清华大学艺术博物馆多维度藏品电子档案的建设与思考

The Practice and Thinking About Construction of Electronic Archives of Multidimensional Collections in Tsinghua University Art Museum

彭 宇
Peng Yu
（清华大学艺术博物馆，北京，100084）
(Tsinghua University Art Museum, Beijing, 100084)

内容提要：随着大数据、知识图谱等信息技术的发展，针对藏品数据资源进行知识体系组织、可视化的表达和互动展示，建立多维度藏品电子档案已经成为大数据时代博物馆藏品管理的新要求。以清华大学艺术博物馆的实践为例，文章详细描述多维度藏品电子档案的数据模型、数据标准、数据采集、维护规则等。

关键词：藏品电子档案　数据模型　数据标准　数据采集　维护规则

Abstract: With the development of information technology such as big data and knowledge atlas, organizing knowledge systems, visualizing the expression and interactive display of collection data resources, and establishing multidimensional collection electronic archives have become a new requirement for museum collection management in the era of big data. Taking the practice of the Tsinghua University Art Museum as an example, this article details the data model, data standards, data collection, maintenance rules, and other aspects of multidimensional electronic archives for collections.

Key Words: Electronic archives of collections; data models; data standards; data collection; maintenance rules

藏品是博物馆的核心，也是博物馆的立馆之本。藏品在其生命周期里，伴随着各项研究、展览等活动产生了海量、多源、异构的数据信息。随着大数据、知识图谱等信息技术的发展，针对藏品数据资源进行知识体系组织、可视化表达和互动展示[1]，建立多维度藏品电子档案已经成为大数据时代藏品管理的新要求。多维度藏品电子档案的建立，确保了藏品数据资源的永久保存和检索利用，以信息化的手段实现藏品科学规范的全生命周期管理，也为学术研究、展览展示提供海量、珍贵的数据支持[2]。

一、多维度藏品电子档案的设计方案

多维度藏品电子档案的设计方案包括数据模型、数据标准、数据采集、维护规则四个方面。

1. 数据模型

藏品从征集鉴定到进入博物馆，在其全生命周期里，会进行展览展示、修复观摩、研究探讨等各种活动。在各项进程中，会产生大量的各种格式形态的数据，它们或许以文件形式存储在电脑上，是一段音频或者是一段视频；也可能以数字形式存在信息系统里，是一串文件或者一些字符[3]。将藏品的这些信息按照一定方式梳理归类，建立数据模型，成为藏品的电子档案，能确保藏品信息完整完善的归纳保存，为后续的研究展示奠定坚实基础。藏品数据主要包括以下六个方面（图1）。

（1）账目信息：账目信息主要是藏品的基本信息，包括藏品年代、作者、尺寸、藏品名称等总账记录信息，也包括鉴定时间、鉴定专家、鉴定结果等鉴定信息，以及修复时间、修复内容、修复人等修复信息等。它覆盖了藏品从征集到入藏、鉴定，以及修复、保养等藏品管理的全过程，是原有藏品总账、藏品卡片的电子化数据化版本，是围绕藏品最基本最核心的数据信息，也是最受人关注的藏品信息。

（2）二维信息：二维信息是藏品的图片信息，主要是藏品的全方位基本图片。不同材质的藏品它的基本图片内容也不同。比如：书画要包括整体图和细节图，家具瓷器等要包括正面、侧面、底面等六个面的信息。此外，有立体构件器物的纹饰应拍摄制作展开图。藏品的图片影像是对藏品基本信息的补充[4]，能比较直观地了解藏品现状。随着技术的发展以及人们对藏品的认识，藏品的二维信息已经成为藏品信息的标准配置项。

