

匯入聯合目錄說明

981119 修訂

一、 定義一筆資料：

典藏機構的藏品匯入聯合目錄，是以物件為基礎單位，一個藏品物件產生一筆資料 XML 檔。若很多數位檔都是出自同一個物件、且僅有一組 metadata 描述，則應該視為同一筆資料。

二、 完整 XML 的定義：

一筆完整的資料 XML 必須包含四個部分：分類架構、對應（聯合目錄所採用的）DC 欄位的 metadata、至少一張圖檔（或為其他種類型的藏品數位檔如影音檔、全文資料等，數量上也可接受更多）、可連回原始資料庫的連結。

三、 對 XML 品質的需求，可分為內容與格式兩部分：

（一）內容方面的需求為：

1. 不可以有以下負面表列條件——
 - a. 數位檔連結重複（表示有多筆資料使用同一個數位檔）；
 - b. 原始資料庫連結重複（表示有多筆資料連結到同一頁面）；
 - c. 原始資料庫連結失效（無法連到原始資料庫）。
2. 數位檔與 XML 比例至少達到 100%。
3. 圖檔或影音檔等數位檔應達到網頁瀏覽等級。有關聯合目錄所需的數位檔等級，請參見下表：

項目	聯合目錄所需數位檔等級
數位影像	尺寸為 500x500pix 以上之 JPEG 格式檔案。
視訊檔	格式為 WMV、WMA、flv，320x240/15fps 以上，資料傳輸率為 150~300Kbps 之視訊檔。
音訊檔	格式為 WMA 或 MP3，資料傳輸率為 64~128Kbps 之音訊檔。

表一：聯合目錄所需數位檔等級

(二) 格式方面的需求為：

1. XML 不可使用 HTML 語法，欄位內容若遇到下列字元時，需改用跳脫字元表示：
 - 「&」 → 「&」
 - 「<」 → 「<」
 - 「>」 → 「>」
 - 「"」 → 「"」
 - 「,」 → 「'」其中需要注意若是遇到類似「&#1234;」的時候，&則不需要轉成跳脫字元，因為&後面帶數字編碼代表本身就已經是跳脫字元。
2. Project 欄位中的 GenDate 屬性，請標註該批 XML 匯出的時間，例如「GenDate="2005-03-24"」，聯合目錄將依據此時間屬性來判斷此筆資料是更新還是新增的 XML。
3. <Record>欄位中的內容主題、典藏機構與計畫需用半形冒號「:」隔開，匯出時需小心類似這樣的錯誤：「內容主題:人類學::::雅美(Yami):器物」，正確範例如下：
 - <Record>內容主題:人類學:臺灣原住民族:高砂族群:雅美(Yami):器物</Record>
 - <Record>典藏機構與計畫:公開徵選計畫:國立政治大學:民族學系:國立政治大學民族博物館數位典藏計畫</Record>
4. 時間請遵照以下任一種格式
 - yyyy-MM-dd hh:mm:ss.S
 - yyyy/MM/dd hh:mm:ss.S
 - yyyy-MM-dd hh:mm:ss
 - yyyy/MM/dd hh:mm:ss
 - yyyy-MM-dd
 - yyyy/MM/dd註：yyyy 代表西元年、MM 代表月份、dd 代表日期、hh 代表 24 小時制的時辰、mm 代表分、ss 代表秒、s 代表毫秒。
5. 若缺少時間或地理資訊，可不需在<Record>欄位中產生時間、地理架構。
6. 匯入前，請測試<Hyperlink>與<ICON>欄位，是否可以與 XML 內容與圖片正確連結。原則上此二欄位不可為空值（數位檔為文字性全文資料除外）。

