

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5章 三维重建与数字典藏体系

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

- 5.1 三维重建技术
- 5.2 数字典藏库构建
- 5.3 知识图谱与语义关联
- 5.4 数字典藏的长期保存
- 5.5 典型数字典藏平台分析
- 知识卡片
- 5.6 知识图谱在文化遗产中的深度应用
- 5.7 数字典藏平台的架构演进
- 5.8 数字典藏的安全与隐私
- 5.9 数字典藏平台的典型案例深度分析
- 5.10 数字典藏的长期保存策略
- 5.11 数字典藏的评价与认证
- 5.6 三维重建技术的前沿进展与文化遗产应用
- 5.7 数字典藏的长期保存策略深化

- 1. 识记：三维重建的核心流程 (点云处理→网格重建→纹理映射) 和数字典藏库的架构组件
- 2. 理解：知识图谱在文化遗产语义关联中的作用机制与构建方法
- 3. 应用：运用 OAIS 模型设计一个数字典藏系统的保存策略
- 4. 分析：比较 Europeana 、数字敦煌和中国国家博物馆三个数字典藏平台的技术架构与服务模式
- 5. 评价：评估当前文化遗产知识图谱构建的技术成熟度与应用瓶颈
- 6. 创造：设计一个面向地方非遗的数字典藏平台功能架构

- 故宫博物院拥有 186 万余件 (套) 藏品和占地 72 万平方米的古建筑群。
- 十多年来, 故宫博物院使用三维数字技术, 采集、加工了大量古建筑和藏品文物的高精度三维数据, 利用数字呈现技术立体再现文化遗产的原貌。
- 故宫的数字典藏系统不仅服务于文物保护和学术研究, 还支撑了 " 数字文物库 " " 数字多宝阁 " " 全景故宫 " 等一系列面向公众的数字化产品, 形成了从典藏到展示到传播的完整链条。

01

5.1 三维重建技术

5.1.1 点云数据处理

- 点云数据处理是三维重建的第一步，主要包括配准、去噪、抽样和补洞四个环节。
- 配准 (Registration) 是将多站扫描获取的点云数据统一到同一坐标系下的过程，常用的算法包括 ICP(Iterative Closest Point) 算法及其变体。
- 去噪 (Denoising) 是去除点云中由环境干扰或设备误差导致的异常点。抽样 (Sampling) 是在保证模型精度的前提下减少点云数据量，提高后续处理效率。
- 补洞 (Hole Filling) 是根据周边点云数据推断缺失区域的三维形貌。
- 朱黎等人在研究文化遗产多分辨率三维重建时，提出了一种基于点云和高清影像数据的融合方法，实现了文物模型和遗址模型的融合展示，在世界文化遗产唐崖土司城遗址的虚拟展示中取得了良好效果。

5.1.2 网格重建与纹理映射

- 网格重建 (Mesh Reconstruction) 是将点云数据转换为连续的三角网格模型的过程。常用算法包括 Delaunay 三角化、Poisson 重建和隐式曲面重建等。
- 网格重建需要兼顾模型精度和网格质量 (三角形形状规整、分布均匀)。纹理映射 (Texture Mapping) 是将高清照片的纹理信息映射到三维网格模型表面的过程。
- 高质量的纹理映射需要解决相机标定、影像配准、色彩管理和接缝处理等技术问题。Barszcz 等人在研究虚拟博物馆三维扫描数字模型时指出, 纹理质量直接影响三维模型的视觉效果和应用价值。

5.1.3 多分辨率模型融合

- 对于大型文化遗产 (如古建筑群、考古遗址)，不同区域可能需要不同精度的三维模型。
- 多分辨率模型融合技术通过 LOD(Level of Detail) 策略，在近景时加载高精度模型，远景时加载低精度模型，实现渲染效率与视觉质量的平衡。
- He 和 Jia 在研究文物 3D 数字化时提出了 Opti3D 方法，解决了低光照环境下的三维数字化和免校准问题，为文化遗产的灵活数字化提供了新技术路径。

02

5.2 数字典藏库构建

5.2.1 数字典藏库架构设计

- 文物数字典藏库的典型架构包括以下层次：数据层（存储三维模型、图像、音视频、元数据等）、服务层（提供检索、浏览、下载、API 等数据服务）、应用层（数字博物馆、虚拟展览、学术研究等应用）、
- 安全层（权限管理、数据加密、日志审计）。

5.2.2 元数据标准与管理

- 元数据管理是数字典藏库的核心功能。文物数字典藏通常采用多标准融合的元数据策略：使用 Dublin Core 进行基础描述，使用 LIDO 进行专业藏品管理，使用 CIDOC CRM 进行语义关联。
- 中国甲午战争博物院在研究博物馆文物数字化保护与管理时指出，元数据管理的规范化是数字化保护工作可持续发展的基础。

5.2.3 数据库技术选型

- 文物数字典藏的数据库选型需要考虑数据类型多样性 (关系型数据 + 三维模型 + 多媒体文件)、查询复杂性和扩展性。
- 传统关系型数据库 (如 PostgreSQL) 适用于元数据管理，对象存储系统适用于大文件存储，图数据库 (如 Neo4j) 适用于知识图谱的语义查询。

03

5.3 知识图谱与语义关联

5.3.1 文化遗产知识图谱构建方法

- 知识图谱是文化遗产数字典藏向智慧化发展的关键技术。
- 任明和贾君枝在研究数字人文领域知识图谱构建时，提出了基于协作智能的 PDC 工作模式，分析了人与 AI 在知识图谱构建中的分工协作，指出协作智能逐渐由人类主导转向 AI 主导。
- De Boer 进一步探讨了知识图谱在文化遗产和数字人文中的应用，指出知识图谱能够支持遗产数据的语义检索、可视化分析和机器推理。

