

提升體育新聞資訊品質之探討— 以 SportsML、NITF、NewsML 整合為例

On improving the information quality of sports news- by integration of SportsML, NITF and NewsML

林信成 (Sinn-Cheng Lin)

淡江大學資訊與圖書館學研究所 副教授

sclin@mail.tku.edu.tw

游忠諺 (Chung-Yen Yu)

淡江大學資訊與圖書館學研究所 研究生

692070039@s92.tku.edu.tw

摘要

SportsML 是 IPTC 利用 XML 所制訂的一套 Metadata 標準，主要用於交換和傳輸各類運動項目紀錄與相關新聞報導，本研究試圖從傳播過程的角度開始，描述詮釋資料在傳播過程中所扮演的角色，並選擇描述體育新聞的詮釋資料-SportsML，探討 SportsML 的技術標準、結構組成、描述能力與 NITF 和 NewsML 的整合，提出範例以供探討，最後描述台灣在描述體育新聞時採用詮釋資料做為文件傳遞格式的現況，並提出建議以供參考。

關鍵字：Metadata、NewsML、NITF、SportsML、詮釋資料

Abstract

SportsML is a Metadata standard based on XML and established by IPTC (International Press Telecommunication) to exchange and transfer sport records and news report. The object of this research is to describe the role of metadata in communication process, to explore the technique standard, structure formation, description ability of SportML, its integration with NITF and NewsML. Also, examples are provided for discussion. This study may have a substantial contribution on providing academics and practitioners with insight into understanding the application of using metadata as a format for transferring document on Sport News in Taiwan.

Keyword : Metadata 、 NewsML 、 NITF 、 SportsML

壹、前言

網路科技發展至今，已經成為一種大眾化的傳播媒介，如同電視、廣播般，對於社會的影響深遠，不亞於十八世紀歐洲的產業革命所帶來的效果。傳播學者麥克魯漢(McLuhan)認為媒介是一種科技、一種型式，它存在的本身就是一種訊息；而「內容」則是科技的運用¹，在載體上規範其傳播的格式，就像採用相同的語言進行溝通一般，有助於在傳播過程中，發訊者與收訊者能夠成功的進行資訊交換，而詮釋資料(Metadata)正是一項有效的傳播格式類型，發訊者與收訊者如採用相同的格式，就能夠在相互溝通上，具有相當大的便利性。大眾媒體透過Metadata與其他媒體進行跨媒體與跨產業的整合，可進一步實現「大編輯台」的理想²。

以標示語言實作Metadata已是舉世公認的共同作法，國際上為了解決數位化新聞的出版、交換、傳遞與共享問題，已發展出數種新聞專用的Metadata格式，例如，國際新聞電訊評議會(The International Press Telecommunications Council，簡稱IPTC)制訂的NITF (News Industry Text Format)³、NewsML (News Markup Language)⁴、SportsML (Sports Markup Language)⁵、ProgramGuideML (Program Guide Markup Language)⁶和EventsML (Events Markup Language)⁷；由IDEAlliance發佈的PRISM (Publishing Requirements for Industry Standard Metadata)⁸；由XMLNews.Org所研擬的XMLNews；還有目前普遍應用於入口引擎、新聞網站、Blog和WiKi的RSS (Really Simple Syndication)⁹等。這些標準皆本於可擴展標示語言--XML (eXtensible Markup Language)，資訊管理者可依需求定義新聞文件的DTD或XML Schema，既規範文件結構又保留擴充彈性；經過XML標示的新聞文件，也可利用程式轉換到其它格式或平台，有助於新聞文件的儲存、檢索以及重複使用¹⁰。

由上述可知，新聞詮釋資料格式目前是百家爭鳴，IPTC於2002年制訂NewsML結合NITF，雖可描述大部份的新聞報導，但一些結構特殊的新聞類型，如體育新聞，由於結構上的差異，為了讓讀者快速的獲得即時且明確的訊息，就

¹ 劉幼琍(民83年12月)。傳播科技對兩岸新聞傳播之影響。台大新聞論壇，1(2)，54。

² 陳百齡(民91年5月)。資料庫裡找新聞：探討「資料庫輔助新聞報導」。圖書與資訊學刊，41，頁1。

³ IPTC. (n. d.). News Industry Text Format Retrieved July 18, 2004, from <http://www.nitf.org/>.

⁴ IPTC. (n. d.). News Markup Language Retrieved July 18, 2004, from <http://www.newsml.org/>.

⁵ IPTC. (n. d.). Sports Markup Language Retrieved October 8, 2004, from <http://www.sportsml.org/>.

⁶ IPTC. (n. d.). Program Guide Markup Language Retrieved October 8, 2004, from <http://www.programguideml.org/>.

⁷ IPTC.(n.d.).Events Markup Language Retrieved July 18,2004, from <http://www.iptc.org/EventsML/>.

⁸ IDEAlliance PRISM Working Group. (n. d.). PRISM: Publishing Requirements for Industry Standard Metadata Retrieved October 9,2004, from http://www.prismstandard.org/PAM_1.0/.

⁹ Berkman Center.(n. d.). RSS 2.0 Specification Retrieved October 9, 2004, from <http://blogs.law.harvard.edu/tech/rss>.

