

# 數位人文的跨領域和基礎性研究： 以文物數位化保護元數據標準規範為例

肖瓏

北京大學圖書館研究館員

E-mail: lxiao@lib.pku.edu.cn

張岩

北京大學圖書館副研究館員

E-mail: zhangyan@lib.pku.edu.cn

張俊娥

北京大學圖書館副研究館員

E-mail: zhangje@lib.pku.edu.cn

馮英

北京大學圖書館副研究館員

E-mail: fengy@lib.pku.edu.cn

關鍵詞：數位人文；文物數位化；元數據框架體系；元數據標準

---

## 【摘要】

數位人文是近年來的研究熱點，其目標之一是通過技術的引進，更加深入地描述人文領域的研究物件，挖掘、組織和傳播人類文明及相關知識。本文以文物數位化保護元數據標準研究為例，通過文獻調查與分析、實地調研、建立模型、規範設計，介紹文物數位化保護元數據（Metadata）規範的總體架構和專業化拓展，闡明數位人文發展領域不可忽略的基礎性、長遠性工作——標準規範研究之必要性和作用，探索人文與技術的共同合作與發展。

## 背景

數位人文是運用資訊技術表現人文內涵的新興學科。隨著電腦和網路技術在人文社會科學領域的廣泛應用，人文社會科學研究引入了大量數位技術在人文資料的搜集、處理、存儲

和展示上，人文社會科學與資訊技術科學的交叉融合，人文教學和科研機構與圖書館、檔案館、博物館等人文資料保存和利用機構的共建共用，都使得跨學科研究、跨組織協作趨勢愈加明顯（熊莉君、張福陽、張燦，2016）。正如美國賓夕法尼亞大學的傑佛瑞·納普（J. Knapp）教授所指出，隨著知識的劇增和學術問題的複雜性，使得閱讀和研究交叉學科及複雜問題的人文學者，越來越需要技術和資料學者來說明他們發現、關聯、轉換、重組不同領域的多形態的知識（Knapp, 2012），技術和資料學者也越來越重視運用資料採擷和視覺化等資料分析方法 and 技術開發特色服務（熊莉君、張福陽、張燦，2016）。

在這種交叉與融合中，人文學者始終堅持認為，數位人文方法是人文社會科學研究的一種補充手段，可以給傳統人文研究提供全新的研究視角、工具和平臺，實現人文研究範式變革，但不能代替人文研究和人文精神。在這種觀點下，人文社會科學基礎研究領域的跨學科合作比較容易為人文學者和技術學者接受，並成為促進人文與技術融合的重要出發點。究其原因，主要是數位人文研究在很大程度上依賴於數位內容，而資料又是其中的基礎和核心。技術和資料學者作為專業的資料處理專家，在文本數位化、資料管理、資料庫建設、文本挖掘、知識組織、資料保存和元數據標準等方面有自己獨特的作用，從發現資訊／知識到利用已發現的資訊／知識來創建新的學術資源，為數位人文使用者提供了高品質的知識服務轉化。而另一方面，對於技術和資料學者來說，跨學科資料的管理活動為其成功地介入到數位人文研究，與人文領域的學者建立密切的合作夥伴關係提供了獨特的機遇（Vandegrift & Varner, 2016）。

數位人文的合作不僅應用於傳統的人文領域，還應用於數位圖書館、數位博物館、數位檔案館建設。目前國際上圖書館、博物館、檔案館「三館合一」的趨勢即體現了數位人文的這種發展，即圖書館、博物館、檔案館進行越來越多的數位化整合；三館可以統稱為「記憶機構」（memory institution）；其數位物件多稱為「文化遺產」（cultural heritage）或者「文化遺產物件」（cultural heritage objects）；三館往往在一起共同開展數位化專案，或者建立統一的數位化入口。比較典型的如「美國記憶」（American Memory）項目、歐洲 Europeana 文化遺產項目等，在這兩個平臺上，有數千所文化機構進行合作，為其提供資料和研究資料。

而隨著數位圖書館的發展，圖書館已經在資料存儲、組織、文本挖掘和元數據標準等資料管理方面積累了豐富的經驗，有能力在圖書館、博物館、檔案館合作的數位人文研究過程中發揮重要的作用（ACRL Research Planning and Review Committee, 2014）。由北京大學圖書館帶領的國家科技支撐計畫「文物數位化保護元數據標準研究」課題（課題編號：2014BAK07B02）就是這方面合作的代表。該課題是國家科技部「國家科技支撐計畫」支援、國家文物局組織的「文物數位化保護標準體系及關鍵標準研究與示範專案」的核心內容，其目標是針對當前文物數位化過程中元數據體系和規範方面的需要，建設文物數位化保護元數據總體框架體系、核心標準、描述標準、管理保存標準及應用技術規範，以及與文物元數據規範語義相關的文

物分類體系等，從而規範文物數位化保護元數據建設，支援和促進文物數位化保護建設、推動文物保護的研究、展示、應用與發展。

「文物數位化保護元數據標準研究」課題的任務包括以下方面：(1) 文物元數據標準總體框架，在面向數位化保護的標準規範體系和文物數位化生命週期模型的基礎上，建立交互操作的、可擴展的文物數位化元數據規範體系與總體框架，包括文物數位化標準規範與元數據、文物數位化物件生命週期與元數據體系、文物數位化保護元數據規範類型及其關係、文物元數據資訊框架。(2) 描述元數據研究，即對文物內容、屬性、外在特徵進行描述的元數據框架體系及相關應用規則，為可移動／不可移動文物元數據規範的設計奠定基礎，制定文物核心元素集及其應用指南、描述元數據應用規範與專門元數據設計規範。(3) 可移動文物專門元數據標準的研製，包括銅器、古籍、甲骨、輿圖、繪畫、雕塑、織繡、傢俱、拓片、陶器、瓷器、玉器等十二大類。(4) 不可移動文物專門元數據標準的研製，包括古文化遺址、古墓葬、古建築、石窟寺、石刻、壁畫、近代現代重要史跡和代表性建築等。(5) 文物數位化保護管理元數據規範與保存元數據規範的研製，支援對數字資源建設各環節的管理需求，包含對技術資訊、智慧財產權、版權保護、許可權控制、來源、數位化過程、服務等與數位資源管理密切相關的內容；研製文物數位化後長期保存元數據規範及相關應用指南。(6) 文物數位化保護元數據標準應用規範的研究，重點包括與元數據輸入、處理、存貯、管理、發佈、檢索、交換等緊密相關的文物數位化保護元數據標識體系、元數據編碼規則、元數據封裝／交換規範、元數據訪問協定與元數據開放機制，為文物數位化保護元數據管理系統的實現以及互聯互通提供標準依據。(7) 文物分類體系研究，對數位化後的文物分類體系框架和內容進行研製。