5.3.2 实体识别与关系抽取

- 文化遗产知识图谱的构建需要解决实体识别 (识别人物、地点、时间、事件、文物等实体) 和关系抽取 (识别实体间的语义关系, 如 " 创作于 " " 出土于 " " 属于 " 等) 两个核心任务。
- Fiorucci 等人对文化遗产机器学习的综述中指出, 机器学习在文物图像识别与分类、实体识别与关系抽取等方面已展现出显著成效。

5.3.3 语义检索与推理

- 知识图谱使文化遗产数据从 " 关键词检索 " 升级为 " 语义检索 "—— 用户可以基于概念关系进行查询，如 " 查找所有明代景德镇窑口的青花瓷器 "。
- 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。

04

5.4 数字典藏的长期保存

5.4.1 数字保存策略 (OAIS 模型)

- OAIS 模型定义了数字保存的完整功能架构，包括摄入、存储、管理、保存规划和访问五个模块。
- 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，数字保存策略的制定需要综合考虑技术、法律和管理三个维度。

5.4.2 数据迁移与格式转换

- 随着文件格式和技术标准的演进，数字典藏数据需要定期进行格式迁移以确保长期可读性。格式迁移需要遵循 " 重要属性保持 " 原则——确保迁移过程不丢失数字对象的核心特征。

5.4.3 灾备与容灾体系

- 数字典藏的灾备体系应包括本地备份、异地备份和云端备份三重保障。Telek 等人在研究 3D 数字重建的存档时强调，多机构协作的灾备策略能够显著提高文化遗产数字资产的安全性。

05

5.5 典型数字典藏平台分析

5.5.1 Europeana 欧洲数字图书馆

- Europeana 是欧洲文化遗产数字平台，整合了来自 4000 余家文化机构的 5800 万件数字资源。
- Borin 和 Donato 在研究 Europeana 的财务可持续性时指出，Europeana 的运营模式依赖欧盟委员会的公共资金支持，正面临着从公共资助向混合收入模式转型的挑战。

5.5.2 数字敦煌资源库

- 数字敦煌资源库是全球最大的石窟壁画数字资源库，其技术架构以高分辨率图像存储和 WebGL 渲染为核心，支持全球用户免费浏览和学术研究使用。
- 陈振旺和樊锦诗在分析敦煌数字化实践时指出，文化科技融合是文化遗产保护的重要路径。

5.5.3 中国国家博物馆数字典藏系统

- 中国国家博物馆的数字典藏系统以三维数字化为核心，建立了馆藏文物的三维数字档案，支撑 " 智慧国博 " 建设。

06

知识卡片

- 知识卡片 5-1：知识图谱构建流程 1. 本体设计 (定义概念体系和关系类型) 2. 数据采集 (从结构化 / 非结构化数据源获取信息) 3. 实体识别 (识别人物、地点、文物等实体) 4. 关系抽取 (识别实体间的语义关系) 5. 知识融合 (消除冲突、去重对齐) 6. 质量评估 (完整性、一致性、准确性检查) 7. 知识推理 (基于规则推导隐含关系) 知识卡片 5-2： LOD(Level of Detail) 策略 | 层级 | 三角面数 | 适用场景 ||-----||-----
- -||-----|| LOD0 | 500 万 + | 近景精细展示 || LOD1 | 50-100 万 | 中景常规展示 || LOD2 | 5-10 万 | 远景概览 || LO
- D3 | <1 万 | 缩略图 / 导航 |

07

5.6 知识图谱在文化遗产中的深度应用

5.6.1 文化遗产知识图谱的构建挑战

- 文化遗产知识图谱的构建面临着独特的挑战。
- 数据异构性——文化遗产数据来源多样，包括结构化数据 (数据库记录)、半结构化数据 (XML、HTML) 和非结构化数据 (文本、图像、三维模型)，如何将这些异构数据统一表示为知识图谱是技术难题。
- 领域复杂性——文化遗产领域的概念关系极其复杂，一件文物可能同时涉及年代、材质、工艺、出土地、收藏机构、文献记载等多个维度，本体设计需要覆盖这些多维关系。
- 多语言性——文化遗产信息往往涉及多种语言 (中文、英文、法文等)，跨语言的实体对齐和关系抽取需要专门的 NLP 技术。
- 任明和贾君枝在研究数字人文领域知识图谱构建时提出了 PDC(Plan-Do-Check) 工作模式，强调人机协作在知识图谱构建中的重要性。
- 在该模式中，人类专家负责本体设计、质量审核和复杂关系判断，AI 负责大规模数据的自动处理和初步知识提取。
- 随着 AI 能力的提升，协作智能逐渐由人类主导向 AI 主导转变，但人类在价值判断和文化理解方面的作用不可替代。

5.6.2 知识图谱的典型应用场景

- 文化遗产知识图谱的典型应用场景包括：语义检索——用户可以基于概念关系进行查询，如 " 查找所有明代景德镇窑口的青花瓷器中纹饰包含龙纹的藏品 " ；
- 关联发现——自动发现文物之间的隐含关联，如同一工匠制作的不同藏品；知识推理——基于已有知识推导新知识，如根据墓葬出土文物的组合推断墓葬等级；
- 可视化分析——以网络图、时间线、地图等多种形式可视化呈现文化遗产知识网络。
- De Boer 在研究文化遗产知识图谱时指出，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。
- 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，使知识图谱与机器学习形成良性互动。

5.6.3 知识图谱的评估与维护

- 知识图谱的质量评估包括：完整性（覆盖了多少实体和关系）、准确性（实体和关系是否正确）、一致性（是否存在矛盾信息）和时效性（知识是否最新）。
- 知识图谱的维护需要建立更新机制，定期补充新发现的实体和关系，修正错误信息。