¹⁰ Hall, Richard. (2000, May.). Why XML is Important for Printing and e-publishing? Retrieved September 5,2003, from http://www.newsandtech.com/issues/2000/05-00/ot/05-00_hall.htm.

必須另外制訂其描述方式。體育新聞的特點有二¹¹：(1) 表格較多，比方像是描述比賽的賽程表、比數表與統計表等；(2) 運動項目種類多元，各種規則與計分方式也不同。因此，體育新聞除了文字稿外，多半都會使用大量的表格與圖片進行新聞描述。況且不同的運動項目，在描述詞彙上也相對的不同，為了在進行資訊傳遞上的便利，以及因應新聞出版方式的多元性，專為描述體育記錄的詮釋資料 — SportsML便因此而誕生了。

我們在一開始執行國科會數位典藏計畫「臺灣棒球運動珍貴新聞檔案數位資料館之建置」¹²時，即遭遇不知選用何種Metadata來描述棒球運動新聞的困擾，後來決定先初步採用與Dublin Core（簡稱DC）相容的NRMF格式，以便能儘速與同採DC格式的數位典藏聯合目錄接軌，未來再陸續加入與其他Metadata規格相容之設計，以提升其資料交換與傳遞能力。¹³而我們在系統實作的過程中亦發現，以DC為藍本的NRMF雖然簡單易用，但當考慮到整個新聞的流程、管理與傳遞時，尚有許多不足之處；且當需要凸顯體育新聞的特色時，亦有許多值得改進之處，故而引發我們研究各式新聞Metadata的動機。本研究以探討SportsML為主，NewsML、NITF為輔，其餘相關的新聞Metadata則擬另以專文探討。在本文中除對SportsML作較深入的說明外，另會提及如何將SportsML與NewsML、NITF兩者相互結合，從而使體育新聞的描述更豐富、完整。

貳、SportsML 標準與內容結構

SportsML是IPTC在2003年3月所推動用來交換體育資料的Metadata。IPTC認為，過去在進行交換跨越不同的體育項目(Cross-sport)與跨語言(Cross-language)的資料與統計時，所需的XML標準制訂仍屬於空窗期的階段，因此在IPTC成員自願提供必要的資源，以設計並完成有關於運動項目的XML DTD，與必要的技術支援文件和範例¹⁴。

SportsML 的特色在於提供了識別與描述大量體育資料的能力，其中強調的重點在於：(1) 得分(Scores)、(2) 賽程(Schedules)、(3) 名次(Standings)、(4) 紀錄統計(Statistics) 和 (5) 新聞報導(News)等五項，SportsML 的 Core DTD 內制訂大量描述體育新聞報導的特性，大多數的體育新聞報導都可透過 Core DTD 撰寫，另外，為了讓新聞記者能報導更有深度的體育新聞，IPTC 亦制訂了美式足球、籃球、棒球、高爾夫、冰上曲棍球、網球與足球等七種不同體育運動的 DTD，可「插入」(Plug in)到 SportsML 中來增加對特定運動項目的描述。

IPTC 制訂 SportsML 主要目標是傳遞運動項目的各項紀錄，同時與新聞報導

¹¹ 那福忠(民91年11月)。新聞的標準語言 NewsML。上網日期：民93年10月1日。網址：http://www.brainnew.com.tw/Article/na2002/n_113002.asp

¹² 淡江大學資圖所數位典藏計畫小組。臺灣棒球運動珍貴新聞檔案數位資料館之建置。上網日期：民91年10月14日。網址：<http://ndap.dils.tku.edu.tw/>

¹³ 林信成、鄭國祥、孫正宜(2004, 8月)。報紙新聞專卷的數位典藏與加值利用。在大連理工大學舉行，2004年海峽兩岸圖書資訊學暨教育發展研討會論文集，(頁11~27)。

¹⁴ IPTC. (n. d.). SportsML: Introduction Retrieved September 9, 2004, from <http://www.sportsml.com/intro.php/>.

進行結合，以做為在新聞交換上的一個重要的規範標準，因此 SportsML 不僅可以獨立使用於描述體育紀錄，更能夠與 NITF 和 NewsML 進行整合，以在進行新聞或是運動資訊的傳遞上更能夠在不同平台、不同的語言上流通交換，因此做為一個描述運動項目的標識語言，SportsML 除了描述運動項目的內容，包括描述選手或比賽隊伍的基本資料、賽程、名次、紀錄統計外，也可先描述單一體育事件，再引用 NITF 描述該事件的相關或詳細新聞內容，最後再以 NewsML 進行新聞封裝，以求整體的結構性與完整性。

IPTC 期望體育協會或是特殊運動項目也能夠積極的參與，除了制訂該運動專屬的 DTD 外，更能夠在未來支持 SportsML 成為全球標準，在共享的目標之下，使得交換運動資料時能夠更加的容易。

