布蘭恰特與李亞</文獻印行資訊>
<文獻日期>1860</文獻日期>
</題名頁>

```

### 19.1.2. 正文前資訊

正文前資訊的主要文字區塊應以元素<文字區段> (<div>) 或<第一層> (<div1>) 標誌。以下建議屬性類型 (**type**) 的屬性值，可用來區分正文前資訊常見的幾種類型：

#### **前言/foreword**

以讀者為對象，由作者、編者或出版者寫作，可能採信件格式。

#### **序文/preface**

以讀者為對象，由作者、編者或出版者寫作，可能採信件格式。

#### **獻辭/dedication**

以讀者之外的某人為對象的文本（通常是信件），作者基本上是以此作向提到的人物致意。

#### **摘要/abstract**

概述作品內容的綱要文章。

#### **謝辭/ack**

謝辭

#### **目次/contents**

目次表（一般而言應以元素<列表> (<list>) 標誌）。

#### **卷頭插畫/frontispiece**

卷頭插畫，其中可能包含文字。

如同任何文本區段，正文前資訊的內容也像其他區段一樣具有結構鬆散或非結構性的元素。通常以某種應以元素<標頭> (<head>) 標誌的標頭或題名起頭。書信格式包含下列附加元素：

#### **<致敬辭>/<salute>**

在信件或序言等文類中，標誌前置的致敬語或問候語、獻詞或其他片段，或是結尾的致敬語。

#### **<簽名>/<signed>**

標誌在結尾處的問候語、前言、獻辭或是文本其他區段的簽名。

#### **<題名頁責任敘述>/<byline>**

標誌在題名頁、作品開頭或結尾處原始責任敘述。

#### **<文件日期>/<dateline>**

標誌信件、報紙報導或其他文件的出版品前置於卷頭或附加為卷尾的地點、日期、時間等的簡短敘述。

#### **<大綱>/<argument>**

針對正文所含的子題做有條理的表列或文字敘述。

<引文>/<cit>

源出其他文獻的引文，含來源的書目參考資料。

<文件開頭>/<opener>

集合開頭部分的文件日期、題名頁責任敘述、致敬辭及其他做為文件序曲的類似段落，特別是用信件上。

<文件末端>/<closer>

集合日期行、責任敘述、致敬詞及其他出現在結尾部分的類似段落，特別是在信件上。

當然，出現在文本其他地方的書信，一樣使用這些元素。

下例所示為米爾頓作品「酒神」(Comus)前面的獻辭，應標誌如下：

```
<文字區段 類型="獻辭">
<標頭>致令人尊敬的<名稱>維斯康特領主，約翰·布拉克利</名稱>，布利里吉瓦
特伯爵的兒子與繼承人。</標頭>
<致敬辭>吾主，</致敬辭>
<段落>此<醒目標示>詩</醒目標示>，承載首度的慶典，自您及您高貴家族中其他
成員的誕生...
此中代表你所握的<名稱>瑟西斯杖</名稱>，於今全然展現。
<文件末端>
<致敬辭>你忠誠，且最卑微的僕人</致敬辭>
<簽名><名稱>拉維斯</名稱></簽名>
</文件末端>
</文字區段>
```

## 19.2. 正文後資訊

### 19.2.1. 正文後資訊的結構區分

因為實際出版的各別差異，上述正文前資訊的元素，正文後資訊實際上都可以使用，同樣的元素在該用的地方就用。此外，正文後資訊可在元素正文後資訊(<back>)中包含下列屬性。就像正文的結構性區段，正文後資訊應以元素<文字區段>(<div>)或<第一層>(<div1>)標誌，並以建議的**類型 (type)**屬性值區分：

**附錄/appendix**

附錄。

**字彙表/glossary**

字與定義的表列，基本上使用「list type=gloss」的型式。

**附註/notes**

一系列的<附註> (<note>)。

### 書目/bibliography

一系列的書目參照，一般而言使用特定的書目表列元素<書目表列> (<listBibl>) 格式，其中的項目是各別的元素<引用書目> (<bibl>)。

### 索引/index

一組索引款目，可能以結構表列或詞彙表來表現，可選用領頭元素<標頭> (<head>)，也或許用某些簡介或結束的文本段落 (TEI P3 定義其他製作文獻時產生索引的特殊元素，在前面的章節「17.3 產生索引」已敘述)。

### 書尾題記/colophon

在書的結尾處描述本書印行的地點、時間及由誰印行，現在的書通常會提供製作的細節與識別使用的字樣。

## 20. 電子題名頁

每一個 TEI 文本都有文件標頭，提供的資訊可類比至印刷本書名頁的資訊。TEI 標頭以元素<TEI 標頭> (<teiHeader>) 標誌，包含四個主要部分：

<檔案描述>/<fileDesc>

標誌電子檔的完整書目描述。

<編碼描述>/<encodingDesc>

記載電子文件與來源文件的關係，或電子文件的源由。

<背景描述>/<profileDesc>

提供文本在非書目性方面的細節描述，特別是使用的語言及次要的語言、文件的緣起、相關人員及其他背景。

<版本描述>/<revisionDesc>

概述檔案更新的歷史。

共用多種特性的文件的全集或合集，可共用一個全集標頭，而每一個組件都可擁有各別的標頭。在這種情況下，屬性可標明其類型。