根據以上內容，從數位人文研究的角度來看，「文物數位化保護元數據標準研究」課題研究及其成果具有以下特點：

1. 符合數位人文的跨學科特點。涉及文物學、博物館學、圖書館學、資訊技術等多個學科，既要利用資訊技術領域的資料科學、資料管理來解決文物數位化過程中的資料規範問題，也要根據文博學科和文博領域的特點，對文物屬性、特徵、應用進行深度研究和表達。
2. 具有數位人文的跨領域合作特點。課題由北京大學圖書館牽頭，北京大學圖書館及考古文博學院、清華大學圖書館、國家圖書館現代技術研究所等元數據研究機構，與故宮博物院、敦煌研究院、中國科技大學、浙江大學等文物機構和文物研究機構共同合作。課題組根據任務組成了六個課題小組，每個小組都由文博學界和資料研究兩方面學者組成，系統性地開展研究。此外，還有約 120 餘位文博和資訊技術領域專家參與了資料規範成果的評審和建議，十餘個文博機構承擔了開放實驗的資料工作。

3. 屬於數位人文領域的基礎性研究工作。數位人文領域的研究是分為多層次的，有資訊收集、處理的基礎性研究，也有用於展示、表達的門戶交互性研究。而無論是數位圖書館還是數位博物館，資料都是最基本和最微粒的組成，因此資料規範的制定屬於數位人文領域的基礎性工作，規範化資料的建設對於文物數位化領域不可或缺。

本文將主要介紹該課題的核心研究內容。

## 文物數位化保護元數據基礎研究

數位人文的學科研究範圍非常廣泛，涵蓋了文學、歷史、考古、藝術、哲學等，文物數位化保護元數據標準規範的研究物件也是包羅萬象，包括建築、繪畫、雕塑、壁畫、手稿、傢俱、服裝、照片、數位文檔、電影、音樂等，同樣涵蓋建築學、文學、文獻學、歷史學、考古學以及藝術學等。因此，這種跨學科的特點使得數位人文研究中的許多方法都可以為文物數位化保護元數據建設所借鑒。

文物數位化保護元數據標準規範體系用來支援文物資訊化系統中資訊資源的產生、加工、組織、發佈、存儲與管理，可以滿足對文物實體物件和數位物件進行描述、組織、管理、服務與保存的需要。在面向數位化保護的標準規範體系和文物數位化生命週期模型的基礎上，根據元數據的結構化屬性、關係特徵以及元數據規範的作用，建立交互操作的、可靈活擴展的文物數位化元數據規範體系與總體框架，包括：(1) 文物數位化標準規範整體結構及其與元數據的關係；(2) 文物數位化物件生命週期與元數據體系的建立；(3) 文物元數據資訊框架；(4) 文物元數據結構框架等。

### 文物數位化保護標準規範體系與元數據

文物數位化保護標準規範體系包括標準體系框架、核心標準規範體系、其他支撐標準和應用示範四部分（見圖 1），文物數位化保護元數據標準體系是其核心內容的重要組成部分。受標準規範管理系統、標準體系管理規範、標準體系擴展規則、標準體系表和文物數位化物件生命週期模型指導，同時與文物數位化技術體系共同構成了核心標準規範體系，在文物分類體系、主題詞表、標準體系編碼規則支援的基礎上，通過文物元數據標準應用示範、可移動文物／不可移動文物數位化保護關鍵標準應用示範進行驗證和規範。

課題1研究內容
  課題2研究內容
  課題3研究內容
  課題4研究內容  
 課題1、2、3、4共同的研究內容
  課題3、4共同的研究內容

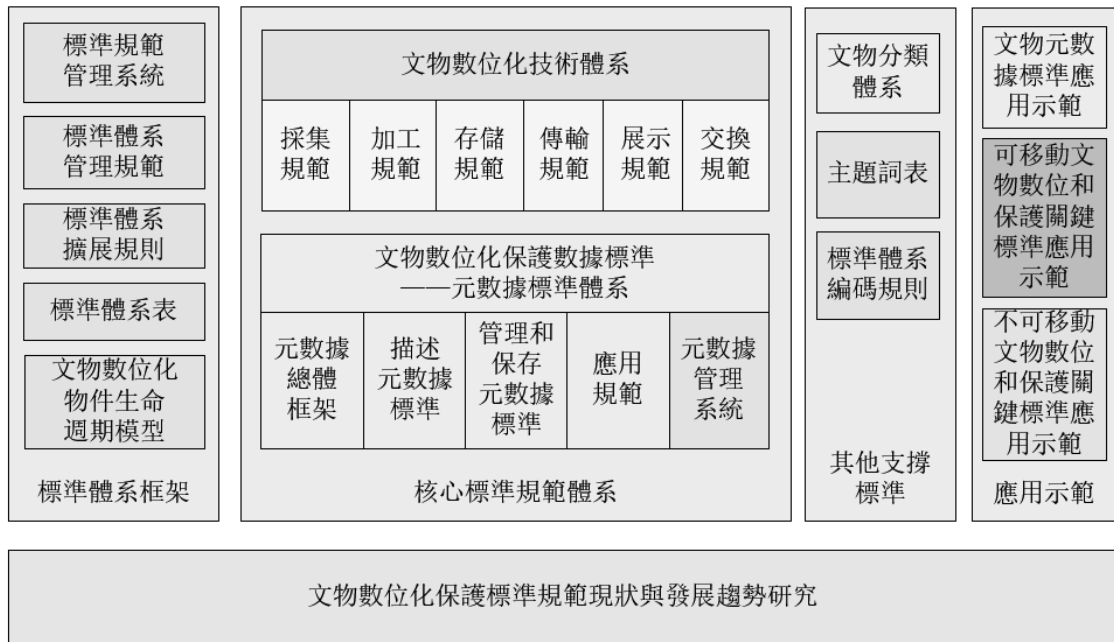


圖 1 文物數位化標準規範體系

如圖 1 所示：文物數位化保護元數據標準體系包括描述元數據標準、管理元數據標準、保存元數據標準、應用規範以及元數據管理系統。該標準體系針對當前文物數位化保護在元數據體系和規範方面的需要，充分借鑒當前元數據標準制定的內容和技術，把握文物數位化保護元數據的特點和性質。從文物數位化保護的應用邏輯出發，建設能夠支撐文物全方位與全過程採集、存儲、管理、展示、共用與應用文物數位化資訊的元數據總體框架與標準體系，為文物數位化保護的整合、共用、有機組織奠定基礎。