SportsML 目前公布的正式版本是為 1.0 版，如同 XML 文件一樣，SportsML 也包括了三方面的內容：(1) 定義 SportsML 文件邏輯結構中的元素、元素的特性與元素之間的關係的 DTD 或 Schema，(2) 定義 SportsML 顯示格式的 XSL (eXtensible Stylesheet Language)，以及 (3) 利用 DTD 或 Schema 中所定義的 SportsML 文件主體內容。原本 IPTC 計畫在未來要制訂符合 SportsML 標準的 Schema，但現行採用的方案則是訂定了專屬的 DTD (Document Type Definition) 做為定義各項元素之用，而 DTD 則包含了 Core DTD、控制檔 (The Control File) 與各項針對不同的體育項目所制訂的 DTD (Plug-in DTD) 三種，如下表所示¹⁵：

表 1 SportsML DTD 類型簡述

名稱	說明
The Core SportsML DTD	針對個別的元素與其屬性進行描述，以適用於在得分、賽程、名次與紀錄統計上的處理，在美式英文裡，Core DTD 是使為表達結構之用，以讓任何國家的新聞記者能夠在閱讀時該份 SportsML 文件時更能加的便利。
The Control File	在 Core DTD 中引用的 DTD，內容除定義公用屬性外，並將其他專屬運動項目的 Plug-in DTD 和 NITF DTD，用 XML 語法引用於內，特點在於不須修改 Core DTD，即可引用不同運動項目的 DTD 與新聞描述格式的 DTD，避免在未來 SportsML 版本更新上的困難。
Plug-in DTDs	針對新聞記者在使用度較高的體育運動項目上所特別制訂的 DTD，並以「插入」的方式和 Core DTD 進行結合。

在 Core DTD 中，主要是針對個別的元素與其屬性進行定義，圖 1 為其結構簡圖。

¹⁵ IPTC. (n. d.). SportsML : Tutorial Retrieved September 9,2004, from <http://www.sportsml.com/tutorial.php/>.

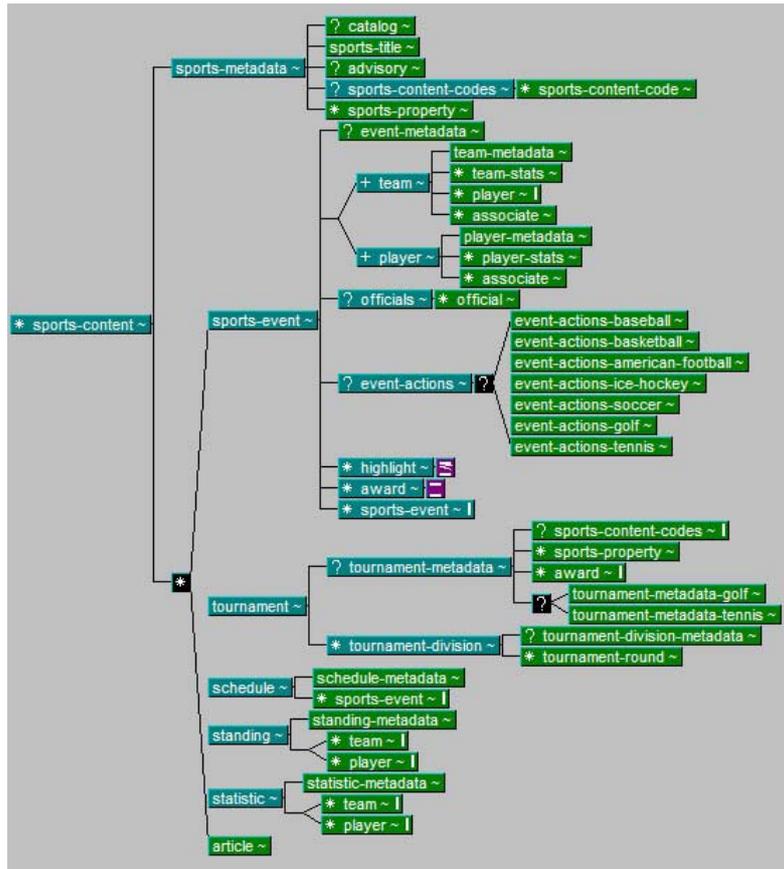


圖 1 SportsML 元素結構簡圖

主要分為幾個部份，簡要說明如表 2¹⁶：

表 2 SportsML-Core DTD 簡述

名稱	說明
Sports-Metadata	主要是描述 SportsML 的文件編號、主題、特性、分類、摘要、區別各項運動的自定編碼等。
Sports-Event	描述各項運動比賽的結果，包括裁判、體育活動的類型、重要事項、獎章類型以及相關比賽結果。
Tournament	描述循環賽的各種細節，包括場地背景、當地資訊、比賽結構等。
Schedule	描述比賽的賽程表
Standing	描述團隊或個別選手的名次表。
Statistic	將團隊或個別選手的紀錄進行統計與比較。
Article	描述該體育活動所參照到相關新聞報導，以 NITF 為推薦格式。
Player	描述個別選手的基本資料，包括姓名、身高、體重、國籍等。
Team	描述團體的基本資料，包括識別是否為主隊或是客隊。

¹⁶ IPTC. (n. d.). Documentation for SportsML Retrieved October 15,2004, from <http://www.sportsml.org/IPTC/1.0/documentation/sportsml-core-documentation.html>.

