

從後設資料概念模式探討知識組織的應用與發展

On Knowledge Organization: A Metadata Conceptual Model

陳亞寧

Ya-Ning Chen

中央研究院計算中心系統分析師兼組長

Systems Analyst

Computing Centre, Academia Sinica

E-mail: arthur@gate.sinica.edu.tw

【摘要 Abstract】

隨著數位圖書館計畫的實施，後設資料 (Metadata) 的應用如何達成有效的知識組織隱然成為一項重要的研究議題。本文一共選取了「書目記錄功能模式」、「CIDOC 概念參考模式」、「記錄連貫性模式」、「開放式檔案資訊系統參考模式」、「應用特徵檔」、「ABC 本體論模式」，以及「後設資料與資訊的交互作用」等七大概念模式為研究對象進行內容分析，研究發現有四：可以同時採用兩個以上的後設資料標準；後設資料的產生時機及涵蓋範圍，往往集中在組織此一方面；數位資訊的永久保存暨典藏的及其後設資料逐漸受到重視；後設資料的類型與關聯性包括了人、事、時、地、物，有其重要性與必要性，必須審慎思索。

How to deploy metadata to achieve knowledge organization becomes an important issue for any digital library projects around the world. This paper examines seven conceptual models to find out how metadata is employed as an effective method for knowledge organization. Four findings are raised as follows: (1) it is a practical and workable approach to use more than two metadata formats for any digital project; (2) the use of metadata formats is always focused on information organization and others such as information creation, access and management, dissemination, and permanent archiving and preservation are often ignored; (3) preservation metadata has become a new research domain for knowledge organization; (4) it is imperative to deliberate on generic metadata attributes such as person, event, time, place, object.

關鍵詞 Keyword

後設資料 模式

Metadata ; Modeling



壹、前言

隨著各項數位圖書館計畫的研究與推動，後設資料已成為一項重要的資訊基礎建設之一，因為後設資料的設計與應用緊密地與知識組織結合在一起。因而，許多數位圖書館計畫對於後設資料的發展與實施方面，產生了許多根本性的問題，例如：是否只能應用一套後設資料標準？如果某一套後設資料標準無法完全符合計畫需求時，有無其他的可行方式？本文試從七項後設資料標準的概念模式為出發點，重新探究後設資料標準的實際應用，以了解知識組織的應用與發展。

貳、研究方法與範圍

在各式數位圖書館計畫發展之際，經常發現使用單位是直接採用某一種後設資料標準，而無法充分了解後設資料標準的全貌，進而有效發揮後設資料標準原始的功能與目的，或是選用最適用的後設資料標準。本文試以後設資料研究領域為範圍，採用內容分析為方法，以後設資料概念模式為研究對象進行探討。在個案的選取方面，就圖書館、博物館、檔案館、資訊科技界各選取一個後設資料概念模式，包括了國際圖書館學會（International Federation of Library Associations and Institutions, IFLA）的「書目記錄功能模式」（Functional Requirements for Bibliographic Records model, FRBR model）、國際文獻委員會（Le Comité international pour la documentation, CIDOC）發展的「概念參考模式」（Conceptual Reference Model）、澳洲檔案館界提出的「記錄連貫性模式」（Records continuum model），以及美國國家航空暨太空總署（National Aeronautics and Space Administration, NASA）太空資料系統諮詢委員會（Consultative Committee for Space Data Systems, CCSDS）提出的「開放式檔案資訊系統參考模式」（Open Archival

Information System Reference Model, OAI reference model）。此外，為了擴大研究範圍的廣度，增加選取了英國線上圖書館資訊網路辦公室（UK Office for Library and Information Networking, UKOLN）的「應用特徵檔」（Application profile）、跨國合作「和諧計畫」（Harmony Project）所發展的「ABC 本體論模式」（ABC ontology model），以及國內中央研究院計算中心後設資料工作組（Metadata Architecture and Application Team, MAAT）的「後設資料與資訊的交互作用模式」（Metadata-information interaction model）。因而本文一共選取了七大概念模式為研究個案進行內容分析，以探討有關知識組織的現況發展與應用。

參、重要的後設資料標準與概念

依據前述概念模式為對象進行個案的研究探討外，並提出作者個人的觀察。

一、書目記錄功能模式

在 1990 年，IFLA 的「書目控制與國際機讀編目計畫」（IFLA Universal Bibliographic Control and International MARC Programme, UBCIM Programme）與「書目控制部門」（IFLA Division of Bibliographic Control）共同舉辦了「Stockholm Seminar on Bibliographic Records」研討會，在會議中提出了一項需求—釐清書目記錄的功能需求。因而，IFLA 開始著手進行研究，以完成上述需求導向的任務。在 1998 年，IFLA 完成並提出了 FRBR 概念模式，其中之一的主要目的在於：建立一個架構，期能清晰、正確地陳述書目記錄的目的，以達成在完成使用者需求時，書目記錄應該提供的資訊要項（IFLA Study Group on the Functional Requirements for Bibliographic Records, pp. 1-2）。

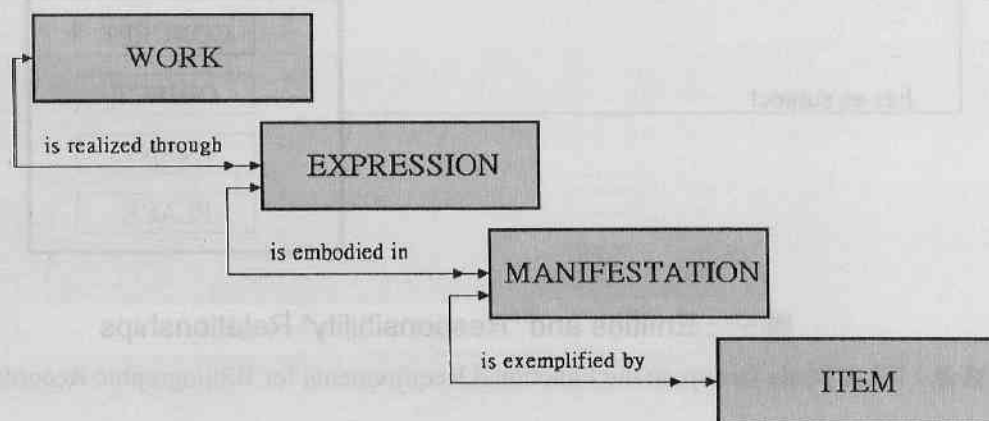
FRBR 概念模式係利用實體關聯模式（Entity-Relationship model, ER model）為基礎進行



設計、發展，預期能達到發現（Find）、辨識（Identify）、選擇（Select）與取得（Obtain）等四項功能目的（IFLA Study Group on the Functional Requirements for Bibliographic Records, pp. 8-9），而模式內涵主要由兩組不同實體（Entity）及其關聯性（Relationship）共同組成（請參見圖一至圖三）。在實體方面，包括了作品（Work）、表達方式（Expression）、媒體（Manifestation）與館藏（Item），以及觀念（Concept）、物件（Object）、事件（Event）與地點（Place）等兩組實體；在關聯性方面，則具備了主要（Primary）、責任（Responsibility）與主題（Subject）等三種核心關係（IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998, pp.12-16），並由此核心關係擴展為更多不同類型的關聯性。

所謂的作品、表達方式、媒體、館藏究竟所指為何？依據 IFLA 的解釋，所謂的作品是指一個抽象的實體，足以代表內涵或藝術上的創作（a distinct intellectual or artistic creation）（IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998, p.16），基本上是一種抽象的意念或概念；表達方式則是充分了解作品後，經由文數

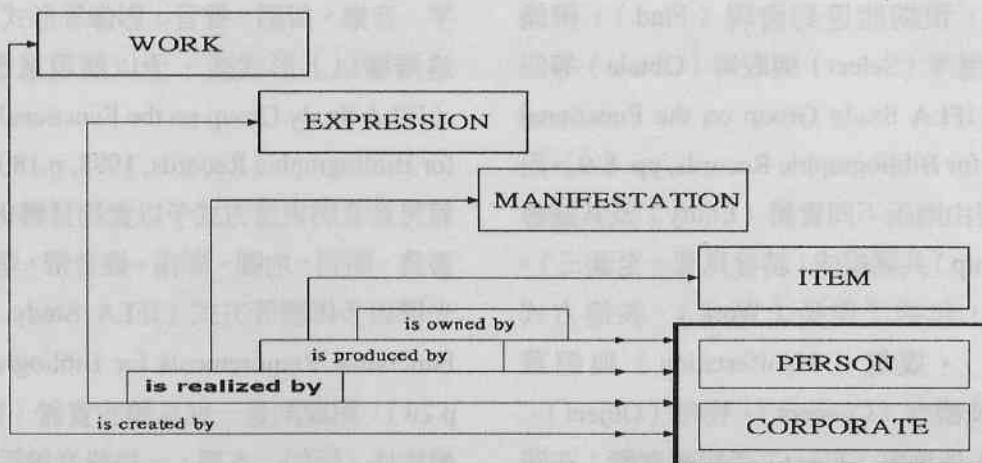
字、音樂、舞蹈、聲音、影像等形式，或是結合前述兩種以上形式後，予以展現意念的一種手法（IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998, p.18）；媒體則是具體將意念與表達方式予以實物具體化，包括手稿、書籍、期刊、地圖、海報、錄音帶、電影、錄影帶、光碟與多媒體等方式（IFLA Study Group on the Functional Requirements for Bibliographic Records, p.20）；館藏則是一種具體的實體，通常是一種實體物件，例如一本書、一卷錄音帶等（IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998, p.20）。如果改由一個實際範例說明上述四種實體，即是：托斯卡（Tosca）是由普契尼（Giacomo Puccini）所創作出來的一種概念性意涵，它被某一劇作家了解後，經由歌劇與電影等不同表達形式予以進行展現，同時被媒體公司製作成光碟、電影片等不同媒體予以發行，最後這些光碟或電影片可以被圖書館或個人蒐集為館藏的一部份。此外，此一托斯卡作品相關主題涉及了：劇作家—普契尼、場景地點—羅馬、姐妹劇作—蝴蝶夫人等。



