

中国科学院国家科学图书馆青年人才领域前沿项目结题报告

电子文件元数据方案研究

中国科学院档案馆

2009年11月

目 录

1.规范性引用文件.....	3
2.国内外研究现状.....	3
2.1 国外电子文件元数据研究.....	3
2.1.1 典型案例.....	4
2.1.2 国际电子文件管理元数据标准.....	4
2.1.3 电子文件元数据项目及其技术实现.....	4
2.2 国内电子文件元数据研究.....	5
2.2.1 关于元数据标准设计框架.....	5
2.2.2 对元数据的管理.....	5
2.2.3 典型案例.....	6
2.3 研究的必要性.....	6
2.3.1 我国档案行业尚未有一个通用的元数据标准.....	6
2.3.2 现有元数据标准不适用中国科学院的档案管理.....	6
3.研究的理论依据.....	6
3.1 文件连续体理论.....	6
3.2 来源原则的“重新发现”.....	8
3.3 前端控制理论.....	8
4.中国科学院电子文件管理现状.....	9
4.1 文件管理流程分析.....	9
4.1.1 三个相关责任者.....	9
4.1.2 多个相关信息系统:.....	9
4.2 文件运动过程分析.....	10
4.3 元素的采集.....	11
4.3.1 院 ARP 电子政务管理系统.....	11
4.3.2 院档案馆档案信息管理系统.....	12
4.3.3 档案实体类型.....	13
5.电子文件元数据方案.....	13
5.1 功能描述.....	13
5.2 元数据总体结构.....	14
5.2.1 内容信息 (Content Information).....	15
5.2.2 保存描述信息 (Preservation Description Information).....	15
5.3 元数据集结构.....	15
6.核心元数据元素定义方法.....	20
6.1 内容信息 (Content Information).....	20
6.1.1 数据对象 (Date Object).....	21
6.2 保存描述信息 (PreservationDescriptionInformation).....	21
6.2.1 内容描述信息 (Content Description Information).....	22
6.2.2 结构信息 (Structure Information).....	25
6.2.3 背景信息 (Context Information).....	33
6.2.4 管理信息 (Management Information).....	41
7 元数据的封装与应用.....	58

1.规范性引用文件

下列文件中的条款通过本规范的引用而成为本规范的条款。

- GB 18030-2005 信息技术 中文编码字符集
- GB/T 7156-2003 文献保密等级代码
- GB/T 7408-2005 数据元和交换格式 信息交换 日期和时间表示法
- GB/T 9704-1999 国家行政机关公文格式
- GB/T 15418-1994 档案分类标引规则
- GB/T 18894-2002 电子文件归档与管理规范
- DA/T 1-2000 档案工作基本术语
- DA/T 12-1994 全宗卷规范
- DA/T 13-1994 档号编制规则
- DA/T 18-1999 档案著录规则
- DA/T 19-1999 档案主题标引规则
- DA/T 22-2000 归档文件整理规则
- DA/T 31-2005 纸质档案数字化技术规范
- 国家档案局, 征求意见稿 电子文件管理细则 第一部分: 文书电子文件元数据方案
- ISO 15489 Information and documentation—Records management
- ISO 23081-1 Information and documentation – Records Management Processes – Metadata for Records—Part 1: Principles
- ISO 14721:2003 Open archives information system
- ISO/IEC 11179-3 :1994 Information technology – specification and standardization of data elements – part3: basic attributes of data elements

随着电子政务的发展, 日益增多的电子文件成为档案部门管理的对象, 如何保证电子文件的真实性和完整性并永久保存具有珍贵价值的电子文件是文件和档案部门亟待解决的问题。而确保电子文件的真实性、完整性和有效性的关键就是元数据的管理, 因此, 电子文件安全科学的管理在很大程度上取决于元数据的管理。

2.国内外研究现状

对于元数据的管理, 国内外相关行业都已经开展或正在开展着相关研究, 并且取得了一定的进展。

2.1 国外电子文件元数据研究

国外对元数据的研究起步比较早, 且已有很多成功的案例为我们开展元数据研究提供标准和参考, 尤其是档案管理的元数据研究。

2.1.1 典型案例

(1)EAD 编码文档描述(Encoding Archival description): SGML 的一个专用数据类型。美国国会图书馆网络开发、MARC 标准办公室是其维护机构,美国档案管理员协会(SAA, the Society of American Archivists)是该标准的所有者。现在已有了 EAD 2002 版,此版本在第一版的基础上新增了<实体技术><phystech>这个元素,由 EAD DTD 1.0 版的 145 个元素变为 146 个元素。EAD 是为数字资源而制定的,其设计原则具有可扩展性、可选择性、可重复性等特点。

(2)美国加州遗产项目(California Heritage Project):加州遗产收藏是一个公共的数字档案馆,收藏有关加州历史的照片、画片、手稿等。他们将大量的这种资料通过扫描的方式数字化,然后直接用 EAD 著录,到 1998 年 3 月,他们已在网上提供了 154 个 EAD 编码的检索工具。

2.1.2 国际电子文件管理元数据标准

(1)《ISO / IEC 11179-1 (1999) 信息技术——元数据注册——第一部分:数据元素的说明及标准化框架》,由国际标准组织制定的国际标准。

(2)IS015489,由国际标准化组织的信息与文献技术委员会(ISO / RC46)下设的档案 / 文件管理分委员会(SC11)起草,于 2001 年 9 月作为国际标准正式发布。IS015489 综合了世界各国文件管理的先进经验,吸收了国际公认的最优化文件管理方法,与 ISO 系列标准和 ICA 国际标准有较好的互操作性、连贯性,适用于指导纸质、电子和网络工作背景下的一切文件的管理。

(3)《ISO / TR23081-1 信息与文件—文件管理过程—文件元数据》,由信息与文献技术委员会(ISO/TC46)下设的档案 / 文件管理分委会(SCLL)起草。该报告服从 ISO 其他委员会的规定,包括 ISOTC10 制定的 IS082045(数据管理,技术文献元数据)和 ISOTC171 等规定。

除上述国际通用标准外,可借鉴的其他标准还包括:澳大利亚联邦电子文件管理元数据标准,澳大利亚维多利亚州电子文件管理元数据标准,英国电子文件管理元数据标准等。

2.1.3 电子文件元数据项目及其技术实现

(1)元数据元素(语义结构):这方面国外发达国家元数据项目更为丰富,以都柏林元数据为例,其第 5 版规定了 15 个元素。

(2)元数据项目的技术实现问题(语法结构):

目前常见的档案元数据格式主要有:对档案内容描述的元数据,如,国际档案著录标准(总则)ISAD(G)(General International Standard Archival Description)、档案编码著录 EAD(En Archival Description)(EAD 是目前全球档案界最有影响的档案著录结构标准之

一)；描述档案背景信息的元数据，如，法人、个人及家庭背景信息国际档案规范文本 ISAAR(CPF) (International standard Archival Authority Record for Corporate Bodies, Persons, and Families)、编码档案背景规范 EAC(Encoded Archival context)；其他，如档案、个人论文和手稿 APPM (Archives, Personal papers, and Manuscripts)等。

2.2 国内电子文件元数据研究

2001年由国家档案局邱晓威研究员主持的“电子文件和电子档案的真实性、完整性保证及其法律地位的认定”国家社科项目的正式启动，标志着电子文件元数据研究开始成为我国档案界研究的最前沿课题。2004年，南京政治学院上海分院信息管理系张正强教授申报的《基于XML的电子文件管理元数据标准研究》获得国家社会科学基金重点课题立项，标志着我国电子文件元数据标准进入了实质性研究阶段。

2.2.1 关于元数据标准设计框架

元数据设计框架，为设计定制管理某种特定资源所用的元数据标准提供了需要遵循的规则和方法，从而确保了元数据标准的科学性和有效性。据文献分析，有关元数据标准设计框架的研究，在我国主要以张晓林和肖珑的研究为代表。

我国以ISO11179系列标准和GB / T18391为基础制定了电子政务数据元设计规范，主要阐述电子政务元数据的设计和管理方法论，规定了电子政务元数据的基本概念和结构、电子政务元数据的表示规范以及特定属性的设计规则和方法，并给出了电子政务元数据动态维护管理机制。到目前为止，我国电子政务标准化总体框架已经建立，并确立了一部分国家标准。

2.2.2 对元数据的管理

(1)集成管理。何嘉荪教授认为，要实现电子文件的集成管理与集成服务，必须以元数据为最基本和最重要的工具，并在此基础上提出了“基于元数据的电子文件集成管理与服务模式”。(2)集中与交换管理。黄蓓光、王申康通过比较研究后提出了元数据管理的另一种策略：通过建立元数据交换途径，使处于异构系统中的元数据可以用这种途径来完成交换管理，从而实现元数据的集成管理；这种元数据交换途径包括元数据桥和元数据交换标准。(3)分布式对象管理。杨晓春提出了分布式环境下的信息合作模型，设计了以区域为单位的分布式对象管理体系结构，以及相应的元数据管理。

2.2.3 典型案例

(1) “浙江大学电子文件管理元数据标准”遵循了文件运动理论以及前端控制和全程管理原则，全面描述了电子文件运动过程中产生的相关元数据，从元数据标准模型设计、原数据标准和新元素定义以及著录实例来进行设计的。

(2) 《基于XML电子文件管理元数据标准研究》是南京政治学院上海分院院长正强教授正在研究的课题，通过对大量现行元数据标准和相关研究成果的分析吸收，根据电子文件管理的总体目标，总结出电子文件元数据标准的设计规则和方法，即电子文件元数据设计框架，并在此设计框架的基础上，制定了基于XML的电子文件元数据标准，开展了关系元素、时间元素、责任者元素等元数据元素的语义结构研究。