### 文物數位化物件生命週期與元數據體系

文物元數據體系與文物數位化物件生命週期息息相關，從文物實體數位化過程的採集、加工整理、服務、保存到處置（主要指資源的剔除／銷毀）的每個階段中，都離不開元數據規範，不同類型的元數據在文物數位化物件生命週期的不同階段產生並在不同環節起著重要作用（肖瓏、申曉娟，2011）。從文物實體被採集加工成為文物數位化物件到進入資訊服務仲介後的加工、服務、保存、處置的生命週期，以及週期中所涉及的事件與文物數位化建設標準規範之間的關係，可建立如圖 2 顯示的文物元數據體系。

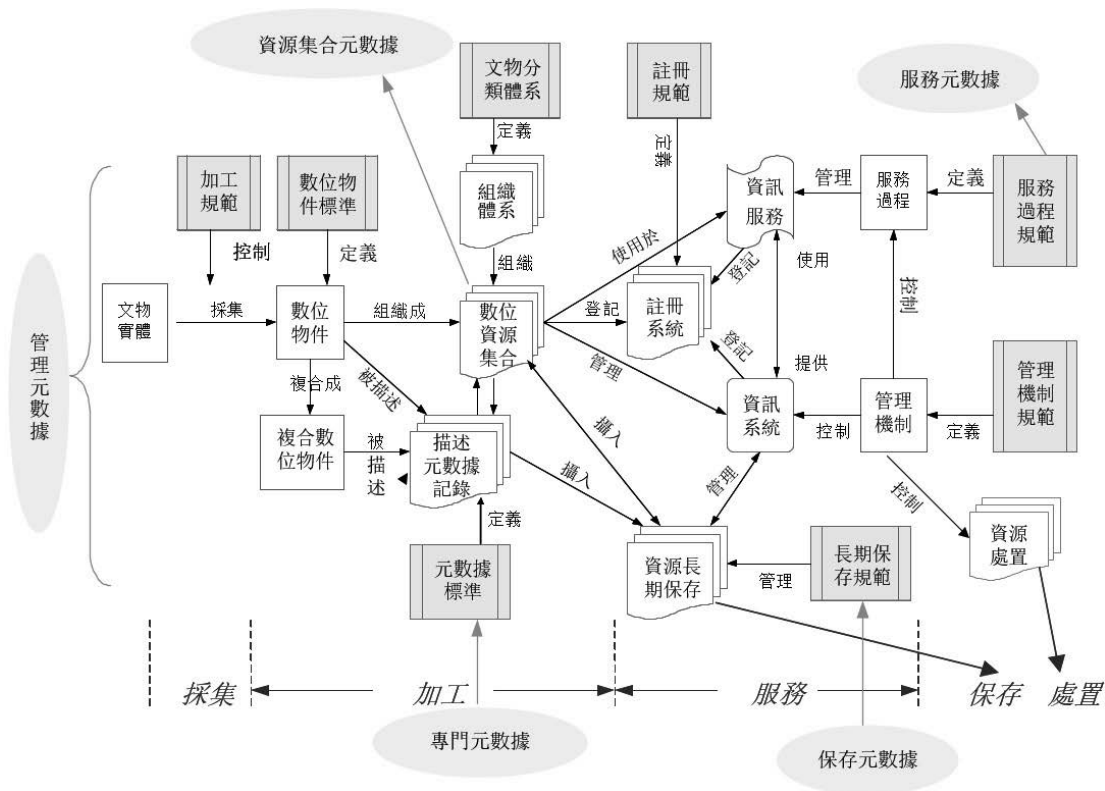


圖 2 文物數位化物件生命週期與元數據體系

資料來源：改編自張曉林（2004 年）。數字圖書館標準與規範建設。

在圖 2 中，從文物實體被採集加工成為數位化對象到被使用者利用，在整個生命週期內經歷了採集、加工、服務、保存、處置等基礎環節，在每個環節都受到管理過程的控制，每個環節都有其相應的管理機制與管理政策。例如在採集加工階段，對文物的類型、數位物件標準、加工規範進行管理；在提供服務階段，對註冊規範、資訊系統、服務過程等設定管理機制；在資源保存、處置階段，使用管理元數據管理資源的長期保存和資源處理方式。每個環節都有特定的管理元數據，它既反映了這一管理環節管理機制的共性，又能根據更加具體的管理功能進行擴展，各個環節的管理元數據是一個相互融合的有機整體。

根據圖 2 文物數位化物件週期中元數據的作用，可以將文物元數據分為以下類型：

1. 專門元數據、資源集合元數據等組成的描述元數據，主要作用於文物實體數位化的採集加工階段，文物實體被採集加工成為數位物件（複合數位物件）的過程，以及對數位物件（複合數位物件）進行加工、描述、檢索，需要由描述元數據來完成。描述元數據在描述不同數位化物件的同一性質時用到的核心元素集，以及描述不同物件時的個性化的專門元數據都需要建立相應的標準規範和應用指南。

2. 服務元數據：主要存在於數位物件提供服務的階段，作用是將採集加工階段完成的文物數位化物件通過描述元數據規範後，再使用服務元數據展示出來，以供讀者檢索、使用，並且實現各資訊資源之間的交換、共用、交互操作。
3. 保存元數據：主要存在於數位物件的保存和處置階段，將描述元數據、管理元數據、服務元數據提供的各種數位化資訊、檢索使用、版本更新、剔除／銷毀等過程進行保存。
4. 管理元數據：作用於整個生命週期，是專門對文物數位資源實施管理的元數據。它描述在管理文物實體物件數位化的過程中所涉及的管理資訊以及對管理政策與管理機制的規範、開放描述，真實記錄業務活動資訊、關聯組織、有效管理數位化物件並進行多維展示。

圖 2 說明了元數據體系中的描述元數據、服務元數據、保存元數據、管理元數據在文物資訊資源從採集、加工整理、利用、保存到處置的生命週期中的作用，以及這些元數據之間的關係，為在不同類型元數據標準的設計中元素的交叉和複用，不同元數據之間的相互連結和調用提供基礎依據。