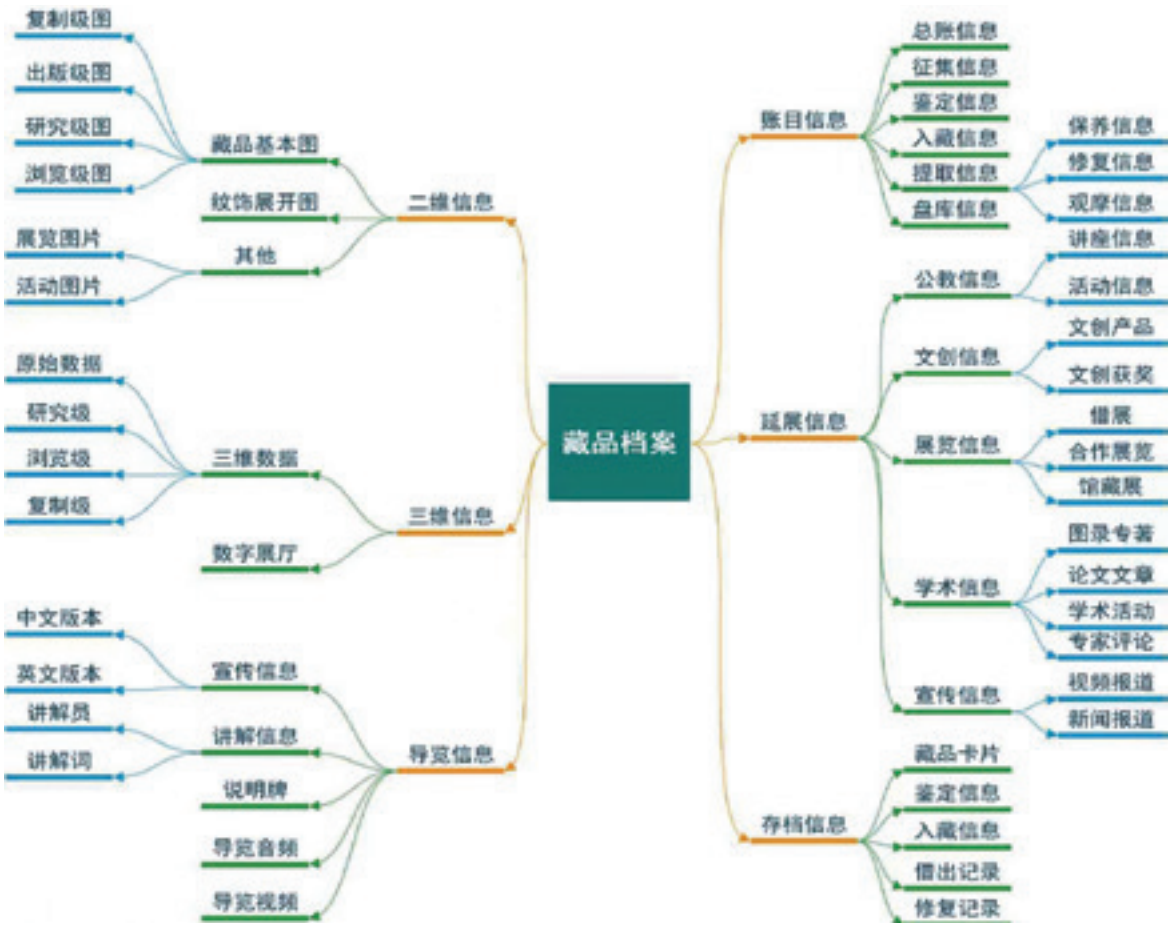

图1 藏品电子档案数据模型

（3）三维信息：三维信息是主要是指藏品的三维模型，是藏品数字影像从静态到动态的进一步发展。采集藏品的三维信息目前主要有两种方法：一种是直接采集，一种是通过二维影像经过一定的算法实现成像。三维信息的采集需要一定的设备、人力、经费的支持，所以目前还未在各馆普及；但是采集三维信息，利用数字化手段制作藏品三维模型，能为藏品展示、保护复制、修复及衍生品开发等应用提供全部素材，也能为馆藏藏品的研究、展示、重建及辅助修复提供数据和技术支持。三维信息还应包括藏品在展厅展出时拍摄的数字展厅信息。

（4）延展信息：延展信息是指围绕藏品进行研究探讨、文创开发、宣传报道等各种活动产生的信息。延展信息包含以下三类。第一，以藏品为原型开发的文创产品以及文创产品参与评奖获奖等信息。第二，藏品到馆外参展的情况：展览时间、展览名称、其他博物馆的藏品参与情况。第三，基于藏品出版的图录专著、发表的论文、举行的研讨活动等学术信息。延展信息记录了藏品在其全生命周期里的各项活动轨迹，是藏品除基本信息等管理信息以外的活动记录。

（5）导览信息：导览信息是指藏品的各类介绍信息，包括宣传信息、讲解信息、说明牌等。宣传信息是官方对外统一介绍藏品的描述词，包含图文等信息，是最全面的藏品介绍词，主要用于网站、海报、宣传册等，分为中文、英文等不同版本。讲解信息主要是用于讲解员现场讲解，它根据藏品宣传信息删减而成。讲解信息是纯文字信息，不包含图片。说明牌主要用于展厅标注藏品，由于受现场环境限制，说明牌内容非常简洁，包含藏品名称、年代、制作者等基本要素，便于参观人员很快了解藏品。同时为提供多元化的导览服务，部分重要藏品也录制了导览音频或视频，更形象全面地介绍藏品。