(3) 国家档案局在2008年下发《电子文件元数据标准》（征求意见稿）和《文书电子文件元数据方案》（征求意见稿）是分别由深圳市档案局和安徽省档案局制定，但目前还没有正式下发我国的元数据标准。

2.3 研究的必要性

2.3.1 我国档案行业尚未有一个通用的元数据标准

我国档案方面尚未有一个通用的元数据规范，无论是深圳市档案局的《电子文件元数据标准》（征求意见稿）还是安徽省档案局的《文书电子文件元数据方案》（征求意见稿）都尚未作为国家正式的元数据标准下发。

2.3.2 现有元数据标准不适用中国科学院的档案管理

现有的元数据规范主要是文书档案和电子公文方面的规范，缺少科研档案元数据的相关规范，而中国科学院档案馆的所保存的档案资源主要是科研档案，因此现有的规范和标准都无法直接应用于中国科学院档案馆的元数据管理的需要。

3. 研究的理论依据

理论是实践的指南，如果档案工作者不能全面地把握并准确地揭示电子文件的生命运动特征，则无法进行科学合理的管理。

3.1 文件连续体理论

澳大利亚档案学者弗兰克·厄普沃德（Frank Upward）于20世纪90年代提出了一种全

新的“文件连续体理论”，从多维的、立体的角度对电子文件生命周期作了描述，即“往复运动于从生成到处置的连续体中的一个过程”¹。该理论构建了一个包括四个坐标轴的多维体系，将电子文件的生命运动过程及其与外界的联系清晰地展现出来，如图1所示。

图1中，文件保管轴代表用来存储人类活动中形成的记录信息的工具，轴上坐标分别为文件、案卷、全宗和全宗集合；凭证轴由行动轨迹、档案所提供的凭证以及它们在机构和集体记忆中所处的地位所组成；业务活动轴由行为、活动、功能和目的组成；实体轴由承办者、部门、部门所属的机构和社会组成。其中，文件保管轴是核心轴，它的变化带动了其它坐标轴的变化：第一围中形成的是单份文件，对应的形成者是某个具体的承办人，反映的业务活动是某一具体行为，表现出的价值是具体行为轨迹；第二围中形成的是案卷，即一组文件的集合，此时对应的形成者是机构内部的某一部门，反映的业务活动是包含若干行为的某一活动，表现出的价值是活动的凭证；第三围中形成的是全宗，即一个机构所有案卷的集合，对

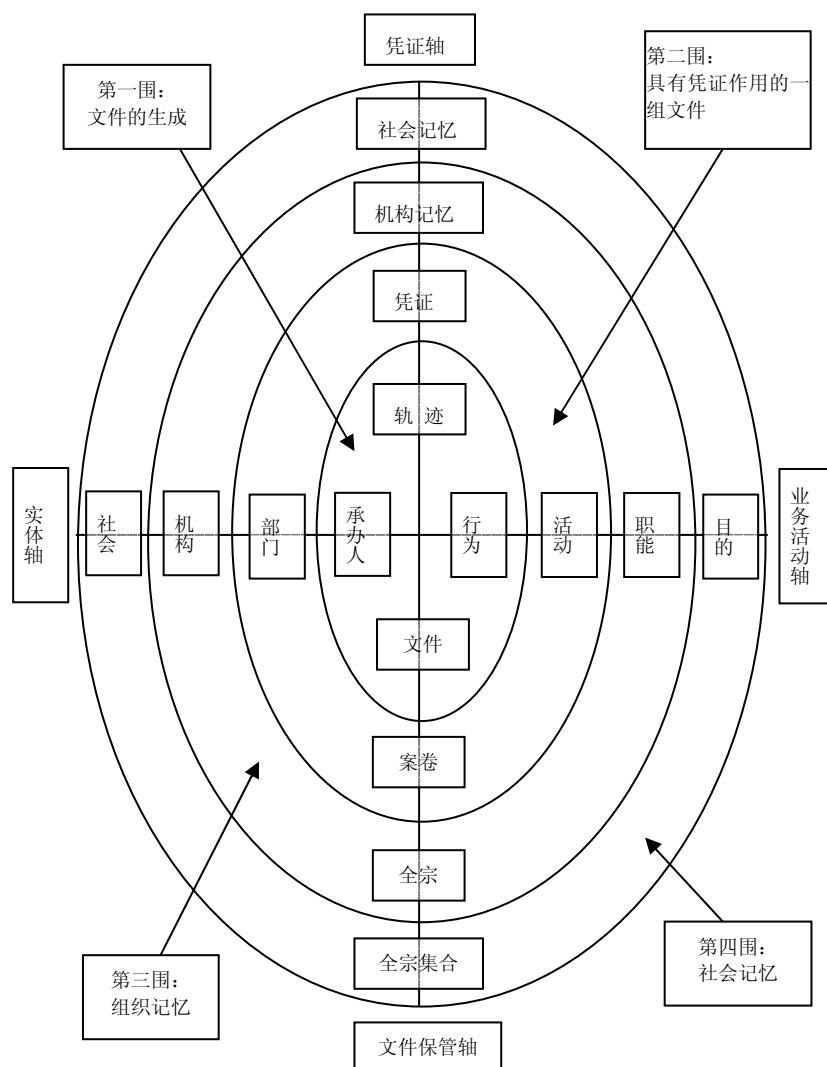


图1：文件连续体模型的图示

¹ Frank Upward. Structuring the records continuum. Archives and Manuscripts. vol. 24. No.26

应的形成者是一个机构，反映的业务活动是包含若干活动的某一职能，表现出的价值是机构记忆；第四围中形成的是全宗集合，即所有全宗的集合，对应的形成者是整个社会，反映的活动是社会的目的，表现出的价值是社会记忆。

3.2 来源原则的“重新发现”

来源原则就是“同一来源的文件/档案不应与其它任何来源的文件/档案相混淆的基本原则”²。一直以来，来源原则都是国际档案界公认的档案专业基本原则。首先，它是档案整理的指导原则，它把全宗确立为档案实体整理和分类的基本单元，为馆藏整理提供了客观真实的立足点；其次，它是档案馆组织的指导原则，它划清了档案与图书的界限，从根本上维护了档案的本质属性和档案专业的独立性；再次，它是档案理论研究的指导原则，它促使研究者深入挖掘档案的历史形成过程和有机联系，以揭示档案的来龙去脉，保持历史的本来面貌。

来源原则同样在电子文件管理上具有一贯的、至高无上的地位。1985年，美国档案学者莱特和比尔曼发表了著名的论文《来源原则的力量》，指出应当运用来源原则指导档案检索，要求档案人员不应局限于关注档案的主题内容，而应重视文件形成者的职能，从而建立一种更有生命力的来源索引。加拿大档案学者T·库克指出电子文件时代，“来源的概念需要重新考虑、酝酿和定义，档案人员的关注焦点应当从以等级结构中原始文件产生部门为中心的实体来源，转变为变动、临时、甚至‘虚拟’的机构中文件形成者的职能和业务活动为重点的来源概念”³，从传统理论对实体保管对象——实态文件的关注转变为对文件、文件形成者及其形成过程的有机联系、目的、意图、相互关系、职能和可靠性的关注。

3.3 前端控制理论

前端控制理论是相对于传统纸质文件的保管模式来说的。在传统环境中，文件管理和档案管理是作为两个独立的系统分开运行的，两者间的交界点为归档。我们通常可以认为，文件（本处指的是“大文件”概念，包括我国特定称谓的文件与档案）所做的是线性运动，文件管理是前端，档案管理是后端，并且鉴于文件工作和档案工作之间泾渭分明的状况，档案工作者对文件（指“大文件”）的干预仅能从后端开始，对归档之前能决定档案质量的文件工作只能心有余而力不足。

这种文件工作与档案工作的分家管理，档案工作者只能进行“后端”控制的管理模式，在电子文件管理上表现出极大的不适应性。对增、删、改不易留痕迹的电子文件，如果还像传统档案那样，按照生成、积累、收集、整理、保管等环节的先后顺序进行管理，其结果很

² 国际档案理事会. 档案术语词典（1984年版）第367条

³ [加拿大]T·库克. 1898年荷兰手册出版以来档案理论与实践的相互影响. 第十三届国际档案大会报告集. 1996年9月. P87

可能是诸多管理工作还未来得及开展，电子文件就已经被更改甚至删除了。

针对这种情况，国际档案界提出了“前端控制”思想，指出“档案工作者干预电子文件的时机，应确定在决定必须保存文件和设计归档系统阶段”⁴，把需要和可能在文件形成阶段实现或部分实现的管理功能尽量在这一阶段实现，使一些传统档案工作中的后工序出现的问题提前介入，以确保有价值的电子文件不被丢失。尤其是为了保证著录信息的准确、标准，提高效率，可以将著录项（即电子文件管理元数据要素）内嵌入电子文件管理系统，由系统自动完成大部分著录工作。

4. 中国科学院电子文件管理现状

为了设计符合中国科学院档案管理工作的元数据规范，我们首先对中国科学院电子文件的管理及其运动过程进行了实际调研和分析。

4.1 文件管理流程分析

文件系统为电子文件的产生、处理、传递、存储、管理和利用提供了相应的环境和工具，离开了文件系统，电子文件就不能有效存在。因此确定电子文件元数据的来源，首先应该明确电子文件的流程。

4.1.1 三个相关责任者

（1）文件的形成者——院属各研究所：中国科学院系统下属80多个研究所，其形成的具有永久保存价值的档案都要向院档案馆移交

（2）文件的处理者——院属各研究所档案管理部门：对本单位所形成的文件进行收集、整理、归档并向中国科学院档案馆移交。

（3）文件的管理者——中国科学院档案馆：对全院具有永久保存价值的档案进行科学安全的管理，并提供便捷高效的利用和最大价值的开发。

4.1.2 多个相关信息系统：

随着中国科学院电子政务的实施，中国科学院电子文件主要在以下系统中形成和管理，如，院ARP电子公文系统负责电子文档的形成与流转；院ARP档案管理系统负责电子文档的预归档、归档以及鉴定移交；院档案馆档案信息管理系统负责各进馆单位重要电子文件的永久保存、管理与利用。若是将中国科学院电子文件的管理流程进行抽象化，可以建立以下文件系统关系模型，如图2：