```
<TEI 標頭 類型="文集">
```

為在標頭帶入全集層次的資訊。

某些標頭的元素由<段落> (<p>) 標誌一或多個篇章的逐頁資訊。其他的可分成下列幾組：

- 元素名稱以 Stmt (表 statement) 結尾者，通常內含一群記錄結構化資訊的元素。
- 元素名稱以 Decl (表 declaration) 結尾者，內含關於使用特殊編碼的資訊。
- 元素名稱以 Desc (表 description) 結尾者，內含篇章敘述。

### 20.1. 檔案描述

元素<檔案描述> (<fileDesc>) 是必備的。以下列元素標誌電子檔完整的書目描述：

<題名與作者敘述>/<titleStmt>

集合關於作品與可為其知識內容負責者的資訊。

<版本敘述>/<editionStmt>

集合文本在某一特定版本中的相關資訊。

<檔案大小>/<extent>

描述電子文件在在某一媒體上儲存的約略大小，可以任何適用的單位記錄。

<出版敘述>/<publicationStmt>

集合與電子或文本相關的出版或發行資訊。

<集叢敘述>/<seriesStmt>

集合任何關於出版品所屬集叢的資訊。

<附註敘述>/<notesStmt>

收集任何與文本相關，但書目描述中其他元素未能記錄的資訊。

<來源描述>/<sourceDesc>

說明電子文件衍生或是製作時來源文件實體的書目描述。

標頭至少含下列結構：

```
<TEI 標頭>
  <檔案描述>
    <題名與作者敘述> ... </題名與作者敘述>
    <出版敘述> ... <出版敘述>
    <來源描述> ... <來源描述>
  </檔案描述>
</TEI 標頭>
```

### 20.1.1. 題名敘述

下列元素可在元素<題名敘述> (<titleStmt>) 中使用：

<題名>/<title>

標誌作品的題名，不論是是文章、書籍、期刊或叢書，其他題名或是副題名也包含在內。

<作者>/<author>

在書目參照中，標誌作品個人或團體作者的名字，或任何書目中的原始責任敘述。

<贊助者>/<sponsor>

標明贊助組織或機構的名稱。

<資助者>/<funder>

標明負責資助計畫或文件的個人、機構或組織的名稱。

<政策研訂者>/<principal>

說明負責建立電子文件的主要研究者。

<責任敘述>/<respStmt>

對知識內容負文字、編輯、記錄或編集之責者的責任敘述，作者、編者等特定功能的元素不符或不適用者。

建議區分電腦與來源文件的題名，例如：

```
[來源文件題名]:機讀轉錄  
[來源文件題名]:電子版  
機讀版:[來源文件題名]
```

元素<責任敘述> (<respStmt>) 包含下列附屬構件：

<責任類型>/<resp>

標誌描述個人智財權性質的慣用語。

<名稱>/<name>

標誌一個正式的名稱或名詞片語。

範例：

```
<題名與作者敘述>  
  <題名>愛德格·愛倫·坡短篇二則：機讀轉錄</題名>  
  <作者>愛倫坡(1809-1849)  
  <責任敘述><責任類型>轉錄者</責任類型>  
  <名稱>詹姆士·班森</名稱></責任敘述>  
</題名與作者敘述>
```

### 20.1.2. 版本敘述

<版本敘述>/<editionStmt>集合文本某版本的相關資訊(同書目中其他處用法)，可包含下列元素：

<版本>/<edition>

描述文本的某一特定版本。

<責任敘述>/<respStmt>

對知識內容負文字、編輯、記錄或編集之責者的責任敘述，作者、編者等特定功能的元素不符或不適用者。

範例：

```
<版本敘述>  
  <版本 識別號="U2">草稿第三版，大幅更動  
  <日期>1987</日期>  
  </版本>  
</版本敘述>
```

電子文件的新版實際內容為何，留給編碼者決定。



### 20.1.3. 檔案大小敘述

<檔案大小>/<extent>

描述電子文件在在某一媒體上儲存的約略大小。

範例：

