



中华人民共和国档案行业标准

DA/T 109—2026

档案专题数据库建设工作规范

Work specification for the construction of archival thematic database

2026-06-22 发布

2026-09-01 实施

国家档案局 发布

目 次

前言 III

引言 IV

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 总体要求 2

 4.1 需求导向 2

 4.2 统筹推进 2

 4.3 科学规范 2

 4.4 安全合规 2

 4.5 可信可用 2

5 组织管理 2

 5.1 组织与人员 2

 5.2 建设方式 3

 5.3 实施模式 3

 5.4 流程管理 3

6 技术保障 3

 6.1 软硬件保障 3

 6.2 系统功能 3

 6.3 技术增效 3

7 选题与工作方案 3

 7.1 提出选题 3

 7.2 分析选题 4

 7.3 确定选题 4

 7.4 制定工作方案 4

8 数据汇集 4

 8.1 汇集规范 4

 8.2 汇集方式 4

 8.3 汇集要求 4

 8.4 数据汇交 5

9 整理与著录 5

 9.1 筛选 5

 9.2 整理 5

9.3 著录 6

10 质量管理与验收..... 7

10.1 数据清洗 7

10.2 数据质量检查 7

10.3 验收 7

11 利用与宣传推广..... 7

11.1 利用共享 7

11.2 公开发布 7

11.3 宣传推广 7

12 管理与维护..... 8

12.1 反馈评价 8

12.2 动态管理 8

12.3 维护备份 8

附录 A（资料性） 扩展著录示例 9

附录 B（资料性） 来源信息著录示例 10

附录 C（资料性） 利用限制信息著录示例 11

附录 D（资料性） 通用著录项目对应关系示例 12

参考文献 13

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由国家档案局提出并归口。

本文件起草单位：国家档案局科技信息化司、云南省档案馆。

本文件主要起草人：王燕民、夏红、郝晨辉、龙岗、王大众、冯剑波、杨健生、林静、何畏、熊爱桃、张懿、罗竞铭、刘璐、周玉倩。

引 言

档案专题数据库建设是提高专项工作、重大活动、突发事件档案管理水平的的重要途径,是高效组织、管理和利用档案资源的手段,是深度挖掘和开发档案资源的方法。本文件提出的档案专题数据库建设的组织管理和实施等要求均以档案专题数据库建设的研究和实践为基础。

档案专题数据库建设工作规范

1 范围

本文件确立了档案专题数据库建设的总体要求,规定了数据库建设的组织管理、技术保障、选题、工作方案制定的要求,数据的汇集、整理、著录要求,质量管理与验收的要求,以及数据库的利用与宣传推广、管理与维护的要求。

注:本文件未涉及数据库、信息系统开发技术。

本文件适用于各级档案部门各类档案专题数据库的建设。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中,注日期的引用文件,仅该日期对应的版本适用于本文件;不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

- GB/T 11821 照片档案管理规范
- GB/T 18894 电子文件归档与电子档案管理规范
- GB/T 35273 信息安全技术 个人信息安全规范
- DA/T 18 档案著录规则
- DA/T 22 归档文件整理规则
- DA/T 31 纸质档案数字化规范
- DA/T 43 缩微胶片数字化技术规范
- DA/T 46 文书类电子文件元数据方案
- DA/T 50 数码照片归档与管理规范
- DA/T 54 照片类电子档案元数据方案
- DA/T 62 录音录像档案数字化规范
- DA/T 63 录音录像类电子档案元数据方案
- DA/T 78 录音录像档案管理规范
- DA/T 89 实物档案数字化规范
- DA/T 99 档案数字资源备份实施规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