（6）存档信息：藏品数字化以后，仍需要有纸质存档的文件；包括手写或者机打的藏品卡片，鉴定、修复时的专家信息以及签署、入藏、借出时的审批报文以及审批签字单据等。这些信息应同时以电子版和纸版存档。

随着信息技术的发展以及人们对藏品研究的深入，还会在不同活动或活动的不同阶段产生更多更丰富的藏品信息。因此藏品电子档案的数据模型也会不断地完善、优化、更新。

2. 数据标准

藏品数据是海量、多源、异构的，为集中存储、维护，我们需要对数据进行一定的标准定义[5]。目前已有标准包括国家行业标准和地方行业标准。国家行业标准在国家文物局官方网站的政府信息公开频道的行业标准栏目可以查询[6]，北京市的地方标准在北京市文物局官方网站政务公开频道的标准化栏目可以查询。表1列举了部分相关标准[7]。

同时，由北京大学承担，敦煌研究院、故宫博物院、国家图书馆现代技术研究所、清华大学、浙江大学、中国科学技术大学6家文博机构、元数据领域研究机构合作研究的国家科技支撑计划项目“文物数字化保护标准体系及关键标准研究与示范”中的研究课题“文物数字化保护元数据标准研究”也取得了丰富的成果，包括62项标准规范（其中59项标准征求意见稿）[8]。相信后续会有更多更细致具体的标准推出，从而规范文物数字化保护元数据建设，支持和促进文物数字化保护建设、推动文物保护的研究、展示、应用与发展。

3. 数据采集

在数字化工作开展前，藏品数据采集工作主要是靠藏品管理员手工完成。那时候的藏品管理员都能写一笔好字，因为藏品的主要信息都靠藏品账本保存，而账本以及藏品卡片都是由藏品管理员手写；但是由于藏品信息项不断增多，而账本、卡片涵盖信息有限，且不便查阅，随着电脑的普及，藏品管理员开始使用软件EXCEL实现藏品数据采集及保存。因此藏品有了电子、纸质两套账目信息。

近20年以来，博物馆内部管理与对外服务的信息系统建设

表1 藏品相关标准示例

序号	标准名称	标准编号	实施时间	发布单位
1	馆藏金属文物保护修复方案编写规范	WW/T 0009—2007	2008-03-01	国家文物局
2	石质文物保护修复方案编写规范	WW/T 0007—2007	2008-03-01	国家文物局
3	馆藏丝织品保护修复方案编写规范	WW/T 0014—2008	2009-03-01	国家文物局
4	馆藏丝织品保护修复档案记录规范	WW/T 0015—2008	2009-03-01	国家文物局
5	馆藏文物出入库规范	WW/T 0018—2008	2009-03-01	国家文物局
6	馆藏文物展览点交规范	WW/T 0019—2008	2009-03-01	国家文物局
7	文物藏品档案规范	WW/T 0020—2008	2009-03-01	国家文物局
8	馆藏出土竹木漆器类文物保护修复档案记录规范	WW/T 0011—2008	2009-03-01	国家文物局
9	陶质彩绘文物保护修复档案记录规范	WW/T 0023—2010	2010-09-01	国家文物局
10	馆藏纸质文物保护修复档案记录规范	WW/T 0027—2010	2010-09-01	国家文物局
11	馆藏文物登录规范	WW/T 0017—2013	2013-08-15	国家文物局
12	古代陶瓷科技信息提取规范方法与原则	WW/T 0053—2014	2014-06-01	国家文物局
13	古建筑壁画数字化测绘技术规程	WW/T 0082—2017	2017-12-01	国家文物局
14	文物艺术品元数据规范	DB11/T 1219—2015	2015-11-01	北京市质量技术监督局
15	文物艺术品数据元规范 第2部分：书画	DB11/T 1219.2—2019	2019-10-01	北京市市场监督管理局
16	文物三维数字化技术规范 器物	DB11/T 1922—2021	2022-04-01	北京市市场监督管理局
17	文物艺术品数据元规范 第3部分：陶瓷	DB11/T1219.3—2022	2022-07-01	北京市市场监督管理局

