

数字文旅专业系列教材 · 教学课件

文化遗产保护概论——非遗数字化与文物数字典藏

第3章 文物数字典藏理论基础

授课对象：安顺学院旅游管理学院 · 旅游管理专业本科生

主讲：祝刚

本章导览

- 3.1 文物数字典藏的概念与范畴
- 3.2 文物数字典藏的理论框架
- 3.3 文物数字典藏的标准体系
- 3.4 文物数字典藏的组织管理
- 3.5 文物数字典藏的发展趋势
- 知识卡片
- 3.6 数字典藏的技术经济学分析
- 3.7 国际数字典藏标准比较
- 3.8 数字典藏的社会维度
- 3.9 数字典藏与学术研究
- 3.10 数字典藏的文化民主化功能
- 3.11 数字典藏的未来形态与智能化发展
- 3.6 文物数字典藏的学科基础深化
- 3.7 元数据标准的比较分析

- 1. 识记：文物数字典藏的定义、学科基础与标准体系
- 2. 理解：信息生命周期理论、数字保存理论与元数据理论在文物数字典藏中的应用
- 3. 应用：运用 OAIS 模型分析一个文物数字典藏系统的功能架构
- 4. 分析：比较 CIDOC CRM、Dublin Core、LIDO 三种元数据标准的异同与适用场景
- 5. 评价：评估当前文物数字典藏从 " 数字化 " 到 " 智慧化 " 转型的成效与挑战
- 6. 创造：设计一个小型博物馆的数字典藏系统架构方案

- 中国国家博物馆作为中国最大的综合性博物馆，馆藏文物超过 140 万件。
- 2020 年，邵壮在研究国博馆藏文物三维数字化实践时指出，国博率先在博物馆藏品保护与研究方面进行了大规模三维数字化探索，使 " 智慧国博 " 建设和博物馆专业品质得到了有效提升。
- 国博的数字化工程采用了高精度三维扫描、多光谱成像、色彩管理等先进技术，对代表性文物进行了毫米级精度的三维数据采集，建立了完整的文物数字档案。
- 这一案例揭示了一个关键问题：文物数字典藏不仅仅是 " 用相机拍一下 " 或 " 用扫描仪扫一下 " 的简单操作，而是一个涉及标准制定、技术选型、工作流程设计、数据管理、
- 长期保存和访问服务的系统工程。

- 如何构建科学的文物数字典藏体系？这正是本章要回答的问题。

01

3.1 文物数字典藏的概念与范畴

3.1.1 数字典藏的定义

- 文物数字典藏 (Digital Archiving of Cultural Relics) 是指利用数字技术对文物进行系统化的信息采集、数据处理、元数据标引、数据库管理和长期保存，
- 以实现文物信息的永久保存、高效检索和开放共享的专业化活动。
- 它不同于简单的 " 文物数字化 "(Digitization of Cultural Relics) ，后者侧重于数字化采集过程，而前者强调的是数字化成果的系统化管理与长期服务。
- Cheng 等人在研究文化遗产数字保存时指出，数字典藏的核心目标是通过系统化的数字管理，确保文化遗物信息得以精确记录、长期保存和有效利用。
- 这一定义强调了数字典藏的三个核心要素：精确记录 (高保真数据采集) 、长期保存 (可持续的数据管理) 和有效利用 (面向用户的访问服务) 。

3.1.2 文物数字典藏与数字图书馆

- 文物数字典藏与数字图书馆 (Digital Library) 在技术架构上有诸多相似之处，但在内容特征、用户需求和服务模式上存在显著差异。
- 数字图书馆主要处理文本和多媒体文献，以检索和阅读为主要服务方式；文物数字典藏则主要处理三维文物数据及其关联信息（年代、材质、工艺、来源等），以多维展示和学术研究为主要服务方式。
- Graham 在研究博物馆藏品数字化的应用时指出，三维数字化扫描设备的商业化普及使博物馆藏品管理、研究和可及性面临范式转变，数字典藏正在从 " 辅助工具 " 发展为博物馆的核心业务系统。

3.1.3 文物数字典藏的学科基础

- 文物数字典藏是考古学、博物馆学、档案学、信息科学和计算机科学的交叉领域。
- 它继承了考古学的类型学和地层学方法，借鉴了博物馆学的藏品管理理论，吸收了档案学的文件生命周期理论，并依托信息科学和计算机科学的技术手段。
- Zhou 等人在《文化遗产数字化》一书中系统阐述了文化遗产数字化的学科基础，指出该领域需要融合人文社科的 " 文化理解 " 与理工科的 " 技术能力 "，是一个典型的跨学科领域。

02

3.2 文物数字典藏的理论框架

3.2.1 信息生命周期理论

- 信息生命周期理论 (Information Lifecycle Theory) 将信息视为有生命周期的有机体，经历创建、采集、组织、存储、利用、归档和销毁 (或永久保存) 等阶段。
- 在文物数字典藏语境中，信息生命周期理论为典藏工作流程设计提供了理论框架。
- 文物数字信息的生命周期包括：创建阶段 (三维扫描、摄影测量等数字化采集)→ 处理阶段 (数据清洗、模型优化、纹理映射)→ 描述阶段 (元数据标引、知识关联)→ 存储阶段 (数据库入库、
备份归档)→ 利用阶段 (检索、展示、下载)→ 保存规划阶段 (格式迁移、技术验证)。
- 每个阶段都有相应的质量控制和管理要求。

3.2.2 数字保存理论

- 数字保存理论关注如何在技术快速变迁的环境中确保数字信息的长期可读性。与纸质档案不同，数字信息面临介质老化、格式过时、软件依赖等多重风险。
- OAIS(Open Archival Information System) 参考模型是数字保存领域的国际标准 (ISO 14721) ，它定义了数字保存系统的功能架构和信息模型。
- Bentkowska-Kafel 和 Macdonald 在探讨文化遗产数字记录与保存技术时指出，
- 数字保存需要关注 " 重要属性 "(Significant Properties) 的识别和维护——即在格式迁移和技术更新过程中，必须确保数字对象的核心特征不被丢失。
- 对于文物三维模型而言，其重要属性包括几何精度、纹理分辨率、色彩还原度和拓扑结构等。