档案专题数据库 archival thematic database

围绕特定对象、范围、内容等主题范畴,描述数据的特征及与数据对应的实体间的关系,按照概念结构组织的档案数据的集合。

注1:档案专题数据库描述这些数据的特征及与数据对应的实体间的关系,并支持一个或多个应用领域。

注2:档案专题数据库的数据包含数据集、文本、图片、图形、音频、视频等结构化、半结构化及非结构化数据。

3.2

数据汇集 data acquisition

根据档案专题数据库建设需要,按照一定规范,获取各种来源、载体形态的档案及其相关信息的过程。

3.3

数据清洗 data cleaning

查找存在不完整、不准确、不一致、格式不规范、重复等错误的数据,并进行纠正的过程。

4 总体要求

4.1 需求导向

应根据服务党和国家工作大局,弘扬时代精神和优秀传统文化,记录专项工作、重大活动、突发事件,服务民生,服务本行业、本单位工作等需求,明确档案专题数据库建设的目标、任务及工作内容,按需建设。

4.2 统筹推进

应统筹目标、任务、档案资源及资金、技术和人员保障,制定档案专题数据库建设工作方案,分步组织实施建设,对工作进度与质量进行全过程管理。

4.3 科学规范

应制定业务规范和技术方案,明确档案专题数据库建设的标准体系、数据规范、技术方法和质量要求等。

4.4 安全合规

档案专题数据库系统应按照国家网络安全等级保护标准建设和运维,采用符合国家信息技术安全要求的电子认证、密码应用技术措施,确保安全。涉及国家秘密的,应符合分级保护有关要求。涉及个人信息保护的,应符合 GB/T 35273 的要求。

4.5 可信可用

应记录数据的形成者、形成时间及数据汇集过程,确保数据来源可靠、可追溯。应落实数据质量管理措施,确保数据内容准确,信息完整。数据存储、元数据维护、转换与迁移应符合 GB/T 18894 的要求,做到数据管理有序,可读可用。

5 组织管理

5.1 组织与人员

5.1.1 应明确档案专题数据库建设的管理、实施部门及人员。

5.1.2 承担管理职责的部门负责监督指导、组织协调、验收评价、人员培训等工作,开展选题策划、建设工作方案制定,组织编制业务规范和技术方案,提供资金、技术等保障条件。可根据需要聘请有关专家和顾问。

5.1.3 承担实施职责的部门负责按照建设工作方案、合同或协议,开展收集、整理、保管、开发利用等工作,负责系统建设与运维、安全保障等。应根据所承担的工作内容,合理配备收集、整理、保管、开发、利

用、数据治理、信息系统管理等方面的专业人员,明确其工作职责,规范开展有关业务工作。

5.2 建设方式

应根据工作实际,采取以下一种或多种方式建设。

- a) 档案汇集。围绕特定主题,筛选馆(室)藏档案数字资源,形成档案专题数据库。
- b) 档案内容提取。围绕特定主题,对馆(室)藏档案内容进行分析,从中提取与主题相关的信息,形成档案专题数据库。
- c) 专项汇集。围绕专项工作、重大活动、突发事件等特定主题,组织开展相关文件材料收集、整理、数字化等,形成档案专题数据库。

5.3 实施模式

根据情况,可采取独立实施,跨区域、跨层级、跨部门合作实施,或服务外包实施等模式。

5.4 流程管理

档案专题数据库建设工作主要包括选题与工作方案、数据汇集、整理与著录、质量管理与验收、利用与宣传推广、管理与维护等。宜结合档案专题数据库建设工作实际,梳理所涉及的业务及其逻辑关系,按照有机衔接、规范管理的要求设计工作流程,确保建设质量、效率和安全。

6 技术保障

6.1 软硬件保障

根据档案专题数据库建设工作需要,宜配备数据汇集处理、存储及备份、安全保密、查阅利用等软硬件。

6.2 系统功能

档案专题数据库的建设宜基于本单位已有电子档案管理系统来实施。不具备电子档案管理系统,或已有电子档案管理系统不满足档案专题数据库管理要求的,应建设满足档案专题数据库管理要求的应用系统或功能模块。系统的功能设置符合以下要求:

- a) 档案专题数据库系统的数据汇集、整理、保存、利用等功能要求见 GB/T 39784 中电子档案接收、整理、保存、利用等的功能要求;
- b) 支持多来源数据整合、多平台发布利用;
- c) 在系统内设置有关模块或使用独立软件实现数据筛选、数据清洗功能。