圖一：Entities and Primary Relationships

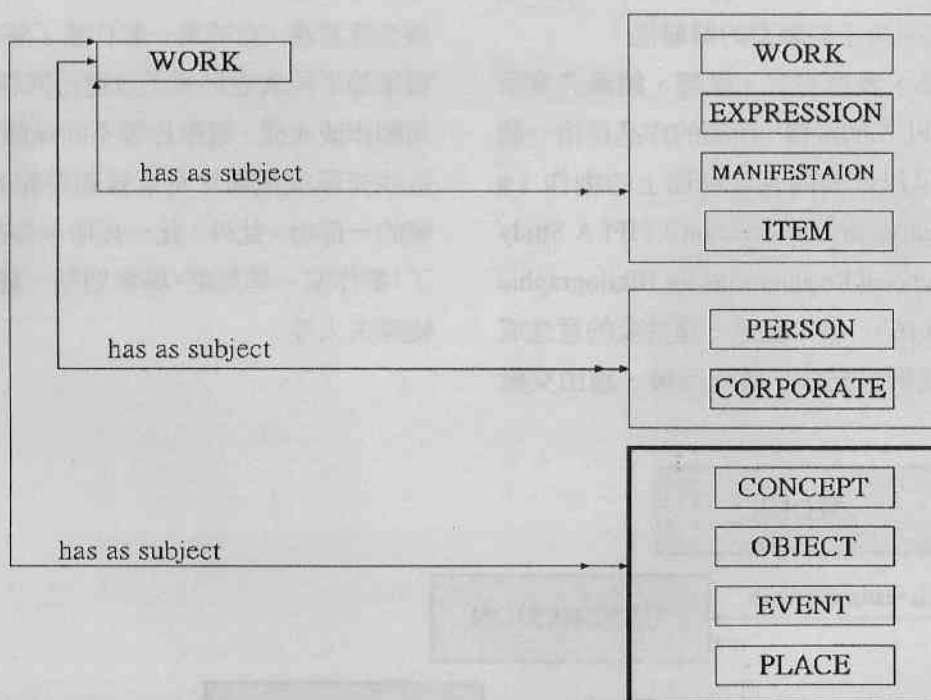
資料來源：IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998. p.13.





圖二：Entities and “Responsibility” Relationships

資料來源：IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998, p.14.



圖三：Entities and “Subject” Relationships

資料來源：IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998, p.15.

近年來，FRBR 模式已被圖書館界視為一項極為重要、創新、突破的觀念，一再被圖書館界應用，同時也廣泛地被其他學術領域引用，包括了影音與

智慧財產權等層面；同時，相關研究與應用計有：歐盟(European Community, EC)的「ECHO Project」(Amato, Castelli, Pisani, Venerosi, Poncin and Vinet,



2000)與「INDECS Project」(Bearman, Miller, Rust, Trant and Weibel, 1999)、W³C 的著作權機制 (Iannella, 2001)、美國機讀編目書目與館藏格式的分析 (Delsey, 2002), 以及國內中央研究院計算中心 MAAT (Chen and Chen 2001; Lin, Chen and Chen, 2001) 等。

從中央研究院計算中心 MAAT 的實際應用與相關研究中, 可以發現 FRBR 概念模式的優缺點分別是 (Lin, Chen and Chen, 2001):

(一)優點:

- 1.釐清不同實體(包括作品、表達方式、媒體與館藏,以及原件與複製品)所屬的各項後設資料元素及實體的雙向繼承關係(Inheritance),乃至於使用需求的傾向。
- 2.可以展延以界定各項實體的功能需求,如索引、展現、著作權、擁有權、管理(含行政與系統等方面)等。

(二)缺點:

- 1.有時候表達方式並無法全然獨立,必須與另外一個實體合用(如作品或媒體),尤其是應用在博物館或美術館藏品時,常有此一現象發生。
- 2.儘管以 FRBR 模式基礎可以擴展出相關的功能需求,但是 FRBR 本身並不針對這些功能導向的內涵機制提出進一步說明,必須由發展單位自行研發,如工作流程(Workflow)、著作權、擁有權、長久保存與典藏(Preservation and archiving)等。

二、CIDOC 概念參考模式

此一模式係由國際博物館委員會(International Council of Museums, ICOM)委託 CIDOC 發展的,而第一版 CIDOC 概念參考模式在

1998 年正式提出,以取代「The International Guidelines for Museum Object Information: The CIDOC Information Categories」(Crofts, 2000)。此一模式預計達成的目的有四,如下(Crofts, Doerr and Gill, 2003):

- (一)涵蓋所有文化遺產的文獻資訊及全球的資訊交換。
- (二)容許局部與對立式知識的文獻記載(Documentation of partial and contradictory knowledge)。
- (三)在資訊整合與交換時,可以達成語意上的通順。
- (四)為未來發展,提供一個定義明確、可展延的架構(Framework)。

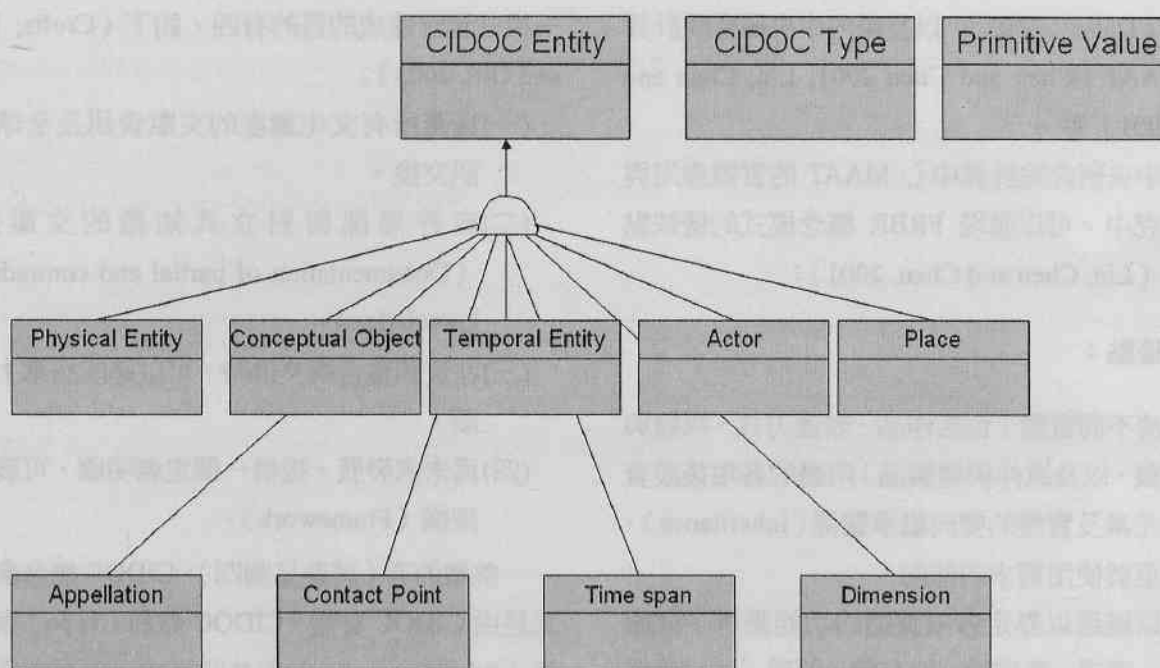
整體而言(請參見圖四),CIDOC 概念參考模式是由 CIDOC 實體、CIDOC 形別(Type)與原始值(Primitive value)所共同組成,主要內涵摘述如下(Crofts, 2000):

(一)CIDOC 實體

在 CIDOC 概念參考模式中,CIDOC 實體是在模式中主要類別(Main class)的最高類別(Parent class)。在實體之下,又可以區分為處理文化資訊所包含主要概念(Primary concepts)的主要實體,以及扮演參考性質的實體。前者包括了物件實體(Physical entity)、概念物件(Conceptual object)、時間實體(Temporal entity)、角色(Actor)、地點(Place),後者則有名稱(Appellation)、連絡方式(Contact point)、期間(Time span)、尺寸(Dimension)。

- 1.物件實體:亦即博物館的館藏藏品,是一切實體物件的最高類別。
- 2.概念物件:用以包含智能或概念上的物件(Intellectual or conceptual objects),有別於實際物件。





圖四：CIDOC 概念參考模式概觀圖

資料來源：Crofts, 2000

3. 時間實體：有關時間屬性的類別，包括一段時間(Period)、事件(Events)與狀態(States)。
4. 角色：可以執行行動的類別，包括了個人、團體 (Groups) 與機構。
5. 地點：表示一個空間實際範圍的概念 (The idea of a physical extent of space)。
6. 名稱：在某一情境下，用以識別某一實例 (Instance) 的名稱、代碼或文字。
7. 連絡方式：包括各式地址的標引方式，如電話、電子郵件或座標系統等。
8. 期間：支援時間實體之用，以決定某一日的日期或期間 (A range of dates or duration)。