3.2.3 元数据理论

- 元数据 (Metadata) 是 " 关于数据的数据 " ，是数字典藏系统的核心组件。
- 元数据理论将元数据分为三类：描述性元数据 (用于资源发现和识别) 、结构性元数据 (描述资源的内部结构) 和管理性元数据 (用于资源管理和技术保存) 。
- 在文物数字典藏中，元数据不仅需要描述文物的物理特征 (尺寸、材质、年代、来源等) ，还需要描述数字对象的 technical 特征 (文件格式、分辨率、采集设备、处理软件等) 。
- Bonn 等人在探讨图书馆和档案馆与非遗保存的关系时指出，元数据标准的制定是文化遗产数字保存的基础性工作。

03

3.3 文物数字典藏的标准体系

3.3.1 国际标准 (CIDOC

- CRM、Dublin Core、
- LIDO) 文物数字典藏的国际标准体系主要包括以下三个标准：CIDOC CRM(Conceptual Reference Model) 是国际博物馆协会 (ICOM) 下属的 CIDOC 委员会制定
- 的概念参考模型，定义了文化遗产领域的核心概念及其关系。
- CIDOC CRM 提供了一个正式的 ontology，能够精确描述文化遗产对象的 " 什么、何时、何地、谁、如何 " 等维度信息，是目前文化遗产领域最权威的概念模型。
- Dublin Core 是一个简洁的元数据标准，包含 15 个核心元素 (标题、创建者、主题、描述、发布者、贡献者、日期、类型、格式、标识符、来源、语言、关系、覆盖范围、权限)，
- 适用于各类数字资源的描述。
- 其优势在于简洁通用，但对于文物的专业描述能力不足。
- LIDO(Lightweight Information Describing Objects) 是专为博物馆藏品设计的元数据标准，基于 CIDOC CRM 开发，支持多语言、

3.3.1 国际标准 (CIDOC

- 多媒体的藏品信息描述，是博物馆数字化展览和藏品交换的推荐标准。

3.3.2 中国文物数字化标准

- 中国文物数字化标准建设起步较晚但发展迅速。国家文物局先后发布了《文物数字化保护标准体系》《可移动文物数字化保护技术规范》《不可移动文物数字化保护技术规范》等标准文件。
- 邵壮在分析国博三维数字化实践时指出，标准的统一是馆藏文物三维数字化工作的关键前提，不同博物馆采用不同的数据格式和技术标准将严重制约数据的互联互通。

3.3.3 互操作性与数据交换

- 互操作性 (Interoperability) 是文物数字典藏标准体系的核心目标。不同机构、不同系统之间的数据交换需要统一的数据格式、协议和接口。
- Europeana 作为欧洲文化遗产数字平台，通过 Europeana Data Model(EDM) 实现了对来自不同成员国、不同类型机构 (博物馆、图书馆、档案馆) 的异构数据的统一整合。
- 这种跨机构、跨国的数据互操作经验，为中国文物数字典藏的联盟化发展提供了重要借鉴。

04

3.4 文物数字典藏的组织管理

3.4.1 典藏工作流程

- 文物数字典藏的标准工作流程包括以下步骤：(1) 规划阶段——确定数字化对象的优先级、制定采集方案、评估技术需求；(2) 采集阶段——现场数据采集，包括三维扫描、摄影测量、色彩管理等；
- (3) 处理阶段——数据清洗、模型重建、纹理映射、质量检查；(4) 描述阶段——元数据标引、知识关联、目录编制；(5) 入库阶段——数据库入库、存储备份、权限设置；
- (6) 服务阶段——检索、展示、下载、API 接口。
- 欧阳宏在介绍故宫博物院文物三维数据采集实践时指出，故宫的数字化工作流程经过了多年的优化和完善，形成了一套覆盖 " 规划—采集—处理—入库—应用 " 全流程的标准化工作体系。

3.4.2 质量控制与评估

- 质量控制是文物数字典藏的关键环节。
- 三维模型的质量评估指标包括：几何精度（与实物误差范围）、纹理分辨率（像素 / 厘米）、色彩还原度（色差 ΔE 值）、模型完整性（是否覆盖文物全部表面）、拓扑质量（网格质量参数）等。
- 俞朝晖在研究馆藏文物三维数据获取及可视化时指出，文物数字化不仅要记录文物的原始三维信息和纹理信息，还需为文物修缮和恢复提供重要的数据和模型支持，因此质量控制必须贯穿数字化全过程。

3.4.3 长期保存策略

- 数字典藏的长期保存策略是数字保存理论的核心实践。主要策略包括：比特保存 (Bit Preservation)—— 确保数字数据的比特级完整；
- 格式迁移 (Format Migration)—— 将过时格式的数据转换为当前格式；仿真 (Emulation)—— 通过软件模拟旧环境来运行过时的数据格式；
- 封装 (Encapsulation)—— 将数据与其描述信息和运行环境信息打包保存。
- Shimray 和 Ramaiah 在分析数字文化遗产保存的问题时指出，版权和相关权利是文化遗产数字保存的主要法律障碍，数字保存活动需要妥善处理知识产权问题。
- Lor 和 Britz 则从伦理视角分析了数字遗产长期保存中的政治经济问题，强调数字保存不仅是技术问题，更是社会和伦理问题。

