



中华人民共和国文物保护行业标准

WW/T XXXXX—XXXX

文物元数据体系

Metadata System of Cultural Heritage

(征求意见稿)

2017-03-16

XXXX - XX - XX 发布

XXXX - XX - XX 实施

中华人民共和国国家文物局 发布

目 次

前言	II
1 范围	1
2 术语与定义	1
3 文物数字化标准规范体系与文物元数据体系	2
4 文物数字化保护元数据规范类型及体系	2
5 文物分类与文物数字化保护元数据体系	3
6 文物数字化对象生命周期与元数据体系	5
7 文物数字化保护元数据的关系模型	6
参考文献	8

前 言

本标准是在面向数字化保护的标准规范体系和文物数字化生命周期模型的基础上,建立了一整套具有可操作性、可适用性和可扩展性的文物数字化元数据规范体系,包括文物数字化保护元数据规范类型及体系、文物分类与文物数字化保护元数据体系、文物数字化对象生命周期与元数据体系、文物数字化保护元数据的关系模型等。

本标准适用于对不同类型元数据标准的设计中元素的交叉和复用,不同元数据之间的相互链接和调用提供基础依据。

本标准由中华人民共和国国家文物局提出。

本标准由全国文物保护标准化技术委员会(SAC/TC 289)归口。

本标准起草单位:北京大学。

本标准起草人:朱强、肖珑、张俊娥、李峰、冯英。

本标准是首次发布。

文物元数据体系

1 范围

本标准在面向数字化保护的标准规范体系和文物数字化生命周期模型的基础上,建立了一整套具有可操作性、可适用性和可扩展性的文物数字化元数据规范体系,包括文物数字化保护元数据规范类型及体系、文物分类与文物数字化保护元数据体系、文物数字化对象生命周期与元数据体系、文物数字化保护元数据的关系模型等。

本标准说明了元数据体系中的描述元数据、管理元数据、保存元数据和服务元数据在文物数字对象从采集、加工整理、利用、保存到处置的生命周期中的作用,以及这些元数据之间的关系。为在不同类型元数据标准的设计中元素的交叉和复用,不同元数据之间的相互链接和调用提供基础依据。

2 术语与定义

本标准规范采用下列术语和定义。

2.1 元数据 metadata

是关于信息资源或数据的一种结构化的数据。

2.2 描述元数据 descriptive metadata

对信息资源本身的内容、属性及外在特征进行描述的元数据。

2.3 管理元数据 administrative metadata

对信息资源实施管理的元数据。

2.4 保存元数据 preservation metadata

对文信息资源实施存储和长期保存时需要的元数据。

2.5 文物核心元数据 core element set for cultural heritage

使用频率高的、共性的、可用于不同类型的文物信息资源描述的元数据。

2.6 文物专门元数据 element set for a certain type of cultural heritage

用来描述某一类文物信息资源的内容、属性及外在特征的元数据。

2.7 服务元数据 service metadata

用于数字资源服务的揭示与表现、服务过程、服务系统等方面的相关信息描述的元数据。

2.8 元元数据 meta-metadata

对元数据的标记语言、格式语言、标识符、扩展机制、转换机制等信息进行描述的元数据，实现对元数据的管理。

3 文物数字化标准规范体系与文物元数据体系

文物元数据体系属于文物数字化标准规范体系中重要的一部分，是文物数字化保护的核心内容。文物数字化标准规范体系包括标准体系框架、核心标准规范体系、其他支撑标准和应用示范四部分（见图1），文物数字化保护元数据标准体系受标准规范管理系统、标准体系管理规范、标准体系扩展规则、标准体系表和文物数字化对象生命周期模型指导，同时与文物数字化技术体系共同构成了核心标准规范体系，在文物分类体系、主题词表、标准体系编码规则支持的基础上，通过文物元数据标准示范、可移动文物/不可移动文物数字化保护关键标准应用示范进行验证和规范。