```
<檔案大小>4532 位元組</檔案大小>
```

### 20.1.4. 出版敘述

元素<出版敘述>/<publicationStmt>是必備的。可包含短文敘或集合下列元素：

<出版者>/<publisher>

提供關於負責文件出版或發行的組織名稱。

<發行者>/<distributor>

補充說明負責文件發行的個人或單位的名稱。

<授權者>/<authority>

補充說明在出版者或發行者之外，負責提供電子檔案的個人或其他單位的名稱。

上面三個元素至要出現一個，除非出版敘述全部以文章敘述。當中可出現下列元素：

<出版地>/<pubPlace>

標誌文件出版地的名稱。

<地址>/<address>

標誌郵寄或其他種類的地址，例如，出版商地址、機構地址、個人地址。

<識別碼>/<idno>

用來辨別文件的任何標準或非標準編碼，屬性包括：

**類型/type**

編碼的種類，例如國際標準書號（ISBN）或其他標準系列。

<取得限制>/<availability>

補充說明關於文件的獲取的資訊，例如任何使用或散佈的限制，版權狀態等等。

屬性包括：

**狀態/status**

說明文件目前的獲取性質的代碼，屬性的參考值包括：*限制 (restricted)*、*不明 (unknown)* 及 *免費 (free)*。

日期/<date>

標誌任何格式的日期。

範例：

```
<出版敘述>
  <出版者>牛津大學出版社</出版者>
```

```

<出版地>牛津</出版地> <日期>1989</日期>
<識別碼 類型="ISBN">0-19-254705-5</識別代碼>
<取得限制>牛津大學出版社版權所有，1989</取得限制>
</出版敘述>

```

### 20.1.5. 集叢與附註敘述

元素<集叢敘述> (<seriesStmt>) 集合關於所屬集叢的資訊。可包含元素<題名> (<title>)、<識別代碼> (<idno>) 或<責任敘述> (<respStmt>)。

使用元素<附註敘述> (<notesStmt>) 包含一或多個內含附註或註解的元素<附註> (<note>)。格式的書目中可見的部分資訊，已被指定為 TEI 架構的特定元素。

### 20.1.6. 來源描述

元素<來源描述> (<sourceDesc>) 為必備，記錄來源的細節，或是電腦檔案衍生的出處資訊。可使用一或多個下列元素包含簡單的短文或引用書目：

<引用書目>/<bibl>

標誌結構鬆散的引用書目，可選擇是否明確標誌其中的構件。

<完整書目>/<biblFull>

標誌結構完整的引用書目，呈現所有 TEI 檔描述的構件。

<書目表列>/<listBibl>

標誌引用書目表列，不限形式。

範例：

```

<來源描述>
  <引用書目>莎士比亞手稿首葉，查爾頓·辛曼提供（牛頓摹本，1968） </
引用書目>
</來源描述>
<來源描述>
  <描述敘述 識別名稱="CNN12">
    <引用書目><作者>美國有線電視新聞網網路新聞
      <題名>頭條新聞
      <日期>1989年1月12日
    </引用書目>
  </描述敘述>
</來源描述>

```

## 20.2. 編碼描述

元素<編碼描述> (<encodingDesc>) 標明規範文本轉錄的方法與編輯原則。強烈推薦使用此元素。可以文章描述或包含下列元素：

<計畫描述>/<projectDesc>

詳述電子檔案編碼的意圖或目標，以及任何關於徵集或收集程序的資訊。

<取樣宣告>/<samplingDecl>

標誌在建立全集或合集時，文件取樣使用原則與方法的敘述。

<編輯宣告>/<editorialDecl>

提供文件在編碼時運用的編輯原則與實作的詳細說明。

<標誌宣告>/<tagsDecl>

提供關於適用於 SGML 文獻標誌細節資訊。

<參考宣告>/<refsDecl>

標明文本的標準參考系統建構的方式。

<分類宣告>/<classDecl>

標誌文件在其他部分使用的一或多個分類法中的類號。

### 20.2.1. 專案與取樣描述

元素<計畫描述> (<projectDesc>) 與<取樣宣告> (<samplingDecl>) 的範例：

```
<編碼描述>
```

```
  <計畫描述>為克拉列蒙特·莎士比亞診所的使用而收集，1990年6月。
```

```
  </計畫描述>
```

```
</編碼描述>
```

```
<編碼描述>
```

```
  <取樣宣告>自文本開始處取二千字為樣本。
```

```
  </取樣宣告>
```

```
</編碼描述>
```

### 20.2.2. 編輯宣告

元素<編輯宣告> (<editorialDecl>) 包含文件編碼實作的描述文章。一般而言這類描述應包含下列主題，每一主題可以簡單的放在不同段落。

#### 更正/correction

在何種情況下如何更正文件。

#### 規格化/normalization

原始文件規格化或標準化的幅度。

#### 引文/quotation

原始文件的引號如何處理——保留或由實體參考所取代，是否區分開放或封閉的引文等。

### 連字符號/hyphenation

原始文件的連字號如何處理（尤其是行尾連字符號）——是被保留或由實體參考所取代等。

### 文件分割/segmentation

文件如何分割，例如分割成句子、語調單位、表列階層等。

### 詮釋/interpretation

文件加入何種分析或詮釋資訊。

範例：

