

二维条码技术在档案管理中的应用研究

石 华

摘 要: 二维条码是一种高密度、高信息含量、价廉可靠的计算机可以直接识别的便携式数据文件。本文提出将电子文件的内容及元数据存储到二维条码中,附着在纸质文件上,形成纸质电子文件,这样在电子文件从产生到永久保存的整个生命周期中的各个阶段都可以通过解码设备扫描条码,读取条码中的信息并将信息保存到本地。可有效解决归档、移交环节中信息的重复录入,避免多次录入中的信息差错,实现电子档案的长期保存、信息共享,提高档案管理水平。本文的研究结果证明将二维条码技术应用到档案管理当中是可行的,必将对档案管理的现代化、标准化以及提高效率产生积极的影响。

关键词: 二维条码; 电子档案; 档案管理

1 背景

1.1 二维条码相关概念

二维条码:区别于传统的一维条码,二维条码是在二维方向上表示信息的条码符号,因此其存储容量比传统的一维条码有了飞跃性的提高,数千个字符能够被存储到一个邮戳大小的条码符号中,是一种简洁而廉价的信息存储方式。

1.2 二维条码技术在档案管理中的实现模型

在档案管理中引入二维条码,实质是将纸质文件、元数据、电子文件封装在一起。步骤如下:

一、将电子文件及其元数据信息存储在二维条码中;

二、将二维条码打印在电子文件对应的纸质文件上,形成电子文件、元数据、纸质文件封装在一起的纸质电子档案;

三、通过扫描设备扫描纸质档案上的二维条码;

四、对扫描得到的二维条码进行解码,还原出对应的电子文件及其元数据。

由于档案管理涉及的单位、软件系统众多,在生成二维条码之前,电子文件应转换为可长期存取的统一格式,而元数据也应遵循统一的标准,以利于在档案管理各阶段各用户的使用。

由于二维条码一经生成,不可更改,而元数据在文

件流转过程中不断增加、变更,所以纸质电子文件在流转到新部门时,使用者(包括档案室工作人员、档案馆工作人员)应根据元数据的不同,将变更、增加的元数据转换为二维条码,附加打印在纸质文件上。

1.3 档案管理中引入二维条码技术的意义

二维条码作为信息的另一种表示形式,是不依赖计算机网络和数据库存储、转移信息的一种方法,是实现文件、档案等信息的存储、携带、传递并识读的理想手段。将之引入档案管理中,可较好地解决当前档案管理中出现的问题。

1.3.1 采用二维条码技术可解决电子档案易更改的问题

将电子档案转化为二维条码,再打印到纸张上,形成纸质电子档案,则可封存电子文件形成的原始状态信息,二维条码一经生成,不可改动,可解决电子档案易更改的问题,是电子文件真实性保障的重要工具。

二维条码因为是直接对原文二进制或ASCII码进行编/解码,所以支持任何语言和文字。而且在电子化的过程中不会丢失任何信息[1]。

1.3.2 采用二维条码使电子档案可长期保存

采用二维条码技术,将电子档案固化在纸质介质上,相当于电子档案的载体与纸质档案相同,可长期保存。纸质电子档案携带方便,不怕折叠,保存时间长,又可复印传真,做更多备份[2]。二维条码有错误修复能

力,即使载体保存不善,印刷的文字、条码已有污损,只要破损面积不超过整个条码符号的50%也可照常恢复数据。

1.3.3 采用二维条码技术可实现纸质档案与电子档案的同步管理

将二维条码引入档案管理,形成纸质电子档案的技术,方便档案人员对纸质档案和电子档案的统一管理。档案部门对纸质档案的管理历史悠久,有许多行之有效的方法,电子档案丢失的可能性就小多了。二维条码相当于在电子档案管理系统与纸质档案管理系统之间建立通道,使人们可以容易地从一种档案找到相同内容的另一种档案,方便利用。