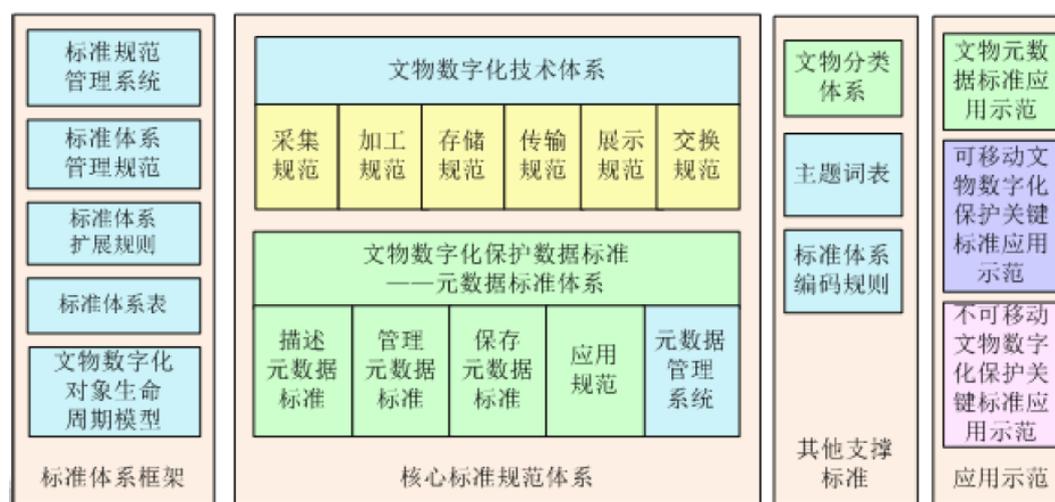


图1 文物数字化标准规范体系

4 文物数字化保护元数据规范类型及体系

从文物数字化保护的内容、作用等角度，文物元数据的类型又可分为专门元数据、资源集合元数据、管理/保存元数据、服务元数据和元数据等，是一个层次分明、结构开放的元数据体系（见图2）。

内容对象 (Content Objects)	专门元数据 描述单一文物数字对象（可移动文物、不可移动文物）的特色内容、属性及外在特征的元数据
内容对象集合 (Collection of Content Objects)	资源集合元数据 按照主题、文物类型、文物年代等形成的数字资源集合（如数据库、知识组织系统等）的描述
对象的管理和保存 (Management and Preservation of Objects)	管理/保存元数据 数字对象的采集、加工、结构、技术处理、存取控制、保存、数据更新与迁移、版权管理以及相关系统等信息
对象的服务、服务过程、服务系统 (Service and process of Objects)	服务元数据 数字资源揭示与表现、服务过程、服务系统等方面的相关信息的描述

元数据的管理 (Management of Metadata)	元元数据 对元数据的标记语 言、格式语言、标识符、扩展机制、转换机制等信息的描述
---------------------------------------	--

图2 文物元数据体系与文物元数据类型¹⁾

在该体系中，专门元数据和资源集合元数据是都是对文物实体资源本身的内容、属性、外在特征的描述，统称为“描述元数据”，描述元数据记录的是文物实体数字化后的最终信息，不记录文物数字化过程中产生的信息。在描述元数据中，文物核心元数据是一组最重要、简明的描述基于文物实体的数字资源对象的元数据元素，主要描述所涉及的所有数字资源对象的共有属性；在文物核心元数据之外，还有根据不同种类的文物资源的特点来建立的个别元数据。资源集合元数据是针对一组基于文物实体的数字资源对象集合的描述。

“管理元数据”是专门对文物数字资源实施管理的元数据。它描述在管理文物实体对象数字化的过程中所涉及的管理信息以及对管理政策与管理机制的规范、开放描述，真实记录业务活动信息、关联组织、有效管理数字化对象并进行多维展示。管理元数据将有助于对文物数字资源实行自动化与网络化的科学管理，有利于文物数字资源管理信息的跨系统交换与共享。

“保存元数据”最初是管理元数据的重要组成部分。但 OAIS (Open Archival Information System, 开放存档管理系统) 参考模型的建立，为构建数字信息长期保存和维护的存档系统提供了一个概念性框架，目前国际上建立了多个比较通用的保存元数据格式如 NLA、NEDLIB、CEDARS、PREMIS，保存元数据已经发展成为一个相对独立的元数据研究领域。保存元数据用来维护文物数字化对象资源的长期保护和长期可存取性，更确切地说，它是支持文物数字资源长期保存过程中的可生存能力、可还原能力与可理解能力的必要信息。保存元数据能够作为保存过程中的输入信息，也可以作为相同过程的输出信息。