05

3.5 文物数字典藏的发展趋势

3.5.1 从数字化到智慧化

- 文物数字典藏正在从 " 数字化 " 阶段向 " 智慧化 " 阶段转型。传统数字典藏以数据采集和存储为核心，而智慧化典藏以知识组织、智能检索和主动服务为核心。
- AI 技术的应用使文物数字典藏具备了自动识别、智能标引、语义检索和个性化推荐等能力。

3.5.2 从单一馆藏到联盟共享

- 文物数字典藏正在从单一机构的封闭系统向多机构联盟的开放平台演进。Europeana 模式展示了跨国文化遗产数据联盟的可行性。
- 在中国，国家文物局正在推进全国文物数字资源平台建设，但跨机构的数据共享仍面临体制机制和技术标准方面的挑战。

3.5.3 从保存导向到服务导向

- 传统数字典藏以 " 保存 " 为首要目标，而当代数字典藏越来越强调 " 服务 " 导向——面向不同用户群体 (研究者、教育者、公众、创意产业) 提供差异化的数据服务。
- 洪闯和李风雪基于 Kano 模型分析了文化遗产数字化服务平台的功能需求，提出了 24 项服务功能项目及其分类，为数字典藏平台的服务优化提供了实证依据。

06

知识卡片

- 知识卡片 3-1： OAIS 参考模型的核心概念 - SIP(Submission Information Package)：提交信息包 - AIP(Archival Information Package)：档案信息包 - DIP(Dissemination Information Package)：分发信息包 - PDI(Preservation Description Information)：保存描述信息
- 知识卡片 3-2： 三种元数据标准对比 | 标准 | 元素数量 | 适用对象 | 优势 | 局限 |

	-----	-----	-----	---
	-----	-----	-----	---
- ----|-----| | Dublin Core | 15 | 通用数字资源 | 简洁、通用 | 文物描述能力不足 | | LIDO | 60+ | 博物馆藏品 | 专为博物馆设计 | 实施复杂度较高 | | CIDOC CRM | 80+ | 文化遗产全域 | 语义丰富、关系完整 | 学习成本高 |
- 知识卡片 3-3： 文物数字典藏质量评估指标 | 指标 | 含义 | 参考标准 |

几何精度	模型与实物的尺寸误差	$\leq 0.1\text{mm}$ (高精度)	纹理分辨率	单位面积的像素数	$\geq 300\text{dpi}$	色彩还原度	色差 ΔE 值	$\Delta E \leq 3$ (专业级)	模型完整性	表面覆盖率	$\geq 95\%$
------	------------	---------------------------	-------	----------	----------------------	-------	-----------------	-------------------------	-------	-------	-------------

07

3.6 数字典藏的技术经济学分析

3.6.1 数字化成本构成

- 文物数字典藏的成本构成复杂，主要包括：设备成本 (三维扫描仪、高分辨率相机、色彩管理设备、计算机等，占 20-30%) ；人力成本 (技术人员、领域专家、项目经理等，占 30-40%) ；
- 时间成本 (采集时间、处理时间、元数据标引时间等，占 20-30%) ；基础设施成本 (存储设备、服务器、网络带宽等，占 10-15%) ；维护成本 (数据迁移、系统升级、安全防护等，占 5-10%) 。
- Manzŭch 指出，数字化项目的高成本使大多数项目依赖外部资金和捐助支持，这种资金模式的可持续性存疑。
- 对于资金有限的中小型博物馆，需要在技术选型和精度要求之间做出权衡——摄影测量是一种成本效益较高的替代方案，虽然精度略低于专业三维扫描，但设备成本仅为后者的 1/5 至 1/10 。

3.6.2 数字化效益评估

- 文物数字典藏的效益可分为直接效益和间接效益。直接效益包括：减少文物搬运和展示带来的物理风险、提高藏品管理效率、支持学术研究。
- 间接效益包括：提升机构品牌影响力、吸引游客和赞助、促进文创产业发展。
- Zorich 在调研数字文化遗产倡议的可持续性时指出，经济挑战是文化遗产数字化项目面临的核心问题，需要建立长期可持续的资金模式。

3.6.3 投资回报分析

- 文物数字典藏的投资回报分析需要考虑时间维度。短期来看，数字化项目通常需要大量前期投入，回报周期较长 (5-10 年)。
- 长期来看，数字资产的价值具有复利效应——随着数据积累和用户增长，数字典藏的价值会持续增长。
- Europeana 的经验表明，数字平台的网络效应可以显著提升长期投资回报，但需要前期持续投入和耐心等待。

08

3.7 国际数字典藏标准比较

3.7.1 元数据标准的适用性分析

- CIDOC CRM、Dublin Core 和 LIDO 三种元数据标准在文物数字典藏中的适用性各有优劣。
- CIDOC CRM 提供了最丰富的语义表达力，能够精确描述文化遗产对象的概念关系，但其复杂性使实施成本较高，适合大型国家级博物馆。
- Dublin Core 简洁通用，易于实施，但对文物的专业描述能力不足，适合小型机构或作为基础描述层。LIDO 专为博物馆设计，在描述深度和实施成本之间取得了较好平衡，是中型博物馆的推荐选择。

3.7.2 数据互操作性的挑战

- 数据互操作性是文物数字典藏面临的核心技术挑战。不同机构使用的元数据标准、数据格式和技术平台各不相同，导致数据交换和整合困难。
- Europeana 通过 EDM 模型和关联数据技术实现了跨国数据互操作，但这一方案的实施需要大量协调和标准化工作。
- 中国文物数字典藏的互操作性问题更为突出——各省级博物馆和文物机构采用不同的系统，缺乏统一的数据交换标准，形成了大量 "数据孤岛"。

3.7.3 标准化发展路径

- 文物数字典藏的标准化发展路径应遵循 " 求同存异、分层实施 " 原则：基础层采用通用标准 (如 Dublin Core) ，确保最低限度的互操作性；
- 专业层采用行业标准 (如 LIDO) ，满足文物描述的专业需求；语义层采用概念模型 (如 CIDOC CRM) ，实现深度语义互操作。这种分层策略可以在保证互操作性的同时，兼顾不同机构的实施能力。