⁴ International Council on Archives. “Guideline on Electronic Records Management”. 1997年2月

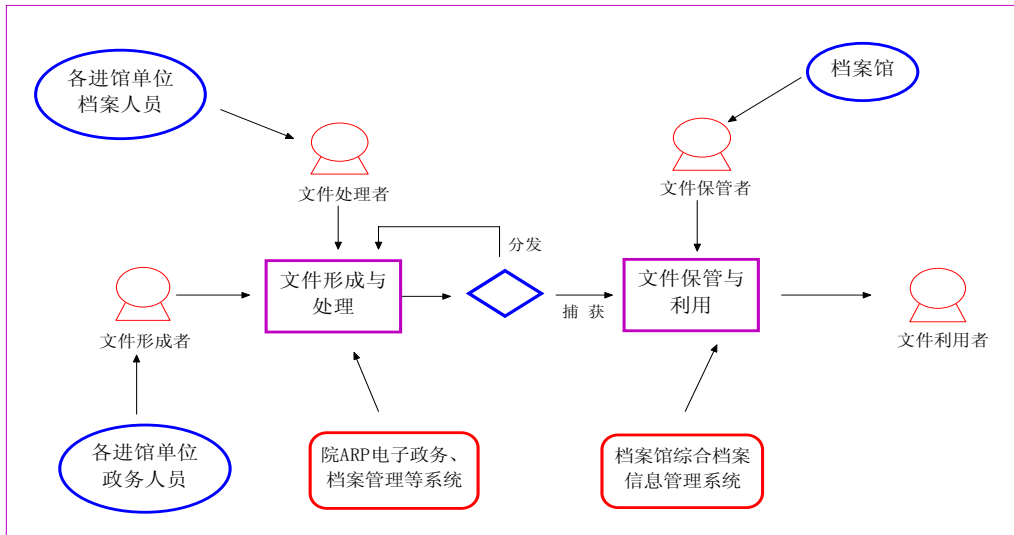


图2：文件管理流程模型图

4.2 文件运动过程分析

明确电子文件运动过程是采集相应元数据元素的主要工作之一，根据文件系统关系模型，按照电子文件全程管理的思想，在电子文件整个生命周期中，其运动过程可以抽象为四个过程：文件形成与处理、文件捕获（收集）、文件鉴定与整理、文件保管和利用。如图3：

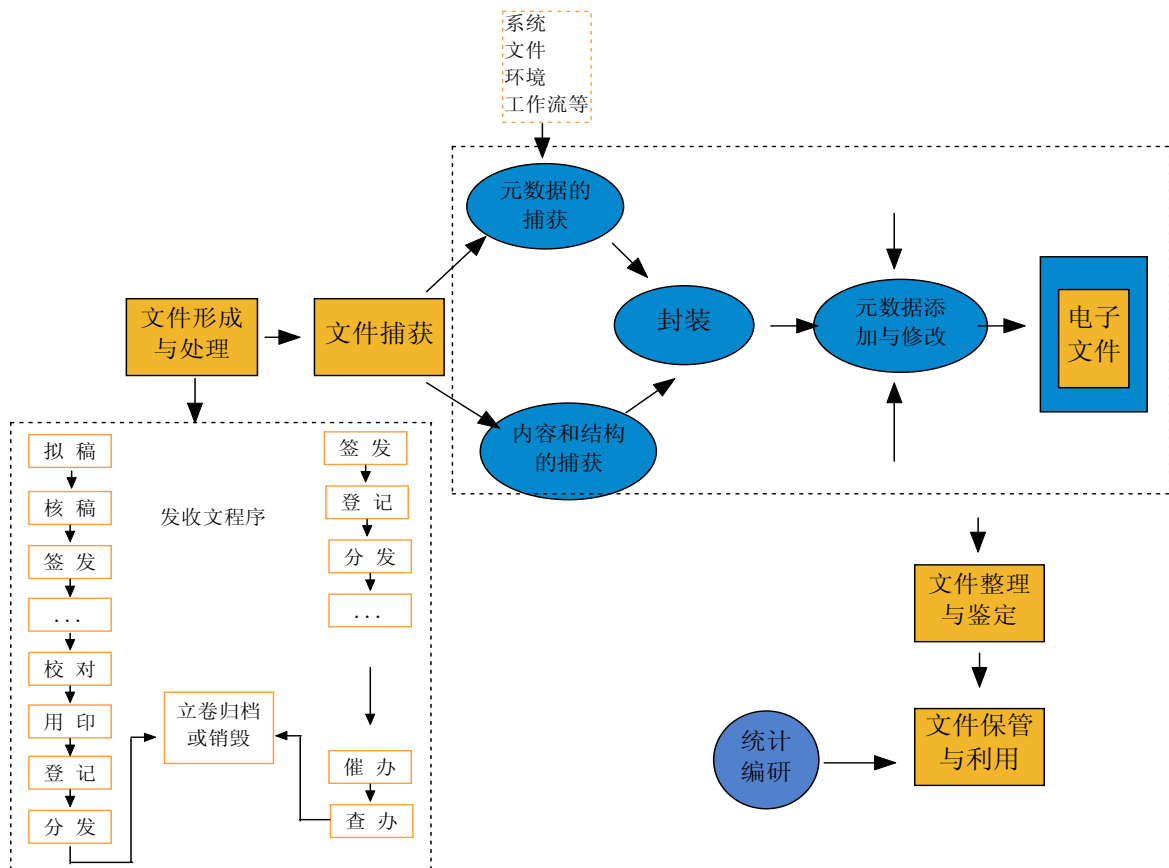


图3：文件运动过程模型图

4.3 元素的采集

根据中国科学院电子文件的运动过程，以及各相关系统的环境分析，借鉴国内外有关标准，采集了以下相关的元数据信息：

4.3.1 院 ARP 电子政务管理系统

包括公文系统、财务系统、档案系统等，共199个元素，根据电子文件管理的需要，采集主要元素如下：

流程	元素名称	
文件形成与处置	文件题名	签发人
	密级	文件编号
	内容摘要	日志文件
	主题词	保管期限
	主送单位	文种
	抄送单位	立档单位
	形成单位	页数
	责任者	归档时间
	成文时间	归档号
	附件信息	文件格式
过程文件	语种	
文件保管与利用	全宗号	文号
	全宗名称	页号
	档案馆档号	页数
	档案馆序号	主题词
	立档单位档号	文本路径
	分类号	原文路径
	分类名称	数据存储格式
	题名	数据存储方式
	立档部门	系统日志
	年度	销毁记录
	起始日期	损坏记录
	终止日期	丢失记录
	文件份数	回收记录
	总页数	结构定义
	保管期限	录入字段
	密级	数据排序
	备考说明	报表维护
	案卷序号	档号定义规则
	立卷人	数据字典

	立卷日期	追加携带项
	检查人	统开发语言
	卷内顺序号	系统运行环境

4.3.2 院档案馆档案信息管理系统

包括档案存储系统、检索利用系统等，共78个元素，根据档案保管和利用的需要，采集主要元素如下：

流程	元素名称	
文件的管理	档案馆档号	文号
	档案馆序号	页号
	立档单位档号	页数
	全宗号	主题词
	全宗名称	获奖情况
	分类号	专利情况
	分类名称	文本路径
	国家分类号	原文路径
	题名	数据存储格式
	项目名称	数据存储方式
	课题名称	系统日志
	立档部门	销毁记录
	年度	损坏记录
	起始日期	丢失记录
	终止日期	回收记录
	文件份数	结构定义
	总页数	录入字段
	保管期限	数据排序
	密级	报表维护
	备考说明	档号定义规则
	案卷序号	数据字典
立卷人	追加携带项	
立卷日期	系统开发语言	
检查人	系统运行环境	
卷内顺序号		
文件的利用	编号	单位名称
	借阅日期	复印页数
	利用者	打印页数
	人次	利用目的
	查阅卷数	所查档案内容及要解决的问题
	利用方式	归还日期

4.3.3 档案实体类型

包括文书档案、科研档案、基建档案、照片档案、声像档案、著名人物档案、科研仪器设备档案、大科学装置工程档案等，共41个元素，根据实体档案管理的需要，除了以上档案信息管理系统中所具有的元素外，还需要采集主要元素如下：

元素名称	
共性	特殊
立档单位档号	国家分类号
案卷名称	项目名称
全宗号	课题名称
全宗名称	室组名称
分类号	获奖情况
分类名称	专利情况
文号	课题（内容摘要）
立档部门	盒（盘）号
年度	册号
起止日期	件号
总页数	目录号
保管期限	照片号
密级	底片号
备考表	参见号
案卷序号	载体类别
立卷人	时间长度
立卷日期	拍摄者
检查人	录制者
卷内序号	
页号	

5. 电子文件元数据方案

5.1 功能描述

电子文件载体的不确定性，使得用来记录来源信息，维系、体现反映电子文件之间有机联系的，只能是元数据。具体而言，元数据在电子文件管理中的应用功能主要体现在以下几个方面。

(1) 确保并且证明电子文件的真实性和凭证价值。

在元数据中将电子文件的内容特征和文件形成、处理和利用的过程忠实地记录下来，就能够帮助知晓电子文件的内容是否与形成时完全一样，从而证明其真实性和完整性。把电子

文件形成时的环境和背景信息记录在元数据中，更有助于烘托和证明电子文件的价值，尤其是证据价值，也使以后的文件利用者便于了解文件形成时的历史环境，从而正确理解其真实的意涵。