“服务元数据”是针对 Web 服务的需要发展起来的一种元数据，用于文物数字化对象资源的服务揭示与表现、服务过程、服务系统等内容的相关信息的描述，规范文件格式、编码规则、统一标识等，在元数据交换、互操作、跨系统共享等方面提供便捷。

“元元数据”是一种最底层、最基础的规范元数据，主要作用是满足最终的互操作要求，元数据领域的各种类型元数据的基本框架、应用原则等都应该满足语义万维网的基础规范。

5 文物分类与文物数字化保护元数据体系

依据文物数字化标准规范体系的划分，文物元数据体系属于其组成部分之一，我们在此针对文物元数据建设框架体系。文物元数据标准框架体系与文物分类体系紧密相关，类型众多的文物为文物数字化及元数据标准研制的科学性、规范性、适用性带来了巨大的挑战。

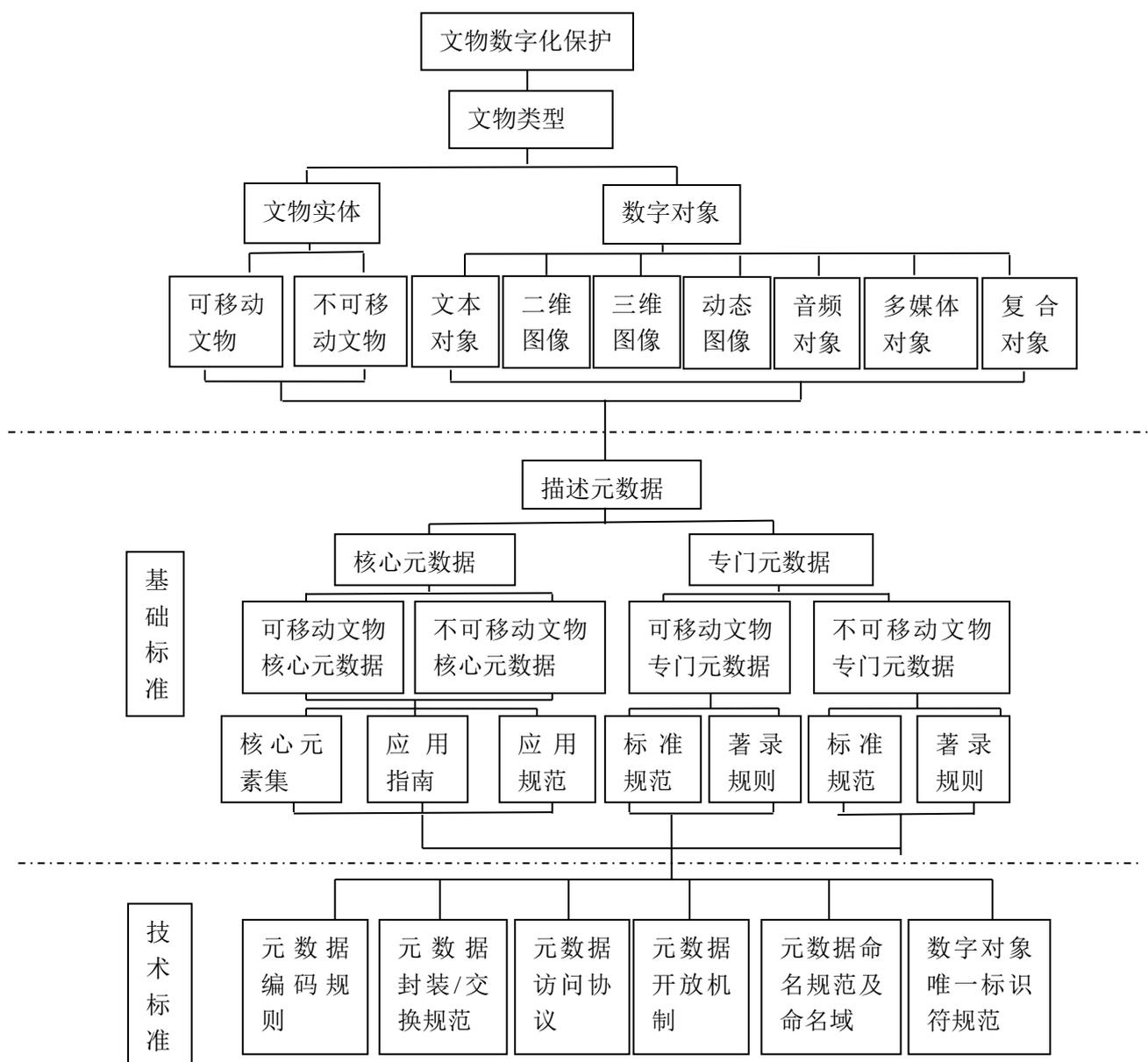
从文物实体的角度，根据文物的存在形态分类，具体可分为可移动文物和不可移动文物。可移动文物主要是指可以移动位置的馆藏文物和流散文物，主要有石器、陶器、铜器、铁器、金银器、玉器、瓷器、漆器、工艺品、书画、古文献等等；不可移动文物基本上都是文物史迹，如古建筑、石窟寺、石刻、古遗址、古墓葬、近代现代重要建筑、纪念地等。