```
<編輯宣告>
  <段落>第四章的全篇的講述分析全部由手動加入，還未檢查。
  <段落>轉錄的錯誤由 WordPerfect 拼字檢查控制。
  <段落>所有的字以韋氏大學字典第九版轉成現代美式拼法。
  <段落>所有的引用符號轉成實體參考。
</編輯宣告>
```

## 20.2.3. 標誌、參考及分類宣告

元素<標誌宣告> (<tagsDecl>) 用來提供關於文件中實際出現 SGML 標誌的細節資訊。可包含使用元素的簡單表列及使用量，使用下列特殊功能的元素：

### <標誌使用>/<tagUsage>

說明採用 TEI 的文件中關於特定元素在元素<文件>(<text>)中統一的使用資訊。

屬性包括：

#### 對應元素名稱/gi

標誌所指元素的名稱（一般識別）。

#### 出現次數/occurs

標明元素在文件中出現的次數。

元素<樣式> (<rendition>) 用來記載來源文本中被標誌部分使用的不同樣式。

### <樣式>/<rendition>

補充說明關於一或多個元素在展現時預定採用樣式的資訊。

### <標誌使用>/<tagUsage>

說明採用 TEI 的文件中關於特定元素在元素<文件>(<text>)中統一的使用資訊。

屬性包括：

#### 出現次數/occurs

標明在文件中元素出現的次數。

#### 識別名稱出現次數/ident

標明具有全域屬性**識別名稱** (id) 識別值的元素在文件中出現的次數。

## 樣式/render

標明定義元素<樣式>使用樣式的識別碼。

範例：

```

<標誌宣告>
  <標誌使用 對應元素名稱="text" 出現次數=1>
  <標誌使用 對應元素名稱="body" 出現次數=1>
  <標誌使用 對應元素名稱=p 出現次數="12">
  <標誌使用 對應元素名稱="hi" 出現次數=6>
</標誌宣告>

```

以上（想象中的）標記宣告可以適用於包含十二個段落的文件，其中標誌了六個元素<醒目標示>（<hi>）。注意，如果使用元素<標誌宣告>，必須包含相關文本中每一個元素的<標誌使用>。

元素<參考宣告>（<refsDecl>）用來記載任何編碼工作內建的標準架構。最簡單的格式由文章敘述組成。

範例：

```

<參考宣告>
<段落>每一個第一層與第二層的屬性 N 包含使用 XX.YYY 這類格式分割的正規的
參考系統，其中 XX 為羅馬數字的冊號，YYY 為阿拉伯數字的章節號。
</參考宣告>

```

元素<分類宣告>（<classDecl>）集合標頭其他部分使用的分類架構定義或來源的描述。這類的架構必須至少提供一種，以下列的元素編碼：

<分類學>/<taxonomy>

定義用在文本的分類法，不論是隱含於書目引用者，或是以明確分類架構呈現者。

<引用書目>/<bibl>

標誌書目構件未明確標記，結構鬆散的引用書目。

<類目>/<category>

標誌使用者定義的分類法中單獨描述的類目，這個類目有可能是位一個更大類目中的一部分。

<類目描述>/<catDesc>

以短文的形式，描述分類法或文件類型中的某些類目。

在最簡單的例子裡，分類可由書目參照定義，如下例所示：

```

<分類宣告>
  <分類學 識別名稱="LCSH">
    <引用書目>國會圖書館標題法</引用書目>
  </分類學>
</分類宣告>

```

此外，編碼者可定義特殊目的的分類架構，如下例所示：

```

<分類學 識別名稱=B>
  <引用書目>布朗語料</引用書目>
  <類目 識別名稱="B.A"><類目描述>出版報導
    <類目 識別名稱="B.A1"><類目描述>每日</類目>
    <類目 識別名稱="B.A2"><類目描述>週日</類目>
    <類目 識別名稱="B.A3"><類目描述>國家</類目>
    <類目 識別名稱="B.A4"><類目描述>地方</類目>
    <類目 識別名稱="B.A5"><類目描述>政治</類目>
    <類目 識別名稱="B.A6"><類目描述>體育</類目>
    ...
  </類目>
  <類目 識別名稱="B.D"><類目描述>信仰
    <類目 識別名稱="B.D1"><類目描述>書籍</類目>
    <類目 識別名稱="B.D2"><類目描述>期刊與追蹤</類目>
  </類目>
  ...
</分類學>

```

這類分類架構中，特殊文本與分類中的類目間的連結是由元素<分類用詞>（<textClass>）中的元素<類目參考>（<catRef>）構成，下面會進一步描述。

### 20.3. 背景描述

元素<profileDesc>允許在單一架構中記錄包含多種特質的描述性資訊。有三種選用構件：

<建置資訊>/<creation>

關於文件建置的資訊。

<使用語言>/<langUsage>

描述在文本中出現的語言、次要語言、語體、方言等。

<分類用詞>/<textClass>

集合以標準分類架構、索引典等來描述文本的性質或主題的資訊。

範例：