6.3 技术增效

可应用大数据、人工智能等新一代信息技术提高数据汇集、整理与著录、质量管理、查阅利用的效率和精准度,增强数据展示的直观性和体验感。

7 选题与工作方案

7.1 提出选题

承担管理职责的部门宜结合工作需要和档案资源实际,提出选题意见,并明确其定位、作用和意义。

7.2 分析选题

应从必要性、可行性、效益等方面对选题进行分析。

- a) 必要性分析。从汇集整合档案资源、深化档案资源开发、促进社会各方面利用等现实需要进行分析,评估选题的必要性。
- b) 可行性分析。调研与主题相关档案的数量、类型、价值、密级、保管单位、应用场景等情况,了解相关文献资料、研究成果等状况,结合政策法规、安全保密、知识产权保护等要求,以及经费、技术、人才保障等因素进行可行性分析。采用专项汇集建设方式的,还应了解专项工作、重大活动、突发事件的工作安排、进展和档案管理情况。
- c) 效益分析。对效益指标进行分析,内容包括但不限于经济、社会、文化、生态、可持续影响等指标,预期成果应符合各项效益指标要求。

7.3 确定选题

应针对分析中存在的不足对选题进行优化,确定选题。可采用专家论证方式辅助确定选题。

7.4 制定工作方案

7.4.1 档案专题数据库建设工作方案应作为档案专题数据库建设工作组织管理、监督指导和实施的依据。

7.4.2 档案专题数据库建设工作方案的编制,宜以深入调研为基础,并遵循立足实际、维护历史真实面貌和便于实施的原则。

7.4.3 档案专题数据库建设工作方案主要包括工作内容和流程、业务规范和技术架构选型、安全合规保障、管理机制、预算和建设周期等内容。

7.4.4 跨区域、跨层级、跨部门协同建设档案专题数据库,应在建设工作方案中对有关知识产权归属、数据发布和共享利用规则、更新维护机制等作出安排。

8 数据汇集

8.1 汇集规范

承担管理职责的部门应组织制定数据汇集业务规范,明确数据汇集的内容、门类、范围、质量要求,规定质量检查流程与方法、数据汇集方式、责任与要求等。

8.2 汇集方式

可采用以下几种汇集方式:

- a) 汇集与主题相关的传统载体档案并进行数字化;
- b) 通过数据复制、在线或离线数据导入等方式,汇集目标系统中与主题相关的档案数据;
- c) 经合规授权,通过网络收集有关用户上传档案数据,或从网站等新媒体平台汇集主题相关档案数据;
- d) 采用文字、语音、图像识别、信息抽取等技术,汇集档案内容中与主题相关的数据。

8.3 汇集要求

应对数据汇集人员进行业务和技术培训,确保按照数据汇集业务规范确定的范围和要求实施数据汇集。汇集的数据应来源可靠、内容准确、信息完整、可读可用、过程规范。

- a) 来源可靠。数据来源可确认、可追溯,处理数据的相关系统安全可靠,能为数据的真实性提供

保障。

- b) 内容准确。数据全面、系统地反映主题或事实,没有逻辑性错误,与档案记载的内容保持一致。
- c) 信息完整。数据的内容、结构、背景信息和管理过程信息等符合数据汇集业务规范规定。
- d) 可读可用。数据符合数据汇集业务规范规定的质量要求,通过病毒检测,能正常读取和显示。
- e) 过程规范。数据汇集过程符合数据汇集业务规范的规定,汇集工作过程记录完整、可追溯。