(2) 有助于对电子文件的保护与长期保存，使其具有长久有效性。

文件可以用作证据的特性，使我们必须特别注意对电子文件进行档案化处理，使之有可能作为档案长期保存。因而除了采取以上措施确保电子文件的真实性和完整性以外，还应该在元数据模型中详尽描述电子文件的格式信息、制作信息、转换方式、技术环境学术探讨等，以便于电子文件的转换与迁移，达到长期保护、保存的作用，使之长期可读、可理解，具有长久有效性。

(3) 有效地管理电子文件（支持电子文件信息的组织和查询）

元数据方案应该包含整个文件管理和利用过程中所需的全部重要元素，它既包括对文件本身的描述元素，又包括对文件的管理信息，它不但是文件著录标引的工具，也是文件鉴别认证 (identification)、鉴定、保护等其它环节的工具。不仅有效管理在异构平台上形成的所有的电子文件，也是完善和发展电子文件管理的首要条件。

(4) 方便查询和利用

元数据方案是文件著录的工具，必须从文件形成之初就能为利用者提供查询途径，使用户清楚了解文件及其形成背景之间的关系，方便利用者发现和检索文件，满足利用者的需求。

5.2 元数据总体结构

根据中国科学院电子文件的构成，及其生成、管理与利用等不同阶段的特点，同时借鉴国内外相关元数据标准的结构，确定电子文件元数据结构由2个元数据模块7个元数据子集构成。

元数据模块	元数据子集	元素
内容信息	数据对象	数字对象
		物理对象信息
	表述信息	结构信息
		语义信息
保存描述信息	内容描述信息	
	结构信息	
	背景信息	创建依据
		全宗信息
		创建过程
		参考信息
数字化过程		

	管理信息	关联信息
		标识信息
		鉴定信息
		归档信息
		权限管理
		维护历史
	利用历史	
固化信息		

5.2.1 内容信息 (Content Information)

指电子文件的正文内容，包括主文件、附件的正文内容。内容信息包括数据对象与数据对象表述信息。数据对象是指表达并保存电子文件正文、附件正文等的比特流。数据对象表述信息是指数据对象的描述信息，由电子文件的结构信息与语义信息构成。

5.2.2 保存描述信息 (Preservation Description Information)

是对电子文件自身最基本特征的描述，能在电子文件自生成直至永久保存全过程中用于显现、理解电子文件内容信息。保存描述信息由内容描述信息、结构信息、背景信息、管理信息、固化信息构成。

5.2.2.1 内容描述信息 (Content Description Information)

指描述、揭示与电子文件内容相关的内部与外部特征的信息。

5.2.2.2 结构信息 (Structure Information)

指表述电子文件内容排列，各构成部分之间的连接方式、相关文件之间的关系以及在存储器中的构成等信息。

5.2.2.3 背景信息 (Context Information)

指描述生成电子文件的职能活动、电子文件的作用、办理过程、结果、上下文关系以及对其产生影响的历史环境等信息。背景信息由来源信息、创建依据、关联信息、技术环境信息构成。

5.2.2.4 管理信息 (Management Information)

指描述电子文件管理过程的信息，由鉴定信息、归档信息、权限管理、维护历史等信息集组成。

5.2.2.5 固化信息 (Fixity Information)

用以保障并确认电子文件及其元数据信息可信性与完整性的信息，如能够描述数字签名的内容值、校验和等。

5.3 元数据集结构

根据中国科学院档案馆电子文件接收、管理以及利用的需要，同时借鉴国内外相关元数

据标准，拟定电子文件元数据集的结构如下：

元数据集	元素		子元素		子元素	
	名称	标识	名称	标识	名称	标识
内容信息 Content Info	数字对象	Digital Object				
	物理对象描述信息	Physical Object				
	数据对象结构信息	DOStructure Info				
	语义信息	Semantic Info				
内容描述信息 Content Des Info	题名	Title	项目名称	ProjectTitle		
			课题名称	programTitle		
			案卷题名	FilesTitle		
			文件题名	DocumentTitle		
	并列题名	Alternative Title				
	副题名及说明题名文字	Title Annotation				
	责任者	Creator				
	文件编号	RecordNum				
	成文时间	Create Date				
	附件	Annex				
	稿本	Manuscript				
	主题词	Subject				
	摘要	Abstract				
	备注	Annotation				
	分类号	Classification	国家分类号	NationClassification		
	分类名称	SeriesName				
	获奖情况					
	专利情况					
	语种	Language				
	内容描述扩展信息	ContentDes Extent				
结构信	文种	RecordType				
	文件集合	Aggregation Level	集合类型	Aggregation Type		

息 Structure Info			集合描述	Aggregation Des		
			集合文件 标识	AGGRecords Identifier		
	格式	Format				
	存储介质	MediumType				
	文件大小	ObjectSize				
	保存历史	Preservation History	操作类型	ActionType		
			操作描述	ActionDes		
			操作时间	ActionTime		
			操作责任 人	Operator		
	原始技术 环境	Original Environment	原始信息 系统描述	OriginalSys temDes		
			原始档案 管理信息 系统描述	OriginalRec ordsDes		
	数字化结 构信息	DPStructure Info	分辨率	AclinicReso lution		
			色彩模式	ColorModel		
			总画幅数	ImageTotal		
			扫描仪型 号	ScannerSN		
			数字化结 构扩展信 息	DPExtent		
	数码照片 结构信息	DPhotoStruct ure Info	大小			
			分辨率	Resolution		
			色彩模式	PColorModel		
			数码相机 型号	CameraSN		
结构扩展 信息	Structur eExtent					
背景 信息 Context Info	创建依据	CreateBG	法律法规 依据	LegalBG		
			行政或行 业依据	AdminBG		
	全宗信息	FondInfo	全宗名称	FondName		
			全宗构成	Composingof Fond		
			全宗维护 历史	HistoryofFo nd		
	创建过程	Create Process	处理类型	ProcessType		
			处理人	Processor		
			接收处理 时间	ReceiveTime		
			实际处理 时间	ProcessTime		

			处理后发出时间	SendTime			
			处理意见	ProcessView			
	参考信息	Reference Info	主送机关	SubmitTo			
			抄送机关	Copyto			
			印发机关	PrintDistribute			
			印发日期	PrintDate			
			关联标识符	RelatesID			
	关联信息	Relates Info	关联类型	RelatesType			
			关联描述	RelatesDes			
			制作者	DPMaker			
	数字化过程	Digital Processing	制作者身份证明	DPCertification			
			制作日期	DPMakeDate			
			制作地	DPMadePlace			
			审核人	DPExaminer			
			审核时间	DPExaminedDate			
			背景扩展信息	Context Extent			
	管理信息 Management Info	标识信息	Identifier Info	标识符	Identifier		
				档案类别	ArchivesClassification		
				档号	RecordID	档案馆档号	Archive sID
						立档单位档号	Creator ID
档案馆序号				ArchivesNO.			
全宗号				FondID			
目录号				CatalogueID			
案卷号				FilesID			
页号				PageNum			
件号				ItemNum			
光盘编号				DiscID			
照片号				PhotoID			
底片号				FilmID			
参见号				ReferenceID			
册号				VolumID			
音视频文件类别				AVClassification			
时间长度				TimeLength			

			开始时间	StartTime		
			结束时间	EndTime		
	鉴定信息	Appraisal Info	鉴定时间	AppraisalDate		
			鉴定人	Appraiser		
			鉴定意见	AppraisalView		
	归档信息	Archiving Info	归档时间	ArchivingDate		
			归档方式	IsOnLine		
			移交责任者	HandoverActor		
			接收责任者	ReceiveActor		
	权限管理	Rights	密级	SecurityLevel		
			保管期限	PreservationPeriod		
			版权说明	CopyrightStatement		
			授权对象	AuthorizationTo		
			授权行为	AuthorizationAct		
			安全分类	SecurityClassification		
	维护历史	History Of Disposal	处置类型	DisposalType		
			处置责任者	DisposalActor		
			处置时间	DisposalDate		
			处置依据	DisposalReason		
			处置过程	DisposalProcess		
处置结果			DisposalResult			
利用历史			UseHistory	利用时间	UseDate	
	利用类型	UseType				
	利用描述	UseDes				
管理扩展信息	Management Extent					
固化信息 FixityInfo	数字签名	Digital Signature				

	数字签名 时间	Digital SignatureDate				
	数字签名 描述	Digital SignatureDes				

6.核心元数据元素定义方法

本规范参照国际标准IS011179和国际上元数据开发应用的最佳实践来定义元数据元素的12个属性：

- 名称(Name)：元素名称
 - 标识(Identifier)：元素的唯一标识
 - 版本(Version)：该元素所属的元数据版本
 - 注册机构(Registration Authority)：注册元素的授权机构
 - 语种(LangIJage)：元素说明的语种
 - 定义(Definition)：对元素概念与内涵的说明
 - 必备性(Obligation)：说明元素是必备的还是可选的
 - 数据类型(Datatype)：元素值中所表现的数据类型
 - 可重复性 (Repeatability)：元素的可重复行与值域重复
 - 值域 (Value)：元素的指定值或允许范围
 - 下层元素 (Sub-elements)：某上位元素的下位子元素
 - 注释(Comment)：元素应用注释
- 其中所定义三个属性是相同的：

- 版本：1. 0
- 注册机构：中国科学院档案馆
- 语种：中文

6.1 内容信息 (Content Infomation)

名称	内容信息
标识	ContentInfo
定义	内容信息是指电子文件的正文内容，包括文件的正文、附件的正文内容等，是电子文件行文或形成的目的。内容信息包括数据对象与数据对象表述信息。
类型	复合型
值域	-
必备性	必备