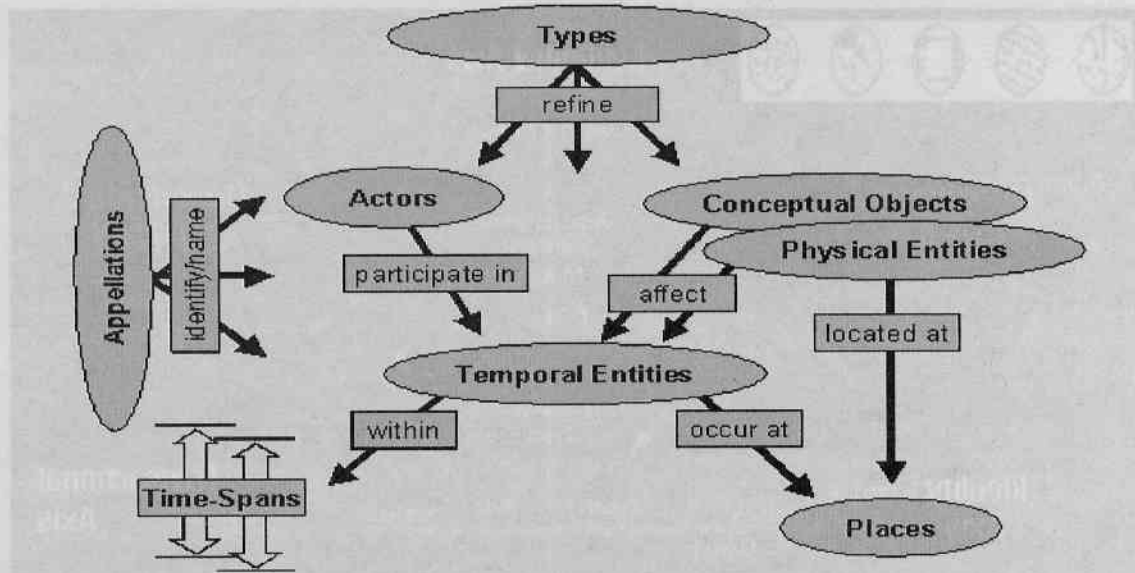
9. 尺寸：記載丈量的相關資訊，如頻率、長度、直徑、重量、明亮度等。

(二)CIDOC 形別：提供一個最高類別，以更進一步精確地限定某一實例的屬性。

(三)原始值：用以記載一些不需要進一步分析的文獻資訊，又可分為字串 (String)、數字 (Number) 與用來處理年代序的資訊 (Time primitive)。

基本上，CIDOC 概念參考模式採用物件導件觀念設計，因而在不同實體間的互動關係就以屬性 (Property) 方式加以標引 (請參見圖五)。





圖五：CIDOC 概念參考模式實體關係圖

資料來源：Crofts, Doerr and Gill, 2003

由上述內容中，我們可以發現 CIDOC 概念參考模式已為知識組織開創出兩項新意：

- (一)強調一種語意上的關聯性 (Semantic association)，同時具備雙向式串聯關係。
- (二)涵蓋的層面包括了物、人、時、地及其關聯。

三、記錄連貫性模式

1996 年 6、7 月份，由澳洲蒙那許大學 (Monash University) 分別在墨爾本 (Melbourne) 與坎培拉 (Canberra) 兩地舉行的研討會中，Frank Upward 提出了「記錄連貫性模式」，以因應數位時代的需求。同年八月，檔案暨博物館資訊學學會 David Berman 針對此一模式提出修正。由 Upward 提出的記錄連貫性模式係由十六個焦點、四大主軸 (Axis) 與四個面向 (Dimension) 所共同組成 (請參見圖六)。此一模式內涵摘述如下：

(一)十六焦點與四大主軸

1. 記錄保存主軸 (Recordkeeping axis)

涉及人類活動所產生的各項記錄及儲存，包括文件 (Document)、描述記載文件的記錄 (Surrogate record)、聚集記錄的單一檔案 (Archive) 及檔案群組 (Archives)。

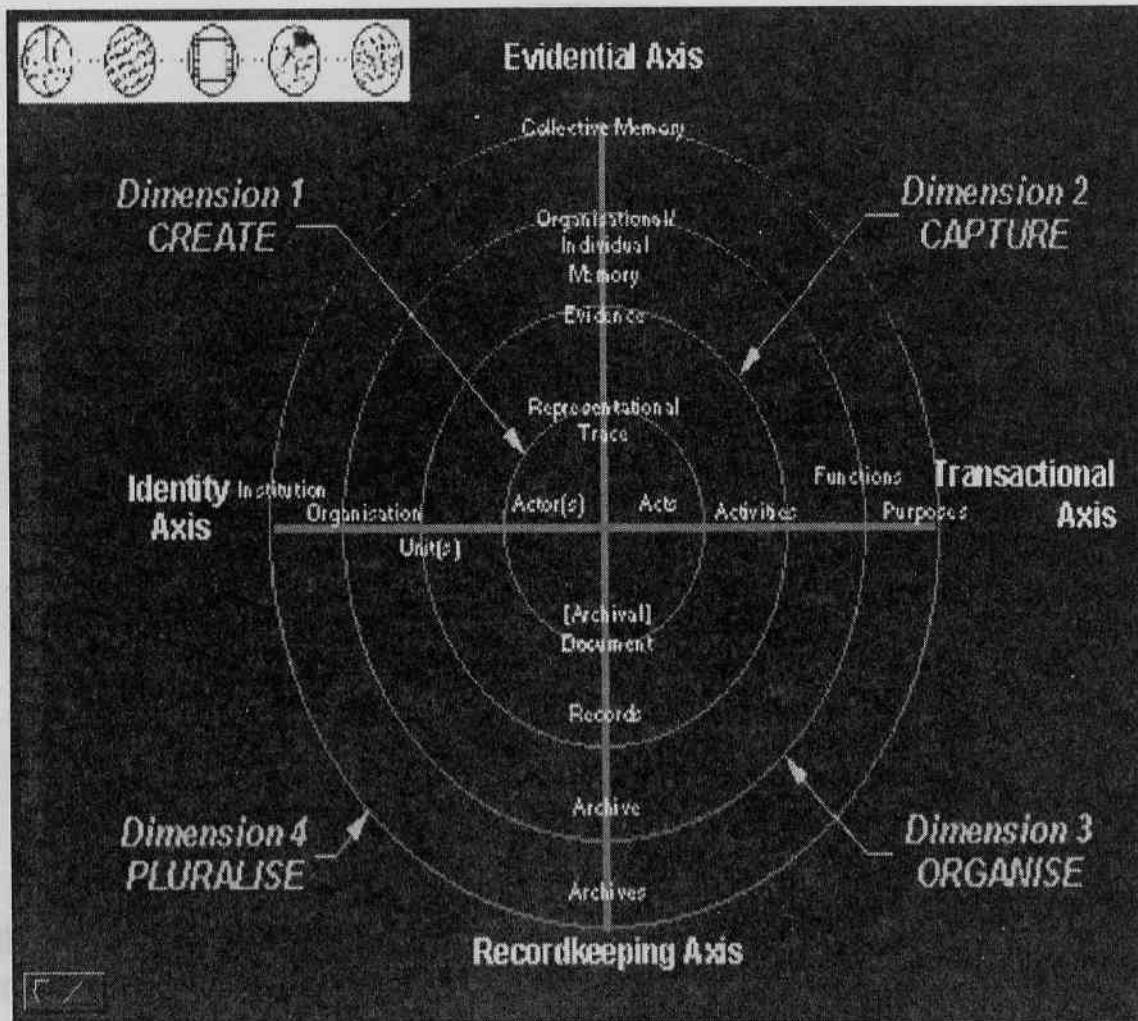
2. 證據主軸 (Evidence axis)

追溯、探究各項活動與措施 (Action)、記錄提供的證據 (Evidence)，乃至於一個機構歷史活動 (Corporate memory) 至社會群體的記憶 (Collective memory)。

3. 活動主軸 (Transactional axis)

標示文件、記錄所透露的各項活動，因層次的不同而有不同的標引名稱，包括行動措施 (Act)、活動 (Activities)、功能 (Functions) 與目的 (Purposes)。





圖六：記錄連貫性模式

資料來源：Upward, 1996

4. 實體主軸 (Identity axis)

顯示相關活動所關聯的人或機構，包括執行者 (Actor)、執行部門 (Unit)、執行機構 (Organization) 與社會群體 (Institution) (Upward, 1996)。

(二) 四個面向 (由內而外，逐一擴大延展)

1. 第一面向：活動的產生 (Create the trace)

經由電子系統產生活動相關的文件，以利系統的交流、互動 (Communication)。

2. 第二面向：活動的追蹤 (Capture trace as record)

各項系統的交流活動，包括註冊、文件的鏈結 (Linking) 與傳播 (Disseminating)、記錄的維護 (如記錄移除) 等。

3. 第三面向：活動的組織 (Organize the record as memory)

涵蓋系統如何建立、調閱與傳播已完成工作任務的各項相關記錄。

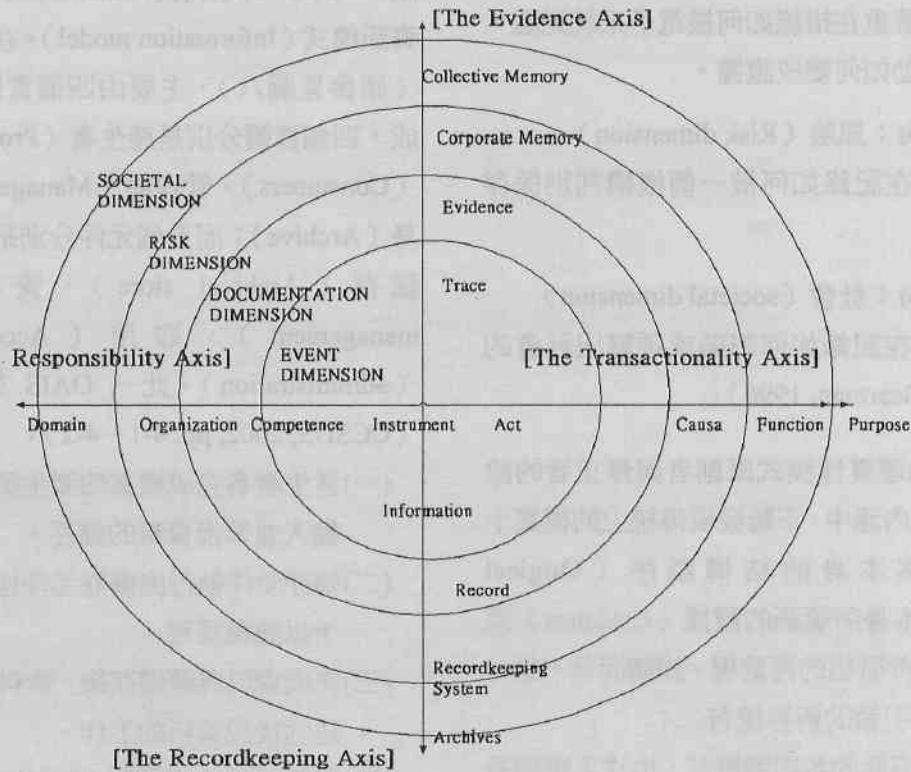


4. 第四面向：活動的累積 (Pluralize the memory)