個別的運動項目都有各自獨特的一套描述方式，SportsML做為一個資料交換的平台，必須考量到不同的運動項目的紀錄方式，因此IPTC在SportsML 1.0 的版本公布時，將冰上曲棍球、美式足球、棒球、籃球、高爾夫、足球與網球的DTD同時包括在內且一併公布，以做為新聞記者在進行新聞描述時的參考，以描述棒球紀錄的DTD¹⁷為例，共有 14 個元素進行描述棒球的相關紀錄，簡述如表 3：

表 3 SportsML Plug-in DTD – 描述棒球紀錄的 DTD 簡述

名稱	說明
Action-baseball-other	棒球比賽進行中其他可能的情況
Action-baseball-pitch	比賽進行時守備的狀況
Action-baseball-play	尚未得分時的進攻紀錄
Action-baseball-score	比賽中的得分紀錄
Action-baseball-substitution	選手的交換紀錄
Event-actions-baseball	按照年代描述有關棒球的事件
Event-metadata-baseball	描述比賽進行所發生的任何事件
Player-metadata-baseball	描述選手的基本資料
Player-stats-baseball	收集有關選手的進攻、防守紀錄的統計
Stats-baseball-defensive	球隊的防守統計資料
Stats-baseball-offensive	球隊的進攻統計資料
Stats-baseball-pitching	投手本身的統計資料
Team-metadata-baseball	描述球隊的資料
Team-stats-baseball	球隊的具體統計，包括進攻與防守

從表 3 可得知，描述棒球的 DTD 內容主要是在描述整個比賽的過程、團隊的攻守紀錄、以及個人的攻守紀錄，以便於在比賽進行的過程中，以「即時戰況」的方式傳遞該比賽所有的比賽過程。

嚴格來說，SportML 並未使用 XML 的 Namespace 作為異質詞彙集的整合，而是以 Control File 插入 (Plug-in) 多重 DTD 的方式達到類似功能。換言之，Core DTD 是描述 SportsML 的核心部份，而 Control File 在 SportsML 中扮演的角色是做為具有擴充性質的 DTD，Control File 是把不同類型運動項目的 DTD 與新聞類型的 DTD，利用 XML 語法整合成為 SportsML 文件結構的一環，Core DTD 中的引用 Control File DTD 語法如例 1 所述：

例 1 SportsML DTD-Core DTD 引用 Control File DTD 之 XML 語法

```
<!-- Control File
```

¹⁷ IPTC. (n. d.). [sportsml-specific-baseball-documentation](http://www.sportsml.org/IPTC/SportsML/1.0/documentation/sportsml-specific-baseball-documentation.html) Retrieved October 1,2004, from <http://www.sportsml.org/IPTC/SportsML/1.0/documentation/sportsml-specific-baseball-documentation.html>

This file that gets read-in can be used to designate which sport-specific DTDs accompany the core SportsML DTD.

-->

```
<!ENTITY % sportsml-control-file SYSTEM "sportsml-control-file.dtd">
%sportsml-control-file;
```

同樣在 Control File DTD 中，引用運動項目與新聞類型的 DTD 語法如例 2 所述：

例 2 SportsML DTD- Control File DTD 引用棒球與 NITF DTD 之 XML 語法

```
<!ENTITY % sportsml-specific-baseball SYSTEM "sportsml-specific-baseball.dtd">
%sportsml-specific-baseball;

<!ELEMENT placeholder-element EMPTY>
<!ENTITY % nitf SYSTEM "../..//NITF/3.2/dtd/nitf-3-2.dtd"> %nitf;
```

因此，SportsML 在不須修改 Core DTD 的前提下，可透過 Control File 擴充各類運動項目的 DTD，彼此之間的關係如圖 2 所示。

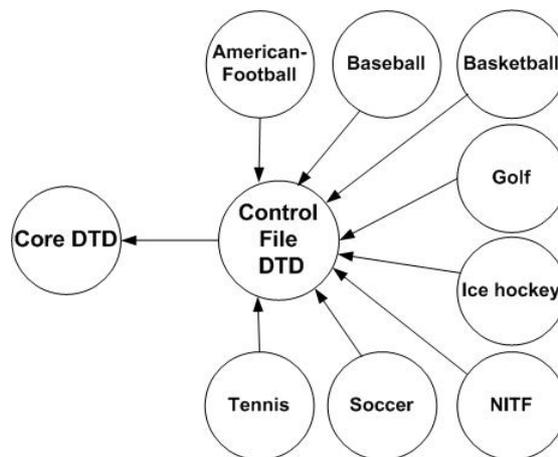


圖 2 SportsML DTD - 三類 DTD 整合示意圖

在利用 DTD 或 Schema 中所定義的 SportsML 文件主體內容中，基本上，SportsML 是以結構化的方式來呈現資料，以 XML 結構進行文件的撰寫與描述，並且可與新聞類型的 Metadata 資料進行整合，從圖 3 可以看出，SportsML 的文件的根元素(Root Element)為<sports-content>，除了必要的<sports-metadata>元素外，其餘的都為可以自行選擇是否撰寫於 SportsML 文件中，並與<team>和<player>元素進行組合，而<article>則是預期使用在描述該運動項目的新聞報導，在規格書中建議使用 NITF。