7. <DigiArchiveID>欄位為必備欄位，且內容必須為唯一碼(Unique Identifier)，若是變更編碼可能造成下次匯入時，筆數無故增加的問題。建議直接採用機構為藏品所編的「館藏號」、「典藏號」、「藏品編號」或「登錄號」等的唯一識別號碼。
8. <MetaDesc>部分請在各 DC 欄位的 tag 中，標註機構資料庫原始欄位名稱，例如：
 - <Subject field="功能">裝飾具</Subject>
9. 創用 CC 的授權資訊標示方式，請分別針對數位檔與 Metadata 標註。數位檔有多值的情況時，也可分別標註不同的創用 CC 授權方式。例如：
 - 同時具有多媒體檔、圖檔以及 Metadata，分別在其 tag 中標註：
 - <Multimedia license="CC2.5:BY-NC-SA">http://.../.../</Multimedia>
 - <ICON license="CC2.5:BY-NC-SA"> http://.../.../</ICON>
 - <MetaDesc license="CC2.5:BY-NC-SA">
 - 不同數位檔採取不同創用 CC 授權方式時：
 - <ICON license="CC2.5:BY-NC-SA"> http://.../.../</ICON>
 - <ICON license="CC2.5:BY-NC-ND"> http://.../.../</ICON>
 - <ICON license="CC2.5:BY-SA"> http://.../.../</ICON>
 - 創用 CC 授權各條款與簡寫請見下表：

授權條款中文	對應英文縮寫
姓名標示	BY
姓名標示—非商業性	BY-NC
姓名標示—非商業性—相同方式分享	BY-NC-SA
姓名標示—非商業性—禁止改作	BY-NC-ND
姓名標示—相同方式分享	BY-SA
姓名標示—禁止改作	BY-ND

表二：創用 CC2.5 條款與縮寫

10. 內容單位匯出的 XML 若是有 DTD 宣告，請附上 DTD 檔案。若每筆 XML 要加上 DTD 驗證的 Tag，也就是文件型別宣告(DOCTYPE)，請注意其 DTD 檔案位置的正確性，尤其當 XML 與 DTD 是放在不同的目錄時，要注意其 XML 是否可正確參考到 DTD 檔案的所在位置。例如：<!DOCTYPE AnthropologyContract SYSTEM "..\AnthropologyContract_v1.0.dtd">
11. 若匯出大量 XML 時，請以五千筆為一個單位放入個別目錄，目錄名稱可自由命名。例：1~5000 可放入名為 01 的目錄中，5001~10000 可放入名為 02 的目錄中，以此類推。

附件一：聯合目錄 XML 的 DTD

```
<?xml version="1.0" encoding="Big5"?>
<!--數位典藏共同目錄資料輸入格式-->
<!ELEMENT DACatalog (AdminDesc, MetaDesc)>
<!--管理性資訊描述段-->
<!ELEMENT AdminDesc (Project, Catalog?, DigiArchiveID, Hyperlink?, ICON?)>
<!--計畫資料描述-->
<!ELEMENT Project (#PCDATA)>
<!ATTLIST Project
  GenDate CDATA #REQUIRED
  Creator CDATA #REQUIRED
>
<!--相對應的資料分類目錄-->
<!ELEMENT Catalog (Record+)>
<!--分類目錄,目錄與目錄之間請以":"符號分隔-->
<!ELEMENT Record (#PCDATA)>
<!--數位藏品編號-->
<!ELEMENT DigiArchiveID (#PCDATA)>
<!--數位藏品資料的超鏈結-->
<!ELEMENT Hyperlink (#PCDATA)>
<!--數位藏品小圖示的超鏈結-->
<!ELEMENT ICON (#PCDATA)>

<!--後設資料描述段-->
<!ELEMENT MetaDesc (Title*, Creator*, Subject*, Description*, Publisher*, Contributor*,
Date*, Type*, Format*, Identifier*, Source*, Language*, Relation*, Coverage*, Rights*)>
<!--標題-->
<!ELEMENT Title (#PCDATA)>
<!--著作者-->
<!ELEMENT Creator (#PCDATA)>
<!--主題 / 關鍵字-->
<!ELEMENT Subject (#PCDATA)>
<!--描述-->
<!ELEMENT Description (#PCDATA)>
<!--出版者-->
<!ELEMENT Publisher (#PCDATA)>
<!--貢獻者-->
<!ELEMENT Contributor (#PCDATA)>
<!--日期-->
<!ELEMENT Date (#PCDATA)>
<!--資料類型-->
<!ELEMENT Type (#PCDATA)>
<!--格式-->
<!ELEMENT Format (#PCDATA)>
<!--資料識別-->
<!ELEMENT Identifier (#PCDATA)>
<!--來源-->
<!ELEMENT Source (#PCDATA)>
<!--語言-->
<!ELEMENT Language (#PCDATA)>
<!--關連-->
<!ELEMENT Relation (#PCDATA)>
<!--範圍-->
<!ELEMENT Coverage (#PCDATA)>
<!--管理權-->
<!ELEMENT Rights (#PCDATA)>
```