08

5.7 数字典藏平台的架构演进

5.7.1 从单体架构到微服务架构

- 传统数字典藏系统多采用单体架构，所有功能模块集中在一个应用中。随着数据规模 and 用户需求的增长，单体架构的局限性日益显现——系统扩展困难、维护成本高、技术栈受限。
- 微服务架构将系统拆分为独立部署的小型服务 (如采集服务、处理服务、检索服务、展示服务)，每个服务可以独立扩展和更新，更适合大规模数字典藏平台。

5.7.2 云原生与容器化

- 云原生技术和容器化部署正在改变数字典藏平台的架构方式。通过 Docker 容器和 Kubernetes 编排，数字典藏平台可以实现弹性扩展、自动恢复和灰度发布。
- 云端存储 (如对象存储 S3) 为海量文化遗产数字资产提供了经济可靠的存储方案。

5.7.3 数据中台与智能服务

- 数据中台理念为数字典藏平台提供了新的架构思路。通过构建统一的数据中台，将采集、处理、存储、分析和
服务全链路打通，实现数据资产的统一管理和复用。
- 在数据中台基础上， AI 驱动的智能服务 (如智能检索、个性化推荐、自动标引) 可以快速迭代部署。

09

5.8 数字典藏的安全与隐私

5.8.1 数据安全风险

- 文化遗产数字资产面临的数据安全风险包括：数据泄露——未授权访问导致敏感数据外泄；数据篡改——恶意修改数字资产内容；数据丢失——硬件故障或自然灾害导致数据永久丢失；
- 数据滥用——数字资产被用于未经授权的商业用途。中国甲午战争博物院在研究博物馆文物数字化保护与管理时指出，数据安全是数字化保护工作的重要组成部分。

5.8.2 安全防护体系

- 数字典藏的安全防护体系应包括：物理安全 (机房安全、设备安全)、网络安全 (防火墙、入侵检测、VPN)、应用安全 (身份认证、权限管理、日志审计) 和数据安全 (加密存储、加密传输、备份恢复)。
- Telek 等人提出的 3D 数字重建存档的协作方法强调了多机构协作的灾备策略对提高文化遗产数字资产安全性的重要作用。

5.8.3 区块链在数字典藏中的应用前景

- 区块链技术的去中心化、不可篡改和可追溯特性使其在数字典藏领域具有广阔应用前景。应用场景包括：数字资产确权——将文物的数字化成果上链，建立不可篡改的权属记录；
- 数据溯源——记录数字资产从采集到使用的完整链条；智能合约——基于智能合约自动执行版权许可和利益分配。
- 2025 年的研究探讨了区块链在武术非遗动作数据确权中的应用，为非遗数字资产确权提供了新思路。

10

5.9 数字典藏平台的典型案例深度分析

5.9.1 Europeana：跨国数字典藏的标杆

- Europeana 是欧洲文化遗产数字平台，整合了来自 4000 余家文化机构的 5800 万件数字资源。
- Europeana 的技术架构以 Europeana Data Model(EDM) 为核心，EDM 基于 CIDOC CRM 和 ORE(Object Reuse and Exchange) 标准，
- 能够统一整合来自不同成员国、不同类型机构 (博物馆、图书馆、档案馆) 的异构数据。
- Borin 和 Donato 在研究 Europeana 的财务可持续性时指出，Europeana 的运营模式依赖欧盟委员会的公共资金支持，年度预算约 3000 万欧元，
- 正面临着从公共资助向混合收入模式转型的挑战。
- Europeana 正在探索的可持续性策略包括：多元化收入来源 (会员费、API 授权、定制服务)、降低运营成本 (技术优化、志愿者参与) 和增强社会价值 (教育服务、文化外交)。

5.9.2 数字敦煌：专题数字典藏的典范

- 数字敦煌资源库是全球最大的石窟壁画数字资源库，其技术架构以高分辨率图像存储和 WebGL 渲染为核心。截至 2024 年，已完成 290 个洞窟的数字化采集，累计数据量超过 300TB。
- 数字敦煌的创新在于：一是建立了壁画数字化的标准作业流程 (SOP)，确保了不同洞窟数字化成果的一致性；二是开发了基于 WebGL 的在线浏览平台，使用户无需安装插件即可在浏览器中高清浏览壁画；
- 三是向全球免费开放，体现了开放获取的理念。陈振旺和樊锦诗从文化科技融合的视角分析了敦煌数字化实践，指出文化科技融合是遗产保护的重要路径，需要将技术创新与文化理解深度融合。
- 这一经验对其他文化遗产数字化项目具有重要借鉴价值。

5.9.3 故宫数字文物库：大规模馆藏数字化的实践

- 故宫数字文物库是中国规模最大的馆藏文物数字平台，收录了超过 78 万件文物的基础数字信息和高清图像。
- 故宫的实践特色在于：一是数字化与藏品管理的深度整合——数字典藏系统与藏品管理系统无缝对接，数字化成果直接服务于日常藏品管理；
- 二是分层次数字化——根据文物的价值和保存状况，采用不同精度的数字化方案，在有限资源下最大化覆盖面；三是面向不同用户群体的差异化服务——为学者提供深度检索和高清下载，为公众提供浏览和分享功能。

11

5.10 数字典藏的长期保存策略

5.10.1 数字保存的风险与挑战

- 数字典藏面临多种长期保存风险：介质老化（硬盘、磁带等存储介质的物理寿命有限）；格式过时（文件格式随技术演进而被淘汰）；软件依赖（数字对象的读取和处理依赖特定软件环境）；
- 硬件依赖（读取数字对象的硬件设备可能不再可用）；元数据丢失（描述数字对象的元数据如果不随数字对象一起保存，数字对象可能变得无法理解）。
- 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。
- Lor 和 Britz 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。