### 文物數位化保護元數據資訊框架

文物數位化保護元數據資訊框架是基於文物數位物件建立的資料模型。該模型根據使用者需求，對應用需求進行抽象概念化，建立概念與概念之間的關係，包含文物實體資訊、參考資訊、數位物件的內容資訊、技術處理、服務管理等。如圖 3 所示：

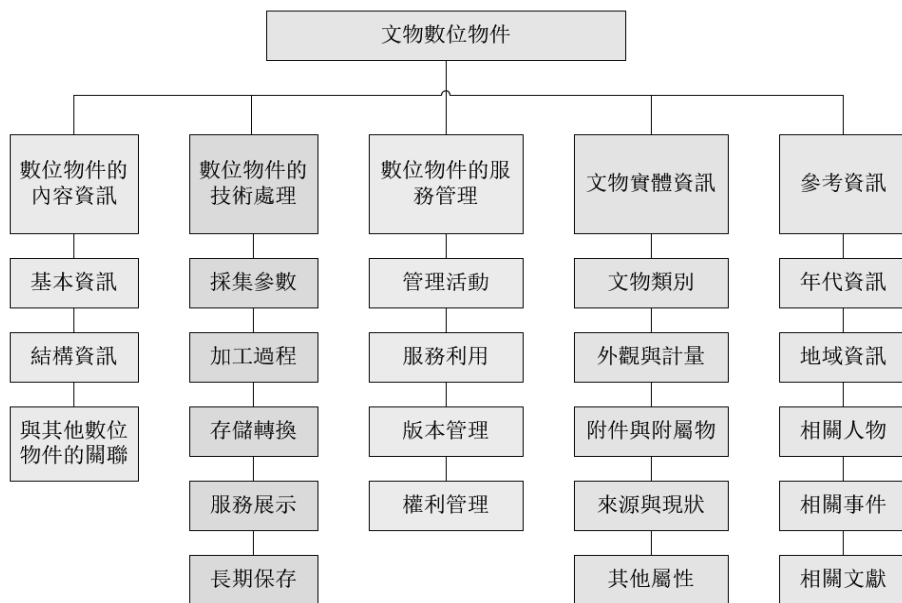


圖 3 文物數位化保護元數據資訊框架

資料來源：改編自 The Forum on Information Standards in Heritage (FISH) (2012). The UK Historic Environment Data Standard.

該模型說明了文物數位物件在數位化過程中所涉及的資料資訊，如果按照文物數位化保護元數據的作用來抽取相關資訊的話，可以歸納為以下三類：

1. 描述資訊（圖 3 之 1、4、5 欄）：即對文物實體的內容、屬性、外在特徵、關聯資訊以及與之相對應的數位物件進行揭示的資訊，以支援資源的查找；
2. 技術資訊：即對文物數位物件在採集、加工、服務過程中為實現長期保存的目的而必須記錄下的技術等資訊，從而支援所描述的數位物件的長期保存；
3. 管理資訊：即關於數位物件使用過程中的存取許可權、智慧財產權、管理等資訊，從而支援對數位物件的有效管理。

### 文物數位化保護元數據結構框架

文物數位化保護元數據的內容結構指的是元數據規範術語間的相互關係（如元素、修飾詞及其屬性等的相互關係）以及元素本身的層級描述（如用於描述性的通用核心元素、用於描述某一類型資源的資源類型核心元素、用於描述某個具體物件的個別元素）等。文物元數據的基本內容結構由三個層次組成——即用於描述性的通用核心元素、用於描述某類文物的文物類型核心元素、用於描述某個特定的文物物件的個別元素。元數據的這種基本內容結構形成各類型資源物件描述元數據的基本框架，用於支持資料交換和檢索，實現資料共用，見圖 4。

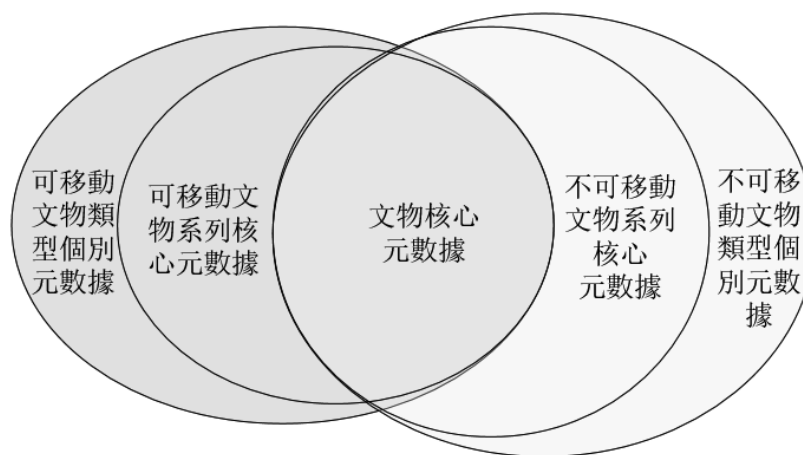


圖 4 文物數位化保護元數據結構框架



文物核心元數據集是根據文物資源共同特點制定、可用於不同類型文物描述的元數據通用元素的集合。從文物實體的角度，根據文物的存在形態分類，分為可移動文物和不可移動文物，因而文物類型核心元數據集分為可移動文物系列核心元數據集和不可移動文物系列核心元數據集。同理，個別元數據集分為可移動文物類型個別元數據集和不可移動文物類型個別元數據集。

## 文物數位化保護元數據應用研究

### 元數據規範設計

文物數位化的資源建設核心內容之一就是建立準確的、規範的、可交互操作的元數據標準規範，在此過程中既要充分考慮到文物的多樣性和獨特性，又要兼顧與國際上通用資料標準的融合，而且為了有利於數位資源整合，也要照顧到國內現有文物資料庫系統的實際情況。

文物元數據規範的設計是一項非常重要的基礎性工作，它是不同類型資料庫在具體元數據方案設計時可依據的一個基礎框架，是不同資料庫系統間實現交互操作和資料共用的基礎。選擇合適的元數據標準規範，是知識組織的重要一步。針對文物的複雜性和多樣性，圖書館與博物館聯合，元數據研究人員、電腦專家與文博專家共同合作，根據文物的特性，建立適於不同類型文物的元數據標準，並努力做到各類型元數據標準之間的相容與交互操作。

