

《录音录像档案数字化规范》解读

文/徐亮

现阶段,档案数字化工作在全国档案部门广泛开展,但档案载体类型较多,不同载体类型的档案数字化开展程度差异较大。特别是由模拟信号形成的录音档案和录像档案数字化工作,由于技术、管理、标准等原因,起步晚、进展慢,这极大地影响了录音录像档案的保存和利用。为保障录音录像档案数字化工作的规范性和科学性,由国家档案局技术部、国家档案局档案科学技术研究所等部门和单位共同起草制定的档案工作行业标准DA/T 62—2017《录音录像档案数字化规范》已于2017年8月2日正式发布,自2018年1月1日起实施。

制定背景

数字技术出现之前,录音录像档案多由模拟信号形成并依赖专门的读取设备。随着时间的侵蚀,受自然衰变和保管条件等诸多因素的影响,录音录像档案正在慢慢地损坏。同时,由于技术发展等原因,许多种类的读取设备正在或者已经消失,如果不及时采取措施抢救这些珍贵的档案信息资源,将给国家和社会造成无法弥补的损失。因此,利用传统读取技术和计算机等数字技术相结合的方式,将录音录像档案进行数字化处理,形成数字档案资源,达到抢救及保护珍贵档案资源并提供方便、有效利用的目的,已成为必然趋势。

但是,在录音录像档案数字化实践中存在着大量的问题,在技术水平、标准化程度等方面都有所欠缺。录音录像档案种类多、技术依赖性强,其数字化技术及管理方法较纸质等类型的档案更为复杂。录音录像档案的载体类型具有多样性,包括开盘式录音带、盒式录音带、钢丝录音带、唱片、录像带等,每种载体类型根据不同的技术规格又可分为不同类别。不同载体、不同技术规格的录音录像档案涉及的读取设备各不相同,增加了数字化工作的复杂性和技术难度。各档案馆采用的录音录像档案数字化技术标准不统一,造成了数字化成果质量很难保证和存储方式、格式不一致,这对数字化成果的存储、交换和利用等各个环节都产生了不利影响,极大地制约了录音录像档案数字化工作的开展。在这样的背景下,2012年,由国家档案局技术部、国家档案局档案科学技术研究所等部门和单位共同提出申请,全国档案工作标准化技术委员审议通过立项,制定《录音录像档案数字化规范》。

制定原则

在标准的制定过程中,课题组以调研和实验数据分析为基础,以满足实际工作需要为出发点,确定了以下制定原则:

1.符合标准编写规则

标准是按照GB/T 1《标准化工作导

则》、GB/T 20000《标准化工作指南》、GB/T 20001《标准编写规范》、GB/T 20002《标准中特定内容的起草》等标准的规定和要求编写的。

2.细化数字化工作技术要求

对录音录像档案数字化技术方面的要求进行细化,使之符合当前数字化工作的需要。

3.强化数字化工作管理要求

录音录像档案数字化工作的开展,其管理与技术起着同等重要的作用。从课题组开展调研得到的反馈情况看,档案部门对数字化管理工作标准化的要求非常强烈。在标准的制定过程中,课题组特别针对数字化组织和管理提出了明确要求。

4.强化安全管理理念

录音录像档案数字化工作的安全问题不容忽视,主要包括档案实体和档案信息安全。在标准的制定过程中,特别注重对档案安全的保障,从管理和技术等不同方面进行规范。

主要内容

标准分13章对录音录像档案数字化工作的全过程提出了普遍适用的规范性要求。

1.适用范围与总体要求

标准的适用对象是由模拟信号形成的录音录像档案,包括录音带(钢丝录音

带、开盘式录音带、盒式录音带)、唱片,以及录像带(U-matic、Betamax、VHS、Betacam、8mm等类型)等。由于标准中包含了管理类要求,因此,范围中明确注明:“本标准规定了模拟录音档案和录像档案数字化的技术和管理要求”。

标准在“总则”中提出了数字化工作应遵循的顶层的、原则性条款,主要包括相关工作的统筹规划、数字化工作基本流程、档案安全管理等内容。数字化工作流程的科学和规范,是确保数字化成果质量和效率的重要条件。课题组在调研的基础上,对各地开展录音录像档案数字化工作的流程进行了总结归纳,择其精华,确定了较为科学合理、具有一定普适性的工作流程(如图1)。标准的结构即按照工作流程进行组织。