对文物实体进行数字化后，得到了文物数字对象。数字对象总体上可以分为：文本对象、图像对象（二

1) 改变自肖珑，赵亮主编. 中文元数据概论与实例[M]. 北京：北京图书馆出版社，2007:11.

维图像对象、三维图像对象、动态图像对象)、音频对象、多媒体对象和复合对象。

- a) 文本对象是指文物资源数字化后以文本的形式保存的对象，如古籍、文档等。文本对象需要记录文物的背景、出土情况、类型、尺寸、外形、重要特征、材质等。
- b) 二维图像对象是指以平面形式展示的数字对象，如图片，需要记录文物的体积、外形、纹理、装饰、图案、颜色、光泽、字形、字体、字和图位置、周围环境等。
- c) 三维图像对象是指以三维立体形式展示的数字对象，如立体古建筑、雕塑、瓷器等，除了具备二维图像对象所需的元数据以外，还需要记录其内部特征、各部件联系、位深、温度变化、地理位置等。
- d) 动态图像对象是指以动态形式展现的数字对象，如动态 gif 图像、流媒体等，具备三维图像对象所需的元数据，还要注意文物各个角度的呈现，需要各个角度的信息。
- e) 音频对象是指以声音形式展示的数字对象，如配乐，一般是配合其它文字、图片、视频等，要贴近文物所处的历史年代、所反映出的文化内涵。
- f) 多媒体对象是指上述多种形式结合展示的数字对象。需要文本、二维图像、三维图像和音频对象的元数据。
- g) 复合对象是资源集合，指上述多种形式、多个对象的集合对象。