可重复性	不可重复		
下层元素	名称	标识	必备性
	数据对象	DateObject	必备
	表述信息	RepresentationInfo	必备
注释	根据 OAIS 的定义，档案信息包的内容信息由数据对象和表述信息构成。		

6.1.1 数据对象 (Date Object)

名称	数据对象		
标识	DateObject		
定义	数据对象是指表达并保存电子文件正文、附件正文等的比特流。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识	必备性
	数字对象	DigitalObject	必备
注释			

6.1.1.1 数字对象 (Digital Object)

名称	数字对象		
标识	DigitalObject		
定义	用于保存、表示电子文件内容信息的比特流 (Bit)，即二进制数据。		
类型			
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	-		
注释	当电子文件在系统中被签发，或纸质档案数字化后，数字对象将自动生成，由系统自动捕获。		

6.2 保存描述信息 (PreservationDescriptionInformation)

名称	保存描述信息		
标识	PreservationDesInfo		
定义	保存描述信息是对电子文件自身最基本特征的描述，能在电子文件自生成直至		

	永久保存全过程中用于显现、理解电子文件内容信息。保存描述信息由内容描述信息、结构信息、背景信息、管理信息、固化信息构成。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识	必备性
	内容描述信息	ContentDesInfo	必备
	结构信息	StructureInfo	必备
	背景信息	ContextInfo	必备
	管理信息	ManagementInfo	必备
	固化信息	FixityInfo	必备
注释	在办公自动化等业务信息系统中、文档一体化管理系统、档案管理信息系统，应设计自动捕获元数据功能。		

6.2.1 内容描述信息 (Content Description Information)

名称	内容描述信息		
标识	ContentDesInfo		
定义	内容描述信息是指描述、揭示与电子文件内容相关的内部与外部特征的信息。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识	必备性
	题名	Title	必备
	责任者	Creator	必备
	文件编号	RecordNum	必备
	文件形成时间	CreateDate	必备
注释			

6.2.1.1 题名 (Title)

名称	题名		
标识	Title		
定义	题名，是档案的主要题名，一般指单份文件文首的标题、文件组合级的标题、案卷封面上的案卷级标题、照片档案与音视频档案的标题。科研档案还包括研		

	究项目的项目名称、课题名称等。		
类型	字符型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识	必备性
	项目名称	ProjectTitle	必备
	课题名称	ProgramTitle	必备
	案卷名称	FileTitle	必备
	文件题名	RecordTitle	必备
注释			

6.2.1.1.1 项目名称

名称	项目名称		
标识	ProjectTitle		
定义			
类型	字符型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素			
注释			

6.2.1.1.2 课题名称

名称	课题名称		
标识	ProgramTitle		
定义			
类型	字符型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素			
注释			

6.2.1.1.3 文件名称

名称	文件题名		
标识	Title		

定义	文件题名,一般指单份文件文首的标题或文件组合级的标题,题名照原文著录。
类型	字符型
值域	-
必备性	必备
可重复性	不可重复
下层元素	
注释	当原文件缺题名,或题名不能准确表达文件主题时,应重拟题名,并用“[]”标出

6.2.1.2 责任者 (Creator)

名称	责任者
标识	Creator
定义	也称作者,是指对档案内容进行创造、负有责任的机构、团体或个人。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	照片档案、音视频档案则著录文件的摄影者或录制者。

6.2.1.3 文件编号 (Record Number)

名称	文件编号
标识	RecordNum
定义	文件编号是文件制发过程中由制发机关,团体或个人赋予文件的顺序号。文件编号包括发文字号、科研试验报告流水号、标准规范类文件的统编号、图号等。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.1.4 成文日期 (Create Date)

名称	成文日期
标识	CreateDate
定义	指文件形成的时间。

类型	日期型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	<ul style="list-style-type: none"> ●照片档案、音视频档案则著录文件的拍摄时间。 ●日期格式 yyyy-mm-dd 著录。

6.2.1.5 摘要或文字说明 (Abstract)

名称	摘要或文字说明
标识	Abstract
定义	对课题（项目）内容的简介，反映其主要内容、研究过程等。对声像档案关于事件、人物、地点等重要内容的说明。
类型	字符型
值域	-
必备性	可选
可重复性	可重复
下层元素	-
注释	

6.2.1.6 分类号 (Classification)

名称	分类号
标识	Classification
定义	类目的简明编码、标记符号，可固定全部类目的先后次序和明确一个类目在分类体系中的特定位置以便于标引、排列和进行检索、组织馆（室）藏。
类型	字符型
值域	依据《中国档案分类法》著录。
必备性	可选
可重复性	可重复
下层元素	-
注释	

6.2.2 结构信息 (Structure Information)

名称	结构信息
标识	StructureInfo
定义	指表述电子文件内容排列、各构成部分之间的连接方式、相关文件之间的关系

	以及在存储器中的构成等信息，是电子文件的重要组成部分。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	可重复		
下层元素	名称	标识	必备性
	文种	RecordType	必备
	文件集合	AggregationLevel	必备
	格式	Format	必备
	文件大小	ObjectSize	必备
注释			

6.2.2.1 文种 (Record Type)

名称	文种
标识	RecordType
定义	按性质和用途确定的文件种类的名称。
类型	字符型
值域	命令、决议、指示、通知、报告、批复、函、会议纪要、说明书、协议书、鉴定书、任务书、判决书、国书、照会、诰、敕、奏折等。
必备性	必备
可重复性	不可重复
下层元素	-
注释	

6.2.2.2 文件集合 (Aggregation Level)

名称	文件集合		
标识	AggregationLevel		
定义	指文件按照同一主题、事由进行著录、管理而形成的物理的、或逻辑的集合。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识	必备性
	集合类别	AggregationType	必备
	集合描述	AggregationDes	必备
注释			

6.2.2.2.1 集合类别 (Aggregation Type)

名称	集合类别
标识	AggregationType
定义	指文件集合的类别。
类型	字符型
值域	单份、请示与批复、计划与总结、文件组合、丛编
必备性	必备
可重复性	不可重复
下层元素	-
注释	<p>●丛编是指在一个总题名下汇集若干同一类型或统一编号发布的系列文件。如关于中共江西省委第九次代表大会的所有归档文件。文件组合是关于同一事件或主题的文件组合。</p> <p>●在办公自动化等业务信息系统、文档一体化管理系统中，应将集合类别的值域嵌入系统之中供用户选择。</p>

6.2.2.2.2 集合描述 (Aggregation Description)

名称	集合描述
标识	AggregationDes
定义	指文件集合类别的名称。
类型	字符型
值域	-
必备性	可选
可重复性	不可重复
下层元素	-
注释	请示与批复集合、文件组合可著录表达主题的简要文字

6.2.2.3 格式 (Format)

名称	格式		
标识	Format		
定义	指电子文件的逻辑格式(电子文件的类别与数据格式)与物理格式(存储介质)。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识	必备性
	电子文件类别	ElecRecordsType	必备

	数据格式	DataFormat	必备
	存储介质	MediumType	可选
	特殊元数据描述	SpecialMetadataDes	可选
	格式扩展信息	FormatExtent	可选
注释			

6.2.2.3.1 电子文件类别 (Electronic Records Type)

名称	电子文件类别
标识	ElecRecordsType
定义	指电子文件的属性类别。
类型	字符型
值域	T、I、G、V、A、O、M、P、D
必备性	必备
可重复性	不可重复
下层元素	-
注释	<p>●T—文本文件，I—图像文件，G—图形文件，V—影像文件，A—声音文件，O—超媒体链结文件，M—多媒体文件，P—程序文件，D—数据文件。●电子文件类别代码应由系统自动获取，并自动赋值</p>

6.2.2.3.2 数据格式 (Data Format)

名称	数据格式
标识	DataFormat
定义	指归档电子文件的原始格式, 或经数字化生成的电子文件格式。
类型	字符型
值域	XML、PDF、TXT、TIFF、JPEG、JPEG-2000、SVG、AVI、WAV、MP3、MPEG-2、MPEG-4、DBF、DWF、DWG
必备性	必备
可重复性	不可重复
下层元素	-
注释	

6.2.2.3.3 存储介质 (Medium Type)

名称	存储介质
标识	MediumType
定义	指存储电子文件的物理载体。
类型	字符型
值域	CD-R、DVD+R、DVD-R、Hard disk、磁带。

必备性	可选
可重复性	可重复，当一份文件分别被存储于两种不同的介质时。
下层元素	-
注释	

6.2.2.4 文件大小 (Object Size)

名称	文件大小		
标识	ObjectSize		
定义	指各类电子文件存储所占空间的大小，计量单位为：BYTE（字节）、KB（千字节）、M（兆）、G（万兆）等。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识	必备性
	归档文件大小	ArchivingSize	
	数字化文件大小	DPSize	
注释			

6.2.2.4.1 归档文件大小 (Archiving Size)

名称	归档文件大小
标识	ArchivingSize
定义	指各类归档电子文件所占存储空间的大小。
类型	字符型
值域	-
必备性	可选
可重复性	不可重复
下层元素	-
注释	主要指原始归档电子文件的大小，应由系统自动捕获并赋值。

6.2.2.4.2 数字化文件大小 (Digital Processing Size)

名称	数字化文件大小
标识	DPSize
定义	馆（室）藏档案经数字化形成的电子文件的大小。
类型	字符型
值域	-
必备性	可选

可重复性	不可重复
下层元素	-
注释	元素值应由系统自动捕获并赋值。

6.2.2.5 保存历史 (Preservation History)