```

<建置資訊>
  <日期 標準格式值="1992-08">1992年8月</日期>
  <名稱 類型="place">塔歐斯，新墨西哥</名稱>
</建置資訊>

```

元素<分類用詞>（<textClass>）參考由元素<分類宣告>（<classDecl>）所定義的系統為文件分類，包含一或多個下列元素：

<關鍵詞>/<keywords>

標誌定義文本主題或性質的關鍵詞或片語。屬性包括：

**分類架構/scheme**



識別控制詞彙定義的相關關鍵字組。

<分類號>/<classCode>

標誌文本使用某種標準分類系統時的分類碼。

**分類架構/scheme**

識別使用的分類法系統或分類學。

<類目參考>/<catRef>

標明在某個分類學或類型學中定義的一個或多個類目。屬性包括：

**目標/target**

識別相關的類目。

元素<關鍵詞>包含識別文件主題或特性的關鍵字或片語。屬性架構 (scheme) 連結到元素分類學 (<taxonomy>) 中定義的分類系統。

<分類用詞>

<關鍵詞 分類架構="LCSH">

<列表>

<項目>英國文學-歷史與評論-資料處理</項目>

<項目>英國文學-歷史與評論-理論等</項目>

<項目>英國語文-風格-資料處理</項目>

</列表>

</關鍵詞>

</分類用詞>

## 20.4. 更新描述

元素<更新描述> (<revisionDesc>) 為文本每一筆更正記錄提供異動登錄。異動登錄可以一系列的元素<更新> (<change>) 記錄，每一個<更新>中都包含：

<日期>/<date>

標誌任何格式的日期。

<責任敘述>/<respStmt>

對知識內容負文字、編輯、記錄或編集之責者的責任敘述，作者、編者等特定功能的元素不符或不適用者。

<項目>/<item>

標誌列表中的一個部件。

範例：

<更新描述>

<更新><日期>1991年6月3日:</日期>

<責任敘述><名稱>EMB</名稱><責任類型>編</責任類型></責任敘述>

<項目>格式更新</項目></更新>

<更新><日期>1990年5月25日:</日期>

<責任敘述><名稱>EMB</名稱><責任類型>編</責任類型></責任敘述>

<項目>史都華的更新鍵入</項目></更新>  
</更新描述>

## 附錄 使用元素表

### 附錄 1.1 全域屬性

下列全域屬性適用於 TEI Lite 中所有的元素：

#### 分析/ana

將元素連結到其特性。

#### 對應元素/corresp

將元素連結到一或多個相對應的元素。

#### 識別名稱/id

元素的識別名稱，命名不可重覆。必須以字母開頭，可使用字母、數字、連字號及時代。

#### 語言/lang

元素所標誌文本的語言；未標註即表示和背景的語言相同。

#### 識別號/n

元素的名稱或編號，字串組成方式沒有限制。常用來記載傳統的參考系統。

#### 下一元素/next

在同一集叢中，連結到下一個元素。

#### 上一元素/prev

在同一集叢中，連結到上一個元素。

#### 樣式/rend

被標誌文本在文件中實際的表現方式，例如：斜體，羅馬字體，區塊等。屬性值可使用任何文字字串。

### 附錄 1.2 TEI 選錄版使用元素

以下表列顯示 TEI Lite DTD 中定義的所有元素與其簡述：

#### <縮寫>/<abbr>

標誌任何形態的縮寫。全稱可使用屬性**全稱 (expand)** 標誌。

#### <加入文字>/<add>

標誌由作者、代筆者、註解者或更正者在文本中加入的字母、字或詞。

#### <地址>/<address>

標誌郵寄或其他種類的地址，例如，出版商地址、機構地址、個人地址。

#### <地址行>/<addrLine>

標誌地址中的一行。

#### <錨點>/<anchor>

標記文件中被連結的位置或特定位點。

<大綱>/<argument>

針對正文所含的子題做有條理的表列或文字敘述。

<作者>/<author>

在書目參照中，標誌作品個人或團體作者的名字，或任何書目中的原始責任敘述。

<授權者>/<authority>

補充說明在出版者或發行者之外，負責提供電子檔案的個人或其他單位的名稱。

<取得限制>/<availability>

補充說明關於文件的獲取的資訊，例如任何使用或散佈的限制，版權狀態等等。

<正文後資訊>/<back>

標誌隨附於正文之後的附錄等。

<引用書目>/<bibl>

標誌結構鬆散的引用書目，可選擇是否明確標誌其中的構件。

<完整書目>/<biblFull>

標誌結構完整的引用書目，呈現所有 TEI 檔描述的構件。