09

3.8 数字典藏的社会维度

3.8.1 开放获取运动与文物数字典藏

- 开放获取 (Open Access) 运动对文物数字典藏产生了深远影响。
- 荷兰国立博物馆的 CC0 开放政策、大都会博物馆的开放获取倡议、史密森尼学会的 3D 模型开放下载等实践，推动了博物馆界从 " 封闭保护 " 向 " 开放共享 " 的范式转变。
- Wallace 和 Euler 在分析公共领域文化遗产的访问权时指出，欧盟在开放获取方面走在了世界前列，通过修订版权法为文化遗产机构的数字化和开放提供了法律保障。然而，开放获取也面临挑战。
- Shimray 和 Ramaiah 指出，版权和相关权利是文化遗产数字保存的主要法律障碍——许多文物的数字扫描是否产生新的版权尚存争议，
- " 孤儿作品 "(无法确定版权人的作品) 的数字化和开放更是法律灰色地带。
- Lykourantzou 和 Antoniou 进一步指出，知识产权问题是文化遗产数字创新可持续性的潜在阻碍点。

3.8.2 数字典藏与文化民主化

- 数字典藏被视为文化民主化的重要工具——通过将原本封闭在库房中的文物数字化并向公众开放，使更多人能够接触和利用文化遗产。
- 然而，Taylor 和 Gibson 指出，数字化互动和社交媒体虽然拓展了遗产的可及性，但技术门槛、设备差异和网络条件等因素形成了新的民主化障碍。
- 数字鸿沟使部分群体被排除在数字化进程之外，反而加剧了文化不平等。

3.8.3 数字典藏的社区参与

- 数字典藏的社区参与是文化遗产保护伦理的重要维度。Manzŭch 指出，数字化项目的伦理要求包括尊重来源社区的文化传统和意愿、确保社区知情同意、实现利益分享。
- 在实践中，数字典藏项目应建立社区咨询机制，在项目规划、实施和评估各阶段纳入社区参与。

10

3.9 数字典藏与学术研究

3.9.1 数字典藏对学术研究的影响

- 数字典藏深刻改变了文化遗产学术研究的方式。传统研究中，学者需要亲赴各博物馆查看实物，受限于开放时间和物理距离。数字典藏使学者可以随时随地访问高清数字资源，大幅提高了研究效率。
- 更重要的是，数字典藏支持计算分析——三维模型的精确测量、大规模图像的对比分析、跨馆藏品的关联发现，这些都是传统实物研究难以实现的。
- De Boer 指出，将遗产数据表示为知识图谱不仅支持语义检索和可视化分析，还能为机器学习和知识发现提供结构化数据基础。
- Fiorucci 等人的综述显示，机器学习在文化遗产领域的应用已从实验室研究发展到实际部署，覆盖了从图像分类到知识发现的完整链路。

3.9.2 开放数据与学术共享

- 开放数据运动正在改变文化遗产学术研究的生态。越来越多的博物馆和档案机构将其数字典藏以开放数据形式发布，使学者可以自由下载和分析数据。
- Europeana 的开放数据政策、大英博物馆的开放馆藏数据库、史密森尼学会的开放 3D 模型等实践，为文化遗产学术共享奠定了数据基础。然而，开放数据也面临挑战。
- 数据质量参差不齐、元数据标准不统一、版权限制等因素制约了数据的可重用性。
- Lee 提出的 "Collections as ML Data" 检查清单为文化遗产数据的机器学习应用提供了标准化框架，强调数据文档化、质量保证和偏见检测的重要性。

11

3.10 数字典藏的文化民主化功能

3.10.1 从 " 精英收藏 " 到 " 公众共享 "

- 博物馆的藏品传统上被视为 " 精英收藏 "——只有少数学者和收藏家能够接触。数字典藏正在打破这一传统，将原本封闭在库房中的文物以数字形式向全球公众开放。
- 这种 " 从精英收藏到公众共享 " 的转型具有深远的社会意义——它使文化遗产真正成为 " 全民遗产 "，而非少数人的特权。
- Wallace 和 Euler 在分析公共领域文化遗产的访问权时讨论了欧盟和国际层面的开放获取发展动态。他们指出，开放获取不仅是技术问题，更是文化权利问题——每个人都有平等获取文化遗产信息的权利。

3.10.2 数字典藏与教育公平

- 数字典藏为教育公平提供了新路径。偏远地区学校可以通过数字博物馆访问世界顶级藏品，经济条件有限的学生可以通过网络获取高质量的文化遗产教育资源。
- 然而，数字鸿沟的存在使这一路径的实现面临挑战——网络条件差、设备不足、数字素养不够等因素制约了数字典藏教育功能的发挥。

3.10.3 参与式数字典藏

- 参与式数字典藏 (Participatory Digital Archiving) 是数字典藏的新发展方向。
- 公众不再仅是数字典藏的被动使用者，而是主动参与数字典藏的建设——通过众包方式为藏品添加标签、转录手写文档、贡献口述历史。
- Díaz 等人在研究增强展览促进文化遗产地参与式文化时指出，数字增强展览能够培养公众对文化遗产的参与式文化。

12

3.11 数字典藏的未来形态与智能化发展

3.11.1 智慧典藏的概念

- 智慧典藏 (Smart Archiving) 是数字典藏的下一代形态，它将 AI 技术深度嵌入典藏的全流程——从自动标引到智能检索到主动推荐。
- 传统数字典藏依赖人工标引和关键词检索，智慧典藏则利用 AI 自动识别文物特征、生成元数据、建立语义关联，使典藏系统能够 " 理解 " 而非仅 " 存储 " 文物信息。
- Wang 等人在《文化遗产智能计算》中系统阐述了智能计算在文化遗产中的应用框架。