#### 一、設計原則

1. 簡單性與準確性原則：設計的元數據標準在著錄實踐時應比較簡單，便於掌握，尤其要考慮到著錄人員除編目員外，更多的是相關專業人員，如書畫專家、古籍專家等。但過於簡單可能導致著錄不夠精確，會降低檢索結果的準確度。因此，同時考慮簡單化和準確度，需要二者兼顧。
2. 專指性與通用性原則：由於文物的各自特性不盡相同，因此無法只用一個元數據標準來描述多種文物，需要根據文物類型和具體的文物實體來確定專門元數據規範。另一方面，必須考慮元數據標準在一定範圍內的通用性，如不同文物之間共用的特性。
3. 實用性原則：制定元數據標準的目的是為了向使用者更充分地揭示資訊資源，在結構與格式的設計、元素的增刪、語法及語義規則的制定等方面都要盡可能從使用者實際需要出發，對廣大使用者的使用需求，包括檢索習慣、對元數據的理解和接受程度、與傳統檢索工具的傳承關係等因素都應當充分考慮。
4. 開放性原則：應盡可能複用或嵌套已有標準、或業界通用的元數據規範，沒有充分的合理理由不宜自行創建新的元數據格式。元數據體系實行開放擴展機制，允許在核心集基礎上按規則進行擴展。

5. 交互操作原則：建立一個適用的開放的元數據體系框架模型，在結構、格式、內容編碼體系等方面進行規範定義，以盡可能達到系統結構的一致性。例如元數據內容描述應使用標準的內容編碼體系，包括中國文物分類主題詞表、國別或地區、日期或時期等，從而保障內容描述方式的標準化和描述內容的可交換；採用 XML 標記語言和「資源描述架構」（Resource Description Framework, RDF）進行開放性描述和標準化封裝，並在其中通過「命名空間」的方式註明元數據來源，使元數據格式在經過這樣的描述和封裝之後，可以方便地被其它系統相容；在具體的資料級別應用上，應易於建立與其他通用元數據格式的轉換機制（肖瓏、申曉娟，2011）。

## 二、核心元數據設計

核心元數據是對文物數位化保護中不同類型文物資源基本特徵進行描述的元數據，是根據文物資源的共同特點確定的元數據集合。作為各領域元數據交互操作的一個基礎元素集，適用於各類文物資訊資源的基本特徵描述。制定核心元素集的目的是為了給資訊資源描述提供通用的元素和規範。核心元素集可作為各類資訊資源分享操作的語義基礎和元數據擴展基礎。

考慮到元數據規範的通用性、可擴展性、元數據之間的易轉換性、系統之間的互通性等需求，以及現有比較通用的 Dublin Core (DC) 元數據是網路環境下資訊資源描述領域影響最大、應用最廣的元數據標準，文物數位化保護元數據的核心元數據集在結構上參考了 DC 元數據，採用元素和元素修飾詞的形式。

但是，DC 元數據標準主要針對的是網路資源的檢索，所描述的是數位資源的共性特徵，無法揭示文物的複雜性和特殊性。因此，核心元數據的元素設計又參考了目前國際上文物方面通用的元數據標準、國內文物普查時頒佈的一系列標準規範、各文博機構自行採用的文物著錄元素集等，以便於在文物數位化過程中更加充分地揭示文物及其數位物件，並與國際國內標準相容，如 CDWA (Categories for the Description of Works of Art)、CCO (Cataloging Cultural Objects)、VRA Core (Visual Resources Association Core Categories)、《館藏文物登錄規範》(WW/T 0017-2013)、《博物館藏品資訊指標著錄規範》、《博物館藏品聲像資訊指標規範》、《第一次可移動文物普查工作標準規範》等相關標準，設立了 20 個元素。見表 1。

表 1 文物數位化保護元數據核心元素集

1	文物類型	11	主題
2	名稱	12	考古發掘
3	文物識別號	13	級別
4	所在位置	14	現狀

(續表 1)

5	創作	15	來源
6	材質	16	權限
7	工藝技法	17	展覽／借展史
8	計量	18	數字對象
9	描述	19	相關文物
10	題識／標記	20	相關知識

## 核心元數據的擴展和關聯

核心元數據的擴展是在核心集基礎上，根據各類型文物特徵進行的專門元數據擴展，以便與已有的文物數位化資源整合，尤其是與文博部門現有的文物資料庫管理系統靈活相容與交互操作。其基本原則如下：

1. 橫向擴展規則：指元素級別上的擴展規則。包括對核心元素不能描述的特性，可以增加元素，即文物類型核心元素，但文物類型核心元素不能與核心元素有任何語義上的重複；文物類型核心元素只擴展描述型元素。例如雕塑類專門元數據根據雕塑類特點橫向擴展了兩個文物類型核心元素：「製作方法」和「紋飾」。古建築類專門元數據橫向擴展了「朝向／佈局」、「局部」兩個文物類型核心元素。
2. 縱向擴展規則：指使用修飾詞的方式向下擴展。修飾詞是對元素的語義進行限定和修飾，以提高元素的專指性和精確性。包括允許擴展核心元素的修飾詞和文物類型核心元素的修飾詞；盡可能遵循向上相容，即自行制定或者複用的修飾詞的語義不能超出被修飾元素的語義，修飾詞只是對被修飾元素的含義範圍作了進一步的限定。例如雕塑類專門元數據核心元素「題識」根據雕塑類特點縱向擴展了元素修飾詞「類型」、「謄寫或描述」、「字體／字形」和「位置」。

各機構在對具體的文物物件描述時可根據本單位實際需要進行本地個別元素的擴展。擴展原則與核心元數據的擴展原則相同。

在核心元素集中，除描述文物實體的內容、屬性、外觀特徵的元素外，還包括與文物實體相關聯內容的其他描述元素，如數字對象、相關文物、相關知識等，可以利用統一資源識別項 (Uniform Resource Locator, URL) 對這些元素進行標識，進而利用資源描述架構 (RDF) 對元素集進行關聯，通過連結可以把這些相關資訊揭示出來，從而實現對數位資源的有效整合。核心元數據的擴展與關聯見圖 5。