<引用範圍>/<biblScope>

定義書目參照的範圍，例如：頁碼表列，或大型作品中被命名的部分。

<正文>/<body>

標誌單一文本的整體部分，不包含正文前及正文後資料。

<題名頁責任敘述>/<byline>

標誌在題名頁、作品開頭或結尾處原始責任敘述。

<類目描述>/<catDesc>

以短文的形式，描述分類法或文件類型中的某些類目。

<類目>/<category>

標誌使用者定義的分類法中單獨描述的類目，這個類目有可能是位一個更大類目中的一部分。

<類目參考>/<catRef>

標明在某個分類學或類型學中定義的一個或多個類目。

<格>/<cell>

標誌表格中的一格。

<引文>/<cit>

源出其他文獻的引文，含來源的書目參考資料。

<分類號>/<classCode>

標誌文本使用某種標準分類系統時的分類碼。分類碼以屬性分類架構 (scheme) 定義。

<分類宣告>/<classDecl>

標誌文件在其他部分使用的一或多個分類法中的類號。

<文件末端>/<closer>

集合文件日期、題名頁責任敘述、致敬詞及其他出現在結尾部分的類似段落，特別是在信件上。

<程式碼>/<code>

標誌使用正規語言（通常是程式語言）的程式片斷。

<更正>/<corr>

標誌對文本中明顯錯誤的更正。

<建置資訊>/<creation>

關於文件建置的資訊。

<日期>/<date>

標誌任何格式的日期，使用屬性標準格式值（value）來規範屬性值格式。

<文件日期>/<dateline>

標誌信件、報紙報導或其他文件的出版品前置於卷頭或附加為卷尾的地點、日期、時間等的簡短敘述。

<刪除>/<del>

標記由作者、代筆者、附註者或更正者在原件中刪除的字母、文字或段落，標記為刪除者，或指明為多餘的或偽造的部分。

<發行者>/<distributor>

補充說明負責文件發行的個人或單位的名稱。

<文字區段>/<div>

標誌正文前資訊、正文及正文後資訊中文本的分項。

<第一層>...<第七層>/<div1>...<div7>

標誌正文前資訊、本文及正文後資訊中文本的第一層至第七層分項。

<自動產生區段>/<divGen>

標示由文字處理應用程式自動產生的文字區段的位置。以屬性類型（type）標明此為索引、目次或其他類型。

<文獻作者>/<docAuthor>

標誌在題名頁所載的文獻作者名字（通常，但並不一定要在元素<題名頁責任敘述>（<byline>）中使用）。

<文獻日期>/<docDate>

標誌題名頁所載的文獻日期。

<文獻版本>/<docEdition>

標誌文獻的題名頁所載的編輯敘述。

<文獻印行資訊>/<docImprint>

標誌在題名頁下方（通常如此）所載的印行敘述（出版地、出版日期、出版者名稱）。

<文獻題名>/<docTitle>

標誌題名頁上所載的文獻題名，包含構成題名的所有組件。必須再標誌成不同的<題名部件>（<titlePart>）。

<版本>/<edition>

描述文本的某一特定版本。

<版本敘述>/<editionStmnt>

集合文本在某一特定版本中的相關資訊。

<編者>/<editor>

文件的第二責任類型敘述，例如編輯、編譯、翻譯等角色的個人、機構或組織（或其他類似者）的名稱。

<編輯宣告>/<editorialDecl>

提供文件在編碼時運用的編輯原則與實作的詳細說明。

<範例>/<eg>

標誌討論中的技術主題所舉的一則短例，例如：程式的片斷或是 SGML 編碼的樣本。

<強調>/<emph>

標誌語言或修詞效果的重音或強調的字或片語。

<編碼描述>/<encodingDesc>

記載電子文件與來源文件的關係，或電子文件的源由。

<卷首引句>/<epigraph>

標誌出現在章節開端或題名頁上具名或不具名的引文。

<檔案大小>/<extent>

描述電子文件在在某一媒體上儲存的約略大小，可以任何適用的單位記錄。

<圖象>/<figure>

標記文獻中圖像插入的位置。屬性可以用來標示包含影象的 SGML 實體（以某些非 SGML 的方式標記）。元素<圖象>（<figure>）中的段落，可記錄圖說。

<圖象描述>/<figDesc>

標誌圖象的出現或內容的文字敘述，只記錄而不顯示圖形時使用。

<檔案描述>/<fileDesc>

標誌電子檔的完整書目描述。

<外國語言>/<foreign>

定義與背景文字不同語言的字或片語。

<公式>/<formula>

標誌數學或化學公式，可選用部分非 SGML 的標記法。屬性**標記法（notation）**標示用來記錄公式的非 SGML 標記法。

<正文前資訊>/<front>