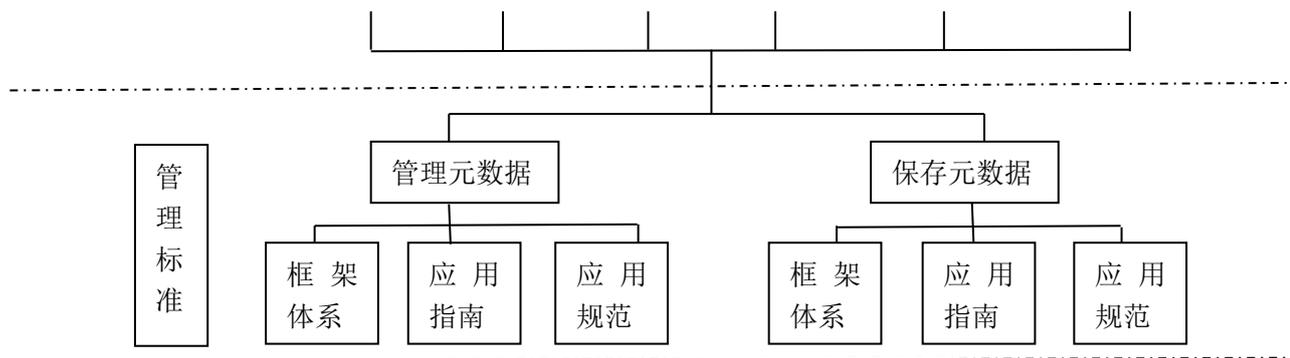


图3 文物分类及文物数字化保护元数据体系

在文物分类及文物数字化保护元数据体系中（图3），第一部分是不同的文物类型及载体形式分类，下面三个部分是基于文物实体对象和数字对象的交叉维度的元数据类型，分别是基础标准、技术标准和管理标准。

基础标准层主要是描述元数据，根据文物资源的共同特点，以可用于不同类型文物描述的元数据通用元素的集合核心元数据为基础，针对文物实体的存在形态分类，扩展为可移动文物核心元数据和不可移动文物核心元数据。在核心元数据之外，根据文物实体的时代、存在形态、质地、功用、属性（性质）、来源等分类体系，针对不同类型的文物设计相应的描述文物对象内容的专门元数据，建立模块化、可扩展的、既通用又可以个性化应用的描述元数据体系及专门元数据标准。描述元数据的内容属性基于文物实体来描述，同时包含对数字化信息的描述和相关链接。

技术标准层，也可以称之为服务元数据，是与元数据输入、处理、存贮、管理、发布、检索、交换等紧密相关的技术体系。在系统层次方面，涉及系统间数字化信息的访问交换、转换等互操作标准；在业务层次方面，涉及输入、存贮、管理过程的标准，在数据层次方面，涉及信息模型、文件格式、编码规则、统一标识等标准，因此包括元数据命名规范及命名域、文物数字对象唯一标识符规范等标识体系以及元数据编码规则、元数据封装/交换规范、元数据访问协议与元数据开放机制。

管理标准层是结合文物的数字化资源建设、服务与管理的需要，建立管理元数据框架和保存元数据模型，定义文物管理/保存元数据元素集，以支持对各类型文物及其数字资源的管理/保存需求，支持对数字资源建设各环节的管理/保存需求，包含对技术信息、知识产权、版权保护、权限控制、来源、数字化过程（包括数字对象经历的技术处理、技术环境、起源环境等）、服务、数据更新与迁移、版本保存等与数字资源管理/保存密切相关的内容。

6 文物数字化对象生命周期与元数据体系

文物元数据体系应该能够全面描述文物信息系统中与数字资源相关的各类特征和属性，包括静态特征、状态特征、生命周期特征、管理特征以及系统在实际功能过程中所要求的其他语义信息等等，由此可见文物元数据与文物数字资源生命周期息息相关，从文物实体数字化过程的创造、采集、加工整理、利用、保存到处置（主要指资源的剔除/销毁）的每个阶段中，都离不开元数据规范，不同类型的元数据在文物数字化对象生命周期的不同阶段产生并在不同环节起着重要作用。从文物实体被采集加工成为文物数字化对象到进入信息服务中介后的加工、服务、保存、处置的生命周期以及周期中所涉及的事件与文物数字化建设标准规范之间的关系可建立如图4。