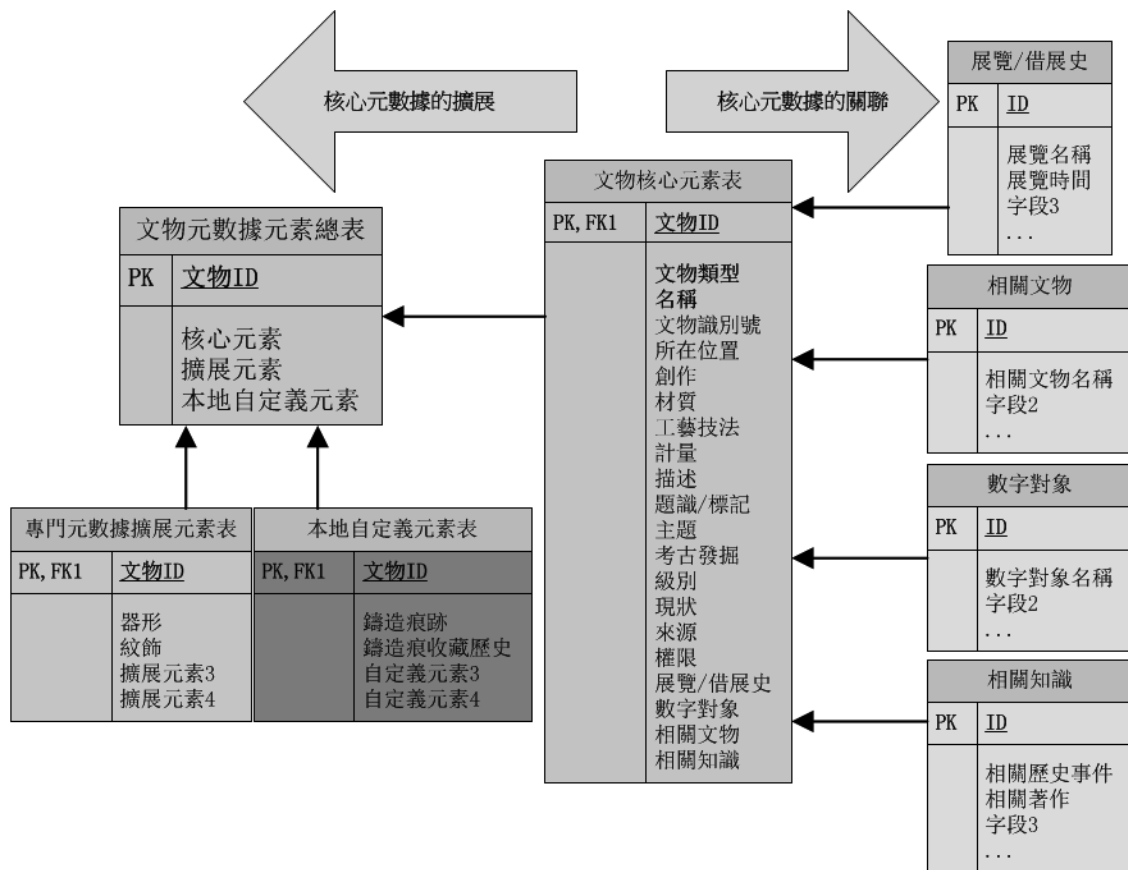


圖 5 文物元數據元素構成圖

如圖 5 所示，文物元數據元素由核心元素、擴展元素和本地自定義元素組成。核心元數據經過擴展，成為一個專門元數據，在本地應用時，還可以繼續擴展增加當地語系化的元素，如銅器類專門元數據根據銅器類特點擴展了兩個文物類型核心元素：「器形」和「紋飾」。本地在對銅器類文物描述時自行擴展了兩個元素：「鑄造痕跡」和「收藏歷史」。

圖 5 中也顯示了核心元數據的關聯，除描述文物實體的內容及外在特徵的元素，如文物類型、名稱、文物識別號、所在位置等，還包括與文物實體相關聯的其他元素，如展覽／借展史、數字對象、相關文物、相關知識，這些元素都有一個識別字 (Identifier, ID)，可通過 URI 形式將這些元素所揭示的資訊與核心元數據關聯。

總之，通過模組化、可擴展的、可關聯的、既通用又可以個性化應用的描述元數據體系及專門元數據規範的設計，整合了現有的各類文物元數據，滿足了不同的文物元數據應用需求，建立了文物數位化保護元數據規範的可持續發展機制。

## 元數據的應用規範

文物元數據的應用規範是對文物資源元數據結構、定義規範、擴展規則、著錄規則、交互操作規則等進行規定，以保持文物元數據結構的一致性，為文物資源的有效組織、元數據之間的交互操作、元數據的廣泛應用和共用奠定基礎，使數位博物館的可持續發展成為可能。文物元數據應用規範基本模型見圖 6。

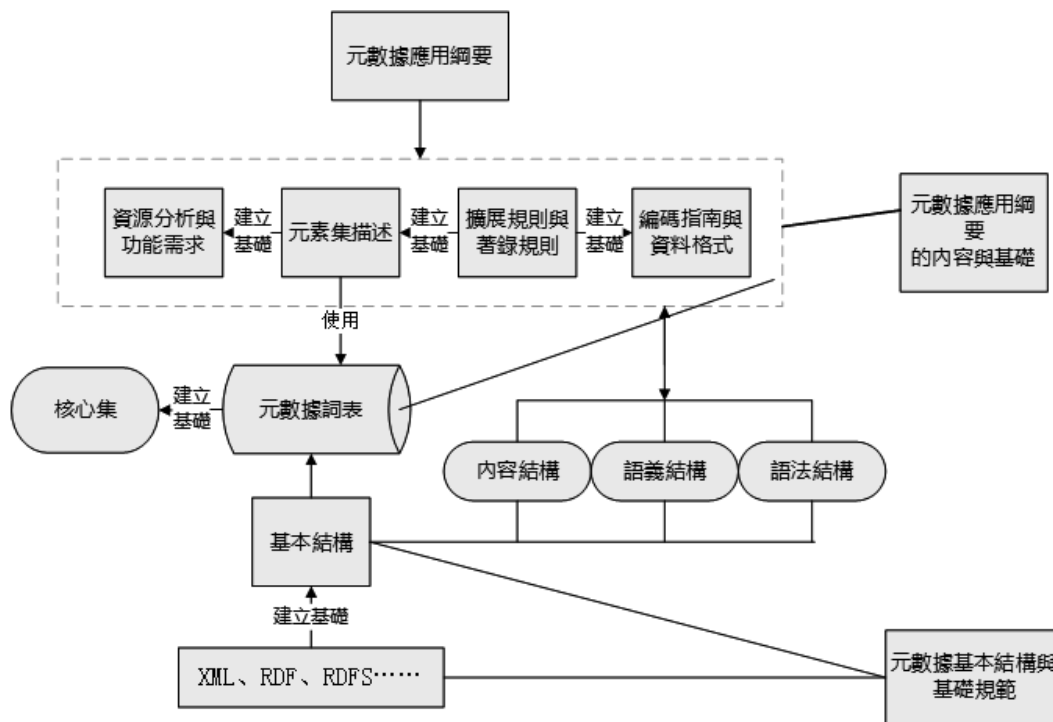


圖 6 文物元數據應用規範基本模型

資料來源：改編自 Baker, T. (2008). The Singapore Framework for Dublin Core Application Profiles.