標誌正文開始之前的任何前置項目（標頭、書名頁、前言、獻詞等）

<資助者>/<funder>

標明負責資助計畫或文件的個人、機構或組織的名稱。

<省略資料>/<gap>



標示資料在轉錄時被省略的部分，可能依 TEI 標頭中敘述的編輯理由刪去，或為抽樣練習的部分，或因為是非法或是限閱的資料。

<類名識別>/<gi>

標誌識別名稱的特殊類型：SGML 的通用識別或是元素名稱。

<註解>/<gloss>

標誌一個字或詞，為其他字或片語提供夾註或定義。

<群組>/<group>

標誌幾篇單一的文本或一組文件。

<標頭>/<head>

可標誌任何的標頭，例如：章節題名，詞彙表或其他表列的表頭。

<醒目標示>/<hi>

標誌與背景文字顯著不同的單字或段落，不論其目的。

<識別詞>/<ident>

標誌某種識別名稱，例如，XML 元素或屬性的名稱及其異稱。

<識別碼>/<idno>

用來辨別文件的任何標準或非標準編碼，屬性**類型 (type)** 定義使用的架構或標準。

<印行資訊>/<imprint>

集合文件出版或發行的相關資訊。

<索引>/<index>

標誌一個因某種目的而被索引的位置。屬性通常給一個主要類型，索引需要時可輸入二至四層的複分。

<詮釋>/<interp>

提供詮釋性的註解，可以連結到一段文字。屬性包括：**標準格式值 (value)**、**責任類型 (resp)**、**類型 (type)**。

<詮釋群組>/<interpGrp>

將<詮釋> (<interp>) 集合在一起。

<項目>/<item>

標誌列表中的一個部件。

<關鍵詞>/<keywords>

標誌定義文本主題或性質的關鍵詞或片語。若關鍵詞取自控制詞彙，可用屬性**分類架構 (scheme)** 來識別。

<關鍵術語>/<kw>

標誌某種正規程式語言中的關鍵詞。

<詩行>/<l>

標誌韻文的一行，可能為未完成句。

<標籤>/<label>

在列表中，標誌與某一項目相對應的標籤；在字彙表中，標誌被定義的詞彙。

<使用語言>/<langUsage>

描述在文本中出現的語言、次要語言、語體、方言等。

<分行>/<lb>

標誌某些編輯或版本上（印刷時）的新起行。

<詩組>/<lg>

標誌形成一個正式單位的一群韻句，如：詩節、疊句、行韻的段落等。

<列表>/<list>

標誌被組織成表的項目排列，不限形式，可為序號列、點列或其他類型。

<書目表列>/<listBibl>

標誌引用書目表列，不限形式。

<指涉>/<mentioned>

標誌被提到的非用字或片語。

<分界>/<milestone>

標明文本分段的界限，如標準參考系統的變更。可用屬性包括：**版本**（**ed**，版本）、**單元**（**unit**，頁碼等）、及**識別號**（**n**）。

<名稱>/<name>

標誌一個正式的名稱或名詞片語。屬性可標示類型、賦予標準格式，或以唯一的識別名稱與特定的人或事做關聯。

<附註>/<note>

包含附註或註釋，以屬性說明附註的類型、位置、及來源。

<附註敘述>/<notesStmt>

收集任何與文本相關，但書目描述中其他元素未能記錄的資訊。

<數字>/<num>

標誌數字，不論其書寫形式。使用屬性**標準格式值**（**value**）來統一格式。

<文件開頭>/<opener>

集合開頭部分的日期行、題名頁責任敘述行、致敬辭及其他做為文件序曲的類似段落，特別是用信件上。

<原始格式>/<orig>

標誌文件材料原始的格式，調整後的格式可放入屬性**採用格式**（**reg**）。

<段落>/<p>

標記文章的段落。

<分頁>/<pb>

以標準參考系統標記文本一頁與下頁間的邊界。

<政策研訂者>/<principal>

說明負責建立電子文件的主要研究者。

<背景描述>/<profileDesc>

提供文本在非書目性方面的細節描述，特別是使用的語言及次要的語言、文件的緣起、相關人員及其他背景。

## &lt;計畫描述&gt;/&lt;projectDesc&gt;

詳述電子檔案編碼的意圖或目標，以及任何關於徵集或收集程序的資訊。

## &lt;指標&gt;/&lt;ptr&gt;

指向本文件中的另一個位置，即一或多個可識別的元素。

## &lt;出版敘述&gt;/&lt;publicationStmt&gt;

集合與電子或文本相關的出版或發行資訊。

## &lt;出版者&gt;/&lt;publisher&gt;

提供關於負責文件出版或發行的組織名稱。

## &lt;出版地&gt;/&lt;pubPlace&gt;

標誌文件出版地的名稱。