3.11.2 联邦学习与隐私保护

- 联邦学习 (Federated Learning) 技术为多个机构联合训练 AI 模型而不共享原始数据提供了技术方案。
- 在文化遗产领域，不同博物馆可以在不共享藏品数据的前提下联合训练文物识别模型，既提升了模型性能又保护了数据主权。这一技术对于解决文化遗产数字化中的 "数据孤岛" 问题具有重要意义。

3.11.3 数字孪生典藏

- 数字孪生技术将数字典藏从 " 静态档案 " 升级为 " 动态镜像 "。传统数字典藏记录的是文物的某一时刻状态，数字孪生典藏则通过传感器网络实时监测文物状态变化，使典藏系统能够动态反映文物的当前状况。
- 这对于文物预防性保护具有重要意义——管理者可以实时了解文物的温湿度环境、结构变形和病害发展情况，及时采取保护措施。

13

3.6 文物数字典藏的学科基础深化

3.6.1 信息科学与数字典藏的交叉

- 文物数字典藏是一个典型的跨学科领域，其理论基础涉及信息科学、博物馆学、档案学、计算机科学等多个学科。
- 从信息科学视角，文物数字典藏可纳入 " 信息生命周期 "(Information Lifecycle) 的理论框架加以理解。
- 信息生命周期理论将信息运动过程划分为创建 / 采集、处置 / 描述、存取 / 使用、维护 / 保存和处置 / 淘汰五个阶段。
- 在文物数字典藏语境下，这一过程对应为：文物数字化采集 (创建)、元数据著录与编目 (描述)、数字典藏平台检索与访问 (使用)、数字保存与迁移 (维护)、数据归档或淘汰 (处置)。
- 每个阶段都涉及特定的技术标准和质量管理要求。
- Hodge 和 Anderson 提出的 " 数字典藏四属性 "—— 内容 (Content)、情境 (Context)、
- 结构 (Structure) 和保存 (Preservation)—— 为文物数字典藏的质量评估提供了有效框架。
- 内容属性关注数字资产本身的完整性和准确性；情境属性关注数字资产的来源、创建条件和关联信息；结构属性关注数字资产的组织方式和内在逻辑；保存属性关注数字资产的长期可读性和可用性。

3.6.2 档案学理论对数字典藏的启示

- 档案学中的 " 来源原则 "(Principle of Provenance) 和 " 原始顺序 "(Original Order) 对文物数字典藏具有重要启示。
- 来源原则要求数字典藏在组织文物信息时保持其来源的完整性，同一考古遗址出土的文物应在数字典藏中保持其出土地、出土层位等来源信息的关联。
- 原始顺序原则要求尊重文物在其原始使用场景中的组织方式，如一套礼仪用器应作为一个整体著录，而非分散为单独条目。
- InterPARES 项目 (真实性数字记录的永久保护国际研究项目) 对数字记录的真实性、可靠性和准确性进行了深入研究，其提出的 " 真实数字记录 " 标准对文物数字典藏具有重要参考价值。
- 根据 InterPARES 的框架，数字典藏中的文物信息必须满足以下条件才能被视为 " 真实的 "：(1) 创建者身份可追溯；(2) 创建过程可复现；(3) 创建后的修改记录完整；(4) 技术环境可迁移。

14

3.7 元数据标准的比较分析

3.7.1 国际元数据标准体系

- 文物数字典藏领域的元数据标准形成了多层次体系：通用元数据标准：Dublin Core(DC) 是应用最广泛的通用元数据标准，定义了 15 个核心元素 (标题、创建者、主题、描述、出版者、贡献者、
日期、类型、格式、标识符、来源、语言、关联、覆盖范围、权限)。
- DC 的优势在于简单易用、跨领域兼容，但语义粒度较粗，对文物的专业信息描述不足。
- 领域专用元数据标准：CIDOC CRM(概念参考模型) 是博物馆领域的国际标准 (ISO 21127) ，定义了 83 个实体类和 281 种属性 / 关系，能够精确描述博物馆藏品的复杂信息。
- LIDO(Lightweight Information Describing Objects) 是 CIDOC CRM 的 XML 序列化格式，专为博物馆藏品的网络交换设计。
- CDWA(艺术信息分类目录) 是艺术类藏品的详细元数据标准，定义了 532 个元素和子元素。
- 技术元数据标准：用于描述数字资源的技术属性，如 PREMIS(保存元数据) 、 MIX(图像技术元数据) 、 TEXTMD(文本技术元数据) 等。
- 这些标准记录了数字资源的技术环境、格式参数和质量指标，对长期保存至关重要。

3.7.2 中国文物元数据标准的发展

- 中国文物元数据标准建设经历了从借鉴国际标准到自主发展的过程：早期阶段 (2000-2010)：主要借鉴 Dublin Core 和 CIDOC CRM，
- 如国家文物局 2003 年发布的《博物馆藏品信息指标体系》参考了 Dublin Core 的框架，但增加了文物特有的指标 (如年代、质地、完残状况等)。
- 发展阶段 (2010-2020)：逐步形成中国特色的标准体系，如《文物数字化保护标准体系建设规划》(2017 年) 提出了覆盖采集、存储、管理、利用全流程的标准框架。
- WW/T 0066-2015 《文物数字化保护技术规范》系列标准对各类文物的数字化采集提出了具体要求。深化阶段 (2020 年至今)：标准建设向精细化、智能化方向发展。
- 如 2022 年发布的《文物数字化成果元数据规范》参考了 LIDO 框架，但增加了中国文物特有的分类体系 (如一级文物、二级文物、三级文物的分级标准) 和管理信息。中国标准与国际标准的互操作仍存在挑战。
- 主要差异包括：(1) 分类体系不同——中国采用文物法定义的分类 (可移动文物 / 不可移动文物)，国际标准采用材质 / 功能分类；(2) 年代体系不同——中国使用朝代纪年体系，国际标准使用公元纪年；
- (3) 权属描述不同——中国文物多为国有，国际标准需兼容多种所有权形式。

