

# 数字图书馆金石拓片资源库建设

龙伟

(国家图书馆信息网络部)

被誉为我国“四大发明”之一的印刷术，对人类文化发展起了不可估量的作用。据记载，从石板上捶拓文字是印刷术的先驱，因此早于印刷术出现的传拓技术也是我国的重要发明之一。金石拓片是一种比较特殊的中国传统文献，它涵盖了丰富的历史文化，是保存文献、考证史传、增补遗闻的重要资料。拓片以 1:1 的比例，忠实反映了原物的面貌，在不能接触原物的情况下，拓片在很大程度上起到替代原物的作用。拓片的史料等多方面价值已为社会和学界所认识。

国家图书馆藏有各类石刻拓片近 3 万种，13 万余件，品种多，质量精，是馆藏中的精品。经过我馆金石组同志的努力，拓片已编制目录的有 2 万条，已制作书目数据库的有 9700 余条。

金石拓片这一反映我国历史文化的文献是国家图书馆数字图书馆中的一个重要的资源库。为更加便于学者的研究，为使金石拓片在互联网上展示原貌，我们制定方案，展开了数字拓片的工作。按照数字图书馆的技术进行资源组织、加工、存储、检索、传递、保护，基于互联网为读者提供服务。

金石拓片资源库按元数据和对象数据制作。在资源库建设初期，我们依循数字图书馆建设原则，确定了拓片建库原则，在总原则的指导下，资源库建设有章可循，很快将此项工作开展起来。

数字图书馆的特点是书目记录与图像、文本的紧密连接，而图像与书目记录应该是一一对应的关系。在元数据制作方面，我们将元数据的采集视为与每个数字影像、数字对象相关联并成为藏品开发过程的一部分。按照数字图书馆技术特点，以著录的书目数据为基础，抽取拓片元数据，在 Dublin Core (简称 DC) 的基础上，我们编制了拓片元数据集，实质上是一个 DC 的扩展集，包括：描述性元数据，结构性元数据和管理性元数据。为了更为全面、准确地完成元数据加工，金石组编目人员对原有的《金石拓片编目规则》进行修订并更名为《中文拓片编目规则》，使其更具合理性、科学性和可操作性，并在一年多的拓片资源库制作的基础上制定《中文拓片机读目录格式使用手册》，现已发行。中文拓片书目数据建立在 CNMARC 基础之上。传统的 MARC 格式，其实也是一种描述性元数据。书目数据记录了拓片的详细内容，如题名、责任者、刻立年代、刻立地点、书体与行款、数量与尺寸、题跋印记、关键词、索书号、石刻类名等著录项目。按照数字图书馆的体系结构，为此，我们开发了“拓片资源整合系统”，解决了 CNMARC 格式向 DC 格式的转换。依据《金石拓片元数据标准》，将

CNMARC 各字段、子字段与 DC 元素建立参照、对照表。运用整合系统，将 CNMARC 形式的记录转换为 DC 元素后，系统记录 DC 元素的属性以及 CNMARC 与 DC 的核心元素和其它元素的映射关系。整合系统功能齐全、方便灵活，提供数据格式的批量转换和单个文件上传，用户可根据工作所需，自由选择。此外，系统具有严格的安全检查，对用户和权限分别控制、管理。在元数据加工过程中，在拓片著录规则相对统一的情况下，整合系统完成的拓片数据制作支持 XML 语言，采用 Unicode 字符集，实现异构操作系统的共享，在互联网上提供检索、查询、浏览使用。

数字影像的定义为，基于光栅的、具有两维效果以及静态数据象素的矩阵，可用于计算机使用文件格式的转换，并能在计算机视频显示器上显示。金石拓片对象数据的制作，我们采取对原拓和拓片出版物影像的捕捉，直接获取数字文件，数字文件如实表现原始对象的影像信息。拓片影像数字文件分为数字主文件和衍生件。数字主文件是通过专用扫描仪对原文献影像捕获的直接结果，其主要作用是作为长期的档案记录和用于产生衍生文件，它可以作为原始对象的代用品。美国加州数字图书馆（CDL）影像藏品标准指出，创建的影像主文件至少要有 50 年的使用寿命。所以，我们考虑到文献影像的保存和未来使用的多样性，使用高精度、高质量捕捉拓片原物。数字衍生件是从数字主文件衍生而来的，用于编辑或增强主文件到不同格式的转换以及网络表现和传输。拓片衍生件不是用于保存的文件，因为数字主文件才是用作保存目的的。拓片衍生件包括缩略图、预览图、中分辨率图和高清晰度图，图像的计算机显示充满计算机显示屏（1024 × 768 象素，标准 PC VGA 分辨率）。衍生件的影像及其尺寸取决于多种因素，包括文献纸张质地、出版物印刷状况、使用用途、计算机网络传输等等。衍生件制作使用特定软件完成，在指定参数下，对主文件进行批量转换和压缩。在衍生件压缩过程中，使用有损压缩进行处理，压缩比选择介于 1-10 级之间。衍生件的数字水印制作目的是为了保护金石拓片的网络使用，所以我们为每张拓片添加显示数字水印，数字水印的添加既不妨碍网上图像的浏览和图像细微部分的观赏，又在一定范围内保护了国家图书馆的文献版权。

金石拓片资源库的建设是国家图书馆数字图书馆建设的重要项目。目前，我们已完成拓片元数据制作 6000 余种，影像数据 8000 余幅，资源库已有一定的规模，运用我们的互联网发布系统，读者在任何时间、地点都可以查阅、获取拓片资源。建设金石拓片资源库的意义是显而易见的，既保护了古文献，也增加了文献的多种表现形式。资源库的建设还在进行当中，我们希望能有更多的文化单位加入我们的行列，丰富拓片资源库的内容，使更多的读者可以分享中国传统文化。