以社會群體為範圍所建立的各项相關記錄，進一步形成檔案資訊系統 (Archival

information system) (Upward, 1996)。

另一方面，Berman 將上述 Upward 的記錄連貫性模式內涵提出修改建議，主要變動如下(請參見圖七)：



圖七：David Berman 修正後的記錄連貫性模式

資料來源：Berman, 1996

(三) 十六焦點與四大主軸

1. 記錄保存主軸 (Record keeping axis)：包括項目更改為資訊、記錄、記錄保存系統 (Record keeping system)、檔案。
2. 證據主軸 (Evidential axis)：同上，詞彙由「Representational trace」變為「Trace」。
3. 活動主軸 (Transactionality axis)：將詞彙「Transactional」與「Activities」分別改為「Transactionality」與「Causa」，強調是一

種變動性的活動。

4. 責任主軸 (Responsibility axis)：包括方法 (Instrument)、能力資產 (Competence)、機構 (Organization) 與領域 (Domain)，強調在不同階段所形成的一種意涵 (Berman, 1996)。

(四) 四大面向

1. 第一面向：證據 (Evidence dimension) 強調活動的異動情形。



2. 第二面向：文件記載 (Documentation dimension)

著墨在一個行動措施如何轉換成變動性的活動、追蹤如何變成記錄、方法如何變成機構的一種能力資產、資料如何形成記錄，著重在措施如何被電子系統見證，以及活動如何變成證據。

3. 第三面向：風險 (Risk dimension)

著重在記錄如何被一個機構判別保存與否。

4. 第四面向：社會 (societal dimension)

注重在記錄如何彰顯或透露出社會的意義 (Bearman, 1996)。

從上述記錄連貫性模式原創者與修正者的詮釋及賦與的意義內涵中，不難發現傳統上的檔案十分重視在檔案本身的結構順序 (Original provenance) 及本身所蘊涵的證據 (Evidence) 意涵，以利歷史事件原相的再重現。總體而言，此一模式為檔案學界引領的新境界有二：

- (一)除了檔案事件的各項過程外，也注入電腦系統化流程的觀念。
- (二)檔案的記錄及其事件是一種持續與連貫性活動，不能只固定在單一時間點、事件點而已。因而，在檔案的描述、內涵與管理上，不能只從單一觀點或層面切入而已，必須從時間與事件的連續性主軸上加以展延，同時以宏觀面切入，從不同的學科領域予以詮釋。

四、開放式檔案資訊系統參考模式

1999 年 5 月，美國 CCSDS 對外發表了 OAIS 參考模式，此一模式是針對檔案系統所設計的一種

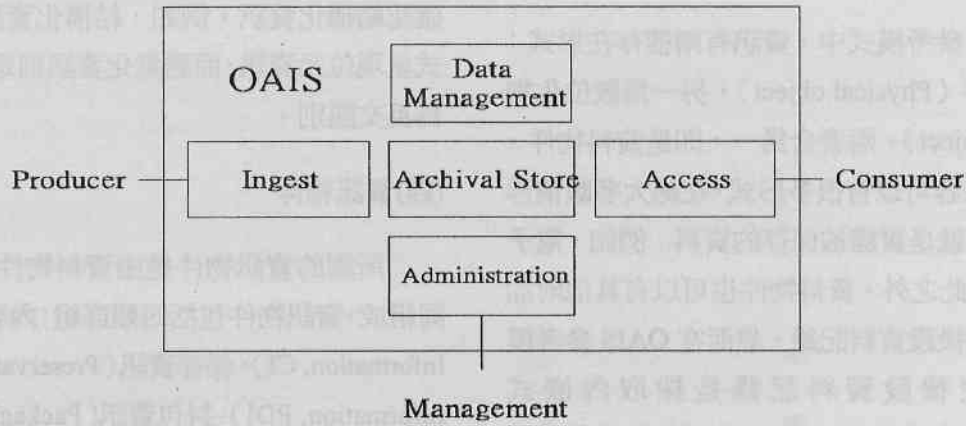
概念架構 (A conceptual framework)，涵蓋了檔案及所屬環境、檔案本身的功能元件 (Components)，以及支援檔案處理的資訊基礎架構。

在 OAIS 參考模式主要由兩個概念共同組成，一為 OAIS 架構 (Framework)，另一個是 OAIS 資訊模式 (Information model)。在 OAIS 架構方面 (請參見圖八)，主要由四個實體與五個元件組成，四個實體分別是產生者 (Producers)、使用者 (Consumers)、管理者 (Management) 與檔案本身 (Archive)；而五個元件分別是攝取 (Ingest)、儲存 (Archival store)、資料管理 (Data management)、取用 (Access) 與管理 (administration)。此一 OAIS 架構主要意涵是 (CCSDS, 2002, pp. 4-1-4-2)：

- (一)當生產者完成檔案的產生後，經由攝取元件輸入並準備資料的儲存。
- (二)儲存元件執行由攝取元件送達備妥資料，並予以典藏管理。
- (三)完成資料典藏儲存後，資料管理元件實施描述性後設資料的工作。
- (四)取用元件主要目的則是協助使用者確認與取得檔案相關的描述資訊，同時將這些資訊傳送給使用者。
- (五)最後，管理元件由管理者負責檔案的日常作業。

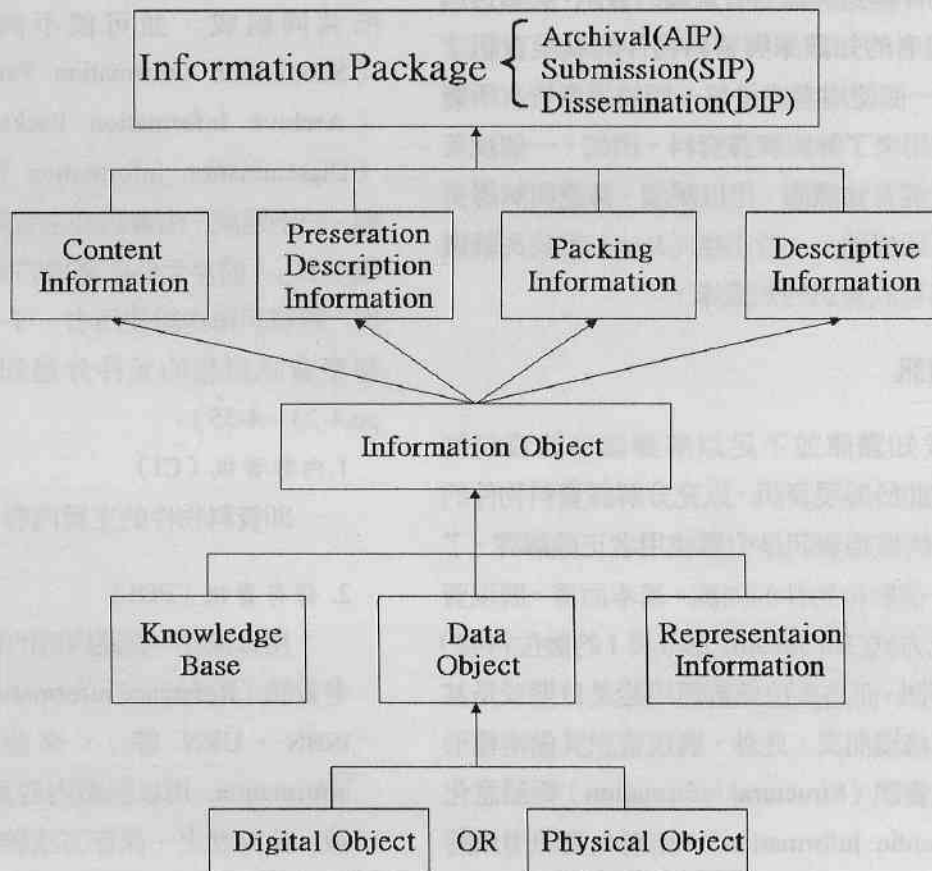
在 OAIS 資訊模式方面 (請參見圖九)，主要由資訊物件 (Information object)、資訊封包 (Information package)、資料物件 (Data object)、知識庫 (Knowledge base)、展現資訊 (Representation information) 五大元件組成，主要內涵摘述如下 (CCSDS, 2002, pp.4-18-4-23; OCLC and RLG Working Group on Preservation Metadata, 2001, pp.13-16)：





圖八：OAIS and its Environment

資料來源：OCLC and RLG Working Group on Preservation Metadata, 2001, p.10.



圖九：OAIS Information Model

資料來源：OCLC and RLG Working Group on Preservation Metadata, 2001, p.11.



(一)資料物件

在 OAIS 參考模式中，資訊有兩種存在形式：一為實體物件 (Physical object)，另一為數位化物件 (Digital object)。兩者合為一，即是資料物件。資料物件的內容可以有很多形式，在絕大多數情形下，資料物件就是實際被保存的資料；例如，電子期刊文獻。除此之外，資料物件也可以有其他附加的物件，例如後設資料記錄。然而在 OAIS 參考模式中，即使後設資料記錄是採取內嵌式 (Embedded) 放在物件之內，這兩者仍被視為不同的兩個資料物件。

(二)知識庫

資料物件經過解譯為有意義的資訊，必須透過每一個使用者的知識庫與資料物件的展現資訊才能達成。每一個使用者或是每一類使用者皆有所屬知識庫，被用來了解與解譯資料。例如，一個說英文的個人就擁有知識庫，用以閱讀、傳達與解譯英文的資料。同樣地，一位爪哇 (Java) 程式員就擁有了了解爪哇程式資訊的知識庫。

(三)展現資訊

有時候知識庫並不足以解譯儲存的資料物件，必須借助於展現資訊，以充分解譯資料物件的內涵。所謂的展現資訊是引導使用者正確解譯、了解與詮釋一個數位物件的內涵。基本而言，展現資訊以位元流方式 (Bit stream, 即 0 與 1 的數位符號) 傳輸這些資訊，而這些位元資訊可能是日期或是其他資訊，因情境而異。此外，展現資訊具備兩種形式：結構化資訊 (Structural information) 與語意化資訊 (Semantic information)。結構化資訊用以將這些位元組織成不同類型 (Types)、群組 (Groups) 與其他更高層次的意義 (High-level meanings)，當然也可以包括資料格式的規格資訊及電腦的軟硬