附件二：聯合目錄 XML 結構範例

```
<?xml version="1.0" encoding="Big5"?>
<DACatalog>
  <AdminDesc>
    <Project Creator="長榮大學視覺藝術學系" GenDate="2008-04-12 09:25:12.000">拓展臺灣數位典藏計畫－屏東縣排灣族群傳統藝術文物數位典藏計畫</Project>
    <Catalog>
      <Record>內容主題:人類學:臺灣原住民族:高砂族群:排灣(Paiwan):器物</Record>
      <Record>典藏機構與計畫:公開徵選計畫:長榮大學:視覺藝術學系:屏東縣排灣族群傳統藝術文物數位典藏計畫</Record>
    </Catalog>
    <DigiArchiveID>pa_d01_001</DigiArchiveID>
    <Hyperlink>http://ptap.cultural.pthg.gov.tw/.../1.html</Hyperlink>
    <ICON license="CC2.5:BY-NC-SA">http://ptap.cultural.pthg.gov.tw/.../1_1.jpg</ICON>
    <Multimedia license="CC2.5:BY-NC-SA">http://vcenter.iis.sinica.edu.tw/.../</Multimedia>
  </AdminDesc>
  <MetaDesc license="CC2.5:BY-NC-SA">
    <Title>門板</Title>
    <Subject>屏東縣排灣族群傳統藝術文物</Subject>
    <Subject>家屋構建</Subject>
    <Description field="文物敘述">此例為門板或壁板構件的一部分。雕刻兩個人像文，風格較近似內文式人像的雕刻文樣，頸部有連杯，眼睛以白色貝殼填裝。未施彩。</Description>
    <Description field="損壞狀況">部分貝殼脫落、木板面略有腐朽及斷裂痕。</Description>
    <Identifier>pa_d01_001</Identifier>
    <Publisher>XXXXXXX</Publisher>
    <Contributor>XXXXXXX</Contributor>
    <Date field="購藏日期">2005.01.20</Date>
  </MetaDesc>
</DACatalog>
```

註解 [C1]: 進行匯出的單位，以及匯出的時間。

註解 [C2]: 該筆數位資料的唯一辨識碼。

註解 [C3]: 該筆資料的原資料庫連結。

註解 [C4]: 若有對該張圖像的創用 CC 授權方式，請標註於此。創用 CC 條款縮寫請參考說明第 9 點。

註解 [C5]: 該筆資料的瀏覽級圖檔位址，500*500 像素

註解 [C6]: 同註解[C4]，若有對該影像檔的創用 CC 授

註解 [C7]: 若有影音、聲音檔等其他多媒體檔案，請提

註解 [C8]: 同註解[C4]，若有對該筆資源 Metadata 的創

註解 [C9]: 以 field="" 標註原始資料庫的欄位名稱。以

```
<Type>XXXXXXXXX</Type>
<Format>150.3*92.1 公分</Format>
<Source>XXXXXX</Source>
<Language>XXXXXX</Language>
<Relation field="參考文獻 (Reference) ">王榮信、林增榮，2000，《關於排灣—排灣族文明巡禮》，屏東：屏東縣政府。</Relation>
<Coverage>屏東縣</Coverage>
<Rights>屏東縣政府文化局</Rights>
</MetaDesc>
</DACatalog>
```

第 5 頁: [1] 註解 [C5]	CANCER	2009/11/19 11:53:00 AM
該筆資料的瀏覽級圖檔位址，500*500 像素以上大小。		
第 5 頁: [2] 註解 [C6]	CANCER	2009/11/19 11:55:00 AM
同註解[C4]，若有對該影像檔的創用 CC 授權方式，請標註於此。		
第 5 頁: [3] 註解 [C7]	CANCER	2009/11/19 11:52:00 AM
若有影音、聲音檔等其他多媒體檔案，請提供檔案給聯合目錄，聯合目錄將批次上傳至 Vcenter，並將多媒體檔案對應的 Vcenter 網址產生於此。		
第 5 頁: [4] 註解 [C8]	CANCER	2009/11/19 11:56:00 AM
同註解[C4]，若有對該筆資源 Metadata 的創用 CC 授權方式，請標註於此。		
第 5 頁: [5] 註解 [C9]	CANCER	2009/11/19 11:56:00 AM
以 field="" 標註原始資料庫的欄位名稱。以下類推。		