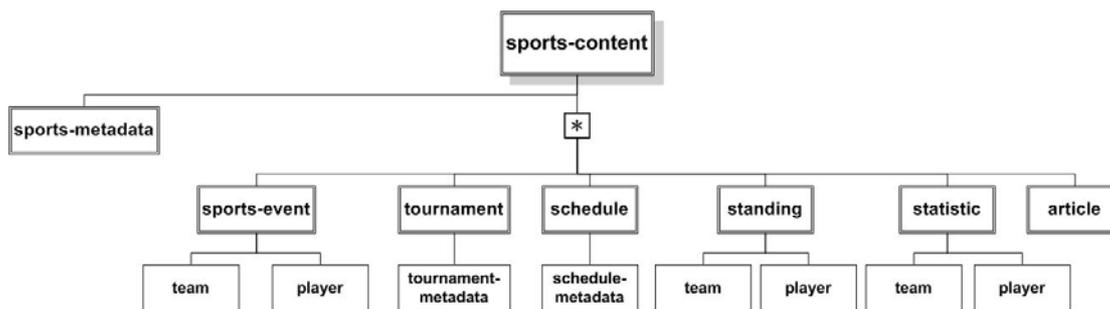


圖 3 SportsML 基本文件架構

參、實例探討

接著我們以國科會數位典藏計畫「臺灣棒球運動珍貴新聞檔案數位資料館之建置」的棒球新聞為例，說明如何單獨使用 SportsML，以及如何與 NITF、NewsML 相互結合描述。

一、SportsML 單獨使用

假設有一則體育新聞，報導某場棒球比賽結果為「統一獅 4:1 擊敗興農牛」，若欲採用 SportsML 來描述此場棒球比賽的戰績，最簡單的情況是只使用最基本的 <sports-event> 與 <team> 兩個元素，即可描述整則新聞內容。原始碼如例 3 所示：

例 3 描述棒球比賽結果－原始 SportsML 部份文件碼

```
<?xml version="1.0" encoding="big5"?>
<sports-content>
<sports-event>
  <sports-metadata event-status="post-event"/>
  <team>
    <team-metadata>
      <name first="統一" last="獅" />
    </team-metadata>
    <team-stats score="4" event-outcome="win" />
  </team>
  <team>
    <team-metadata>
      <name first="興農" last="牛"/>
    </team-metadata>
    <team-stats score="1" event-outcome="loss" />
  </team>
</sports-event>
```

</sports-content>

在此例中，文件首先是宣告遵守XML 1.0 規格，接下來則是開始表達這份文件主要是描述一場比賽結束後的結果，其次則對兩個參賽隊伍(term)進行描述，文件內可看出第一個隊伍名為「統一獅」，「得分」(score) 為「4」，而「事件結果」(event-outcome) 為「win」，另一隊伍名為「興農牛」，「得分」(score) 為「1」，而「事件結果」(event-outcome) 為「loss」。在這份文件中，可看到在宣告事件類型時的屬性為”post-event”，也就是描述最後的結果。若欲將 SportsML轉換為HTML格式，IPTC也定義了數個XSL(eXtensible Stylesheet Language)樣式，可呈現成表格形式¹⁸，如圖 4 所示。當然使用者也可自訂符合自己所要的樣式或另行撰寫SportsML剖析程式來進行轉換。

統一獅 - 興農牛	
隊伍	總分
統一獅	4
興農牛	1

圖 4 描述職棒比賽結果—以 XSL 將 SportsML 轉換為 HTML

此例僅是描述該場比賽的最終結果，並未包含新聞報導的內容。SportsML 內的 <article> 元素便是用來描述與該體育紀錄的相關新聞報導用的，預設建議是採用 NITF 格式，下文則對 SportsML 與 NITF 的整合進行探討。

二、SportsML 與 NITF 結合使用

NITF為IPTC於90年代的一項計畫，目的在於提供在描述新聞時的一種結構化架構，且可以用來標示新聞文件的專門格式，並且加強新聞重覆使用的便利性，不同的通訊社、報社、網站都可藉由NITF轉換新聞內容為不同的資料格式，而達到新聞多重發行的目的，NITF同樣也是由XML語言所建構，因此在資料傳遞與描述上繼承XML的優點，其優點在於(1)新聞內容與實際呈現的方式是分離，並不互相統屬，也就是說，NITF在標示新聞之後，新聞內容的標示與如何呈現是分開進行，NITF只描述新聞資料的內容與語意結構，而不處理新聞資料的呈現方式，所以新聞編輯人員在處理新聞資料時，只需要專注於新聞處理，而不需要了解如何呈現在不同的平台上(2)採用NITF來描述新聞，可提高檢索的準確度。¹⁹

¹⁸ IPTC. (n. d.). [sportsml2html.xml](http://www.sportsml.org/IPTC/SportsML/1.0/examples/sportsml2html.xml) Retrieved October 3,2004, from <http://www.sportsml.org/IPTC/SportsML/1.0/examples/sportsml2html.xml>.