15

3.8 文物数字典藏的平台架构与互操作

3.8.1 分布式数字典藏架构

- 当代文物数字典藏平台正从集中式架构向分布式架构转型。集中式架构将所有文物数字资源存储和管理在单一系统中，虽然管理简便，但面临扩展性瓶颈、单点故障风险和数据主权问题。
- 分布式架构通过联邦式数据管理，使各机构保留对自身数据的控制权，同时支持跨机构检索和数据交换。
- 联邦检索架构：各文博机构维护独立的数字典藏系统，通过统一的检索协议（如 OAI-PMH、SRU/CQL）实现跨机构检索。
- 用户提交检索请求后，系统向所有参与机构的数据库发送查询，汇总并去重后返回结果。Europeana 即采用联邦检索架构，聚合了来自欧洲 4000 余家文化机构的数字资源。
- 数据聚合架构：各机构将元数据（不包含原始数据文件）提交到中心聚合平台，用户在聚合平台检索后，通过链接跳转到源机构访问原始数据。这种模式降低了对中心平台的存储压力，同时确保了源机构的数据主权。
- 区块链联邦架构：利用区块链的分布式账本技术，建立去中心化的数字典藏联盟。各机构的数字资产指纹（哈希值）存储在区块链上，确保数据完整性和可溯源性。
- 这种架构尚处于实验阶段，但代表了数字典藏互操作的前沿方向。

3.8.2 语义互操作技术

- 语义互操作要解决的是 " 不同系统对同一概念使用不同术语 " 的问题。核心技术手段包括：本体对齐 (Ontology Alignment)：在不同本体之间建立概念映射关系。
- 如将中国文物分类体系中的 " 青铜器 " 映射到 AAT(Art & Architecture Thesaurus) 中的 "bronze(visual works)"，使不同系统的数据可互相理解。
- 链接数据 (Linked Data)：采用 RDF 三元组格式描述文物信息，通过 URI(统一资源标识符) 实现概念的全局唯一标识和互链接。
- 如大英博物馆将其馆藏数据发布为链接数据，与 DBpedia、GeoNames 等外部数据源建立了语义关联。知识图谱融合：将多个机构的知识图谱进行实体对齐和关系融合，形成更大规模的统一知识图谱。
- 如 ARIADNEplus 项目将欧洲多个考古数据集的知识图谱融合为统一的考古知识网络。

16

3.9 文物数字典藏的智慧化转型

3.9.1 从 " 数字化 " 到 " 智慧化 " 的演进路径

- 文物数字典藏正在经历从 " 数字化 " 到 " 智慧化 " 的转型。
- 这一转型的核心标志是从 " 数据存储与管理 " 转向 " 知识发现与服务创新 "：
 - 数据层智慧化：利用 AI 技术自动提取文物的特征信息 (纹样、器型、材质、工艺标记)，生成丰富的描述性元数据，
 - 减少人工著录的工作量。
 - 如基于深度学习的青铜器纹饰自动识别系统，能自动识别饕餮纹、夔龙纹、云雷纹等典型纹饰，准确率达到 90% 以上。
- 知识层智慧化：通过知识图谱将分散的文物信息关联为结构化知识网络，支持语义检索、关联推荐和知识推理。
- 如 " 中国古代铜镜知识图谱 " 将铜镜的形制、纹饰、铭文、出土地、年代等信息关联为统一知识网络，支持用户从任意维度进行探索。
- 服务层智慧化：基于用户画像和知识推理，提供个性化的文物知识服务。
- 如智能导览系统根据用户的兴趣偏好和知识水平，自动生成定制化的参观路线和讲解内容。

3.9.2 智慧博物馆的概念与实践

- "智慧博物馆"(Smart Museum) 是数字典藏智慧化在博物馆领域的具体体现。
- 国际博物馆协会 (ICOM) 将智慧博物馆定义为 " 利用物联网、大数据、云计算、人工智能等技术, 实现藏品管理、展览展示、观众服务和场馆运营的全面智能化的博物馆 "。
- 智慧博物馆的核心特征包括: (1) 全面感知——通过传感器网络实时感知藏品状态、环境参数和观众行为; (2) 互联互通——各信息系统之间数据互通, 形成统一的数据平台;
- (3) 智能决策——基于数据分析自动优化运营决策; (4) 个性化服务——为每位观众提供定制化的参观体验。
- 中国智慧博物馆建设的典型案例包括: 故宫博物院的 " 数字故宫 " 智慧平台、上海博物馆的 " 智慧上博 " 系统、南京博物院的 " 智慧南博 " 项目等。
- 这些项目在藏品管理数字化、观众服务智能化和场馆运营数据化方面进行了积极探索。

17

3.10 文物数字典藏的数据治理与质量管理

3.10.1 数据治理框架

- 文物数字典藏的数据治理需要建立覆盖数据全生命周期的管理框架：数据所有权治理：明确数字典藏中各类数据的所有权归属——文物的数字扫描数据所有权属于博物馆还是数字化服务提供商？
- 元数据的著录权归属是机构还是著录者？这些问题需要通过制度安排明确界定。数据访问治理：建立分级访问控制机制——基础元数据和低分辨率预览图对所有用户开放；高清图像和 3D 模型对注册用户开放；
- 原始采集数据（如点云数据）仅对授权研究者开放。
- 访问控制应基于角色 (Role-Based Access Control, RBAC) 和属性 (Attribute-Based Access Control, ABAC) 相结合的方式。
- 数据质量治理：建立数据质量评估和改进机制。
- 质量评估维度包括：完整性（必填字段是否齐全）、准确性（数据内容是否正确）、一致性（不同系统间的数据是否一致）、及时性（数据是否及时更新）、唯一性（是否存在重复记录）。
- 质量改进通过数据清洗、数据补录和数据审核等手段实现。数据生命周期治理：制定数据保留和处置策略——哪些数据需要永久保存？哪些可以定期清理？过时数据如何迁移或归档？数据删除需要遵循什么审批流程？