附件三、Dublin Core（都柏林核心集，DC）欄位定義

名稱	定義	說明
標題 (Title)	資源被賦予的名稱	典藏品的名稱，例如：動植物的學名、中英俗名，物件名稱，文件資料標題…等。
創作者 (Creator)	創造資源之主要負責實體	編輯、創造該筆資源內容的主要負責人、團體、機構等，例如：作者，照片攝影者，地圖測繪者，標本鑑定者…等等。
主題 (Subject)	資源之主題	例如：相關族群，礦物類別…等等，也可能是原典藏單位直接著錄之關鍵字。
描述 (Description)	對資源的說明	資源內容的說明，例如：器物用途解釋，生物習性，圖畫流派意境，古文字、圖像釋文或解說，新聞報導內容…等等，通常是一小段文字。亦可提供摘要與目次等資訊。
出版者 (Publisher)	使資源可供使用之實體	包括出版的個人與團體、機構。
貢獻者 (Contributor)	對創造資源有貢獻之實體	對於資源有特別知識內容貢獻的個人或機構，但這些貢獻相較於創作者中指定的個人或機構而言是次要的。例如：編輯者、翻譯者、插圖畫家…等。
日期 (Date)	在資源生命週期中，與事件相關的時間點或時期	例如：製造日期，書寫日期，測繪日期，出版品發行日期，發掘日期…等等。
型式 (Type)	資源之性質或人文藝術作品類型	建議使用控制詞彙 DCMI Type Vocabulary，例如：圖像、文字、聲音、合集、原件、替代品…等。
格式 (Format)	資源的檔案格式、實體媒體或維度	包括媒體類型或資源的度量資料，例如：jpg 檔，長寬高 50x10x13，放映時間 5 分鐘…等等。
識別碼 (Identifier)	資源的明確參照	該典藏品資源的明確辨識資料，例如：登錄號，典藏號，URI…等等。
來源 (Source)	衍生出描述資源的相關資源	敘述衍生目前資源的來源資源，例如，剪報的掃描數位檔，是衍生自紙本報紙。
語言 (Language)	資源的語言	例如，內閣大庫檔案使用的語言類型是「漢文」或「中文」，部分文書可能是「原文：英文，譯文：中文」。圖畫或影像無語言時，則不需著錄此欄位。
關聯 (Relation)	說明相關的資源	例如：參考書目、其他版本、其他媒體形式…等。

名稱	定義	說明
涵蓋範圍 (Coverage)	資源之空間或時間屬性的主題、空間應用性，或與資源相關的政治轄區	包括與資源內容有關，但不適宜放置主題的時、空元素，例如：時期（時間名稱、日期或期間），空間地點（地理座標或地名），或行政區域。
權利 (Rights)	有關資源的權利資訊，包括資源原有的與被賦予的權利	例如：著作權擁有者、取用限制、授權條款、典藏單位…等。

附註：

1. 以上各欄位標籤及定義採用 [Dublin Core Metadata Element Set, Version 1.1](http://dublincore.org/documents/dces/) ¹。
2. 各計畫亦可參閱由「後設資料工作組」於網路上所提供 [Dublin Core 的參考規範](http://metadata.teldap.tw/standard/standard-frame.html)²，或者直接參考[聯合目錄目前的用法](http://catalog.digitalarchives.tw/dacs5/System/Search/Advanced_search.jsp)³。
3. 對比時，請依據「一對一原則」，釐清描述的資源，一筆後設資料勿同時描述原件及數位檔。

¹ <http://dublincore.org/documents/dces/>。

² <http://metadata.teldap.tw/standard/standard-frame.html>。

³ http://catalog.digitalarchives.tw/dacs5/System/Search/Advanced_search.jsp。