8.4 数据汇交

承担工作实施职责的部门,应将所汇集的数据按照数据汇集业务规范规定的时间、频次和方式等要求,向指定的部门汇交。

9 整理与著录

9.1 筛选

应从与主题契合度、汇集要求的符合性、重要性方面对汇交的数据进行筛选,剔除不符合要求的数据及重复数据。

9.2 整理

9.2.1 数据整理

数据整理符合以下要求。

- a) 电子档案应符合 GB/T 18894 对电子档案真实性、可靠性、完整性和可用性的要求。其格式、元数据、档案信息包符合 GB/T 18894、DA/T 46、DA/T 50、DA/T 54、DA/T 63、DA/T 78 的要求。
- b) 传统载体档案的整理应符合 GB/T 11821、DA/T 22、DA/T 78 的要求,其数字化工作应符合 DA/T 31、DA/T 43、DA/T 62、DA/T 89 的要求。
- c) 声像档案宜主题鲜明,图像画面清晰、完整、端正,声音清晰,尽可能保证较高的技术质量;除不影响真实性的格式转换、裁剪外,不进行其他技术处理;数码照片如非必要,不应采用缩略图。
- d) 从档案内容中分析提取的数据应进行结构化整理,保证电子目录与档案原文关联正确。

9.2.2 数据组织

宜考虑档案专题数据库的管理和利用需求,按档案门类、内容、时间、来源、密级等对数据进行分类,对分散、多元、异构的数据进行有序组织。

9.2.3 数据编号

应制定数据编号规则,对数据以件为单位进行编号,确保编号在同一档案专题数据库内的连续性和唯一性。数据编号格式根据实际自定义。以专项汇集方式建设的档案专题数据库,其档案独立设置全宗或按专门类别管理的,可将档号作为数据编号。

示例 1:

数据编号格式定义为“档案专题数据库名称代码-分类代码-顺序号”,如:贸易博览会档案专题数据库,按年度分类,2024 年第 105 件,数据编号为:MYBLH-2024-00105。

示例 2:

数据编号格式定义为“档案专题数据库名称代码-顺序号”,如:劳模档案专题数据库,不分类,第 148 件,数据编号为:劳模-00148。

9.3 著录

9.3.1 一般要求

一般要求包括以下内容。

- a) 以专项汇集方式建设档案专题数据库,按照 DA/T 18、DA/T 46、DA/T 54、DA/T 63 等实施著录。
- b) 以档案汇集、档案内容提取等方式建设档案专题数据库,宜根据建库需求,选择所需著录项目。
- c) 根据需要扩展设置著录项目,制定著录规范,对主题、地名、人名、其他专门信息,以及来源信息和利用限制信息进行规范著录。扩展著录示例见附录 A。

9.3.2 主题信息著录

主题信息著录包括以下内容。

- a) 每个档案专题数据库应建立主题词表或关键词表,所设置的主题词或关键词应反映特定主题的构成内容。
- b) 主题词或关键词可动态增加,但总数不宜过多。
- c) 应根据每件档案内容,选取主题词或关键词进行著录。多个主题词或关键词并列标识,用“;”(中文分号)分隔。

9.3.3 地名信息著录

根据档案内容,对与主题相关的事件、人物所涉及的地名进行著录。按“国别(地区)—行政区划”分级著录。地名用全称或规范简称,在同一个专题数据库内,同一地名的著录信息应一致。多个地名应逐个著录,用“;”(中文分号)分隔。历史地名照原文著录,可将考证后的当今地名附后注明。

9.3.4 人名信息著录

根据档案内容,选择与主题直接相关的人物,按其在该件档案中呈现的姓名与身份,进行人名信息著录。可根据需要设置多个著录项目,著录姓名、别名、曾用名、籍贯、职衔或身份等信息。

9.3.5 其他专门信息著录

对反映主题、数据管理和提供利用有重要作用的其他专门信息进行著录。

9.3.6 来源信息著录

应对所收集数据的来源信息进行著录,示例见附录 B。来源信息著录应符合以下要求:

- a) 来源信息反映提供数据的机构或个人的基本信息、获取方式、获取时间;
- b) 以档案汇集、档案内容提取方式获取的数据,著录原档号信息;
- c) 以专项汇集方式获取的数据,著录移交者信息。