如火如荼。EXCEL文件被藏品管理信息系统等相关软件系统所取代，信息系统完成藏品数据采集、保存的同时，还可提供信息查询、数据共享等服务。

博物馆不同业务流程都会产生藏品相关信息，因此信息的采集也会涉及博物馆内部多个业务系统。

第一，账目、存档信息：这是藏品的基本数据项信息，也是藏品管理信息系统建设的数据核心。这些信息基本都由藏品管理信息系统采集、维护，并通过标准接口对外提供查询等服务。

第二，二维信息、三维信息：这是藏品的影像数据。由于每个博物馆信息化建设规划和实施进度不同，有的馆在藏品管理系统中建立了图片信息和三维信息模块，负责维护藏品二维、三维信息。有的馆由于藏品管理系统建设在前，藏品二维、三维信息采集在后，所以藏品二维和三维信息采集系统独立于藏品管理系统。还有的馆把藏品二维、三维信息与馆内其他资源信息（比如：场馆图片、参观图片、导览视频等）一起作为数字资源管理系统的一部分。

第三，导览信息：藏品导览信息是在藏品基本信息的基础上，依据展览主题、观众需求加工创作后形成的，由馆内导览系统采集和保存。藏品导览信息为多种导览手段提供基础信息，不同导览手段均可共享。

第四，延展信息：藏品的延展信息，是由馆内的其他业务活动产生的。馆内的文创开发售卖，就会反馈藏品相应的文创信息；馆内展览以及向外借展活动就能反馈藏品参与每个展览的信息。学者对藏品研究和讨论，就能反馈相应的学术信息。因此藏品的延展信息由对应业务活动的支撑系统采集维护。

4. 维护规则

数据的维护规则是：谁产生谁维护，维护工作具有唯一性，任何数据只有唯一的数据采集源头和维护方。比如：藏品的日常管理保护工作由藏品保管部门负责，因为藏品的基本信息维护权利在藏品保管部门，任何藏品信息的增加、删除、修改、共享等操作均需要通过藏品管理系统的授权，且信息维护权限需与操作系统人员的岗位职责一致。

二、清华大学艺术博物馆多维度藏品电子档案的建设实践

从2016年开馆以来，至2023年，经过近7年的建设，清华大学艺术博物馆（以下简称清华艺博）的藏品电子档案从无到有，并根据运营情况，不断推出新的数字化服务，提高数字化管理水平。

1. 数据模型

目前清华艺博的多维度藏品电子档案的数据模型集中在藏品账目信息、二维信息、三维信息、存档信息、导览信息等部分。

第一，账目信息、存档信息：清华艺博已经完成全部馆藏品的账目整理及录入工作，并在官方网站对外提供馆藏品信息的实时查询和下载。清华艺博从藏品管理工作实际需要出发，保留了电子、纸质两种方式存档信息，并保证两个档案内容一致。

第二，二维信息：目前对馆藏品的二维影像拍摄在按照计划进行，已经完成拍摄的馆藏品二维影像同步在官方网站开放浏览下载。

第三，三维信息：清华艺博已完成150件藏品的三维扫描，后续会在官方网站开放浏览。清华艺博官方网站上为每个展览建立虚拟展厅，其访问信息以及拍摄完成的全景素材信息会作为藏品的数字展厅信息归类保存。

第四，导览信息：清华艺博开馆时就建有导览系统，对外提供租赁导览机、自助语音导览、人工讲解三种服务。通过近7年的积累，导览系统里保存了馆藏品以及外来参展展品的导览信息，随着馆藏展览的展出不断更新和完善。

由于清华艺博的学术管理系统、展览管理系统等业务系统还在建设过程中，因此藏品的延展信息还未做到统一采集和使用。后续将通过不同业务系统的建设来不断采集和完善藏品延展信息，并通过数据接口实现数据共享和整合查询。

2. 采集标准与数据维护

清华艺博的信息系统建设均参照国家或者北京市的标准进行，但是在二维影像和三维数据的采集方面，根据清华艺博藏品的实际情况与工作需求，提出了更具体的要求。

（1）二维信息：二维影像数据主要包括藏品的实物图和纹饰连续展开图。根据使用需求，清华艺博对藏品格式、分辨率的要求如下。藏品纹饰的连续展开图包含藏品的每个立体构件，单张影像像素不低于5000万，保存为RAW格式及JPG文件。