元數據應用規範的基本模型主要由元數據基本結構與基礎規範、元數據應用綱要兩大部分組成。元數據基礎規範主要是指形式化描述語言，即識別字、標記語言、格式語言，是與互聯網的底層規範相符合的。元數據的基本結構包括內容結構、語義結構和語法結構。元數據應用綱要主要由資源分析與功能需求、元素集描述、擴展規則與著錄規則以及編碼指南與資料格式四部分組成。

建立文物元數據應用規範模型的目的是保證根據此模型建立的元數據規範在資料結構、格式、語義等方面的一致性和整體性，從而在不同層面上為各種資訊資源的檢索、整合、交換以及其他應用提供支援。因此文物元數據應用規範的核心內容包括確定元數據的基本結構、建立元數據核心集、制定元數據應用綱要。

文物元數據的應用流程是從元數據需求和應用環境分析、元數據標準的選擇與制定、應用綱要的設計、編制元數據一直到在數位環境中應用元數據。在整個流程中，首先對文物資源的數位物件及已有的文物元數據標準進行分析，建立元數據應用的基本框架，選擇適用的元數據標準，再根據具體應用需求，規定元數據結構（包括內容結構、句法結構、語義結構）、元數據核心元素集及命名空間、擴展規則、著錄規則、本地應用規則以及所有元素的編碼規則。在此基礎上創建元數據記錄，形成元數據倉儲，對使用者提供統一的檢索介面，同時對其維護和遷移以保證資料倉儲中元數據的品質與交互操作。最後，通過在數位博物館中的應用，實現文物資訊資源的檢索與展示、管理與保存，支援和促進文物數位化保護建設，推動文物數位化保護的研究、應用與發展（肖瓏、申曉娟，2011）。

## 結語

文物數位化保護元數據的基礎與應用研究為文物數位化提供了基礎保障，滿足了文物數位化和知識資料庫建設的需求，為技術與人文的合作構築了非同尋常的新平臺。數位人文方法的運用，為文物數位化後的深度開發提供了創造性的基礎理論和實踐性的基本方法。從文物資訊的簡單組織、檢索與利用到深度分析和挖掘，從單一的研究手段到多維度的科學分析和智慧介入，文物數位化後的研究工作可以實現質的飛躍。

文物數位化保護元數據作為數位博物館資訊資源的資訊組織方式，也是知識組織的基礎。知識組織是資訊組織的延伸，不僅包括詞表、分類法等傳統的資訊組織技術，也包括了語義網、本體等現代資訊和知識組織技術，可以使眾多的資源集成一張巨大的知識網路，使得面向使用者的知識檢索與知識服務成為可能（蔡璐、熊擁軍、劉燦姣，2016）。基於本體的知識組織以概念為基礎，而概念是從元數據中的元素中抽取出來的，例如：文物元數據中的文物類型、創作者、地理位置、主題等，其取值和取值的規範對知識組織都是非常重要的基礎性工作。未來可在此基礎上，通過對元數據的整合，明確其各種概念及主題關係，用關聯資料的形式發佈資源，形成層次清晰的網狀結構知識圖譜體系。

未來，文物數位化保護元數據設計工作完成後，還有一系列重要工作要進行，如與時俱進地更新內容，推廣文物數位化保護元數據規範的應用，多維度聚合現有數位化資源以實現知識發現，通過關聯資料整合文物數位資源等等。這一系列的工作仍然需要文物工作者、資料和技術工作者攜手合作，充分運用數位人文方法，大力促進文物數位化後的深度研究、開發與應用，助推人類文明文化的傳承。

## 參考文獻

ACRL Research Planning and Review Committee (2014). Top trends in academic libraries: A review of the trends and issues affecting academic libraries in higher education. *College & Research Libraries News*,

75(6), 294-302. Retrieved from <http://crln.acrl.org/content/75/6/294.full>

Knapp, J. A. (2012). Plugging the 「whole」: Librarians as interdisciplinary facilitators. *Library Review*, 61(3), 199-214. doi: 10.1108/00242531211259328

Vandegrift, M. & Varner, S. (2013). Evolving in Common: Creating Mutually Supportive Relationships Between Libraries and the Digital Humanities. *Journal of Library Administration*, 53(1), 67-78. doi: 10.1080/01930826.2013.756699

熊莉君、張福陽、張燦 (2016)。圖書館在數字人文領域的傳播功能與服務研究。《圖書館》(2)，88-93、99。

【Xiong, Li-Jun, Zhang, Fu-Yang & Zhang, Can (2016). On the Dissemination Function and Service of Library in the Digital Humanities. *Library*, (2), 88-93, 99.】

肖瓏、申曉娟 (主編) (2011)。《國家圖書館元數據應用總則規範彙編》。中國北京市：國家圖書館出版社。

【Xiao, Long & Shen, Xiao-Juan (Eds.) (2011). *Kuochiatushukuan yuanshuchu yingyung tsungtse kuifan huipien*. Beijing, China: National Library of China Publishing House.】

蔡璐、熊擁軍、劉燦姣 (2016)。基於本體和元數據的非遺資源知識組織體系構建。《圖書館理論與實踐》(3)，39-43。

【Cai, Lu, Xiong, Yong-Jun & Liu, Can-Jiao (2016). The Construction of Intangible Cultural Heritage Information Knowledge Organization System Based on Ontology and Metadata. *Library Theory and Practice*, (3), 39-43.】

# ***A Cross-disciplinary and Fundamental Study of Digital Humanities: Taking “Metadata Standards for Digital Cultural Heritage” for Example***

## **Long Xiao**

Professor, Peking University Library

E-mail: lxiao@lib.pku.edu.cn

## **Yan Zhang**

Associate Professor, Peking University Library

E-mail: zhangyan@lib.pku.edu.cn

## **June Zhang**

Associate Professor, Peking University Library

E-mail: zhangje@lib.pku.edu.cn

## **Ying Feng**

Associate Professor, Peking University Library

E-mail: fengy@lib.pku.edu.cn

Keywords: Digital Humanities; Digital Preservation of Cultural Heritage Standards; Metadata Framework and System; Metadata Standards

---

### **【Abstract】**

Digital humanities is one of the research hotspots in recent years. One of its goals is to describe the object of study in the field of humanities more deeply, excavate, as well as to organize and disseminate human civilization and related knowledge by the technology. This paper takes the research on the Metadata for Digital Cultural Heritage as an example to introduce the general framework and professional development of Metadata Standards for Digital Cultural Heritage. It clarifies that standards and norms are necessary and the metadata research is a basic, long - term work in the development of digital humanities.