9.3.7 利用限制信息著录

数据利用限制信息的著录信息来源包括有关政策规定、法律法规、协议或移交记录,著录示例见附录 C。

9.3.8 通用著录项目对应关系

可根据跨库、跨类检索利用的需要,明确有关著录项对应关系,示例见附录 D。

10 质量管理与验收

10.1 数据清洗

10.1.1 根据需要,可选择在数据汇集、筛选、整理、著录等环节进行数据清洗。

10.1.2 应对数据缺失、重复、乱码、多余字符、逻辑错误、表述不规范、格式不规范、文件损坏问题,以及电子目录与档案原文关联错误情况进行检核纠正。

10.1.3 开展数据清洗,应制定操作规程,明确数据清洗规则、流程、方法和要求,记录数据清洗过程。数据清洗前,应进行数据备份。

10.2 数据质量检查

在数据汇集、筛选、整理、著录等环节,应按照数据汇集业务规范的要求,采用人工或技术辅助的方法,对数据进行100%质量检查,合格后方可进入下一工作环节。

10.3 验收

档案专题数据库建设完成后,经测试和试运行,承担管理职责的部门依据档案专题数据库建设工作方案、有关合同或协议、相关标准规范等,组织开展验收。

11 利用与宣传推广

11.1 利用共享

档案专题数据库的利用共享符合以下要求。

- a) 档案专题数据库利用参照本单位档案开放审核、利用等制度执行。档案专题数据库共享按照档案专题数据库建设工作方案或共享协议实施。
- b) 应建立档案专题数据库利用共享渠道,包括但不限于档案查阅系统、多媒体交互展示系统、政务信息共享系统、网站系统、社交媒体等。
- c) 可开发档案专题数据库专题片、短视频、电子相册等产品提供利用。

11.2 公开发布

档案专题数据库的公开发布符合以下要求:

- a) 档案专题数据库公开发布应遵守档案专题数据库建设工作方案对数据发布和知识产权归属的安排;
- b) 宜根据安全保密要求、数据使用限制要求等,合理选择数据向有关发布利用平台或渠道分发投放,当发布环境不符合安全保密、个人信息保护要求时,应进行脱敏脱密处理。

11.3 宣传推广

宜配合档案专题数据库发布,制作宣传资料、文创产品等,以及利用电视、广播、报刊、网络、融媒体等多种渠道,开展形式多样的宣传推广活动。

12 管理与维护

12.1 反馈评价

适时开展用户满意度调查,收集用户反馈意见,对档案专题数据库的用户满意度、成本效益、可持续性、安全性等绩效指标进行综合评价。

12.2 动态管理

宜根据用户反馈意见和综合评价结果,及时优化档案专题数据库功能;根据专题内容、性质及馆(室)藏档案增长变化情况,适时对数据库进行扩展、更新。

12.3 维护备份

档案专题数据库应纳入馆(室)藏档案资源安全体系,开展日常管理维护和应急工作。档案专题数据库的备份应符合 DA/T 99 的规定。

附 录 A
(资料性)
扩展著录示例

表 A.1 给出了扩展著录的示例。

表 A.1 扩展著录示例

扩展著录项目名称	内容	著录示例
主题信息	根据内容和主题归属,选取主题词或关键词进行著录	1. 反映特定主题的主题词或关键词:脱贫攻坚,乡村振兴…… 2. 反映特定主题构成内容主题词或关键词:基础设施,产业发展,保障措施…… 3. 反映具体事物的主题词或关键词:传统民居,古村落……
地名信息	对与主题相关事件、人物所涉及的地名进行著录	1. ××省××市××区××街道××社区 2. ××省××市 3. ××省××县(××省××市××区) ^a ……
人名信息	选择与主题内容直接相关的人物,按其在档案中呈现的姓名和身份,规范进行人名信息著录	1. 姓名:张三 2. 身份:××县县长,××局局长,养殖大户,非物质文化遗产传承人…… 3. 籍贯:××省××市……
其他专门信息	对反映主题、数据管理和提供利用有重要作用的其他专门信息进行著录	1. 民族:傣族,景颇族…… 2. 区域:革命老区,边境地区…… 3. 类别:建筑,桥梁……
^a “××省××县”为历史地名,“××省××市××区”为考证后的当今地名		