二维基本图：是每件藏品必备的图片信息，根据使用需求，我们将它分为四个等级，用于不同的场景。复制级：是原始拍摄图片，包含藏品的整图以及细节图，存为tif格式。该级别图片未进行任何后期处理，是最大格式最高清的图片数据，能达到藏品1：1的复制要求。出版级：是经过裁剪处理的图片，一般是5—30兆左右的jpg图片。主要用于图录、宣传册等纸质出版物的印刷。以上这两类图片，由于涉及版权问题，不会通过网络或其他渠道公开发布，使用需有授权许可。研究级：是经过裁剪、压缩、换底色处理的图片，一般是5兆以内的jpg图片。这类图片在清华艺博的官方网站开放浏览和下载。浏览级：是用于网络宣传或日常办公的图片。大小基本在1兆以内，格式为jpg或者png，便于网络浏览和下载。

（2）三维数据：目前清华艺博委托第三方公司对藏品进行三维扫描，建立藏品三维模型，并提供复制级、浏览级、研究级和原始数据四个等级的藏品三维数据。四个等级的具体数据要求如下。

复制级：主要应用于3D打印和复制藏品。数据不带纹理映射，但完整性及模型精度要求较高，要达到95%以上的数据完整性，瓷器类整体精度达到0.03mm，家具类整体精度达到0.1mm以内，数据格式为obj。

研究级：为带纹理映射的三维模型数据，主要用于藏品研究、病害分析等。该项数据需保证纹理映射的准确性，其纹理映射误差不得大于0.2mm，采用原分辨率照片进行纹理贴图，数据格式为obj。

浏览级：主要应用于移动端展览展示等，模型数据量一般在10M以内。纹理贴图精度不得大于0.2mm，数据格式为obj。

原始数据：三维扫描仪采集的未经任何处理的点云网格化数据，格式是stl。

3. 数据采集与维护

如前文所述，根据“谁产生谁维护”的数据维护原则，清华艺博的藏品电子档案主要由典藏部维护，同时涉及到馆内所有业务部门及主要业务系统，具体情况如表2所示。

表2 清华艺博现有支撑系统情况[①]

序号	数据项	支撑系统	管理维护部门
1	账目信息	藏品管理系统	典藏部
2	存档信息	藏品管理系统	典藏部
3	二维信息	藏品管理系统	典藏部
4	三维信息	藏品三维数据平台	典藏部
5	导览信息	藏品管理系统（宣传信息）	信息部
		藏品管理系统（说明牌）	典藏部
		导览系统（讲解信息、导览音频、导览视频）	公共教育部

三、小结

随着大数据、信息技术的发展，还需要持续进行藏品信息的挖掘、收集、归纳等工作，并根据数据收集情况不断优化调整藏品多维度电子档案数据模型，结合既有研究成果，构建藏品信息知识体系。

同时应基于博物馆宣传矩阵，在门户网站、新媒体平台等进行藏品的可视化的表达和互动展示，让文物活起来，发挥博物馆教育职能，满足日益增长的广大观众与用户的需求。

参考文献

[1] 戴畋. 博物馆智慧化背景下文物知识图谱构建的实践探索[J]. 科学教育与博物馆, 2022, 8(1): 41−49.
[2] 杨焱. 高校博物馆文物藏品档案建设思考: 以西藏民族大学博物馆为例[J]. 科学教育与博物馆, 2022, 8(5): 90−96.
[3] 柴秋霞, 姜琳馨. 博物馆藏品数字化展示的现存问题及对策建议[J]. 科学教育与博物馆, 2022, 8(2): 40−49.
[4] 毛若寒, 郑宽. 试论博物馆数字影像技术应用的现状与转向[J]. 中国博物馆, 2022(3): 55−60.
[5] 张元成, 杨敬. 山西博物院数字化技术与可及性创新的实践[J]. 中国博物馆, 2022(2): 10−13.
[6] 国家文物局. 行业标准[S/OL]. [2025−08−28]. http://www.ncha.gov.cn/col/col2423/index.html.
[7] 北京市文物局. 标准化[S/OL]. [2025−08−28]. https://wwj.beijing.gov.cn/bjww/362690/bzh/index.html.
[8] 北京大学图书馆. 文物数字化保护元数据标准规范征求意见稿发布[EB/OL]. (2017−05−11)[2025−08−28]. https://lib.pku.edu.cn/portal/cn/news/0000001494.

① 如前文所述，因藏品延展信息尚未统一采集和使用，故未列入表2中。