1.3.4 采用二维条码系统可提升档案信息录入的质量与效率

在办公系统中加入二维条码生成控件,将文件的著录信息生成标准的二维条码,在档案部门只需要扫描设备就可将电子文件信息识别出来,不需铺设网络、不需考虑软件对接,对公文系统影响最小,容易实现,而且二维条码一次生成,可多次使用,可以保证著录信息的准确,减少重复劳动。

1.4 本文选择PDF417码

二维条码有许多不同的编码方法,本文决定采用PDF417码作为档案管理二维条码技术应用软件的码制。主要原因是在公文系统中公文系统已使用PDF417码[3]。而许多机关也已购置了二维条码扫描设备。为节约硬件、软件投入,与公文系统保持一致,档案部门也采用PDF417码是合理的选择。

1.5 二维条码只适合以件为整理单位的文书档案

PDF417码存储空间虽然比一维条码大得多,但也有限,只有1K左右,而且存储的信息越多,尺寸越大,不利于识读器识别。所以二维条码只适合以件为整理单位的文书档案,其信息较少,有可能存储在二维条码中。

1.5.1 以页为单位存储

电子文件以页为单位,将每一页的内容存储在二维条码中,附着在每一页纸质文件上。这样不但可以降低文件大小,而且有利于使用者只选择需要的页进行扫描利用。

1.5.2 可采用宏PDF417码

宏PDF417码是一种变形的PDF417码制形式,当文件内容太长,无法用一个PDF417码表示时,可用包含多个(1~99999个)条形码分块的宏PDF417码来表示。

2 研究内容

本文重点研究三个问题:根据二维条码特点,选择电子档案应转换为何种格式,制定《电子文件元数据二维条码使用标准》,考虑文件形成部门、档案馆、档案利用者的实际需求,设计二维条码档案管理软件的功能。

2.1 电子档案长期存取格式选择

二维条码一经产生,就不可更改,所以在生成前,先将电子文件转换为可长期存取的文件格式。本文选择PDF文件格式。

2.2 制定《电子文件元数据二维条码使用标准》

采用二维条码作为电子档案存储与共享的载体,就必须对电子档案数据进行标准化。档案管理涉及的系统众多,二维条码起着在各系统间进行有效的数据交换的作用。只有实现数据的标准化,才能真正实现信息共享和业务协同,电子档案数据标准化已成为档案管理中应用二维条码的关键环节。在制定元数据标准时,应与已有的相应的国家标准保持一致。本标准直接使用《电子文件管理细则 第一部分:文书电子文件元数据方案》,兼容文书部门所使用的《机关公文二维条码使用标准》。

制定标准时,需要在通用与易用间取得平衡。二维条码容量有其限制,所以存储的元数据应尽可能简练。从目前已经制定的国家标准中选取通用性、基础性和共性的数据元素,作为本标准数据元素的主要内容。

本文设定电子档案已经转换为PDF格式了,可将电子签名、文档创建环境等无用数据元素去除。

将本标准的版本号暂定为“GB0626-2008”,凡按此标准生成的二维条码,存储的数据的前十一位必须是“GB0626-2008”。这样在读取数据时,就可按此规范转换数据了。

本标准采取动态存储方式,以“^”作为数据元素的分隔符,以“\”作为条码内容的结束符。这与《机关公文二维条码使用标准》一致。

本标准在数据框架结构上借鉴了国家标准《GB/T20163-2006中国档案机读目录格式》,把一条记录划分为四个区:记录头标、记录目次区、数据字段区、记录分隔符。

数据样例如下:

GB0626-20080010004000000050004000064006000400009^文件^2008^永久\

其中,前11个字节“GB0626-2008”是记录头标,表明遵循《电子文件元数据二维条码使用标准》。从第12个字节到第一个“^”之间是记录目次区,每12个字节是一个目次项。

001 0004 00000, 005 0004 00006, 006 0004 00011

001 0004 00000表示是编号为001的数据元素,数据长度为4,起始地址从“0”开始,起始地址就是第一个“^”后。005 0004 00006表示是编号为005的数据元素,数据长度为4,起始地址从“6”开始;006 0004 00011表示是编号为001的数据元素,数据长度为4,起始地址从“11”开始。