5.10.2 OAIS 参考模型的实践应用

- OAIS 参考模型定义了数字保存的完整功能架构，包括摄入 (Ingest)、存储 (Storage)、管理 (Administration)、
- 保存规划 (Preservation Planning) 和访问 (Access) 五个模块。
- 在数字典藏实践中，OAIS 模型的具体应用包括：制定提交协议 (明确数据提交者的责任和数据格式要求)、建立档案信息包 (AIP，包含数字对象及其保存描述信息)、
- 定期执行保存规划 (评估格式风险并制定迁移策略)、提供分发信息包 (DIP，面向用户的访问格式)。

5.10.3 数据迁移与仿真

- 数据迁移 (Format Migration) 和仿真 (Emulation) 是应对格式过时的两种主要策略。
 - 数据迁移将过时格式的数字对象转换为当前格式，优点是用户可以直接使用，缺点是迁移过程可能丢失信息。
 - 仿真通过软件模拟旧环境来运行过时的数据格式，优点是保持原始体验，缺点是开发和维护仿真器成本高昂。
- 对于文化遗产数字典藏，通常采用数据迁移策略，同时保留原始格式副本以备将来需要。

12

5.11 数字典藏的评价与认证

5.11.1 评价指标体系

- 数字典藏平台的评价应从以下维度进行：数据规模（数字资源数量、覆盖面）、数据质量（精度、完整性、准确性）、标准化程度（元数据标准采用情况、数据互操作性）、服务质量（检索效率、响应速度、
- 用户满意度）、可持续性（资金保障、技术更新、组织稳定性）和影响力（学术引用、公众访问、社会影响）。
- 洪闯和李中明基于可拓云模型对文化遗产数字化平台信息服务质量进行了评价研究，以“云游敦煌”和“数字故宫”为实证案例，从还原文化遗产细节、提升用户满意度、响应用户诉求、
- 重视移动端体验四个方面提出了优化建议。
- 这一研究为数字典藏平台的评价提供了方法论参考。

5.11.2 认证体系

- 建立数字典藏认证体系有助于推动行业标准化发展。
- 认证可以包括：技术认证（数字化技术参数是否符合标准）、流程认证（工作流程是否规范）、数据认证（数据质量和元数据完整性是否达标）和服务认证（用户服务是否满足要求）。
- 认证应由独立的第三方机构执行，定期复检。

13

5.6 三维重建技术的前沿进展与文化遗产应用

5.6.1 神经辐射场与 3D 高斯泼溅技术

- 2023 年以来，
- 神经辐射场 (Neural Radiance Fields, NeRF) 和 3D 高斯泼溅 (3D Gaussian Splatting, 3DGS) 技术为文化遗产三维重建带来了革命性变化。
- 与传统摄影测量和激光扫描相比，这两种基于 AI 的技术具有显著优势： NeRF 技术通过多层感知机 (MLP) 隐式表示三维场景的辐射场，从多视角照片中学习场景的几何和外观信息。
- 在文化遗产应用中， NeRF 的优势体现在： (1) 对反光、半透明等传统方法难以处理的材质有更好表现，如陶瓷釉面、玻璃器皿等文物的数字化；
- (2) 生成的模型具有视角连续性，不存在传统网格模型的拼接缝问题； (3) 模型文件更紧凑，便于网络传输和在线展示。如 Yu 等将 NeRF 应用于漆器的三维重建，成功复原了漆器表面的复杂光泽变化。
- 3DGS 技术以 3D 高斯基元显式表示场景，兼顾了 NeRF 的高质量 and 传统网格方法的实时渲染性能。
- 其训练速度比 NeRF 快约 10 倍，渲染速度可达 100 FPS 以上，为文化遗产的实时交互展示提供了可能。如斯坦福大学将 3DGS 应用于斯坦福雕塑库的重建，实现了在网页端实时浏览高质量 3D 雕塑模型。
- 然而，这两项技术在文化遗产领域也面临挑战： (1) 对拍摄条件要求较高，需要充足且均匀的光照； (2) 对大规模场景 (如整个建筑群) 的重建效率仍需优化；

5.6.1 神经辐射场与 3D 高斯泼溅技术

- (3) 生成的模型精度 (绝对精度通常在毫米级) 尚不能完全替代激光扫描在精确测量领域的应用。

5.6.2 多源数据融合的三维重建

- 单一技术手段往往难以满足文化遗产数字化对精度、色彩、材质等多维度信息的需求。
- 多源数据融合已成为文化遗产三维重建的主流方向：激光扫描 + 摄影测量融合：激光扫描提供高精度的几何数据，摄影测量提供高质量的纹理信息。
- 融合策略包括：基于几何约束的纹理映射、基于法线一致性的网格优化等。
- 如 " 数字兵马俑 " 项目采用 Faro 激光扫描仪获取几何数据 (精度 0.5mm) ，配合 Canon 5DSR 相机拍摄纹理照片 (5000 万像素) ，最终生成的三维模型同时具备高几何精度和高纹理保真度。
- 多光谱 + 三维数据融合：多光谱成像可揭示文物表面不可见的信息 (如底层素描、修复痕迹、颜料成分) ，将这些信息与三维几何数据融合，可生成 " 信息增强型 " 三维模型。
- 如荷兰国家博物馆将 X 射线荧光光谱 (XRF) 数据与三维模型叠加，使观众可在 3D 模型上直接查看每一区域的元素成分。
- 历史数据 + 现状数据融合：对于已损毁或变化的文化遗产，可将历史照片、测绘图纸与现状三维扫描数据融合，重建文化遗产的历史面貌。
- 如巴黎圣母院火灾后，Andrew Tallon 在 2010 年拍摄的激光扫描数据成为重建的关键参考，与历史文献和现状扫描数据融合，指导了精确的修复工程。