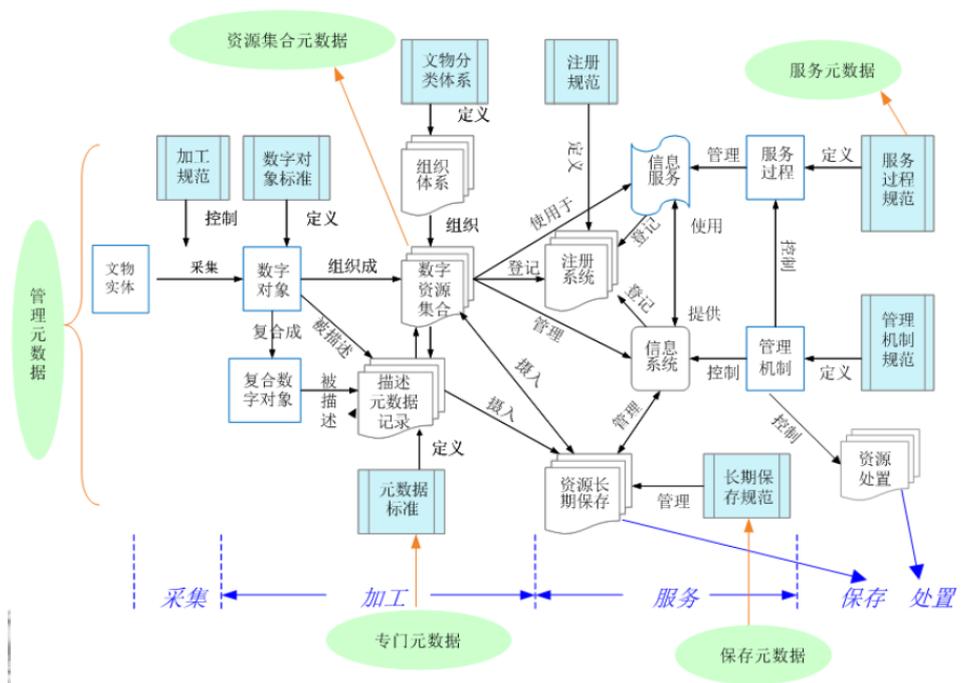


图4 文物数字化对象生命周期与元数据规范²

在图4中，从文物实体被采集加工成为数字化对象到被用户利用，在整个生命周期内经历了采集、加工、服务等基础环节，在每个环节都受到管理过程的控制，每个环节都有其相应的管理机制与管理政策，例如在采集加工阶段，对文物的类型、数字对象标准、加工规范进行管理，在提供服务阶段，对注册规范、信息系统、服务过程等设定管理机制，在资源保存、处置阶段，使用管理元数据管理资源的长期保存和资源处理方式。每个环节都有特定的管理元数据，它既反映了这一管理环节管理机制的共性，又能根据更加具体的管理功能进行扩展，各个环节的管理元数据是一个相互融合的有机整体。

以专门元数据、资源集合元数据等组成的描述元数据，主要作用于文物实体数字化的采集加工阶段，文物实体被采集加工成为数字对象（复合数字对象）的过程，以及对数字对象（复合数字对象）进行加工、描述、检索，需要由描述元数据来完成。描述元数据在描述不同数字化对象的同一性质时用到的核心元素集，和描述不同对象时的个性化的专门元数据都需要建立相应的标准规范和应用指南。

服务元数据主要存在于数字对象提供服务的阶段，作用是将采集加工阶段完成的文物数字化对象通过描述元数据规范后，再使用服务元数据展示出来，以供读者检索、使用，并且实现各信息资源之间的交换、共享、互操作。

保存元数据主要存在于数字对象的保存和处置阶段，将描述元数据、管理元数据、服务元数据提供的各种数字化信息、检索使用、版本更新、剔除/销毁等过程进行保存。

7 文物数字化保护元数据的关系模型

文物数字化对象的描述元数据、管理元数据、保存元数据、服务元数据不仅在文物资源的采集、加工、服务、保存和处置的生命周期的各个环节都起着不同的作用，而且它们之间的关系也是既有交叉重合又有独立的部分。文物数字化对象（如数字对象、复合数字对象、资源集合、知识库等等）与对数字化对象的描述、管理、保护和服务的信息共同构成了关于该文物的完整数据。文物数字化保护元数据的

2) ² 本图改编自：张晓林“数字图书馆标准与规范建设”PPT，并增加了元数据所起作用的环节。

关系模型见图 5，元数据之间信息交叉的部分是针对文物数字化对象的最基本的描述和管理信息，实现数字化对象的核心描述和管理，元数据各自独立的部分是该元数据特有的作用与功能，体现了不同元数据类型的特色。但是所有这些元数据都需要在规范的框架体系、应用指南、著录规则和应用纲要构建的开放的信息环境中才能实现应有的作用，才能够被不同的数据系统无缝衔接和利用。