¹⁹ 陳百齡(2002)。讓電腦也能辨識新聞：新聞內容標記問題初探。上網日期：民93年10月5日。網址：<http://ccs.nccu.edu.tw/oldccs/con2002/conworks/11B-1.doc>

NITF 現行的版本為 3.2 版，其主要的結構分為<head>與<body>兩個部份，而<body>又分為<body.head>、<body.content>與<body.end>，說明如表 4：

表 4 NITF 內容結構簡述

名稱	說明
<head>	記錄有關新聞管理的 Metadata，包括新聞檔頭、詮釋資料、主題詞表、格式、出版日期及版本修訂記錄
<body.head>	主要存放新聞標題組件，包括新聞標題、作者、版權、貢獻者、日期與摘要。
<body.content>	存放新聞內文，包括文字、表格、清單、圖片與其他項目與可參照其他相關領域的網址。
<body.end>	新聞的結尾，包括註錄或其他書目資料。

我們仍以上述「統一獅 4:1 擊敗興農牛」為例，並引用聯合報 89/10/23 對該場比賽的體育新聞報導²⁰，原新聞內容節錄一段如下：

羅敏卿雙響砲 統一獅封王【記者李國彥／台中報導】這回統一獅隊真的要去歐洲旅遊了，能夠成行，群獅要感謝強打羅敏卿；昨天中華職棒總冠軍賽第七戰，羅敏卿擊出包括致勝全壘打的單場雙響砲，貢獻三分打點，助獅隊以四比一擊敗興農牛隊，以四勝三敗戰績獲中華職棒十一年總冠軍，並且追平前味全龍隊紀錄，添得隊史上第四面總冠軍錦旗。

如欲在 SportsML 描述比賽結果之外，另加入 NITF 來描述新聞內容，則兩者可整合如例 4 所示。

例 4 描述職棒比賽結果－SportsML 與 NITF 整合描述

```
<sports-content>
<sports-event><sports-metadata event-status="post-event"/>
  <team><team-metadata><name first="統一" last="獅" />
    </team-metadata><team-stats score="4" event-outcome="win" />
  </team>
  <team><team-metadata><name first="興農" last="牛" />
    </team-metadata><team-stats score="1" event-outcome="loss" />
  </team>
</sports-event>
<article format="NITF">
<nitf version="-//IPTC//DTD NITF 3.2//EN" >
```

²⁰ 李國彥(民 89 年 10 月 23 日)。羅敏卿雙響砲 統一獅封王。聯合報，第 30 版，

```

<body>
  <body.head>
    <headline>
      <h1>羅敏卿雙響砲 統一獅封王</h1>
    </headline>
    <byline> By
      <person>李國彥</person>
      <location><city>台中</city></location>報導
    </byline>
  </body.head>
  <body.content><p>這回統一獅隊真的要去歐洲旅遊了，能夠成行，群獅要感謝強打<person>羅敏卿</person>；昨天<org>中華職棒</org>總冠軍賽第七戰，<person>羅敏卿</person>擊出包括致勝全壘打的單場雙響砲，貢獻三分打點，助獅隊以四比一擊敗興農牛隊，以四勝三敗戰績獲<org>中華職棒</org>十一年總冠軍，並且追平前味全龍隊紀錄，添得隊史上第四面總冠軍錦旗。</p>
    <media media-type="text/xml" source="596757"></media>
  </body.content>
</body>
</nitf>
</article>
</sports-content>

```

其中，我們可以清楚的看到在<article>標籤中，以 format 屬性指明新聞報導的描述格式為 NITF，接著便是一連串 NITF 的標籤，<nitf version="-//IPTC//DTD NITF 3.2//EN" >、<body>、<body.head>、<headline>、<byline>、</body.head>、<body.content>、<person>、<location>、<city>、<org> ...等。其中，新聞標題包夾在<body.head>標籤中，而新聞內容則置於<body.content>內。

由此可知在 IPTC 所建議的架構中，體育記錄與相關報導是以獨立形式進行描述，平時各不統屬，但由於各類運動有其獨特的記錄方式，SportsML 只做為提供一個交換資料的平台，同時 NITF 也只做為描述新聞之用，所以可將兩者視為描述一項體育活動資料時各自獨立的一塊拚圖，將其整合在同一份文件中，賦予其新的意義。

三、SportsML 結合 NITF 並與 NewsML 整合

NewsML與NITF和SportsML同樣則是由IPTC所發展出的一套讓新聞傳播媒體互通使用的新聞標示語言標準，它可有效記錄與管理多媒體新聞資訊的編採、儲存、傳遞、檢索與使用，把雜亂無章的新聞資料，整合於一結構裡，讓任何的

儲存裝置都能夠進行解讀²¹，NewsML的概念是隨著詮釋資料來描述各種不同的媒體，如文字、影像、照片等，主要是用來描述電子出版、傳送、典藏的新聞文件，同時對新聞內容的格式並不預設立場，因此NewsML是可以用在文字、圖片、語音、影視，以及未來發展的媒體型態²²。