14

5.7 数字典藏的长期保存策略深化

5.7.1 OAIS 参考模型的实践应用

- 开放档案信息系统 (OAIS) 参考模型 (ISO 14721:2012) 是数字保存领域最重要的国际标准，已在文化遗产数字典藏中得到广泛应用。
- OAIS 模型定义了数字保存的六个核心功能模块：摄取模块 (Ingest)：接收数字内容并进行质量验证，生成保存描述信息 (PDI)。
- 在文化遗产数字典藏中，摄取环节需要特别关注：文件格式验证 (确保格式合规)、完整性校验 (MD5/SHA-256 校验和)、技术元数据提取 (分辨率、色彩空间、编码格式等) 和保存描述信息生成 (来源、权限、保真度要求等)。
- 保存模块 (Archival Storage)：实现数字内容的长期可靠存储。文化遗产数字典藏通常采用 "3-2-1" 备份策略——3 份副本、2 种不同介质、1 份异地存储。
- 存储介质选择需要考虑容量、成本、寿命和可读性等因素。当前，LTO 磁带 (寿命 30-50 年) 和企业级硬盘 (寿命 5-10 年) 的组合是最常见的选择。
- 数据管理模块 (Data Management)：维护描述性元数据和管理元数据，支持数字内容的发现和管理。
- 文化遗产数字典藏通常采用 Dublin Core 或 CIDOC CRM 作为元数据框架，配合关系型数据库 (PostgreSQL) 或图数据库 (Neo4j) 进行管理。

5.7.1 OAIS 参考模型的实践应用

- 访问模块 (Access)：为用户提供数字内容的检索和获取服务。文化遗产数字典藏的访问需要平衡开放共享与版权保护，通常采用分级访问策略——基础元数据完全开放，高清图像按权限分级开放。
- 管理模块 (Administration)：协调整个系统的运行，包括政策制定、标准执行、质量监控等。
- 保存规划模块 (Preservation Planning)：监控技术环境变化，制定格式迁移和保存策略调整方案。
- 这一模块在文化遗产数字典藏中尤为重要，因为数字格式的技术过时是长期保存的最大威胁之一。

5.7.2 数据迁移与格式规范化

- 数字格式过时 (Format Obsolescence) 是数字保存面临的核心挑战。
 - 为确保数字文化遗产的长期可读性，
 - 需要实施主动的数据迁移策略： 格式选择原则： 优先选择开放标准格式 (如 TIFF for images, OBJ/PLY for 3D models, WAV for audio) ，
 - 避免专有格式。
 - 对于无法避免的专有格式，需要保存格式规范文档和读取工具。迁移触发机制： 建立格式风险评估机制，当某种格式的支持度下降到预警线时启动迁移。
 - 风险评估指标包括： 软件支持度 (有多少软件能读取该格式) 、 硬件依赖度 (是否需要特定硬件) 、 社区采用度 (同行机构是否仍在使用该格式) 。
- 迁移验证： 迁移后需要验证数据的完整性和保真度。
- 常用方法包括： 文件比对 (哈希值校验) 、 视觉比对 (人工或自动化的图像对比) 、 功能测试 (关键功能是否正常) 。

15

5.8 知识图谱在文化遗产典藏中的深度应用

5.8.1 文化遗产知识图谱的构建流程

- 文化遗产知识图谱的构建是一个系统工程，涉及以下关键步骤： 本体设计：基于 CIDOC CRM 设计领域本体，定义文化遗产领域的概念体系、属性体系和关系体系。
- 本体设计需要平衡形式化程度与实用性——过度形式化难以维护，形式化不足则丧失推理能力。数据采集与预处理：从博物馆藏品管理系统、考古发掘报告、学术文献等多种数据源采集结构化和非结构化数据。
- 非结构化数据需要通过信息抽取技术转化为结构化数据。实体识别与消歧：识别数据中的实体 (人物、机构、地点、文物、事件等) ，并将不同表述的同一实体合并。
- 文化遗产领域的实体消歧面临特殊挑战——同一文物可能有多个名称 (如官方名称、俗称、学术名称) ，同一人物在不同时期可能有不同身份。
- 关系抽取与验证：从数据中识别实体间的语义关系 (" 创作 - 被创作 " " 收藏 - 被收藏 " " 影响 - 被影响 " 等) ，并通过人工校验或自动化方法验证关系的正确性。
- 知识融合与推理：将多个数据源的知识进行融合，利用推理规则发现隐含关系。如通过 " 某文物出土于某墓葬 " 和 " 某墓葬属于某时期 " 的已知关系，可推断 " 某文物属于某时期 " 的隐含关系。

5.8.2 知识图谱驱动的智能检索与推荐

- 文化遗产知识图谱可实现超越关键词匹配的智能检索和推荐：语义检索：理解用户的查询意图，返回语义相关的结果。
- 如用户搜索 " 明代青花瓷 " ，系统不仅返回直接匹配的藏品，还推荐相关的窑口信息、工艺技术、贸易路线等关联知识。
- 可视化探索：将知识图谱以网络图、时间线、地图等可视化方式呈现，支持用户通过交互式探索发现知识关联。
- 如 Europeana 的 "Collections" 平台利用知识图谱支持用户从一件藏品出发，探索相关的历史事件、人物和文化主题。
- 个性化推荐：基于用户画像和知识图谱推理，为用户推荐可能感兴趣的文化遗产内容。如根据用户浏览过的藏品，推荐同时期、同地域、同类型的相关藏品。

16

5.9 数字典藏平台的用户体验设计

5.9.1 用户研究与需求分析

- 数字典藏平台的用户群体多样，包括研究者、教育工作者、学生、文化爱好者、文博从业人员等，不同群体的需求差异显著：
研究者需求：高分辨率图像下载、详细元数据、关联文献、批量数据接口 (API)、