## &lt;引句&gt;/&lt;q&gt;

標誌引文或顯然是引文的部分。

## &lt;參照&gt;/&lt;ref&gt;

指向本文件中的另一個位置，即一或多個可識別的元素，可能為額外的修正文字或意見。

## &lt;參考宣告&gt;/&lt;refsDecl&gt;

標明文本的標準參考系統建構的方式。

## &lt;採用格式&gt;/&lt;reg&gt;

標誌文件材料因某種原因調整或是規格化後的格式，原始的格式可放入屬性**原始格式 (orig)**。

## &lt;樣式&gt;/&lt;rendition&gt;

補充說明關於一或多個元素在展現時預定採用樣式的資訊。

## &lt;責任類型&gt;/&lt;resp&gt;

標誌描述個人智財權性質的慣用語。

## &lt;責任敘述&gt;/&lt;respStmt&gt;

對知識內容負文字、編輯、記錄或編集之責者的責任敘述，作者、編者等特定功能的元素不符或不適用者。

## &lt;版本描述&gt;/&lt;revisionDesc&gt;

概述檔案更新的歷史。

## &lt;欄&gt;/&lt;row&gt;

標誌表格的一列。

## &lt;關聯字串&gt;/&lt;rs&gt;

標誌一般性質的名稱或是關聯字串。屬性可標示類型、賦予標準格式，或以唯一的識別名稱與特定的人或事做關聯。

## &lt;成句單位&gt;/&lt;s&gt;

標誌文獻中的成句單位 (*s-unit*)。目的在建立一個適用於文本整體的簡單標準參考架構。

## &lt;致敬辭&gt;/&lt;salute&gt;

在信件或序言等文類中，標誌前置的致敬語或問候語、獻詞或其他片段，或是結尾的致敬語。

<取樣宣告>/<samplingDecl>

標誌在建立全集或合集時，文件取樣使用原則與方法的敘述。

<分割段落>/<seg>

標示文件中被連結的文字區塊或片段。可用屬性類型(**type**)可為連結段落分類。

<集叢>/<series>

標誌書籍或其他類型文件被集叢收錄的資訊。

<集叢敘述>/<seriesStmnt>

集合任何關於出版品所屬集叢的資訊。

<原文照錄>/<sic>

標誌雖有明顯錯誤卻仍照原件所錄者。

<簽名>/<signed>

標誌在結尾處的問候語、前言、獻辭或是文本其他區段的簽名。

<另稱>/<soCalled>

標誌作者或編寫者標示為放棄責任的個字或詞，如：使用醒目的引號或斜體者。

<來源描述>/<sourceDesc>

說明電子文件衍生或是製作時來源文件實體的書目描述。

<講述>/<sp>

標誌劇本中的單一演說，或在散文和詩中以講述方式表現的段落。可用屬性人物(who)識別講者。

<講者>/<speaker>

標誌在劇本或文章的片段中一或多個講者的特殊標頭或標籤。

<贊助者>/<sponsor>

標明贊助組織或機構的名稱。

<表演動作>/<stage>

標誌劇本或文章片斷中的任何表演指示。

<表格>/<table>

標誌以欄與列組成的表格形式呈現的文本。

<標誌宣告>/<tagsDecl>

提供關於適用於 SGML 文獻標誌細節資訊。

<標誌使用>/<tagUsage>

說明採用 TEI 的文件中關於特定元素在元素<文件>( <text>)中統一的使用資訊。

<分類學>/<taxonomy>

定義用在文本的分類法，不論是隱含於書目引用者，或是以明確分類架構呈現者。

<專有名詞>/<term>

標誌被視為技術性詞彙的單字、複合字或設計圖象。

<分類用詞>/<textClass>

集合以標準分類架構、索引典等來描述文本的性質或主題的資訊。

<時間>/<time>

標誌定義一天之內時間的詞彙，不限格式，使用屬性標準格式值/value 來記錄標準格式。

<題名>/<title>

標誌作品的題名，不論是是文章、書籍、期刊或叢書，其他題名或是副題名也包含在內。

<題名頁>/<titlePage>

標誌文本的題名頁，出現在正文前或正文後的資訊之中。

<題名部件>/<titlePart>

依題名頁上所示，標誌作品題名的部分或區段，也可以標誌題名頁上不屬文獻標題、著作身分等無目的性的片段。

<題名與作者敘述>/<titleStmnt>

集合關於作品與可為其知識內容負責者的資訊。

<卷末>/<trailer>

標誌出現在文本結尾處的結尾辭或尾聲。

<不明>/<unclear>

標誌因來源文件中非法或是限制公開的內容，而不可被確實轉錄的字、詞或段落。

<外部指標>/<xptr>

定義連向本文件或外部文件中另一位置的指標。

<外部參照>/<xref>

定義連向本文件或外部文件中另一位置的指標，可能為額外的修正文字或意見。