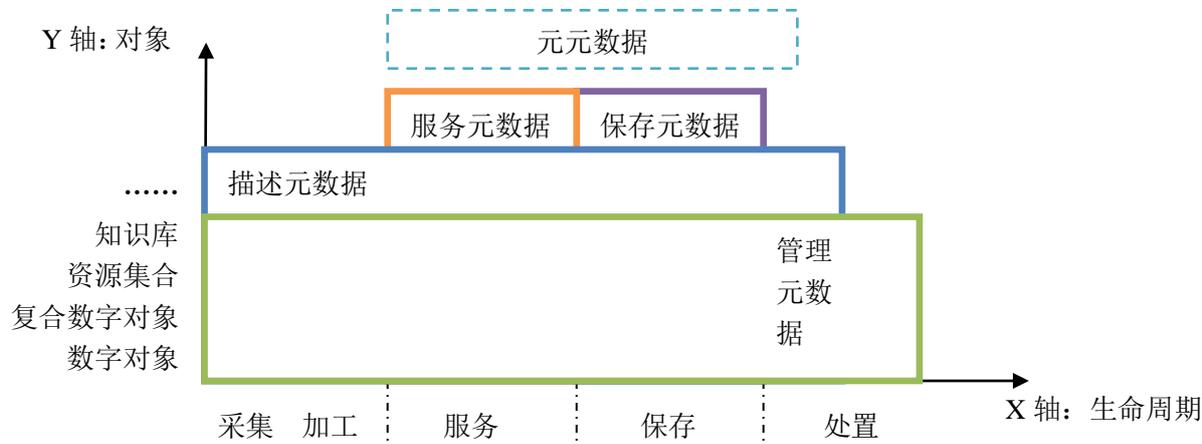


图5 文物数字化对象元数据关系模型

“元元数据”是一种存在于各个阶段元数据的基础规范，在整个文物元数据体系中是管理其他元数据的“元数据”标准。

描述元数据和管理元数据都存在于文物数字化对象资源的整个生命周期，基本是并行存在的。描述元数据主要存在于采集加工阶段，但对在其他元数据中资源描述的功能也不可或缺，而管理元数据在文物数字化对象的采集、加工、服务、保存、处置的各个阶段，都起着重要的管理、控制功能。描述元数据、管理元数据通过服务元数据进行格式、编码、标识等的规范，有助于用户或第三方系统方便地、无障碍地共享本系统的信息资源，以实现不同系统的数据交换共享。

服务元数据主要在文物数字化对象资源被采集加工后，面向读者提供资源检索利用的服务阶段起作用。它是将描述元数据进行展示、对描述元数据、管理元数据、保存元数据的互操作、共享进行规范的元数据。

保存元数据存在于文物数字化对象资源的保存阶段，也是对描述、管理、服务环境等信息的保存，以便于信息资源的长期保存、未来的使用和阅读。在文物数字化对象资源的生命周期中，信息资源，尤其是数字信息资源的长期保存是其有效利用的基本保证。我们没有把保存性元数据纳入管理性元数据框架之内，在信息资源管理的整体架构中建立长期保存机制，可采用已有的、国际通用的长期保存元数据方案，但需要在管理性元数据的相应模块设立特定的元素以与保存性元数据之间建立有机的联系。

以上不同类型的元数据在信息资源生命周期中实现的主要功能虽然不同，但彼此是互相关联的一个整体，都是为信息资源的建设提供标准化数据，从而实现信息资源的共享。

不同类型元数据之间的界定，以它们在信息资源生命周期中的不同功能为标准进行区分，例如文物数字化过程中产生的信息由管理元数据记录，而文物数字化的结果信息由描述元数据记录。在具体的原始设计上允许交叉和复用，但这种重复应以最小限度为前提，应通过元数据的模块化和开放性使不同的元数据可以相互链接和调用，形成不同的元数据应用方案。

参 考 文 献

- [1] 张晓林. 元数据研究与应用. 北京: 北京图书馆出版社, 2002
 - [2] 肖珑, 赵亮. 中文元数据概论与实例. 北京: 北京图书馆出版社, 2007
 - [3] 肖珑, 申晓娟. 国家图书馆元数据应用总则规范汇编. 北京: 国家图书馆出版社, 2011
-