样例所表示的元数据如下表:

数据元素编号	数据元素名称	数据	数据长度
001	集合层次	文件	4
005	年度	2008	4
006	保管期限	永久	4

采用这样的结构,变固定的字段长度为变长,变固定的字段数量为可变的数量,变固定的记录长度为不定长的记录长度,实际著录是多长字节就是多长字节,而不会造成存储空间的浪费,这对于存储空间有限的二维条码尤其有利。而采用关系型数据库格式、XML格式、HTML格式、文本格式等时,都有多余字符。

2.3 建立基于二维条码的档案管理系统

研究二维条码的应用技术也就是帮助用户解决如何建立基于二维条码的应用系统、怎样将二维条码技术整合进现有应用系统之中等问题。目前二维条码生成识读技术已经相当成熟,在文书、档案管理系统中实施二维条码应用关键点在于如何与原有的系统结合。

2.3.1 二维条码测试平台

二维条码测试平台的功能是:测试软件开发商开发的各二维条码档案管理软件及硬件设施是否符合国家标准。只有通过档案部门网站上的测试,各档案管理软件、硬件设施才可以投入市场。该平台的使用者是档案部门,以保证软件符合国家标准,数据统一规范。

2.3.2 文档一体化系统相关环节功能设计

2.3.2.1 生成正式文件

使用二维条码的生成技术,将文件转换为PDF格式并与著录信息一起转换为二维条码,再加入到电子文件的相应位置,最终生成正式的带有二维条码的纸质文件。为保证著录信息的准确完整、文件信息的长期可读,应将以上流程设计到办公自动化系统中,由系统给形成者提要求,不完成此工作,无法进入下一个流程。

2.3.2.2 收文登记

利用二维条码的识读技术,自动采集电子文件及其著录信息,减少重复劳动,提高效率与质量,是原

收文登记的“改良”,在管理上实现电子文件将与纸质文件的合拍与一致。如果收到的纸质文件上的二维条码是遵循《机关公文二维条码使用标准》,也应该能识读。但《机关公文二维条码使用标准》数据项不全,应按照《标准》补充著录。

2.3.2.3 归档

档案室在将文件整理、装盒之后,确定了盒号、件号、类别之后,按《标准》补充著录相关元数据信息,将档案室所著录的元数据转换成二维条码,打印在标签上,粘在文件上。如归档的文件是从外单位来的,无二维条码,无电子文件,应进行数字化,将数字化的信息及元数据转换为二维条码,粘贴于文件上。

2.3.2.4 移交入馆

档案馆在接收档案时,如果是有二维条码的纸质文件,对二维条码扫描,还原为电子文件进入原文数据库,还原为元数据进入目录数据库。将档案进行整理(查重、重新定盒号),上架,在档案入库房确定上架位置后,再按《标准》补充著录元数据,打印二维条码,以保证二维条码的内容足够完整,方便以后利用管理。

3 总结

由于二维条码具有成本低、信息可随载体移动、不依赖于数据库和计算机网络等优点,非常适合我国档案管理部门数据来源复杂、底子薄、计算机网络投资难度较大的特点,可以预见它将在文档一体化领域有着极为广阔的应用前景和巨大的市场空间。

参考文献:

- [1]张伟.文字图表等一般载体上的信息电子化的方法及其系统[P].中国发明专利, CN1396538.2003-02-12
- [2]二维条码的应用[DB/OL].[2005-28-27].<http://www.enpot.com.cn/zhishi/barcodeC01.htm>.
- [3]政府行业文印中心“二维条码制作”解决方案政府工作报告政府工作报告 [DB/OL].[2008-02-09].http://www.founder.com.cn/Solution/2008-01/19/content_5483.htm.

(作者单位:郑州市档案局 来稿日期:2009-10-20)