體環境。至於語意化資訊則是提供額外的意義，以確認結構化資訊。例如，結構化資訊以 ASCII 格式呈現位元資訊，而語意化資訊則是指定資訊內容為英文語別。

(四)資訊物件

所謂的資訊物件是由資料物件與展現資訊共同組成，資訊物件包括四類群組：內容資訊 (Content Information, CI)、保存資訊 (Preservation Description Information, PDI)、封包資訊 (Packaging Information, PI)、描述資訊 (Description Information, DI)。

(五)資訊封包

由內容資訊、保存資訊、封包資訊與描述資訊所共同組成，並可依不同目的區分為投遞 (Submission Information Package, SIP)、儲存 (Archive Information Package, AIP) 與傳遞 (Dissemination Information Package, DIP) 等三種，分別達成「由資訊產生者將生產的資訊傳送至儲存處」、「儲存元件將傳送的資訊予以保存、典藏」與「將資訊傳送給使用者」等不同目的。至於四大類型資訊封包的元件分述如下 (CCSDS, 2002, pp.4-23-4-35)：

1. 內容資訊 (CI)

即資料物件的主旨內容。

2. 保存資訊 (PDI)

用以保存、典藏內容的相關資訊，包括了參考資訊 (Reference information, 即辨識碼，例如 ISBN、URN 等)、來源資訊 (Provenance information, 用以記載內容資訊的歷史，例如起源、保存歷史、保存方法與效果等)、情境資訊 (Context information, 用以標示內容的關聯資訊，例如生產的原因與其他內容的關係)，以及資訊的固著性 (Fixity, 用以註明資訊的真實性，



例如資訊檢查碼、數位簽章等)。

3. 封包資訊 (PI)

就是組成封包的內容單位及相關資訊,如前述的 SIP、AIP 與 DIP 等皆是。

4. 描述資訊 (DI)

有利於使用者查詢、檢索的資訊,亦即描述性的目錄資訊。

由上述 OAIS 參考模式的內涵說明中,可以明顯發現此一模式已為知識組織奠定幾項新的里程碑,如下:

- (一)對於資料物件區分為實體與數位物件兩種,與前述 FRBR 模式有相同的見解與發現。
- (二)擴大知識組織的面向,除了資料或檔案本身外,尚包括了生產者、管理者與使用者等三大面向。
- (三)除了著重在描述性後設資訊外,亦開拓出永久保存/典藏與資訊封包/粒度 (Package or granularity) 的後設資料資訊。
- (四)也特別注重在關聯性資訊,與前述 FRBR 模式有相同的著力焦點。
- (五)除此之外,OAIS 參考模式已深深地影響全球永久保存/典藏後設資料的發展,包括 OCLC、RLG、澳洲國家圖書館 (National Library of Australia, NLA), 以及 EC 贊助的「典範數位檔案計畫」(CURL Exemplars in Digital ArchiveS project, CEDARS) 與「網路化歐洲寄存圖書館計畫」(Networked European Deposit Library, NEDLIB) 等。另一方面,後設資料標準的既有範圍也從全球公認的描述 (Descriptive)、管理 (Administrative) 與結構化 (Structural) 等三大層面,新增了永久保存/典藏 (Preservation and archiving) 的層面。

五、應用特徵檔

此一概念主要源自於英國 UKOLN 執行「DESIRE Project」計畫時所萌芽、發展的一個觀念,並於 2000 年正式提出。所謂的應用特徵檔是一組綱要 (Schemas),為了因應某一種特定目的,結合兩種以上的名稱領域 (Namespace, 即綱要),以符合某一情境需求。就圖書館界而言,其實對應用特徵檔並不陌生。在圖書館實施自動化系統時,各圖書館在應用機讀編目格式 (MACHINE Readable Catalog, MARC) 過程中,圖書館人員會發現有些欄位並無法適用於 MARC,因而在 MARC 中規訂了一個彈性空間—自由使用段 (Local tag),如 XX9 (Heery and Patel, 2000) 與 X9X。因為應用範圍從一個單位擴展至網際網路的範圍時,此種需求日益增多,也愈趨複雜,例如資訊交換。同樣地,在發展 Z39.50 檢索協定後,圖書館界除開始尋求改善之道外,也思索如何讓各使用單位可以因應本身情形而有所客製化 (Customization),而朝向此一概念發展,其中最著名者分別是由英美發展的特徵檔案:「The Bath Profile」、「Z Texas Profile」。在應用特徵檔內容中,具備下列四項主要特質:

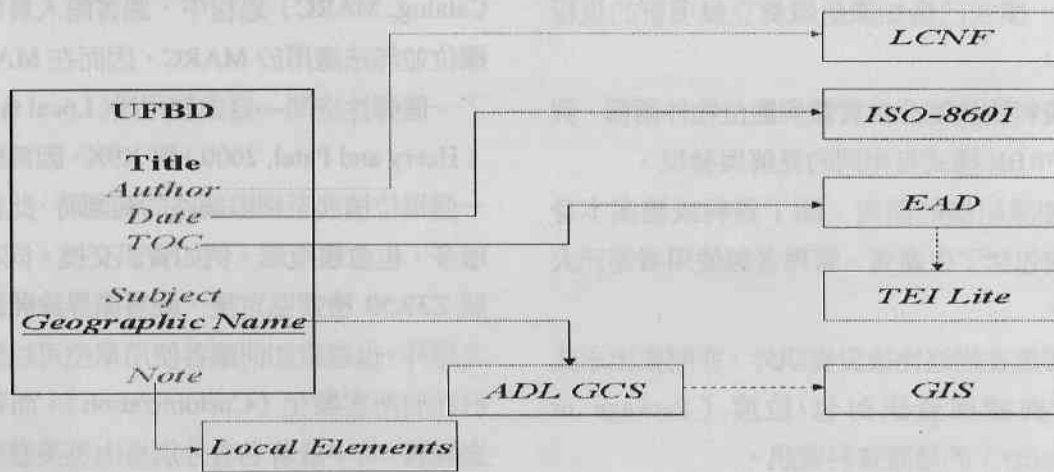
- (一)可以引用兩種以上的名稱領域或綱要。
- (二)不需額外引介新的資料元素 (Data element),各單位可自行發展新的元素或元素集 (Sets),但必須負責對外公開宣告與維護,以利其他單位的使用。
- (三)可以指定某些元素只能使用特定的編碼與內容值 (Schemes and values),如控制詞彙。
- (四)以既有的元素為基礎,加以精緻化 (Refine) (Heery and Patel, 2000)。

以都柏林核心集 (Dublin Core, DC) 為例,若要著錄地名方面的元素,只能利用「Coverage - Spatial」的元素加以描述。另外一種方式即是運用應用特徵檔觀念的方式導入,加以利用現有的地名的



相關標準，例如美國亞歷卓 (Alexandria Digital Library, ADL) 地名標準 (Gazetteer Content Standard, GCS)、亞提信託基金會 (J. Paul Getty Trust) 的地名索引典 (Thesaurus of Geographic Names, TGN)，或是美國國會圖書館 (Library of Congress, LC) 發展的地名權威標目 (Geographic subject headings of authority record format)。換言之，在應用特徵檔的概念架構下，可以使用兩種以上的後設資料標準 (請參見圖十)，而此種情形也日趨增多，

可從各國數位圖書館計畫獲致驗證。例如，美國加州線上檔案 (Online Archive of California, OAC) 計畫，就同時應用了機讀編目格式與檔案描述編碼格式 (Encoded Archival Description, EAD)；美國藝術影像聯盟 (Art Museum Image Consortium, AMICO) 則是平行應用了 DC 與藝術品描述類目 (Categories for the Description of Works of Art, CDWA) 兩套標準。



圖十：應用特徵檔

如果仔細探究應用特徵檔的內涵，其實還深受「華威架構 (Warwick Framework) 中的封包 (Package)」(Lagoze, 1996)、「資源描述架構」(Resource Description Framework, RDF) 及「可延伸標誌語」(eXtensible Markup Language, XML) 中的名稱領域等觀念的影響。

六、ABC 本體論模式

此一概念乃是由英、美、澳洲等國跨國合作，國際計畫名為「和諧計畫」。此計畫有鑑於許多後設資料標準缺乏模式原則 (Modeling principles)，以及只著墨於靜止不動的物件與單一時間點，而忽略了對時間軸及其演變的處理。此一概念預計達成