名称	保存历史		
标识	PreservationHistory		
定义	指电子文件被捕获、收集、登记进入文件管理系统或档案管理系统后,对每一次保存、备份、迁移等操作的时间、类型、操作人等方面的描述,以确保电子文件的可读性与可利用性。		
类型	复合型		
值域	-		
必备性	可选		
可重复性	不可重复		
下层元素	名称	标识符	必备性
	操作类型	ActionType	必备
	操作描述	ActionDes	必备
	操作日期	ActionDate	必备
	操作责任人	Operator	必备
注释			

6.2.2.5.1 操作类型 (Action Type)

名称	操作类型
标识	ActionType
定义	对一份电子文件执行保存操作的类型。
类型	字符型
值域	备份、介质更新、硬件平台迁移、操作系统平台迁移、数据库系统平台迁移、格式迁移、应用软件迁移、压缩、镜像。
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.2.5.2 操作描述 (Action Description)

名称	操作描述
标识	ActionDes
定义	对保存操作细节的描述。

类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	主要著录某种操作被执行的具体情况，如“备份 X035 全宗所有文本类电子文件。备份介质：DVD+R，SONY，光盘编号：436001T2007012。”等

6.2.2.5.3 操作日期 (Action Date)

名称	操作日期
标识	ActionDate
定义	指执行保存操作的日期与时间。
类型	日期型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	应由系统自动捕获操作日期与时间。

6.2.2.5.4 操作责任人 (Operator)

名称	操作责任人
标识	Operator
定义	执行保存操作的责任人姓名。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	由系统自动捕获操作责任人的用户名，系统中的用户名应以真实姓名赋值。

6.2.2.6 数字化结构信息 (Digital Processing Structure Infomation)

名称	数字化结构信息
标识	DPStructureInfo
定义	各类档案经扫描数字化加工形成的电子文件的部分结构信息。
类型	复合型
值域	-
必备性	可选

可重复性	不可重复		
下层元素	名称	标识	必备性
	光学分辨率	OpticalResolution	必备
	色彩模式	ColorModel	必备
	色彩深度	BitDepth	必备
	总画幅数	ImageTotal	必备
注释	<ul style="list-style-type: none"> ●本元素主要针对纸质档案数字化而设置。 ●除总画幅数、数字化结构扩展信息外，其余 5 个下层子元素都应在扫描应用软件中设为系统自动获取。 		

6.2.2.6.1 光学分辨率（Optical Resolution）

名称	光学分辨率
标识	OpticalResolution
定义	指对档案数字化时采用的实际光学分辨率，计量单位为 dpi（Dot Per Inch），如 600dpi。
类型	字符型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	

6.2.2.6.2 总画幅数（Image Total）

名称	总画幅数
标识	ImageTotal
定义	指一份文件级档案经扫描形成的图像总数。
类型	整数型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	

6.2.2.7 数码照片结构信息（Digital Photo Structure Information）

名称	数码照片结构信息
标识	DPhotoStructureInfo
定义	指表示数码照片尺寸、分辨率、色彩深度等技术环境的结构信息。

类型	复合型		
值域	-		
必备性	可选		
可重复性	不可重复		
下层元素	名称	标识符	必备性
	照片宽度	PhotoWidth	必备
	照片高度	PhotoHeight	必备
	水平分辨率	AclinicResolution	必备
	垂直分辨率	VerticalResolution	必备
注释	数码照片结构信息应在应用软件中设计为自动获取。		

6.2.2.8 音视频文件结构信息 (Audio Video Structure Information)

名称	音视频文件结构信息		
标识	AVStructureInfo		
定义	可由用户自定义的声音文件或影像文件结构信息。		
类型	复合型		
值域	-		
必备性	可选		
可重复性	可重复		
下层元素	-		
注释			

6.2.3 背景信息 (Context Information)

名称	背景信息		
标识	ContextInfo		
定义	背景信息是指描述生成电子文件的职能活动、电子文件的作用、办理过程、结果、上下文关系以及对其产生影响的历史环境等信息。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识符	必备性
	创建依据	CreateBG	必备
	全宗信息	FondInfo	必备
	创建过程	CreateProcess	必备

注释	
----	--

6.2.3.1 创建依据 (Create BackGround)

名称	创建依据		
标识	CreateBG		
定义	创建电子文件的相关法律、法规或行业依据信息。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识符	必备性
	法律法规依据	LegalBG	必备
	行政或行业依据	AdminBG	必备
注释	本元素主要针对立档单位自身形成、签发的文件而设置。收文是否著录本元素可自选。		

6.2.3.1.1 法律法规依据 (Legal BackGround)

名称	法律法规依据
标识	LegalBG
定义	指文件形成时所处的法律体系，具体指的是有关的法律法规和制度，规范工作流程的规章制度。即文件责任机构、组织据以成立或者得到授权的有关法律法规或者行政文件中的有关规定，以及它们具有的职权。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	此元素内容可由档案管理人员手工著录并录入系统后由系统自动赋值。

6.2.3.1.2 行政或行业依据 (Administration BackGround)

名称	行政或行业依据
标识	AdminBG
定义	是文件形成时所处的行政/行业体系，具体是依照行政/行业部门的职能及职权对电子文件进行的处置。
类型	字符型
值域	-
必备性	必备

可重复性	可重复
下层元素	-
注释	此元素内容可由档案管理人员手工著录并录入系统后由系统自动赋值。

6.2.3.2 全宗信息 (Fond Information)

名称	全宗信息		
标识	FondInfo		
定义			
类型	复合型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	名称	必备性
	全宗名称	FondName	必备
	全宗类型	FondType	必备
	立档单位组织机构代码	ORGCode	必备
注释			

6.2.3.2.1 全宗名称 (Fond Name)

名称	全宗名称
标识	FondName
定义	档案部门赋予档案全宗 (或者文件全宗) 的名称。
类型	字符型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	元素可由系统自动赋值。

6.2.3.2.2 全宗类型 (Fond Type)

名称	全宗类型
标识	FondType
定义	描述全宗的类型。分为主体全宗与客体全宗两种。
类型	字符型
值域	全宗、联合全宗、汇集全宗。默认值为全宗。
必备性	必选
可重复性	不可重复

下层元素	-
注释	<p>联合全宗：由两个或两个以上立档单位的互有联系不易区分全宗而作为一个全宗对待的档案实体。</p> <p>汇集全宗：由若干个文件数量很少且具有某些共同特征或联系的小全宗组成的作为一个全宗对待的档案整体。</p> <p>●元素可由系统自动赋值。</p>

6.2.3.2.3 全宗构成 (Composing of Fond)

名称	全宗构成
标识	ComposingofFond
定义	指本全宗内档案门类、档案数量、档案形成起止时间、珍贵档案等构成情况。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	元素可由档案管理人员手工著录，由系统自动赋值。

6.2.3.3 创建过程

名称	创建过程		
标识	CreatesProcess		
定义	电子文件形成或产生过程中的相关信息。发文创建过程包括拟稿、审核、校对、签发、会签、发文等流程。收文创建过程指收文、拟办、批办、承办、催办、注办、其他等工作流程。		
类型	字符型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识符	必备性
	处理类型	ProcessType	必备
	处理人	Processor	必备
	接收处理时间	ReceiveDate&Time	必备
	实际处理时间	ProcessDate&Time	必备
	处理后发出时间	SendDate&Time	必备
	处理意见	ProcessView	必备
注释	●创建过程主要指办公自动化等业务系统中，发文或收文文件的文书处理过		

	<p>程。</p> <ul style="list-style-type: none"> ●针对创建过程中捕获的每一个处理类型值，都应有相对应的处理人、处理时间及处理意见。 ●收文需经数字化再进入办文流程的，相关数字化过程信息由背景信息的下位元数据“数字化过程”表述。
--	--

6.2.3.3.1 处理类型（ProcessType）

名称	处理类型
标识	ProcessType
定义	指文书处理程序的类型。
类型	字符型
值域	发文：拟稿、审核、校对、签发、会签、发文 收文：收文、拟办、批办、承办、催办、注办、其他
必备性	必备
可重复性	可重复
下层元素	-
注释	元素值应由系统自动捕获。

6.2.3.3.2 处理人（Processor）

名称	处理人
标识	Processor
定义	指各文书处理流程责任人的姓名。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.3.3.3 接收处理时间（Receive Date/Time）

名称	接收处理时间
标识	ReceiveDate&Time
定义	指文书处理人员在业务系统中接收某一文书处理程序的日期与时间。
类型	日期型
值域	-
必备性	必备
可重复性	不可重复

下层元素	-
注释	系统对接收处理时间的捕获点应文书处理人员点击鼠标接收当前流程的时间。

6.2.3.3.4 实际处理时间 (Process Date/Time)

名称	实际处理时间
标识	ProcessDate&Time
定义	用于实际撰写处理意见的日期与时间。
类型	日期型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	系统对实际处理时间的捕获点应为实际撰写处理意见的开始时间。

6.2.3.3.5 处理后发出时间 (Send Date/Time)

名称	处理后发出时间
标识	SendDate&Time
定义	文书处理结果的发出日期与时间。
类型	日期型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	

6.2.3.3.6 处理意见 (Process View)

名称	处理意见
标识	ProcessView
定义	与同一文书处理类型相对应的处理意见。
类型	字符型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	可对机关文书处理过程中经常使用的处理意见进行分析、归纳，提炼出常用的处理意见短语或句子，并嵌入系统之中供用户选择。

6.2.3.4 关联信息 (Relates Information)

名称	关联信息		
标识	RelatesInfo		
定义	标识并定义文件之间、文件集合之间、文件与其它职能机构之间的关联关系。		
类型	复合型		
值域	-		
必备性	可选		
可重复性	可重复		
下层元素	名称	标识符	必备性
注释	<p>●关联信息元数据使用目的是为一份归档电子文件建立大的职能背景信息,维护文件的凭证性。但本元数据不用于描述责任者之间的关联,包括机构内部或不同机构。</p> <p>●关联信息元数据与文件集合元数据不同,文件集合表述的是相同事由文件的集合。</p>		