附 录 B
(资料性)
来源信息著录示例

表 B.1 给出了来源信息著录的示例。

表 B.1 来源信息著录示例

来源信息著录项	内容	示例
来源名称	反映提供数据的机构或个人信息	1. ××省档案馆 2. ××省发展与改革委员会 3. ××卫视 4. 张三,××村农民 5.“http://www.yunnan.cn,××网”
获取方式	获取数据的途径	接收,××馆(室)藏档案,获赠,购买,寄存,下载, 收录.....
获取时间	获取数据的时间	1.2013-11-25 2.20131125
移送者	反映专项汇集方式的数据移送者信息	1. ××组委会××部门××人 2. ××会议××组××人 3. ××省应急管理厅××部门××人 4. ××市××事件处置工作组××人

附 录 C
(资料性)
利用限制信息著录示例

表 C.1 给出了利用限制信息著录的示例。

表 C.1 利用限制信息著录示例

利用限制信息著录项	内容	示例
控制标识	根据数据安全利用需要设定的相应管理标识	公开,档案馆内部使用,档案涉及人员或机构使用,政务网使用,经审批后使用,互联网公开使用……
授权行为	需经有关单位批准,或需著作权、肖像权使用授权等限制使用要求	报批有关单位使用,作者授权使用,肖像授权使用……
密级和保密期限	反映数据保密程度和保密期限	秘密、机密、秘密 5 年、机密 10 年……
……		

附 录 D

(资料性)

通用著录项目对应关系示例

表 D.1 给出了通用著录项目对应关系示例。

表 D.1 通用著录项目对应关系示例

通用著录项目	文书类档案元数据		照片类档案元数据		录音录像类档案元数据	
	编号	元数据	编号	元数据	编号	元数据
题名	M22	题名	M6	题名	M7	题名
档号	M8	档号	M5	档号	M6	档号
档案馆代码	M4	档案馆代码	M1	档案馆代码	M1	档案馆代码
内容描述	M21	内容描述	M15	主题	M19	主题
来源	M2	来源	M29	来源	M23	来源
格式信息	M47	格式信息	M51	格式信息	M48	格式信息
责任者	M32	责任者	M7	责任者	M8	责任者
日期	M33	日期	M11	时间	M13	时间
密级	M38	密级	M35	密级	M29	密级
保管期限	M12	保管期限	M33	保管期限	M27	保管期限
.....						

参 考 文 献

- [1] GB/T 5271.1—2000 信息技术 词汇 第1部分:基本术语
- [2] GB/T 15418 档案分类标引规则
- [3] GB/T 20273—2019 信息安全技术 数据库管理系统安全技术要求
- [4] GB/T 35274 数据安全技术大数据服务安全能力要求
- [5] GB/T 39362 党政机关电子公文归档规范
- [6] GB/T 39784 电子档案管理系统通用功能要求
- [7] GB/T 42727 政务服务事项电子文件归档规范
- [8] DA/T 2 科学技术研究项目档案管理规范
- [9] DA/T 13 档号编制规则
- [10] DA/T 22 归档文件整理规则
- [11] DA/T 32 公务电子邮件归档管理规则
- [12] DA/T 50 数码照片归档与管理规范
- [13] DA/T 68 档案服务外包规范
- [14] DA/T 70 文书类电子档案检测一般要求
- [15] DA/T 77 纸质档案数字复制件光学字符识别(OCR)工作规范
- [16] DA/T 78 录音录像档案管理规范
- [17] DA/T 80 政府网站网页归档指南
- [18] DA/T 88 产品数据管理(PDM)电子文件归档与电子档案管理规范
- [19] DA/T 103 招标投标电子文件归档规范
- [20] DA/T 104 ERP系统电子文件归档和电子档案管理规范
- [21] 中华人民共和国个人信息保护法(中华人民共和国主席令第91号)
- [22] 网络数据安全管理条例(中华人民共和国国务院令第790号)
- [23] 重大活动和突发事件档案管理办法(国家档案局令第16号)
- [24] 国家档案馆档案开放办法(国家档案局令第19号)
- [25] 电子档案管理办法(国家档案局令第22号)
- [26] 国家档案局办公室关于印发《电子档案管理系统基本功能规定》的通知(档办发〔2017〕3号)