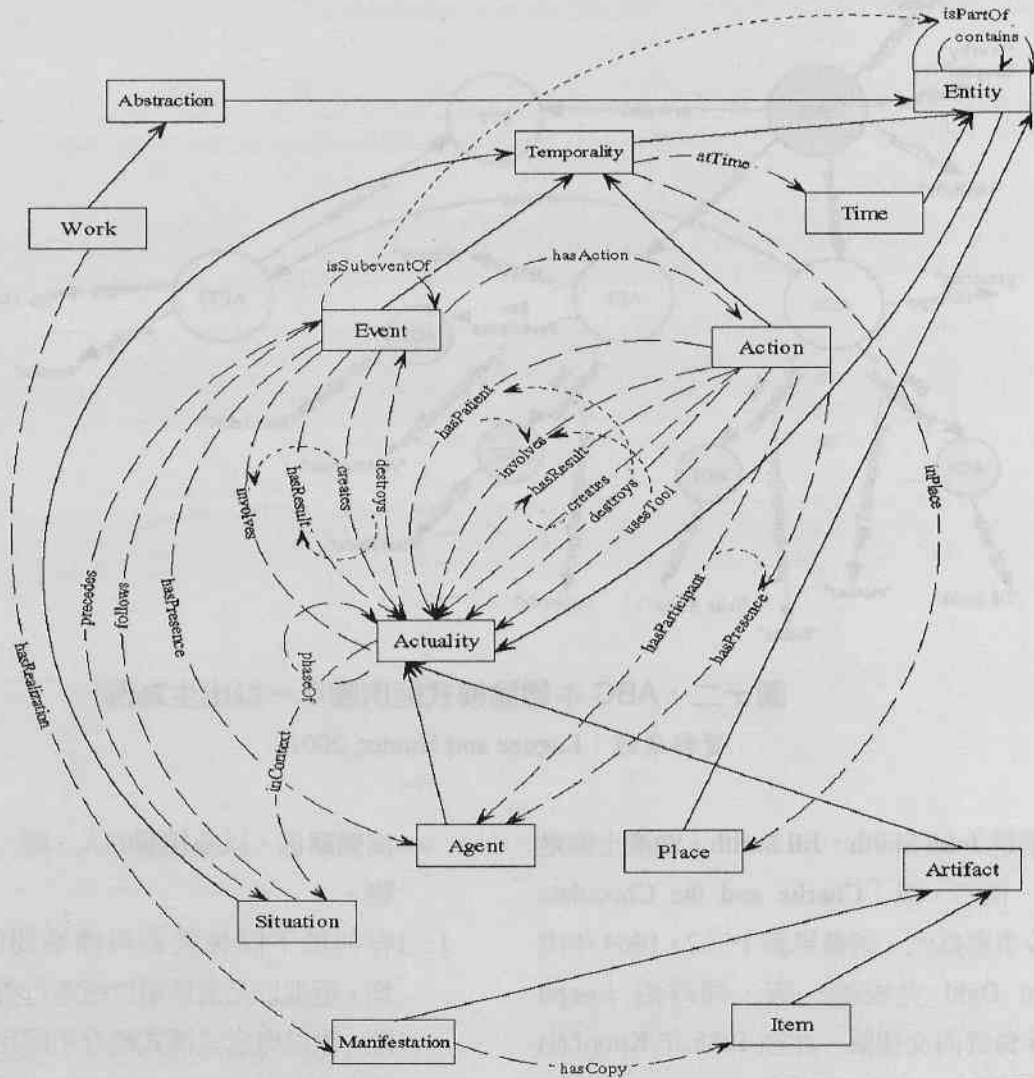
三大目標，如下：

- (一) 提供一個概念基礎，以利於了解與分析既有後設資料的本體與實例 (Ontologies and instances)。
- (二) 提供指引，以利檢視與發展描述性本體論。
- (三) 發展一個概念基礎，以利不同後設資料本體間的自動對照 (Lagoze and Hunter, 2001)。

ABC 本體論模式主要由實體、類別 (Class)、屬性 (Property) 三者共同組成 (請參見表一)，以建立起 ABC 本體論模式的內涵 (請參見圖十一)，要點如下：

- (一) 時間實體 (Temporality)：包括了情境 (Situation)、事件 (Event) 與動作 (Action)，





圖十一：ABC 本體論模式內涵示意圖

資料來源：Lagoze and Hunter, 2001

分別用來標引「一般的時間情境與特定的時間情境」、「不同情境間的轉移」與「事件相關人 (Agent) 的責任與參與」。

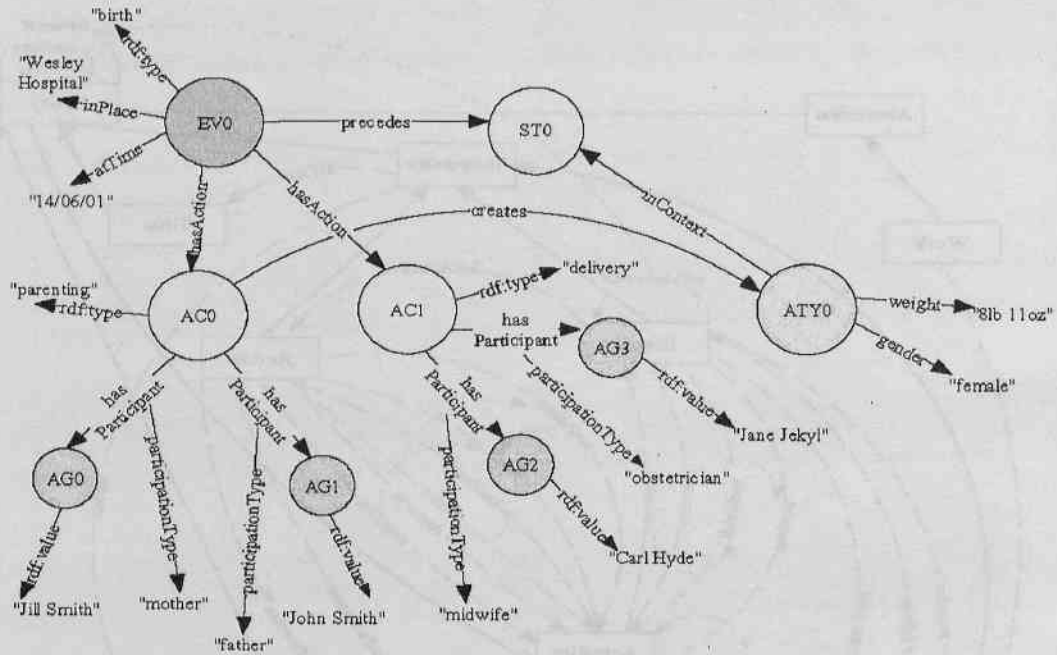
(二)具象實體 (Actuality)：泛指可以察覺到的實體，可以被聽、看、聞與摸到等；如何林頓 (Bill Clinton) 特定性具像實體是：美國總統。

(三)抽象實體 (Abstraction)：泛指一個觀念或概念，通常具備下列兩項特徵，亦即「抽象實體不能有情境」，以及「抽象實體必須經由具象

實體的「has Realization」來展現」(Lagoze and Hunter, 2001)。

除了上述三個實體的基礎，ABC 本體論模式又由許多類別與屬性組成 (請參見表二)，來形成以時間為主軸的各式關聯性，以完整描述各式實體間的關係。至於 ABC 本體論模式內涵如何應用至實際的後設資料標準與知識組織，就從下列範例得知一二。例一，以名為 Wesley Hospital 女孩的出生為例 (請參見圖十二)，出生時 81 磅 11 盎司，





圖十二：ABC 本體論模式範例圖示一以出生為例

資料來源：Lagoze and Hunter, 2001

父母親分別是 John Smith、Jill Smith，助產士則是 Carl Hyde。例二，以「Charlie and the Chocolate Factory」故事書為例（請參見圖十三），1964 年由作者 Roald Dahl 完成第一版，同時由 Joseph Shindleman 負責內文插圖，並在 1985 年 Knopf 出版商出版精裝書。之後，第二版在 1998 年由 Puffin 出版商出版為平裝書，改由 Quentin Blake 負責插圖工作。至 1995 年，出版商 Caedmon 將此書製作為三個小時的錄音帶，由 Robert Powell 負責口述，Sam'n Ella's Catering 負責行銷（Caterer）（Lagoze and Hunter, 2001）。

鑑於時間軸演化的重要性與關聯性，ABC 本體論模式已達成下列具體成果，如下：

(一)突破以單一時間點為主的後設資料典型觀念，以時間的連續性與關聯性為主軸，進行事件、物件、時間與人物的個別描述。

(二)以時間為主軸，精確描述事件在時間先後的

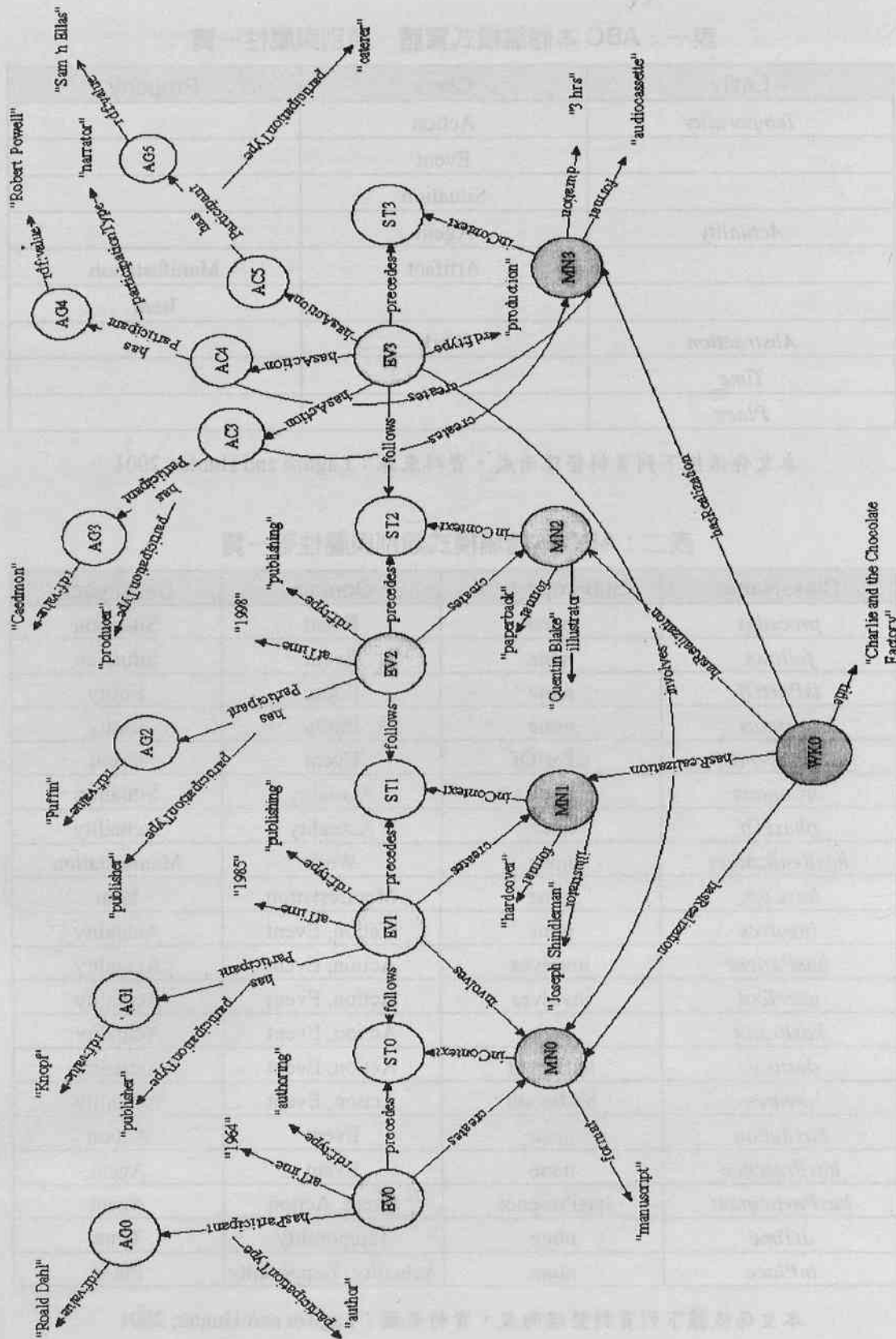
演變關係，以及相關的人、時、物等屬性關聯。

(三)有利於不同後設資料標準間的對照與交換，超越以元素為單位所進行資訊轉換的範圍，改以概念式模式建立不同元素與實體間的關聯，達到更精準的資訊共享目的。

七、後設資料資訊交互作用模式

就學術傳播系統（Scholarly communication system）而言，基本上一份出版品的產生是經過一定的程序與流程，包括作品初稿（Preprint）、會議論文、投稿至出版商正式成為出版品、圖書館等典藏單位購入蒐藏、索引摘要服務、讀者取閱資料與建議回饋等。從檔案界而言，文獻資料進入檔案界前後也有一定的流程，包括徵集、鑑定、登記、描述、管理、處置、移轉與銷毀等，檔案界稱之為「檔案的生命週期」。傳統上，圖書館界是針對已成為