6.2.3.5 数字化过程 (Digital Processing)

名称	数字化过程		
标识	DigitalProcessing		
定义	指档案经数字化输入设备处理生成电子文件的过程。		
类型	复合型		
值域	-		
必备性	可选		
可重复性	不可重复		
下层元素	名称	标识符	必备性
	数字资源制作者	DPMaker	必备
	数字资源制作时间	DPMakeDate	必备
	数字资源制作地	DPMadePlace	必备
	数字资源审核人	DPExaminer	必备
	数字资源审核时间	DPExaminedDate	必备
注释	数字化输入设备包括扫描仪、数码相机等。		

6.2.3.5.1 数字资源制作者 (Digital Processing Maker)

名称	数字资源制作者		
标识	DPMaker		
定义	档案数字化的实际操作者,如扫描者、拍摄者等。		
类型	字符型		

值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.3.5.2 数字资源制作日期 (Digital Processing Make Date)

名称	数字资源制作日期
标识	DPMakeDate
定义	指档案数字化处理日期与时间。
类型	日期型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	●应由系统自动捕获制作日期值并赋值。

6.2.3.5.3 数字资源制作地 (Digital Processing Made Place)

名称	数字资源制作地
标识	DPMadePlace
定义	指开展档案数字化的实际地点。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.3.5.4 数字资源审核人 (Digital Processing Examiner)

名称	数字资源审核人
标识	DPExaminer
定义	经数字化生成的电子文件质量审核责任人。
类型	字符型
值域	-
必备性	可选
可重复性	可重复
下层元素	-

注释	如果需进行多次质量审核，应著录一审、二审、三审责任人姓名。
----	-------------------------------

6.2.3.5.5 数字资源审核时间 (Digital Processing Examined Date)

名称	数字资源审核时间
标识	DPExaminedDate
定义	经数字化生成的电子文件质量审核日期与时间。
类型	日期型
值域	-
必备性	可选
可重复性	可重复
下层元素	-
注释	●审核时间应设置为业务信息系统或应用系统自动获取。

6.2.4 管理信息 (Management Information)

名称	管理信息		
标识	ManagementInfo		
定义	描述归档电子文件鉴定、归档、权限、利用、维护等方面的管理信息。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	可重复		
下层元素	名称	标识符	必备性
	标识信息	IdentifierInfo	必备
	鉴定信息	AppraisalInfo	必备
注释			

6.2.4.1 标识信息 (Identifier Information)

名称	标识信息		
标识	IdentifierInfo		
定义	根据来源信息，为科学、有序地管理归档电子文件及馆（室）藏档案而赋予的标识符。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识符	必备性

	标识符	Identifier	必备
	档号	RecordID	必备
	档案馆序号	ArchiveNO.	必备
	全宗号	FondID	必备
	目录号	CatalogueID	必备
	案卷号	FilesID	必备
	页号	PageNum	必备
	国家分类号	NationSeriesID	必备
注释			

6.2.4.1.1 标识符 (Identifier)

名称	标识符
标识	Identifier
定义	根据特定规则编制的一组具有全国唯一性的代码。
类型	字符型
值域	-
必备性	必备
可重复性	不可重复
下层元素	
注释	

6.2.4.1.2 档号 (Record ID)

名称	档号		
标识	RecordID		
定义	档案馆（室）在整理和管理档案的过程中，以字符形式赋予档案的一组代码。		
类型	字符型		
值域			
必备性	必备		
可重复性	不可重复		
	名称	标识符	必备性
下层元素	档案馆档号	Archives ID	必备
	立档单位档号	RecordID	必备
	归档号	Erecord ID	必备
注释			

6.2.4.1.2.1 档案馆档号 (Archives ID)

名称	档案馆档号
----	-------

标识	ArchivesID
定义	用于标识档案馆的具有唯一性的代码。
类型	字符型
值域	根据《编制全国档案馆名称代码实施细则》所赋予的档案馆代码著录。
必备性	必备
可重复性	不可重复
下层元素	-
注释	应由系统自动赋值。

6.2.4.1.2.2 立档单位档号

名称	立档单位档号
标识	RecordID
定义	档案室在整理和管理档案的过程中，以字符形式赋予档案的一组代码。
类型	字符型
值域	各种门类的档号根据 DA/T13-94《档号编制规则》进行编制。
必备性	必备
可重复性	不可重复
下层元素	
注释	

6.2.4.1.2.3 归档号（Erecord ID）

名称	归档号
标识	ERecordID
定义	电子文件的名称，即文件名，用电子文件的档号命名。
类型	字符型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	应在形成档号后由系统设置的人-机互动方式命名，或者完全自动命名。

6.2.4.1.3 档案馆序号（Archives NO.）

名称	档案馆序号
标识	Archives NO.
定义	档案馆
类型	字符型
值域	

必备性	必备		
可重复性	不可重复		
下层元素			
注释			

6.2.4.1.4 全宗号 (Fond ID)

名称	全宗号		
标识	FondID		
定义	档案馆（室）给定每个全宗的代码。		
类型	字符型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	-		
注释	应由系统自动赋值。		

6.2.4.1.5 案卷号 (Files ID)

名称	案卷号		
标识	FilesID		
定义	馆（室）藏档案案卷排列的序号。		
类型	字符型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	-		
注释			

6.2.4.1.6 页号 (Page Number)

名称	页号		
标识	PageNum		
定义	一份文件在案卷内的起始页号。		
类型	字符型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	-		
注释	页号项也可以为张号。本元数据适用于以卷为保管单位的馆（室）藏档案。		

6.2.4.1.7 件号 (Item Number)

名称	件号
标识	ItemNum
定义	在一个事由下、一个案卷内或一张光盘内，以件为单位按顺序为文件编制的流水号。
类型	字符型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	本元数据适用于文书档案、音频档案、视频档案的管理。

6.2.4.1.8 分类号 (Series ID)

名称	分类号
标识	SeriesID
定义	全宗内按机构或问题设置的分类方案类目代码。
类型	字符型
值域	-
必备性	可选
可重复性	不可重复
下层元素	-
注释	

6.2.4.1.9 分类名称 (Series Name)

名称	分类名称
标识	SeriesName
定义	全宗内按机构或问题设置的分类方案类目名称。
类型	字符型
值域	-
必备性	可选
可重复性	不可重复
下层元素	-
注释	

6.2.4.1.10 光盘编号 (Optical Disc ID)

名称	光盘编号
标识	OpticalDiscID

定义	存储、备份电子文件光盘的规范性编号。
类型	字符型
值域	
必备性	可选
可重复性	不可重复
下层元素	-
注释	

6.2.4.1.11 照片档案扩展标识信息 (photo ID Extent)

名称	照片档案扩展标识信息		
标识	photoIDExtent		
定义	为有序管理照片档案而赋予的标识代码。		
类型	复合型		
值域	-		
必备性	可选		
可重复性	不可重复		
下层元素	名称	标识符	必备性
	照片号	PhotoID	必备
	底片号	FilmID	必备
	参见号	ReferenceID	必备
	册号	VolumID	必备
注释	各子元素需由照片拍摄者或档案管理人员手工著录。		

6.2.4.1.11.1 照片号 (Photo ID)

名称	照片号
标识	PhotoID
定义	是固定和反映每张照片在全宗内分类与排列顺序的一组字符代码。
类型	字符型
值域	-
必备性	必备
可重复性	不可重复
下层元素	-
注释	照片号即档号，参见“档号”元数据。

6.2.4.1.11.2 底片号 (Film ID)

名称	底片号
标识	FilmID

定义	是固定和反映底片在全宗内排列顺序的一组字符代码。
类型	字符型
值域	依据 GB/T 11821—2002《照片档案管理规范》编制。
必备性	可选
可重复性	不可重复
下层元素	-
注释	底片号的格式为：全宗号-保管期限代码-张号。

6.2.4.1.11.3 参见号 (Reference ID)

名称	参见号
标识	ReferenceID
定义	指与本张照片有密切联系的其他载体档案的档号。
类型	字符型
值域	-
必备性	可选
可重复性	可重复
下层元素	-
注释	本元素与关联信息密切相连。

6.2.4.1.11.4 册号 (Volum ID)

名称	册号
标识	VolumID
定义	在某一全宗某一保管期限内照片册的排列从“1”开始的顺序编号。
类型	字符型
值域	-
必备性	可选
可重复性	不可重复
下层元素	-
注释	

6.2.4.1.12 音视频档案扩展标识信息 (Audio&Video ID Extent)

名称	音视频档案扩展标识信息
标识	AVIDExtent
定义	标识信息中其它可自定义的有关音视频档案的标识信息。
类型	复合型
值域	-
必备性	可选

可重复性	不可重复		
下层元素	名称	标识符	必备性
	音视频文件类别	AVClassification	必备
	时间长度	TimeLength	必备
注释	各子元素应由音视文件制作系统自动捕获。		

6.2.4.1.12.1 音视频文件类别 (Audio & Video Classification)

名称	音视频文件类别		
标识	AVClassification		
定义	指定音频文件或视频文件。		
类型	字符型		
值域	A、V		
必备性	必备		
可重复性	不可重复		
下层元素	-		
注释	“A”指代音频文件，“V”指代视频文件。		

6.2.4.1.12.2 时间长度 (Time Length)

名称	时间长度		
标识	TimeLength		
定义	指音频文件或视频文件的时间长度。		
类型	时间型		
值域	-		
必备性	可选		
可重复性	不可重复		
下层元素	-		
注释	时间长度的应精确到秒。		

6.2.4.2 鉴定信息 (Appraisal Information)