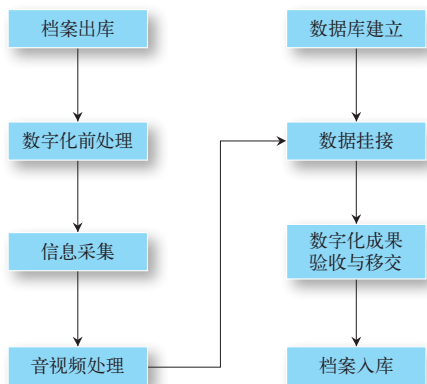


图1 录音录像档案数字化流程示例

2. 组织与管理

为强化数字化工作的管理要求,标准专门设立“组织与管理”章节,从机构、人员、基础设施、工作方案、管理制度、工作流程控制、工作文件管理及档案数字化外包等方面,系统化地提出了管理要求。例如,目前很多档案馆采用外包方式进行数字化工作,但如何确认承包商的资质、从哪些方面对承包商进行考察等问题,是需要在标准中给予明确的。针对这些问题,课题组对数字化外包资质、相关标准和规范、目前较为正规和完善的数字化加工企业情况等进行了详细调研。据此,在标准中,对档案行业数字化外包时对外包企业资质的认定及档案部门在数字化过程中的参与和管理等提出了明确要求。

3. 档案出(入)库

档案出(入)库管理是确保录音录像档案数字化过程中档案安全的关键一步。标准将档案出库(第6章)、档案入库(第13章)分别独立成章进行描述,提出档案出(入)库过程中的管理要求。

4. 数字化前处理

数字化前处理是保障成果质量和利用效果的基础性工作,标准主要从确定信息采集范围、档案检查项目等方面提出要求。

5. 数据库建立

标准提出,在数字化工作开展前期应建立数据库,并规定了建立数据库的相关要求。课题组在调研中发现,很多档案馆在进行数字化工作前就已经建立了数据库,那么数字化工作过程中只需对其进行规范化处理即可。数据库的建立相对较为独立,在数字化工作前期开展会更加灵活。

6. 信息采集

信息采集是标准中关于对技术参数要求的重要部分,也是标准制定的重点和难点之一。其中,“基本要求”这一部分对录音录像档案的信息采集提出了原则性要求。同时,以目前的技术现状和实际调研情况为基础,对信息采集设备、技术参数、文件切分与著录、文件命名、质量检查、档案恢复等提出了具体要求。

关于数字化技术参数的建议是本章的关键内容,针对各项技术参数的选择,课题组进行了充分调研,并选取了有代表性的录音录像档案类型,重点针对典型、共性技术问题开展实验,主要包括:10余种类型的录音录像档案数字化实验、编码格式对比实验、文件存储格式对比实验、音视频处理实验等,并对实验结果进行分析与总结,结合多年的工作和科研经验,对数字化参数的选择进行了研究和确定。

(1) 录音档案数字化主要技术参数

① 文件格式

标准提出,音频文件格式应采用WAVE格式。WAVE格式是最早的数字音频格式,支持多种音频位数、采样频率和声道,支持无损压缩,是公认的数字音频文件通用的长期保存格式。

② 采样率

采样率是每秒从连续声音信号中提取离散信号的数量,是数字音频声音质量的关键因素。标准建议采样率应不低于44.1k Hz,珍贵或有特别用途的录音档案采样率不低于96k Hz。人耳对采样率达到44.1k Hz的声音差异已不再敏锐,这个一般经验在课题组所做的实验中也得到了验证。日常生活中,CD音质为44.1k Hz取样的16bit声音,作为需要长期保存的档案,其采样率不应低于CD。从实验中的信息采集情况看,对于人声采集可以采用44.1k Hz或更高的采样率;如果需要辅助信息,如人耳难以识别的噪声进行保存,建议采用48k Hz或更高的采样率;对于高保存价值的档案或者是含有音乐的档案,可以采用96k Hz或更高的采样率,能够准确捕捉噪声以及人耳难以感知的声音,如听不见但影响音质效果的和声、可在未来被新算法捕捉的辅助信息等。

③ 量化位数

量化位数是描述数字音频波形的二进制数据为多少位的数据,也是数字音频声音质量的重要指标。标准建议量化位数应为24bit。课题组所做的实验结论为:量化位数为8bit的音频文件无论采样率多高,其音质普遍较差,能听到较明显和较强的噪声;用于保存档案时,量化位数为16bit的音频文件无法完整地保存声音细节,量化位数为24bit或以上的音频文件则可以满足需求。

(2) 录像档案数字化主要技术参数

① 文件格式

在计算机文件中,视频文件是比较复杂的一类,并且为适应存储和利用视频文件的需要,人们设定了不同的视频文件格式,把视频、音频放在一个文件中以方便同时播放。标准规定视频文件格式应采用AVI或MXF。

AVI(Audio Video Interleave,音频视频交叉存取)是在20世纪90年代初出现的文件格式,是常见的视频文件格式。AVI将视频和音频封装在一个文件里,音频和视频按帧交错排列,以此达到音频同步于视频的播放效果。AVI本身只是提供了一