• 引文工具集成。

• 研究者关注数据的学术可用性——能否用于学术分析和发表，而非仅用于浏览。教育工作者需求：可下载的教学素材、课程整合指南、适合课堂展示的高清图像和 3D 模型、不同教育层级的内容适配。

• 教育工作者关注内容的可教学性——能否直接用于课堂教学活动。普通公众需求：直观的浏览界面、引人入胜的故事化叙述、简洁的文字说明、便捷的分享功能。

• 普通公众关注体验的愉悦性——能否在一次访问中获得有意义的发现和情感共鸣。文博从业者需求：同行经验分享、技术标准参考、管理工具模板。文博从业者关注平台的工具性和参考性。

5.9.2 界面设计原则

- 基于用户需求分析，数字典藏平台的界面设计应遵循以下原则：渐进式信息披露：首页提供简洁的视觉入口（精选藏品、主题展览），随用户深入逐步展示更多信息和功能。避免在首页堆砌过多信息导致认知超载。
- 多模态检索支持：提供关键词检索、分类浏览、时间线浏览、地图浏览、颜色检索、以图搜图等多种检索方式，满足不同用户的检索习惯和需求。
- 响应式设计：确保平台在桌面电脑、平板和手机上都有良好的浏览体验。考虑到移动端用户占比已超过 50%，移动优先的设计策略尤为重要。
- 无障碍设计：遵循 WCAG 2.1 AA 级标准，确保残障人士能够无障碍访问。包括屏幕阅读器支持、键盘导航、色彩对比度达标、替代文本等。

17

5.10 典型数字典藏平台的深度案例分析

5.10.1 Europeana 的运营模式与数据治理

- Europeana 是欧洲文化遗产数字聚合平台，自 2008 年上线以来已聚合超过 5800 万条数字文化遗产记录，是全球最大的文化遗产数字平台之一。
- 其运营经验和数据治理模式具有重要参考价值：数据聚合模式：Europeana 不直接存储原始数字资源（高清图像、3D 模型等），而是聚合各机构的元数据和预览图。
 - 原始资源存储在各成员机构的服务器上，用户通过 Europeana 检索后链接到源机构访问原始资源。这种“元数据聚合 + 链接访问”模式既降低了中心平台的存储压力，又保障了源机构的数据主权。
- 数据质量治理：Europeana 建立了多层质量控制体系：
 - (1) 入库前验证——检查元数据完整性、格式合规性和链接有效性；
 - (2) 自动质量评估——利用 AI 工具检测元数据错误、重复记录和低质量预览图；
 - (3) 人工抽检——编辑团队定期抽检数据质量，向数据提供者反馈改进建议；
 - (4) 数据清洗——定期执行批量数据清洗，统一格式、消除重复、补充缺失信息。
- EDM 数据模型：Europeana 数据模型 (EDM) 是其互操作性的核心。
 - EDM 基于链接数据原则，采用 RDF 格式描述文化遗产资源。
 - EDM 的核心设计理念是“语境化”——不仅描述资源本身 (What)，还描述其来源 (Where/Who)、时间 (When) 和主题 (Which)，形成丰富的语境信息网络。

5.10.2 数字敦煌资源库的技术创新

- "数字敦煌"资源库是中国文化遗产数字典藏的标杆项目，其技术创新体现在：超高精度壁画数字化：采用定制的轨道自动拍摄系统，配合1亿像素数码后背，实现每平方米300 DPI的采集精度。
- 色彩管理采用从采集到显示的全流程ICC色彩配置文件，确保色彩还原度 $\Delta E < 3$ 。
- 全景漫游与虚拟洞窟：利用全景摄影技术构建洞窟的360度虚拟漫游环境，用户可在Web浏览器中"走进"洞窟，自由观察壁画细节。系统支持热点标注——点击壁画特定区域可弹出详细解说和高清局部图。
- 多光谱数据集成：除了可见光图像，还采集了紫外荧光、红外反射、X射线等多模态数据，并支持用户在Web界面中切换不同光谱视图，揭示壁画肉眼不可见的信息。
- 开放API与数据共享：数字敦煌资源库提供了RESTful API，支持研究者和开发者程序化访问元数据和低分辨率图像。高清图像需通过授权获取，既促进了学术利用，又保护了数字资产的商业价值。

18

5.11 数字保存的风险管理与灾备体系

5.11.1 数字遗产风险分类

- 数字文化遗产面临的风险可分为以下几类：技术风险：(1) 格式过时——文件格式被技术淘汰，无法被现有软件读取；(2) 介质老化——存储介质 (硬盘、磁带、光盘) 物理老化导致数据丢失；
- (3) 硬件故障——服务器、存储设备故障导致数据损坏或丢失；(4) 软件漏洞——管理系统漏洞导致数据被篡改或泄露。自然风险：(1) 地震、火灾、洪水等自然灾害摧毁数据中心；
- (2) 电磁脉冲 (如太阳风暴) 导致存储设备数据损坏；(3) 温度湿度异常加速存储介质老化。人为风险：(1) 网络攻击——黑客入侵系统窃取或破坏数据；(2) 内部威胁——员工误操作或恶意行为；
- (3) 资金中断——数字化项目资金中断导致系统无法维护和更新；(4) 机构变更——文博机构合并、撤销或重组导致数字资产归属不清。