名称	鉴定信息		
标识	AppraisalInfo		
定义	指对电子文件或馆(室)藏档案的鉴定,既包括对文件内容价值的鉴定,也包括对电子文件真实性鉴定和技术鉴定。鉴定元数据记录了对归档电子文件的历次鉴定信息。		
类型	复合型		
值域	-		
必备性	必备		

可重复性	可重复		
下层元素	名称	标识符	必备性
	鉴定类型	AppraisalType	必备
	鉴定时间	AppraisalDate	必备
	鉴定人	Appraiser	必备
	鉴定意见	AppraisalView	必备
注释			

6.2.4.2.1 鉴定类型 (Appraisal Type)

名称	鉴定类型
标识	AppraisalType
定义	指不同类型的鉴定方式。
类型	字符型
值域	归档鉴定、公开与非公开鉴定、进馆鉴定、开放与控制鉴定、期满鉴定、解密鉴定、销毁鉴定
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.4.2.2 鉴定时间 (Appraisal Date)

名称	鉴定时间
标识	AppraisalDate
定义	鉴定行为开展的时间。
类型	日期型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	以责任人签署鉴定意见的时间为准

6.2.4.2.3 鉴定人 (Appraiser)

名称	鉴定人
标识	Appraiser
定义	负责本次鉴定的责任者姓名。
类型	字符型
值域	-

必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.4.2.4 鉴定意见 (Appraisal View)

名称	鉴定意见
标识	AppraisalView
定义	本次鉴定结果及鉴定意见描述。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.4.3 归档信息 (Archiving Information)

名称	归档信息		
标识	ArchivingInfo		
定义	描述电子文件归档过程的相关信息。		
类型	复合型		
值域	-		
必备性	可选		
可重复性	可重复		
下层元素	名称	标识符	必备性
	归档时间	ArchivingDate	必备
	归档方式	IsOnLine	必备
	移交责任者	HandoverActor	必备
	接收责任者	ReceiveActor	必备
	逻辑归档标识	LIsArchiving	必备
	物理归档标识	PIsArchiving	必备
注释			

6.2.4.3.1 归档时间 (Archiving Date)

名称	归档时间
标识	ArchivingDate
定义	对电子文件进行逻辑归档或物理归档的日期与时间。

类型	日期型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.4.3.2 归档方式 (Is OnLine)

名称	归档方式
标识	IsOnLine
定义	指电子文件物理归档的方式。
类型	字符型
值域	在线、离线
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.4.3.3 归档类型 (Archiving Type)

名称	归档类型
标识	ArchivingType
定义	指电子文件的归档类型。
类型	字符型
值域	逻辑归档、物理归档
必备性	可选
可重复性	可重复
下层元素	-
注释	

6.2.4.3.4 移交责任者 (Handover Actor)

名称	移交责任者
标识	HandoverActor
定义	进行电子文件物理归档移交操作的执行者。
类型	字符型
值域	-
必备性	必备
可重复性	可重复

下层元素	-
注释	

6.2.4.3.5 接收责任者 (Receive Actor)

名称	接收责任者
标识	ReceiveActor
定义	进行电子文件物理归档接收操作的执行者。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.4.4 权限管理 (Rights)

名称	权限管理		
标识	Rights		
定义	归档电子文件及馆(室)藏档案数字化资源自身所具有的或被赋予的权限信息。		
类型	复合型		
值域	-		
必备性	可选		
可重复性	可重复		
下层元素	名称	标识符	必备性
	密级	SecurityLevel	必备
	保管期限	PreservationPeriod	必备
	安全分类	SecurityClassification	必备
注释			

6.2.4.4.1 密级 (Security Level)

名称	密级
标识	SecurityLevel
定义	密级是指归档电子文件、馆(室)藏档案保密程度的等级。
类型	字符型
值域	绝密、机密、秘密
必备性	必备
可重复性	不可重复
下层元素	-

注释	
----	--

6.2.4.4.2 保管期限 (Preservation Period)

名称	保管期限		
标识	PreservationPeriod		
定义	经鉴定后确认的归档电子文件、馆（室）藏档案保管期限。		
类型	复合型		
值域	永久、长期、定期		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识符	必备性
	保管期限名称	PPeriodName	必备
	保管期限代码	PPeriodCode	必备
注释			

6.2.4.4.2.1 保管期限名称 (Preservation Period Name)

名称	保管期限名称		
标识	PPeriodName		
定义	保管期限的类别名称。		
类型	字符型		
值域	永久、长期、定期		
必备性	必备		
可重复性	不可重复		
下层元素	-		
注释	应将值域先期嵌入系统供用户选择。		

6.2.4.4.2.2 保管期限代码 (Preservation Period Code)

名称	保管期限代码		
标识	PPeriodCod		
定义	保管期限类别的代码。		
类型	字符型		
值域	Y、C、D		
必备性	必备		
可重复性	不可重复		
下层元素	-		
注释			

6.2.4.4.2.3 保管期限值 (Preservation Period Value)

名称	保管期限值
标识	PPeriodValue
定义	保管期限为定期时所对应的保管年限值。
类型	字符型
值域	30 年、10 年
必备性	可选
可重复性	不可重复
下层元素	-
注释	

6.2.4.4.3 安全分类 (Security Classification)

名称	安全分类
标识	SecurityClassification
定义	为保证归档电子文件、馆（室）藏档案安全利用的安全分类信息。
类型	字符型
值域	开放、控制、公开、内部、未解密、已解密
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.4.4.5 维护历史 (History of Disposal)

名称	维护历史		
标识	HistoryofDisposal		
定义	对归档电子文件、馆（室）藏档案所做的维护历史情况的描述。		
类型	复合型		
值域	-		
必备性	可选		
可重复性	可重复		
下层元素	名称	标识符	必备性
	处置类型	DisposalType	必备
	处置责任者	DisposalActor	必备
	处置时间	DisposalDate	必备
	处置结果	DisposalResult	可选
注释			

6.2.4.5.1 处置类型 (Disposal Type)

名称	处置类型
标识	DisposalType
定义	指处置行为的类型。
类型	字符型
值域	开放、解密、保管期限更新、销毁、修改录入错误
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.4.5.2 处置责任者 (Disposal Actor)

名称	处置责任者
标识	DisposalActor
定义	指根据鉴定意见实施处置行为的责任者姓名。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.4.5.3 处置时间 (Disposal Date)

名称	处置时间
标识	DisposalDate
定义	实施处置行为的日期与时间。
类型	日期型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.4.5.4 处置结果 (Disposal Result)

名称	处置结果
标识	DisposalResult
定义	关于处置行为的结果描述。

类型	字符型
值域	-
必备性	可选
可重复性	可重复
下层元素	-
注释	

6.2.4.6 利用历史 (Use History)

名称	利用历史		
标识	HistoryofUse		
定义	自电子文件登记进入文件管理系统直到销毁的生命历程中，对归档电子文件、馆（室）藏档案数字化资源的合法访问与利用，或试图非法访问与利用日期等方面的描述信息。		
类型	复合型		
值域	-		
必备性	可选		
可重复性	可重复		
下层元素	名称	标识符	必备性
	利用时间	UseDate	必备
	利用类型	UseType	必备
	利用描述	UseDes	可选
注释			

6.2.4.6.1 利用时间 (Use Date)

名称	利用时间
标识	UseDate/Time
定义	对归档电子文件或馆（室）藏档案数字化资源访问、利用行为发生的日期与时间。
类型	日期型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.4.6.2 利用类型 (Use Type)

名称	利用类型
----	------

标识	UseType
定义	对归档电子文件或馆（室）藏档案数字化资源访问或利用的类型描述。
类型	字符型
值域	复制、阅读、下载、非法访问等。
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.5 固化信息（Fixity Information）

名称	固化信息 o		
标识	FixityInfo		
定义	用以封装档案信息包，确认电子文件及其元数据信息真实性、完整性和可靠性的信息。		
类型	复合型		
值域	-		
必备性	必备		
可重复性	不可重复		
下层元素	名称	标识符	必备性
	数字签名	DigitalSignature	必备
	数字签名时间	DigitalSignatureDate	必备
	数字签名描述	DigitalSignatureDes	必备
注释			

6.2.5.1 数字签名（Digital Signature）

名称	数字签名
标识	DigitalSignature
定义	确保电子文件真实性、完整性、有效性的方法。
类型	字符型
值域	-
必备性	可选
可重复性	可重复
下层元素	-
注释	

6.2.5.2 字签名时间 (Digital Signature Date)

名称	数字签名时间
标识	DigitalSignatureDate
定义	指实施数字签名的日期与时间。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

6.2.5.3 字签名描述 (Digital Signature Description)

名称	数字签名描述
标识	DigitalSignatureDes
定义	指对数字签名相关信息的描述。
类型	字符型
值域	-
必备性	必备
可重复性	可重复
下层元素	-
注释	

7 元数据的封装与应用

本规范对元数据表结构的定义依据XML规范进行,适用于基于XML规范的办公自动化等业务信息系统对元数据的设置、捕获与存储,以及最终生成基于XML的档案信息包。

在文档一体化管理系统中生成、运行、维护的结构化数据库中,档案管理人员可依据“电子文件元数据核心元素的定义和描述”中关于元数据注释属性的定义,进行元数据项的著录。

封装 (Encapsulation),是指元数据与电子文件一起被存储、传递、紧密联系的过程。⁵经元数据封装的电子文件,具有实际的自我说明能力与凭证价值。这也是本课题下一步研究的重点。

8.参考文献

1. 电子文件全程管理研究,硕士学位论文,肖英,2003.5

⁵ <http://www.prov.vic.gov.au/vers/published/final/6impleme.htm#6.2.3>

2. 浙江大学电子文件元数据规范, 浙江大学档案馆, 2005. 6
3. ARP 项目组, 中国科学院 ARP 系统数据说明书, 2004. 8
4. 中国科学院办公厅, 院属单位综合档案管理系统培训手册, 2004. 10