3.10.2 FAIR 原则在文物数字典藏中的应用

- FAIR 原则 (Findable, Accessible, Interoperable, Reusable) 是国际科学界提出的数据管理指导原则，
- 对文物数字典藏具有重要参考价值： 可发现 (Findable)： 数字资源应能被人 and 机器容易发现。
- 实现方式包括： 为每个数字资源分配持久性标识符 (如 DOI、 ARK) ； 提供丰富的元数据支持多维度检索； 在搜索引擎和聚合平台中注册资源。
- 可访问 (Accessible)： 数字资源应能通过标准协议访问。实现方式包括： 提供 RESTful API 和 OAI-PMH 接口； 支持 HTTP 协议访问； 明确访问条件和授权流程。
- 可互操作 (Interoperable)： 数字资源应能与其他系统交换和使用数据。实现方式包括： 使用国际标准元数据格式 (如 Dublin Core、 LIDO) ； 提供本体映射和概念对齐；
- 采用开放数据格式。可重用 (Reusable)： 数字资源应能被重新利用。实现方式包括： 提供清晰的使用许可 (如 Creative Commons 许可) ； 记录数据来源和处理历史 (溯源信息) ；
- 提供足够详细的数据文档。

18

3.11 文物数字典藏的组织管理深化

3.11.1 数字典藏团队的能力建设

- 文物数字典藏工作需要复合型团队，核心角色包括：数字典藏主管：负责数字典藏的整体规划、标准制定和质量监督。需要兼具文化遗产领域知识和数字技术理解力，是连接文博专业和 IT 技术的桥梁。
- 数字化工程师：负责数字化采集的技术实施，包括 3D 扫描、摄影测量、多光谱成像等。需要掌握多种数字化设备的操作和数据处理软件的使用。
- 元数据专家：负责元数据标准的制定和实施，确保数字资源的描述信息准确、完整和标准化。需要熟悉国际元数据标准 (CIDOC CRM、Dublin Core 等) 和文物专业知识。
- 数据管理员：负责数字资源的存储管理、数据备份、系统维护和用户权限管理。需要具备数据库管理和网络安全的技能。
- 数字策展人：负责数字资源的在线展示和虚拟展览策划，将数字典藏转化为公众可理解和欣赏的数字内容。需要兼具策展能力和数字媒体设计能力。

3.11.2 数字典藏项目的管理流程

- 数字典藏项目的管理应遵循系统化的流程：规划阶段：(1) 需求评估——明确数字化的范围（哪些文物需要数字化）、优先级和目标用途；(2) 资源评估——评估可用的人员、设备、资金和时间；
(3) 技术选型——根据文物类型和精度要求选择数字化技术方案；(4) 标准制定——制定采集、处理、存储和管理的标准规范。
- 实施阶段：(1) 试点先行——选择代表性文物进行试点数字化，验证方案可行性；
(2) 批量实施——在试点成功的基础上推广到批量数字化；(3) 质量控制——每批次进行质量检查和抽样验收；(4) 进度管理——跟踪项目进度，及时调整资源配置。
- 运维阶段：(1) 数据更新——定期更新元数据和数字资源；(2) 系统维护——维护数字典藏平台的运行和升级；(3) 用户支持——为研究者、教育者和公众提供数据访问支持；
(4) 持续改进——根据用户反馈和技术发展持续优化系统和服务。

19

3.12 文物数字典藏的经济分析

3.12.1 成本效益分析框架

- 文物数字典藏项目的经济评估可采用成本效益分析框架：直接成本：数字化采集成本 (设备、人员、时间)、平台开发成本 (软件开发、系统集成)、运营维护成本 (服务器、带宽、人员)、质量控制成本 (质检、审核、返工)。
- 直接效益：数字化服务的直接收入 (图像授权费、定制服务费)、效率提升 (减少实体文物的频繁取出和搬运、降低物理损害风险)。
- 间接效益：学术研究促进 (数字资源支持更多学术研究)、文化旅游引流 (数字展示激发实地参观意愿)、文创产业带动 (数字资源为文创开发提供素材)、品牌价值提升 (数字化增强机构的品牌形象和社会影响力)。
- 社会效益：文化教育促进 (数字资源支持远程教育)、文化认同增强 (数字传播扩大文化影响力)、保护意识提升 (数字化使更多人关注文化遗产保护)。
- 由于文化遗产的公共物品属性，其社会效益远大于直接经济收益。因此，数字典藏项目的经济评估不能仅以直接投资回报率为标准，而应纳入社会效益的综合评估。

20

3.13 数字典藏的未来图景

3.13 数字典藏的未来图景

- 展望未来，文物数字典藏将呈现以下发展趋势：从机构典藏到社会典藏：传统数字典藏以博物馆为核心机构，未来将向更开放的社会化模式发展。
- 社区数字典藏、个人数字收藏、公民科学项目等将丰富数字典藏的来源和视角，形成多元主体共同参与的文化遗产数字资源建设生态。
- 从数据典藏到知识典藏：当前数字典藏以 " 数据 " 为核心——存储和管理文物的数字文件。
- 未来将向 " 知识典藏 " 转型——不仅存储数据，还通过知识图谱和 AI 技术将数据关联为结构化知识网络，支持智能检索、知识发现和推理。
- 从孤岛典藏到联邦典藏：各机构的数字典藏系统将通过标准化接口和互操作协议实现互联互通，形成联邦式的数字典藏网络。用户可在单一入口检索和访问全球数字典藏资源，打破机构壁垒。
- 从静态典藏到动态典藏：数字典藏将从 " 一次采集、永久存储 " 的静态模式转向 " 持续更新、动态演化 " 的模式。文物的数字孪生体将随物理实体的变化而实时更新，确保数字典藏与物理遗产保持同步。