5.11.2 灾备体系建设

- 针对上述风险，需要建立完善的灾备体系： 3-2-1 备份策略：保留 3 份数据副本，存储在 2 种不同介质上，其中 1 份存放在异地。这是数字保存领域最基本的灾备策略。
- 异地灾备：在地理上相距较远的不同城市建立灾备中心，确保单一自然灾害不会同时摧毁所有数据副本。
- 异地灾备可采用 " 主 - 备 " 模式 (灾备中心平时不运行，主中心故障时启用) 或 " 双活 " 模式 (两个中心同时运行，互为备份)。冷数据归档：对于不常用的历史数据，迁移到低成本的冷存储介质 (如 LTO 磁带)。
- 冷存储虽然访问速度慢，但寿命长 (30-50 年)、成本低，适合长期归档。定期数据校验：定期检查存储介质上的数据完整性 (通过校验和比对)，及时发现和修复数据损坏。
- 对于磁带介质，应定期 (如每 5 年) 进行数据迁移，防止介质老化导致数据丢失。应急响应计划：制定详细的应急响应计划，包括风险预警机制、应急处理流程、数据恢复程序和业务连续性方案。
- 定期进行应急演练，确保团队成员熟悉应急流程。

19

5.12 数字典藏中的数字孪生技术

5.12.1 文化遗产数字孪生的概念

- 数字孪生 (Digital Twin) 是指物理实体的虚拟镜像，不仅复制其几何形态，还实时同步其状态数据。
- 在文化遗产领域，数字孪生为遗产地的保护监测和管理提供了全新的技术范式：
静态数字孪生：通过 3D 扫描和摄影测量建立文化遗产的高精度三维模型，这是数字孪生的基础层。
动态数字孪生：在静态模型基础上集成实时传感器数据，使数字孪生体能实时反映物理遗产的状态。
- 传感器监测结构变形、环境参数、游客流量等数据，通过物联网实时传输到数字孪生平台，管理者可在虚拟环境中实时观察遗产状态。
- 智能数字孪生：在动态数字孪生基础上叠加 AI 分析能力，实现预测性保护和智能决策。AI 模型基于历史数据和实时数据预测遗产状态变化趋势，自动生成保护建议和预警信息。

5.12.2 数字孪生在遗产地管理中的应用

- 结构健康监测：通过在遗产地安装结构传感器（应变计、倾斜仪、裂缝计等），实时监测结构状态。
- 数字孪生平台将传感器数据映射到 3D 模型上，管理者可在 3D 模型上直观查看各区域的结构状态，红色标记表示高风险区域。
- 环境管理：集成温湿度、光照、空气质量等环境传感器数据，实时评估环境对遗产的影响。
- 当某区域环境参数超出安全阈值时，系统自动发出预警并建议调整环境控制措施（如开启空调、调整照明、限制游客数量等）。游客管理：通过人流计数器、Wi-Fi 探针等技术实时监测游客分布和流量。
- 数字孪生平台可在 3D 模型上显示游客密度热力图，帮助管理者优化游客动线和限流策略。如故宫博物院的数字孪生系统可实时显示各展厅的游客密度，当某展厅密度超过阈值时自动推送其他展厅推荐路线。
- 灾害应急管理：数字孪生可模拟灾害（如地震、洪水、火灾）对遗产地的影响，帮助管理者制定应急预案。如通过数字孪生模拟不同震级下古建筑的结构响应，识别最脆弱的区域并优先加固。

20

5.13 数字典藏与文化遗产研究的数字化转型

5.13 数字典藏与文化遗产研究的数字化转型

- 数字典藏不仅是保护手段，也是文化遗产研究方法的变革推动力：数据驱动的研究范式：传统文化遗产研究以 " 个案研究 " 为主——研究者选择少量典型文物进行深入分析。
- 数字典藏提供了大规模数据资源，使 " 数据驱动 " 的研究范式成为可能——通过对数万件文物的定量分析，发现传统个案研究无法发现的模式和规律。
- 如对数千件青铜器的纹饰数据进行统计分析，可能揭示纹饰风格的地域传播路径和时代演变趋势。可重复的研究流程：数字典藏中的数据可被不同研究者重复使用和分析，提高了研究的可重复性。
- 研究者应公开其数据处理和分析方法，使其他研究者能够验证和扩展其结论。
- 跨学科协作平台：数字典藏平台可作为跨学科协作的基础设施——历史学家、考古学家、艺术史家、材料科学家、计算机科学家等不同学科的研究者可在同一平台上共享数据、交流方法和协作研究。
- 开放科学的实践：数字典藏推动文化遗产研究向开放科学转型——开放数据、开放方法、开放发表、开放同行评审。开放科学提高了研究的透明度和公信力，也加速了知识传播和创新。

21

5.15 延伸阅读与技术实践

5.15 延伸阅读与技术实践

- 延伸阅读：建议阅读 OAIS 参考模型 (ISO 14721:2012) 原文，深入理解数字保存的功能架构。
- 同时推荐阅读 Europeana 的技术文档和 API 文档，了解大规模文化遗产数据聚合的技术实现。对于知识图谱方向，推荐阅读 CIDOC CRM 教程和 Linked Data 相关文献。
- 技术实践：建议学生使用开源工具进行实践操作：(1) 使用 MeshLab 处理 3D 点云数据，体验点云去噪、网格重建和纹理映射的完整流程；
- (2) 使用 Gephi 构建小型文化遗产知识图谱，理解实体 - 关系建模的可视化；(3) 使用 Omeka 或 CollectionSpace 搭建小型数字典藏原型系统，体验元数据著录、检索和展示的完整流程。
- 前沿关注：NeRF 和 3D Gaussian Splatting 技术正在快速发展，建议关注相关学术论文和开源项目 (如 Nerfstudio)，了解这些技术如何改变文化遗产三维重建的方式。
- 同时关注 AI 在数字保存中的应用——如 AI 驱动的模式迁移自动化、AI 辅助的数据质量评估等。