圖十三：ABC 本體論模式範例圖示——以出版品為例

資料來源：Lagoze and Hunter, 2001



表一：ABC 本體論模式實體、類別與屬性一覽

Entity	Class	Property
<i>Temporality</i>	Action	
	Event	
	Situation	
<i>Actuality</i>	Agent	
	Artifact	Manifestation
		Item
<i>Abstraction</i>	Work	
<i>Time</i>		
<i>Place</i>		

本文係依據下列資料整理而成，資料來源：Lagoze and Hunter, 2001

表二：ABC 本體論模式類別與屬性等一覽

ClassName	Subproperty of	Domain	Description
<i>precedes</i>	none	Event	Situation
<i>follows</i>	none	Event	Situation
<i>isPartOf</i>	none	Entity	Entity
<i>contains</i>	none	Entity	Entity
<i>isSubEventOf</i>	isPartOf	Event	Event
<i>inContext</i>	none	Actuality	Situation
<i>phaseOf</i>	none	Actuality	Actuality
<i>hasRealization</i>	none	Work	Manifestation
<i>hasCopy</i>	none	Manifestation	Item
<i>involves</i>	none	Action, Event	Actuality
<i>hasPatient</i>	involves	Action, Event	Actuality
<i>usesTool</i>	involves	Action, Event	Actuality
<i>hasResult</i>	none	Action, Event	Actuality
<i>destroys</i>	hasPatient	Action, Event	Actuality
<i>creates</i>	hasResult	Action, Event	Actuality
<i>hasAction</i>	none	Event	Action
<i>hasPresence</i>	none	Event	Agent
<i>hasParticipant</i>	hasPresence	Event, Action	Agent
<i>atTime</i>	none	Temporality	Time
<i>inPlace</i>	none	Actuality, Temporality	Place

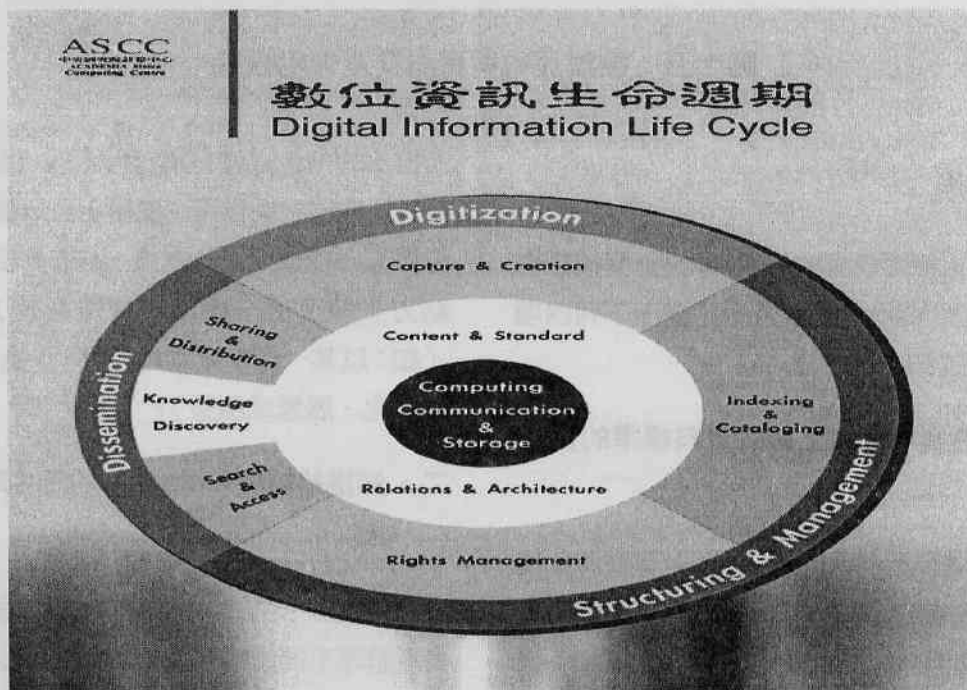
本文係依據下列資料整理而成，資料來源：Lagoze and Hunter, 2001



正式出版品的資訊進行知識組織，也就是分類編目作業與目錄 (Catalogues)；另一方面，檔案界所謂的「目錄指引」(Finding aids)，也是側重在檔案資料的描述，亦即前述檔案生命週期的描述階段。

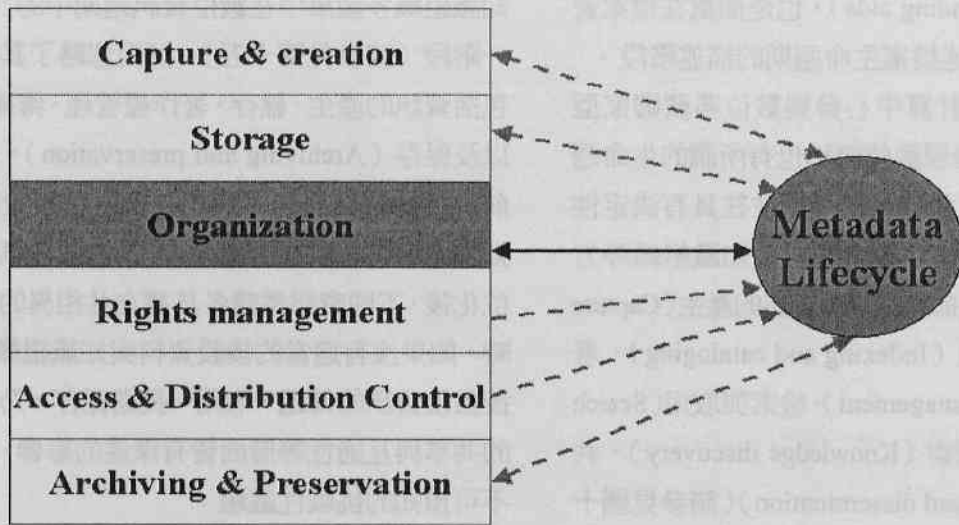
在中央研究院計算中心參與數位典藏國家型科技計畫時，明顯發現數位資訊也有所謂的生命週期 (Lifecycle)，而此一生命週期往往具有決定性的影響力，乃至擴及到後設資料與知識組織等方面。所謂數位資訊生命週期係由資訊的產生 (Capture and creation)、組織 (Indexing and cataloging)、著作權管理 (Rights management)、檢索與取用 (Search and access)、知識探索 (Knowledge discovery)、共享與傳遞 (Sharing and dissemination) (請參見圖十

四)。從中央研究院計算中心 MAAT 支援數位典藏國家型科技計畫各項計畫中，發現許多後設資料與知識組織多數集中在數位資訊週期中的「組織」此一階段 (請參見圖十五)，反而忽略了其他階段，包括資訊的產生、儲存、著作權管理、傳播與取用，以及保存 (Archiving and preservation)。事實上，前述這些數位資訊週期的各個階段與後設資訊及知識組織有著密切的關係，尤其是傳統媒體經過數位化後，不同資訊載體各具獨立且相異的生命週期時，如果沒有適當的後設資料與知識組織，對於日後數位資訊的傳遞、取用、長期保存，乃至於資訊的共享與互通性等層面皆有深遠的影響，也會產生不可預知的挑戰性議題。



圖十四：數位資訊生命週期





圖十五：後設資料與資訊交互作用模式

肆、研究發現

從上述七大概念模式的探討,本文針對後設資料標準的應用與知識組織等方面,提出下列四項研究發現,闡述如下:

一、可以同時有兩個以上後設資料標準的實際應用

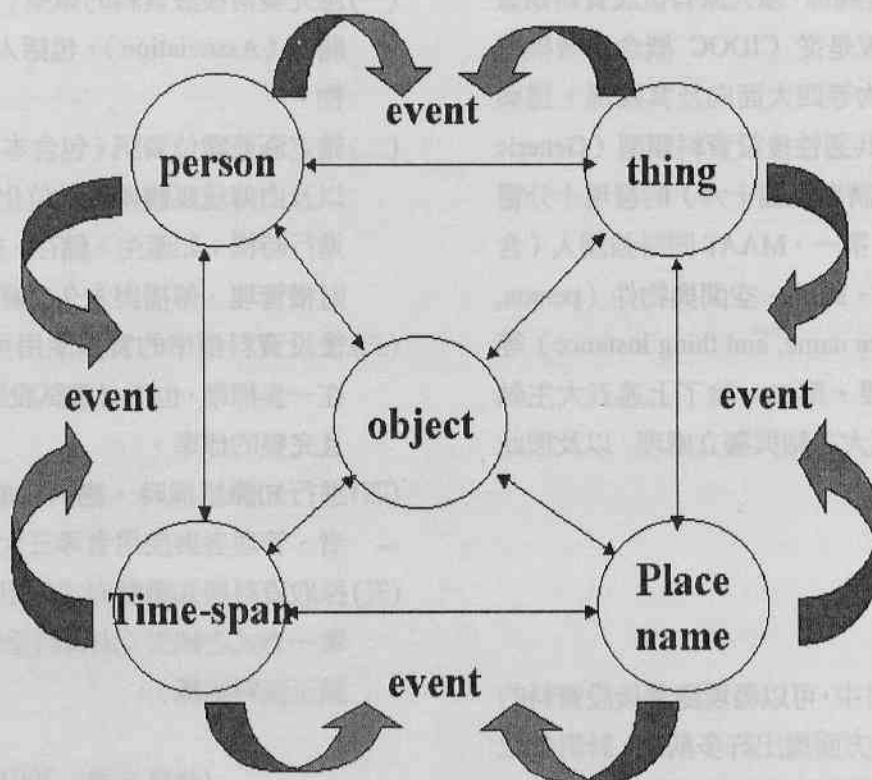
在數位圖書館計畫領域中,各項計畫在應用後設資料標準於知識組織與描述時,常有下列兩項疑問:是否只能應用一套後設資料標準?如果某一套後設資料標準無法完全符合計畫需求時,有無其他的可行策略?由應用特徵檔的觀念與 MAAT 實際應用經驗個案中,可以證明同一個數位圖書館計畫或系統可以同時使用兩種以上的後設資料標準,但是標準之間必須具備兩大顯著特性:第一,不能將不同標準混雜為一個單一標準。第二,不同標準間

所扮演的角色定位必須有所不同,呈現一種相互支援、互補的功能作用,使用上可以直接採用(如:借用某一標準元素的描述,或是直接引用,並以鏈結方式連至另一標準的實際系統)或間接採用(如:以某一標準元素為基礎,以進行元素結構上的簡化、展延或調整)。

二、知識組織與 Metadata 的產生時機及涵蓋範圍

傳統上,無論是圖書館、檔案館、博物館或是美術館等不同領域,本文發現即使邁入數位圖書館之際,後設資料仍然被定位在知識組織的階段中,並不涉及其他資料及其內容的原始起點處與完成終點處。在資訊數位化之際,如何跳離既有的窠臼,從一個更全面性的數位資訊生命週期與後設資料資訊交互作用等不同觀點切入,求取更完整、深入的知識組織層面,是一項十分值得深思的研究議





圖十六：共通性後設資料類型

題。此外，上述五大概念模式（除應用特徵檔與 OAIS 參考模式外）皆從知識的組織觀點去探究原本知識的結構，以及知識與資料間的關聯，這也建議從事知識組織的各專業領域（如圖書館、檔案館、博物館與美術館等），除了致力於資料至知識的發展方向外，也應該同時採取從知識到資料的方式，以補強現有方式的缺憾。

三、永久保存與典藏方面的後設資料日形重要

鑑於資訊數位化，許多研究計畫除了留意傳統媒體資訊的保存外，也已經開始注意到數位資訊的保存問題，包括電腦軟硬體、平臺與作業系統、數位儲存媒體，乃至於數位資訊的展現方式等，在在顯示數位資訊的保存有別於以往的媒體資訊，尤其

隨著資訊科技的日新月異，永久保存與典藏的後設資料標準也益發重要，且不可或缺，而 CCSDS 發展出的 OAIS 參考模式更值得從事知識組織與後設資料標準應用者再三思索。

四、後設資料的類型 (Metadata attributes) 與關聯性 (Association)

環顧全球現有後設資料的發展情形而論，多數後設資料標準皆集中在物件 (Thing instance)，如各式的檔案、圖書與期刊、藝術品、標本、物種等。從 OAIS 參考模式中，CCSDS 建議應從不同人的角度思索後設資料的層面，包括管理者、使用者與生產者等三方面。再由記錄連貫性模式與 ABC 本體論模式中，又可以發現以時間為主軸衍生的時



間、人、事件 (time, person, and event instance) 與物件等四大面向及其關係，擴大原有後設資料涵蓋的面向與關聯性；或是從 CIDOC 概念參考模式中，人、時、地、物等四大面向及其關係。這與 MAAT 發展出來的共通性後設資料類型 (Generic metadata attributes, 請參見圖十六) 的發現十分雷同，不同之處有二：第一，MAAT 同時強調人 (含家族與機構)、事件、時間、空間與物件 (person, event, time-span, place name, and thing instance) 等五大主軸的平等重要。第二，除了上述五大主軸外，MAAT 也著重五大主軸與獨立處理，以及彼此間的互動關聯關係。

伍、結語與建議

從本文前述探討中，可以發現隨著後設資料的興起，已為知識組織方面提出許多創見。針對數位圖書館計畫在設計、應用後設資料或進行知識組織

前，本文提出下列建議：

- (一) 應先釐清後設資料的類型 (Attributes) 及其關係 (Association)，包括人、事、時、地、物。
- (二) 確定涵蓋數位資訊 (包含本身即為數位化，以及由傳統媒體轉為數位化兩種) 的範圍與進行時機；如產生、儲存、組織、取用與智財權管理、傳播與永久典藏等。
- (三) 後設資料標準的實際應用可以不用只局限在一套標準，也不必嘗試設計一套完美無缺且完整的標準。
- (四) 進行知識組織時，應同時兼顧資料的生產者、管理者與使用者等三大層面。
- (五) 採取資料與知識雙向式的組織，以補強墨守單一方式之缺失 (由資料至知識組織、由知識至資料組織)。

(收稿日期：2003 年 12 月 15 日)

備註：本文係依據下列會議論文增修而成，相關資訊如下：

陳亞寧。(民 91 年)。從後設資料標準概念模式探討知識組織應用與發展。在佛教知識管理研討會論文集 (頁 6-1—6-23)，中央圖書館臺灣分館、中國圖書館學會專門圖書館委員會、財團法人伽耶山基金會聯合主辦，香山尼眾佛學院圖書館承辦，91 年 9 月 13-14 日。臺北市：中央圖書館臺灣分館中正廳。

參考書目：

- Amato, G., Castelli, D., Pisani, S., Venerosi, P., Poncin, P., & Vinet, L. (2000). Metadata modeling report. Retrieved 1 April 2001 from <http://pc-erato2.iei.pi.cnr.it/echo/public/deli/D3-1-1%20ECHO%20Metadata%20Modelling.pdf>
- Bearman, D. (1996). Item level control and electronic recordkeeping. Archives and Museum Informatics, 10(3), 195-245. Retrieved 6 August 2002 from <http://www.archimuse.com/papers/nhprc/item-lvl.html>
- Berman, D., Miller, E., Rust, G., Trant, J., & Weibel, S. (1999). A common model to support interoperable metadata. D-Lib Magazine, 5(1). Retrieved 22 January 1999 from <http://dlib.ejournal.ascc.net/january99/bearman/01bearman.html>



- Chen, Y.-N., & Chen, S.-J. (2001). A practical metadata approach towards Chinese rubbings, Paper presented at The Digital Resources for the Humanities, 8-10 July 2001. London: SOAS. Retrieved 8 August 2002 from <http://www.sinica.edu.tw/~metadata/bibliography/bibliography-frame.html>
- Consultative Committee for Space Data Systems. (2002). Reference model for an open archival information system (OAIS). Retrieved 1 August 2002 from <http://www.ccsds.org/documents/pdf/CCSDS-650.0-B-1.pdf>
- Crofts, N. (2000). Introduction to the ICOM/CIDOC conceptual reference model. Retrieved 24 December 2003 from <http://www.niso.org/international/SC4/CIDOC-rf.pdf>
- Crofts, N., Doerr, M., & Gill, T. (2003). The CIDOC conceptual reference model: A standard for communicating cultural contents. Cultivate Interactive, 9, Retrieved 24 February 2003 from <http://www.cultivate-int.org/issue9/chios/>
- Delsey, T. (2002). Functional analysis of the MARC21 bibliographic and holdings formats. Retrieved 28 March 2002 from <http://www.loc.gov/marc/marc-functional-analysis/analysis.pdf>
- Heerry, R., & Patel, M. (2000). Application profiles: Mixing and matching metadata schemas. Ariadne, 25, Retrieved 27 September 2000 from <http://www.ariadne.ac.uk/issue25/app-profiles/>
- Iannella, R. (2001). Digital rights management architectures. D-Lib Magazine, 7(6). Retrieved 2 January 2002 from <http://dlib.ejournal.ascc.net/dlib/june01/iannella/06iannella.html>
- Lagoze, C. (1996). The Warwick framework: A container architecture for diverse sets of metadata. D-Lib Magazine. Retrieved 25 November 1998 from <http://dlib.ejournal.ascc.net/dlib/july96/lagoze/07lagoze.html>
- Lagoze, C., & Hunter, J. (2001). The ABC ontology and model. Journal of Digital Information, 2(2). Retrieved 1 February 2002 from <http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Lagoze/lagoze-final.pdf>
- Lin, S.C., Chen, Y.-N., & Chen, S.-J. (2001). A Metadata Case Study for the FRBR model based on Chinese painting and calligraphy at the National Palace Museum in Taipei. Paper presented at DC-2001: International Conference on Dublin Core and Metadata Applications, 22-26 October 2001. Tokyo, Japan: National Institute of Informatics. Retrieved 24 December 2001 from <http://www.nii.ac.jp/dc2001/proceedings/product/paper-08.pdf>
- OCLC, & RLG Working Group on Preservation Metadata. (2001). Preservation metadata for digital objects: A review of the state of the art. Retrieved 9 June 2002 from http://www.oclc.org/research/pmwg/presmeta_wp.pdf
- IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998). Functional requirements for bibliographic records: Final report. München, Saur. Retrieved 24 December 1998 from <http://www.ifla.org/VII/s13/frbr/frbr.pdf>
- Upward, F. (1996). Structuring the records continuum, part one: Postcustodial principles and properties. Archives and Manuscripts, 24(2). Retrieved 1 August 2002 from <http://www.sims.monash.edu.au/research/rcrg/publications/recordscontinuum/fupp1.html>