而NewsML如同SportsML與NITF，同樣也從XML語言所發展，在NewsML的架構中，最主要是分為新聞封裝(NewsEnvelope)與新聞條目(NewsItem)兩個部份，NewsEnvelope主要是用來記錄新聞傳送過程中所需要的各種訊息，就如同一個將新聞打包的過程，將要傳遞的新聞內容裝在一個信封裡，並在信封註明何時要由何地將新聞資料送往何處等資訊。NewsItem則記錄了新聞內容、新聞素材及相關資訊，也可說是包裹在信封裡，實際要傳送的東西，包含新聞的實際內容、各種管理訊息和識別資訊等，分為以下幾個部份²³：

表 5 NewsItem 結構簡述

元素	說明
識別號 (Identification)	新聞條目 ID，其中包括提供者 ID、日期 ID、版本 ID、公眾 ID 等識別號的記錄資訊。
新聞管理 (NewsManagement)	有關新聞管理的相關記錄，包括新聞類型、初建時間、最後修改時間與新聞狀態等資訊。
新聞組件 (NewsComponent)	由多個內容素材(Content Item)所構成，成並描述所有新聞物件在結構上之相互關係。

NewsML 的概念，最底層的結構為內容素材(ContentItem)，在描述 NewsML 的文件結構上，SportsML 和 NITF 所描述的內容同樣被視為內容素材的一部份，再與其他相關的內容素材進行結合，而成為新聞組件(NewsComponent)的一部份，並可與其他的新聞組件再進行嵌合，其次再註明何時要傳遞到何地，以一層層包裹的方式進行新聞封裝(NewsEnvelope)，最後再進行傳遞。

同樣，若仍以上述的棒球比賽新聞為例，將包含 NITF 的新聞內容和 SportsML 比賽紀錄，與 NewsML 進行整合之後的結果如例 5 所示。

例 5 利用 NewsML 將 SportsML 與 NITF 進行新聞封裝部份文件碼

```
<?xml version="1.0" encoding="big5"?>
```

²¹ 張榕容。新聞標示語言標準-NewsML。

http://river.glis.ntnu.edu.tw/homework/information-1/general/newsml_D.pdf

²² 孫正宜(民 93)。新聞專卷的數位化與增值應用—以台灣棒球報紙新聞數位典藏為例，未出版之碩士論文，淡江大學資訊與圖書館學所碩士論文，頁 26

²³ 同註 21

```

<NewsML>
  <NewsEnvelope></NewsEnvelope>
  <NewsComponent>
    <ContentItem>
      <DataContent>
        <sports-content>
          <sports-event><sports-metadata event-status="post-event"/>
          <team><team-metadata><name first="統一" last="獅" />
          </team-metadata><team-stats score="4" event-outcome="win" />
          </team>
          <team><team-metadata><name first="興農" last="牛" />
          </team-metadata><team-stats score="1" event-outcome="loss" />
          </team>
        </sports-event>
      </DataContent>
    </ContentItem>
  </NewsComponent>
</NewsML>

```

```

<article format="NITF"><nitf version="-//IPTC//DTD NITF 3.2//EN" >
  <body>
    <body.head>
      <headline>
        <h1>羅敏卿雙響砲 統一獅封王</h1>
      </headline>
      <byline> By
        <person>李國彥</person>
        <location><city>台中</city></location>報導
      </byline>
    </body.head>
    <body.content><p>這回統一獅隊真的要去歐洲旅遊了，能夠成行，群獅要感謝強打<person>羅敏卿</person>；昨天<org>中華職棒</org>總冠軍賽第七戰，<person>羅敏卿</person>擊出包括致勝全壘打的單場雙響砲，貢獻三分打點，助獅隊以四比一擊敗興農牛隊，以四勝三敗戰績獲<org>中華職棒</org>十一年總冠軍，並且追平前味全龍隊紀錄，添得隊史上第四面總冠軍錦旗。</p>
      <media media-type="text/xml" source="596757"></media>
    </body.content>
  </body>
</nitf>
</article>

```

```
</NewsComponent>
</NewsML>
```

從以上各例的描述可得知，體育活動的紀錄與相關新聞報導等，在結構上皆為各自獨立的一部份，相對而言，運動紀錄員與新聞記者都只要單純的進行資料處理的工作，而不須了解整合過後的文件是如何被呈現，或是如何的被傳遞，無論是 SportsML、NITF、NewsML 三者，都只要按照既定所制訂好的元素填入即可，以結構化的文件進行傳遞，以達到跨平台的目的。

肆、數位典藏計畫與中文新聞詮釋資料之發展概況

國科會推動的「數位典藏國家型科技計畫」中，與新聞相關的內容發展分項計畫皆歸屬「新聞主題小組」²⁴。由於國際上新聞詮釋資料格式眾多，加上新聞報導涵蓋政治、財經、體育、娛樂、科技 ... 等多元面向，內容差異性頗大。因此，目前新聞主題小組並未訂定統一的中文新聞詮釋資料格式，而是由各分項計畫自行依實際需求進行規劃。今年度（2004 年）正在執行的計畫中，歸屬「新聞主題小組」的有兩件，分別是：(1) 世新大學的「北平『世界日報』內容數位化開發計畫」；(2) 淡江大學的「臺灣棒球運動珍貴新聞檔案數位資料館之建置」計畫。其中，前者因該報為早期之舊式報紙，在內容、版面、廣告等各方面之編排皆迥異於今日之報紙，以致於目前現有的 Metadata 格式不盡然適用，因此，該計畫預計制訂一份適用舊式報紙的標示語言，目前正在研擬中。而後者則初步先採用與 DC 相容的 NRMF 格式，以便能儘速與同採 DC 格式的數位典藏聯合目錄接軌，未來擬再陸續加入與其他規格相容之設計，如 NewsML、NITF 或 SportsML 等，以提升其資料交換與傳遞能力。