22

5.16 数字典藏的数据迁移实践

5.16 数字典藏的数据迁移实践

- 数据迁移是数字保存中最关键也是最危险的操作之一——迁移过程中数据丢失或损坏的风险最高。
- 以下是数据迁移的实践指南： 迁移触发条件： (1) 格式过时风险——当某种文件格式的支持软件数量下降到预警线时 (如仅剩 1-2 款软件支持) ；
(2) 存储介质老化——当存储介质的预估剩余寿命低于安全阈值时 (如硬盘使用超过 5 年) ； (3) 系统升级——当管理系统升级换代，旧格式不被新系统支持时。
- 迁移策略选择： (1) 格式转换——将数据从旧格式转换为新格式 (如 TIFF→JPEG2000) ， 需验证转换后的数据保真度；
(2) 仿真——通过模拟旧格式运行环境来保持数据可读性，适用于无法无损转换的格式； (3) 封装——将数据与其读取软件打包保存，确保数据在未来环境中可运行。
- 迁移验证：迁移后必须进行严格的验证—— (1) 文件级验证——比对迁移前后的文件属性 (分辨率、色彩空间、几何精度等) ； (2) 内容级验证——通过人工或自动化方法检查内容是否完整和准确；
(3) 功能验证——验证迁移后的数据在新系统中能正常检索和展示。迁移文档：每次迁移应完整记录迁移过程，包括：迁移时间、源格式、目标格式、转换工具、验证结果、异常情况处理等。
- 这些文档构成了数字资源的技术溯源链，对长期管理至关重要。

23

5.17 数据治理中的隐私与安全

5.17 数据治理中的隐私与安全

- 文物数字典藏中的数据安全和隐私保护需要特别关注以下风险：(1) 文物位置信息泄露——对于未公开的考古遗址或窖藏，精确位置信息属于敏感数据，泄露可能导致盗掘。
- 数字典藏系统应对位置信息进行脱敏处理，仅授权人员可访问精确坐标。(2) 高精度 3D 数据滥用——文物的高精度 3D 模型可能被用于 3D 打印赝品。
- 应根据使用场景控制 3D 模型的数据精度，公开版本适当降低精度。
- (3) 网络攻击——数字典藏平台可能成为网络攻击目标，需要部署防火墙、入侵检测系统和 Web 应用防火墙等安全措施，定期进行安全审计和漏洞扫描。在数字孪生技术的实践中，数据实时同步是核心技术挑战。
- 物联网传感器的数据采集频率、网络传输延迟、数据处理吞吐量等因素都会影响数字孪生体的实时性。对于结构监测等需要快速响应的场景，数据延迟应控制在秒级以内；
- 对于环境监测等允许一定延迟的场景，数据采集间隔可以放宽到分钟级。此外数字孪生平台还需要处理多源异构数据的融合问题。
- 建议采用模块化的数字孪生架构，将数据采集、数据处理、模型渲染和决策支持分别构建为独立模块，通过标准化接口实现互操作，便于系统的渐进式建设和持续优化。
- 在数字典藏的实践中，还应特别关注数据的版本管理。文物的数字资产可能经历多次修改和更新——如 3D 模型经过去噪优化、纹理经过色彩校正、元数据经过补充完善等。

5.17 数据治理中的隐私与安全

- 每次修改都应保存为独立版本，并记录修改原因、修改内容和修改人员。版本管理系统应支持版本回溯（查看任意历史版本）、版本对比（比较两个版本的差异）和版本恢复（回滚到历史版本）。
- 这确保了数字资产的可追溯性，也为学术研究提供了数据演变的完整记录。建议采用 Git 或类似版本控制系统的理念管理数字资产版本。
- 数字典藏的长期价值取决于数据的持续可用性和质量保障，这需要制度性投入而非一次性建设。在数字孪生和知识图谱等前沿技术的推动下，文物数字典藏正在从静态的数据存储向动态的知识服务转型。
- 这一转型要求从业者不仅掌握传统的数据管理技能，还需要具备 AI、语义网、物联网等新技术的理解和应用能力。教育机构和行业组织应加强相关人才培养，为数字典藏的智慧化转型提供人才保障。
- 同时，文博机构应在战略层面规划数字典藏的发展路径，将数字典藏纳入机构核心业务而非边缘化的辅助功能，确保数字化建设的持续投入和长期运营，真正实现数字典藏作为文化遗产保护基础设施的战略定位。
- 总之，文物数字典藏作为文化遗产数字化保护的核心基础设施，其建设质量和管理水平直接关系到文化遗产数字资源的长期保存与有效利用，需要持续投入和不断创新。

- 本章从三维重建技术、数字典藏库构建、知识图谱与语义关联、长期保存策略、典型平台分析、用户体验设计、数字孪生技术、风险管理、数字人文研究转型等多个维度，
- 系统论述了文物数字典藏的理论与实践。
- 数字典藏是 G.U.A.R.D. 框架中 A(Archiving) 维度的核心内容，是连接数字化采集 (G) 和数字化呈现 (R) 的关键环节。
- 高质量的数字典藏不仅确保文化遗产数字资源的长期保存，也为学术研究、教育传播和活化利用提供了基础数据支撑。

- 点云数据处理的四个核心环节是什么？ 2. (理解) 解释知识图谱如何将文化遗产数据从 " 关键词检索 " 升级为 " 语义检索 "。
- 3. (应用) 运用 OAIS 模型分析 " 数字敦煌 " 资源库的保存策略。 4. (分析) 比较 Europeana 和数字敦煌在数据来源、技术架构和服务模式上的差异。
- 5. (评价) 评估协作智能 (人 +AI) 在文化遗产知识图谱构建中的优势与风险。 6. (创造) 为一个省级非遗中心设计数字典藏平台的功能架构，包含知识图谱模块。

- 1. Caciara 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. Montusiewicz J, Barszcz M, Korga S. Preparation of 3D Models of Cultural Heri

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