专题·提质增效 标准先行

个框架,内部的图像数据和声音数据可以采用多种编码格式。

MXF (Material Exchange Format, 文件交换格式) 是SMPTE (Society of Motion Picture and Television Engineer, 美国电影与电视工程师协会) 定义的一种专业音视频媒体文件格式。MXF的作用是为数据的发送者和接收者建立不同数据格式转换的通用标准。MXF广泛应用于影视行业媒体制作、编辑、发行和存储等环节。

其他的视频文件格式由于存在依赖特定编码格式、使用范围较小等问题,标准中未建议采用。

② 视频编码格式

标准建议视频编码格式采用H.264、MPEG-2 IBP。

H.264是由ITU-T和ISO/IEC(国际标准化组织/国际电工委员会)联合发布的标准。ITU(国际电信联盟)是联合国在电信领域的一个专门机构,ITU-T是ITU的一个常设机构,进行国际电信标准的制定。H.264标准是被广泛使用的高精度视频录制、压缩和发布的视频压缩标准。

MPEG-2是ISO/IEC成立的MPEG (Moving Pictures Experts Group, 运动图像专家组) 发布的标准,适用于高质量的图像和声音的压缩编码,也是数字电视广播数字视频存储和传送的基础标准,在图像、声音的存储和传输相关领域得到

广泛应用。为保证压缩效率和质量,标准建议采用IBP的压缩方式。

其他视频编码格式,由于存在编码过程信息损失较多、成熟度不高等问题,标准中未建议采用。如H.265标准是在H.264标准的基础上发展起来的,已经在很多方面实现了对H.264的超越,但是由于H.265是最近几年才发布的,目前可以提供技术支持的软硬件很少,没有通过大规模的应用并验证其可用性,所以,课题组认为H.265的成熟度不高,不具备在档案部门普遍使用的条件。

③ 帧率、画面宽高比

为了真实反映档案原件的原始内容,标准规定不应对视频的帧率、画面宽高比进行转换,应与档案原件相同。

7. 音视频处理

标准规定了音视频文件在提供利用前,可针对原始音视频文件(即通过档案原件数字化后未处理过的音视频文件)的拷贝文件进行文件转换、降噪、校色及画面稳定等处理。课题组认为,用于长期保存的音视频文件应是原始音视频文件,而非处理后的文件。主要考虑是,随着技术进步,降噪、校色等处理技术会更加完善,若原始文件仍然留存,便可在未来获得质量相对更佳的处理后文件。

8. 数据挂接

通过数据挂接,实现目录数据与音视

频文件的关联,并通过档案号或者原始介质索引号等形式建立音视频文件与档案原件的关联。

9. 成果验收与移交

标准从验收方式、验收内容、验收指标、验收结论等方面对录音录像档案数字化成果的验收环节提出了要求。档案数字化工作任务量巨大,仅靠人工方式不能全面验收,验收质量也难以保证。应采用计算机自动检验与人工检验相结合的方式对录音录像档案数字化成果进行验收检验。课题组对数字化过程形成的成果进行了详细总结,提出了目录数据、音视频文件、数据挂接、工作文件、存储载体等几方面的验收内容。验收方式的选择对数字化成果质量的保障具有关键性作用。随着信息技术的发展,计算机自动检测的方法更为多样,范围更为广泛。计算机自动检测不仅使数字化成果100%全面检测成为可能,同时可很好地确保检测的准确性。为此,标准针对档案数字化成果验收指标进行科学划分,对于能够通过计算机程序自动化方式验收的指标,提出100%验收比率和100%合格率的要求。对于无法采用自动化验收的指标,也根据具体情况提出了不同的合格率要求。■

作者单位:国家档案局科研所

责任编辑:黄佳音

动态

北京市进一步提高档案工作依法行政水平

近日,北京市档案局制发了《关于进一步提高档案工作依法行政水平的意见》。提出了未来5年北京市档案法治建设工作的总体目标:一是档案地方法规制度建设的质量进一步提高;二是档案依法履职能力进一步增强,档案执法力度不断加大;三是档案普法宣传领域不断拓展,档案法治意识更加深入人心;四是档案执法人员依法行政水平全面提升,素质能力能够适应新形势的要求,达到与首都地位相适应的一流档案法治水平。

《意见》从着力提高档案法规制度建设的质量、严格规范档案执法、大力提升全民档案法治意识、加强档案执法队伍建设4个方面明确了11项具体措施。加强城乡规划建设管理、京津冀协同发展、2022年北京冬奥会冬残奥会等方面档案制度建设;依法加强对社会档案事务的行政监管,对法律法规规章明确规定的执法检查事项实施随机抽查;引导公民通过档案行政复议、档案行政诉讼、举报档案违法案件等多种手段维权,妥善化解档案行政争议;档案执法工作人员必须取得“北京市行政执法证”等措施被列入其中。

北京市档案局 马秋影