有感於中文新聞詮釋資料之重要性，海峽兩岸新聞業界近年也開始著手研擬相關標準或開發相關軟體。例如台灣方面，有聯合報系正在進行中的「聯合新聞標示語言」(UDN Markup Language, 簡稱 UdnML)；大陸方面則有新華社的「新華標示語言」(Xinhua Markup Language, 簡稱 XinhuaML)²⁵，中國報業協會的「中國報業電子新聞文稿格式」(Chinese News Text Format, 簡稱 CNTF)²⁶等；香港方面，「中文新聞標示語言協會」²⁷不但加入 IPTC 成為會員，亦著手開發了一套遵循 NewsML 標準的網上文檔處理及發佈系統 CIDAX²⁸。

目前中文世界在新聞詮釋資料的發展，尚不若西方國家成熟，以 NITF 和

²⁴ 國科會。數位典藏國家型科技計畫。上網日期：民 93 年 9 月 18 日。網址：<http://www.ndap.org.tw/>

²⁵ 新華社技術局標準工組(2003 年 1 月 18 日)。[XinhuaMLv1.0 功能說明書](#)。

²⁶ 中國報業協會規範工作組(2000 年，5 月)。[中國報業電子新聞文稿格式](#)。

²⁷ 中文新聞標示語言協會。[中文新聞標示語言協會首頁](#)。上網日期：民 93 年 10 月 14 日。網址：<http://cnewsml.org/>

²⁸ 中文新聞標示語言協會。[應用演示 CIDAX 企業資訊數碼系統](#)。上網日期：民 93 年 10 月 14 日。網址：<http://cnewsml.org/tradchi/cidax.shtml>

NewsML為例，因其發展已有一段時日，目前已有法新社（AFP）、美聯社（AP）、紐約時報（The New York Time）、義大利ANSA通訊社 ... 等，高達一、二十家國際知名的通訊社或報社已成為其用戶。^{29,30}但就描述體育新聞的SportsML而言，不論國內外目前皆處於剛在起步的階段，未來仍有努力的空間。

伍、結論與建議

SportsML 為 IPTC 為了傳遞體育活動資訊而所制訂的一項標準，目前現行所公布的 SportsML 1.0 版，是以提供一個資料交換的平台，讓不同的體育活動能夠跨越格式的限制，以 XML 語法結構進行資料的交換，同時結合其他相關的詮釋資料，如 NITF 與 NewsML，以標準化的文件結構，增強資料儲存與檢索的能力。

SportsML 的優點在於可讓使用者自行以 Control File 擴充控制詞彙與 DTD，在已公布的版本中，制訂了美式足球、籃球、棒球、高爾夫、冰上曲棍球、網球與足球等七種不同體育運動的 DTD，對在 SportsML 上推廣上有其助益。但就使用的便利性上，仍有努力的空間，因仍有更多的運動項目未自行制訂其專屬的 DTD，或甚至並未發展到具備能制訂 DTD 的層級，這也是 SportsML 在發展上所面臨的困境。因此，IPTC 鼓勵其他的運動項目，無論是成熟度高的或是仍在發展的運動，都能夠邊發展邊制訂其專屬的 DTD 或 Schema，如此在體育紀錄的儲存與檢索上，能更加的快速與具便利性。

本研究建議，如媒體欲採用 SportsML 做為體育新聞的詮釋資料格式，在加值利用上可與其他同類型的媒體進行合作，報業媒體可派出文字記者與攝影記者進行平面上的記錄，包括體育活動的紀錄、新聞內容與圖像，而電視記者則拍攝體育新聞的影音，廣播記者則為以聲音記錄該新聞內的內容，如此以平行整合的方式，將文字、影音、圖像的進行整合，並以 NewsML 結合 SportsML，讓新聞傳遞上更具結構化，在解讀上也較為容易，並配合資料庫的儲存，讓新聞記者在新聞檢索上能更加的便利。

我國在新聞類型的詮釋資料的發展上，未來仍有相當大的努力空間，網路與全球化仍是無可避免趨勢，在新聞傳遞上如要與國際接軌，最好能採取國際上所普遍認可的格式，以具備能與國內外媒體進行資料交換的能力，才能更有競爭力。

致謝

承蒙國家科學委員會經費補助（計畫編號：NSC 93-2422-H-032-001），使本研究得以順利進行，特此致謝。

²⁹ IPTC.(n. d.). Who's Using NITF ? Retrieved October 18,2004, from <http://www.nitf.org/users.php>.

³⁰ IPTC.(n. d.). Who's Using NewsML ? Retrieved October 18,2004, from http://www.newsml.org/pages/whouse_main.php.