21

3.14 延伸阅读与思考

3.14 延伸阅读与思考

- 延伸阅读：建议阅读 CIDOC CRM 官方文档 (ISO 21127:2014) ，深入理解文化遗产信息模型的设计逻辑。
- 同时推荐阅读 InterPARES 项目的研究成果，了解数字记录真实性的理论基础。拓展思考：数字典藏的 " 典藏 " 一词暗含了 " 权威性 " 和 " 永久性 " 的意涵。
- 但数字技术本质上是易变和脆弱的——文件格式会被淘汰，存储介质会老化，软件系统会过时。如何在技术的不确定性中追求典藏的确定性？这是数字典藏领域需要持续面对的根本性挑战。
- OAIS 参考模型提供了技术层面的框架，但更深层的答案是制度层面的——建立持续投入、定期迁移、多方备份的制度保障，使数字典藏的 " 永久性 " 建立在制度韧性而非技术不变性之上。
- 教学建议：本章可采用项目式学习法，让学生分组设计一个小型文物数字典藏系统的方案 (选择一个博物馆的 5-10 件藏品) ，从元数据设计、数据库选型、存储策略到用户界面进行完整规划，
- 通过实践理解数字典藏的复杂性。

22

3.15 关键技术辨析

3.15 关键术语辨析

- 在数字典藏领域，以下术语容易混淆，
- 需要明确辨析：数字化 (Digitization) 与数字转型 (Digital Transformation)：数字化指将模拟信息转换为数字格式的过程 (如扫描纸质文献为 PDF)，
- 是技术层面的操作。
- 数字转型指机构利用数字技术全面变革业务模式、组织结构和服务方式的系统性过程，是战略层面的变革。数字典藏是数字化的核心内容，但数字典藏的深层目标应服务于机构的数字转型。
- 数字保存 (Digital Preservation) 与数字归档 (Digital Archiving)：数字保存强调确保数字资源的长期可读性和真实性，涉及格式迁移、数据校验、灾备等技术措施。
- 数字归档强调将数字资源按规范存储和管理，涉及分类、编目、检索等管理措施。数字保存是数字归档的技术基础，数字归档是数字保存的管理框架。
- 元数据 (Metadata) 与本体 (Ontology)：元数据是描述数字资源的数据 (如标题、作者、日期)，通常以扁平化的字段形式存在。
- 本体是描述特定领域概念体系的形式化模型 (如 CIDOC CRM)，定义了概念之间的层次和关系。元数据是本体的实例化表达，本体是元数据的理论框架。

3.15 关键术语辨析

- 此外，数字典藏的发展还需要关注 " 暗数据 "(Dark Data) 问题——即机构内部已采集但未发布、未编目或不可检索的数字资源。
- 据估算，全球文博机构中约 60% 的数字资源处于 " 暗数据 " 状态，这些资源的学术和文化价值尚未被充分释放。激活暗数据需要投入编目和发布资源，但这往往是资金紧张时最先被削减的环节。
- 建立可持续的数字资源发布机制，是数字典藏从 " 保存导向 " 走向 " 服务导向 " 的关键一步。
- 总之，文物数字典藏是一个持续演进的过程，需要在技术、标准、管理和制度多个层面协同推进，才能实现文化遗产数字资源的长期保存和有效利用。
- 通过建立完善的治理框架和持续投入机制，数字典藏将为文化遗产保护提供坚实的数据基础和技术支撑。

- 本章系统阐述了文物数字典藏的理论基础与标准体系。首先界定了文物数字典藏的概念范畴，区分了数字典藏与文物数字化、数字图书馆的概念边界；
- 其次梳理了信息生命周期理论、数字保存理论和元数据理论三大理论基础，其中 OAIS 参考模型为数字典藏系统设计提供了功能架构；
- 然后详细介绍了 CIDOC CRM、Dublin Core、LIDO 三种国际元数据标准以及中国文物数字化标准的发展现状，分析了互操作性的核心挑战；
- 接着从工作流程、质量控制和长期保存三个方面阐述了数字典藏的组织管理；最后指出了从数字化到智慧化、从单一馆藏到联盟共享、从保存导向到服务导向三大发展趋势。
- 本章对应 G.U.A.R.D. 框架的 A 维度（数字典藏），为后续技术章节提供了理论支撑。

- 何区别？ 2. (理解) 试比较信息生命周期理论与数字保存理论在关注焦点上的异同。 3. (应用) 运用 CIDOC CRM 的概念框架，为一个青铜器文物设计元数据描述方案。
- 4. (分析) 分析 Dublin Core、LIDO 和 CIDOC CRM 三种标准在文物描述能力上的梯度差异。
- 5. (评价) 评估 "从保存导向到服务导向" 转型对文物数字典藏机构的影响，讨论其利弊。
- 6. (创造) 为一个县级博物馆设计一个分阶段的文物数字典藏系统建设方案，涵盖技术选型、标准采用、工作流程和人员配置。

- 1. Caciora T, Jubran A, Ilies D C, et al.. Digitization of the Built Cultural He
- 2. Espina-Romero L, Guerrero-Alcedo J. Fields Touched by Digitalization: Analysi
- 3. Cheng H M, Yen Y, Yang W B. Digital Archiving in Cultural Heritage Preservati
- 4. Graham C A. Applications of digitization to museum collections management, re
- 5. Wachowiak M J, Karas B V. 3D Scanning and Replication for Museum and Cultural
- 6. 欧阳宏 . 故宫院藏文物的三维数据采集与应用 . 万方数据 .

谢谢聆听 · 下章再见

安顺学院旅游管理学院 · 数字文旅专业系列教材