## **【Long Abstract】**

### **Background**

This study used the Ministry of Science and Technology of China funded project “Metadata Standards for Digital Cultural Heritage,” jointly implemented by Peking University Library, National Palace Museum, Dunhuang Academy of China, Institute of Advanced Technology, National Library of China, China University of Science and Technology, Zhejiang University, and Tsinghua University Library, as the background to explain the research on standards of overall structure of cultural heritage digitization preservation and metadata and their application. The said research is a study with interdisciplinary features which belongs to the fundamental field of digital humanities, as well as a case of digital humanities development.

### **Fundamental study of Metadata Standards for Digital Cultural Heritage**

The standard system for cultural heritage digital preservation and metadata is used to support the generation, processing, organization, release, storage, and management of information resources in cultural heritage information systems, in order to meet the needs for describing, organizing, managing, serving, and preserving physical objects and digital objects of cultural heritage. It includes the following:

#### **Metadata Standards for Digital Cultural Heritage**

The standard system for digital cultural heritage includes four parts: standard system framework, core standard system, other supporting standards, and application demonstrations. Metadata standards are the main component of the core standard regulation system of the system.

#### **Lifecycle of cultural object digitization and metadata system**

The lifecycle of cultural objects’ digitization is the foundation of standard for metadata. It not only confirms the different functions of metadata in the cultural objects’ digitization process, but also divides cultural heritage metadata into the following types:

- (1) Descriptive metadata: composed of specific metadata, resource collection metadata, etc. mainly function at the acquisition and processing stages of digitization of physical objects of cultural heritage. The processes of acquiring and processing physical objects of cultural heritage into digital objects (composite digital objects), as well as the processing, description, and retrieval of digital objects (composite digital objects), have to be completed using descriptive metadata.
- (2) Service metadata: mainly function at the stage of providing of service for the digital objects.
- (3) Preservation metadata: mainly exist at the stage of the preservation and handling stages of the

digital objects.

- (4) Management metadata: function at the entire lifecycle and are the metadata specifically for the management of digital resources of cultural objects.

### **Information framework of Metadata Standards for Digital Cultural Heritage**

The information framework of Metadata Standards for Digital Cultural Heritage include (1) descriptive information: the content, properties, external features, associated information of physical objects of cultural heritage and the corresponding digital objects; (2) technological information: information on technologies, etc. that has to be recorded in order to achieve the purpose of long-term preservation during the processes of acquisition, processing, and service of digital objects of cultural heritage; (3) management information: information concerning access permission, intellectual property rights, and management during the use of digital objects.

### **Structural framework of Metadata Standards for Digital Cultural Heritage**

The basic content structure of cultural heritage metadata is the common core elements used for description that are used to describe the cultural heritage type core elements of certain type of cultural heritage, as well as the components of three hierarchies of individual elements of specific cultural objects.

## **Study on application of Metadata Standards for Digital Cultural Heritage**

### **Design of metadata standards**

The design of the metadata standards is a very important fundamental work, as well as a fundamental framework serving as the basis for different types of databases during the design of specific metadata programs. It is also the basis for different database systems to achieve interoperability and data sharing, as well as an important step of knowledge organizing. For the complexity and diversity of cultural objects, libraries and museums cooperate with metadata researchers, computer experts, and creative expo experts to establish metadata standards applicable to different types of cultural objects according to the characteristics of cultural objects, and endeavor to achieve the compatibility and interoperability among various types of metadata standards. Therefore, it is necessary to follow the basic principles, such as simplicity & accuracy, specificity & versatility, practicality, openness, and interchangeability.

### **Core metadata design**

Because Dublin Core (DC) metadata standards are mainly established for the retrieval of network resources, they describe the common features of digital resources and cannot uncover the complexity and particularity of cultural objects. Therefore, the design of elements of core metadata consulted the

metadata standards that are commonly used for current international cultural heritage. In China, a series of standard were promulgated during the general survey on cultural objects. The records and element sets of cultural objects adopted by various creative expo institutions are compatible with international standards and standards in China. 20 elements were established, as shown in Table 1.

**Table 1 Core Element Sets of Cultural Heritage Digitization Protection and Metadata**

1	workType	11	Subject
2	title	12	ArchaeologicalInformation
3	Identifier	13	Level
4	CurrentLocation	14	CurrentCondition
5	Creation	15	Source
6	Materials	16	CopyrightOrRestrictions
7	Techniques	17	ExhibitionOrLoanHistory
8	Measurements	18	Digital object
9	Description	19	RelatedWorks
10	InscriptionsOrMarks	20	RelatedKnowledge

### Expansion and association of core metadata

In addition to describing the elements of content, property, and appearance features of physical objects of cultural heritage, core element sets also include other descriptive elements concerning physical objects of cultural heritage, such as digital objects, relevant cultural heritage, and relevant knowledge. Uniform Resource Identifier (URI) can be used to mark and identify these elements, and Resource Description Framework (RDF) can be further used to associate with element sets.

Cultural objects metadata elements are composed of core elements, expansion elements, and locally established elements. Each element has an Identifier (ID), which can link with the information uncovered by these elements with the core metadata in the form of a URI. For example, two cultural object type core elements are expanded from specific metadata of brassware according to the characteristics of brassware: “appliance shape” and “pattern and decoration.” two elements are expanded from local descriptions of cultural objects brassware: “casting traces” and “history of collection.”

### Rules and best practices of metadata application

They are to establish rules and best practices on cultural heritage resource metadata structure, definition rules, expansion rules, cataloguing rules, and interoperability rules, in order to maintain the structural consistency of cultural heritage metadata. These rules lay foundation for the effective organization of cultural heritage resources, interoperability among metadata, and comprehensive application and sharing of metadata. Their basic model is mainly composed of two major parts: basic

structure and basic rules of metadata and metadata application guidelines.

### **Conclusion**

The investigation on the application of digital humanities methods in the study on metadata standards for digital cultural heritage provides basic methods of creative fundamental theories and practicability for the in-depth development after cultural objects' digitization. As the basis of knowledge organization of digital museum information resources, metadata can release resources and form network structure knowledge map systems with clear hierarchies in the form of linked data through integration and confirmation of various concepts and thematic relationships in the future.

**【Romanization of Chinese references is offered in the paper.】**