



International Federation of  
Library Associations and Institutions

# 馆藏善本与手稿数字化计 划指南

国际图联善本与手稿组撰稿

杭州图书馆组织编译

中文版

翻译：吴忠超（主译）、彭喜双、沈 静

日期：2015年2月

注：措辞以英文原件为准。

国际图联

馆藏善本与手稿数字化计划指南

---

## 馆藏善本与手稿数字化计划指南

国际图联善本与手稿组撰稿

2014年9月

国际图联专业委员会批准



国际图书馆协会和机构联合会，2014年

©2014 年国际图书馆协会和机构联合会。本文件采用知识共享“署名 3.0（未本地化版本）”许可协议授权。要查看此许可证的副本，请访问：<http://creativecommons.org/licenses/by/3.0>

国际图联

邮箱号 95312

2509 CH 海牙

荷兰

<http://www.ifla.org/>

## 目录

引言和适用范围.....	5
感谢 .....	6
1 设计项目.....	8
2 选择原稿.....	9
3 创建数字馆藏的工作流程.....	10
4 元数据.....	13
5 显示.....	14
6 传播、推广和再利用.....	15
7 评估.....	16
8 数字馆藏的长期保存.....	16
9 建议摘要.....	17
参考文献 .....	18

# 引言和适用范围

## 引言

图书馆馆藏数字化工程正在逐渐改变人们发现信息并进行研究的方式。图书馆有责任将他们的数字化馆藏提供给全球，这是公众的要求，也是学者的期待。互联网的广泛应用已经消除了地理距离、经济环境、政治边界和文化敏感性等传统障碍。研究人员正在开发新的查询领域，将不同的来源和学科中的研究资源排在一起的新技术和处理数字文本和图像资源的新方式，也对该项研究有促进和激励作用。

在可发现性和使用方面，善本和特殊馆藏的数字化，比一般馆藏更有意义。一旦这些馆藏易于获取，它们就成为核心资源。如果不进行数字化，善本和特殊馆藏将不为世人所知，其真面目将一直模糊不清。

在参与“大规模数字化”项目时，虽然许多图书馆已经制定了相关操作程序，但对孤本、善本和非印刷形式的馆藏进行数字化时，需要特别加以考虑，并制定不同的操作程序。这些指南的目的之一是要填补这一空白。

这些指南适用于所有参与规划善本和特藏数字化项目的人，包括提议项目的图书馆领导，规划和执行项目的图书馆员和研究人员，以及正在考虑对特藏数字化进行支持的基金组织。

## 适用范围

数字化已经深刻地改变了我们看待图书馆馆藏、服务和战略规划的方式。其结果是，在过去 15 年中，在国际、国家、地方和机构各个层面，数字化指南以多种形式激增，而所有这些指南都反映了当下不断发展的最佳实践的情况。其中一些指南试图涉及方方面面<sup>1</sup>，另一些则范围较小，适用于其他某些特殊文献资源<sup>2</sup>。还有一些专注于特定媒体的数字采集标准<sup>3</sup>，而另一些则专注于更强大的数字保存程序的需求<sup>4</sup>。每个相继的文件往往是对以往指南的改进，而不是完全取而代之。为了给善本与特藏数字化项目规划的特定需求提供解决方案，准备这些指南的工作组研究了这些之前的文本，尝试撰写一份补充性文件。

这些指南并不是专注于技术问题、数字采集的具体方法或数字保存方法，而是着重于概念性规划和与潜在用户的合作，以达到预期效果和取得可持续的成果。他们是从特藏管理人、图书馆馆长和研究者的角度而撰写的。他们把这些物理对象当作既承载有知识内容又蕴涵有内在历史证据的手工艺品来研究。他们还试图预示那些基于数字化开展研究的用户的需求，即要求全文分析或“大数据”的大聚合。它们也将帮助专业人士和非专业人士创建可持续发展的且适应性强的数字馆藏，这些馆藏可以重复使用，并且在未来会继续发挥其研究价值。

要强调的是数字化的馆藏，而不是个别对象的创造和发现。善本和孤本馆藏的数字化往往是复杂的，因为既要维持单独实体对象的各个部分又要保持整体馆藏各个部分在实体、知识、和上下文中的原有联系是很重要的。对用户而言，要维持这些联系常常会涉及特殊的元数据，以及不同的发现和使用平台。此外，用户必须具有足够的上下文信息和元数据，以便能够评价这些数字化替代品的真实性

和可靠性。在保持人工复制品的真实性且确保沟通信息具备的同时，要在数字背景下呈现整个馆藏，还需要仔细思考和规划，超前准备，有效实施大量相关活动，以及和很多不同图书馆之间的合作。

为了特殊用途而对部分文献进行数字化，为文献保护分析或技术驱动型研究的需求而采用先进的光谱成像处理，不在本指南考虑范围之内。它们也无意指导旨在长期保存的成像处理，这种技术需要关于成像技术、文件格式、图像分辨率和专用设备的更多细节的信息。

### 价值、一般原则与预期效果

善本与手稿组的成员在商讨这些指南时，得出了一些共同的专业价值观，并希望能够取得一些预期效果，撰写出这个最后文本。这里把它们表述为指导决策的一般原则：

- 继续坚定不移地捍卫图书馆的责任，保护并提供原始馆藏文献。
- 向世界各地提供免费的研究资源，并在可能的情况下，让用户享有下载数字化对象和馆藏的能力。
- 关注学者的需求；鼓励学者与用户间的对话；邀请他们参与规划过程。
- 为可实现的效果而努力，开放存取，改善馆藏保存条件，只要有可能，都要尽力提高馆藏价值。
- 在不断发展的最佳实践案例和成功项目基础上，建立质量控制体系，贯串于项目最初规划直至最后评估的各个阶段。
- 数字化复制件的结构和上下文关系要和原始馆藏文献保持一致，并要保持数字化复制件回到目录记录的链接。
- 通过添加越来越多的可存取数字馆藏的方式，“按需应变”管理需求。
- 借助数字化项目提供那些过去一直隐藏着的馆藏的存取。

## 感谢

国际图联善本与手稿组衷心感谢制定指南的工作组同仁，他们为制定指南奋斗了三年；感谢那些帮助修改和完善之前几个版本草案的其他人士。

2010年8月，在哥德堡会议上，制定这些指南被确定为优先战略任务。2011年4月，工作组主席，伊莎贝尔·加西亚-蒙赫（西班牙），在马德里举办了为期两天的中期会议，此次会议，工作组形成了第一稿。2011年8月，在波多黎各又再次召开了为期一天的工作会议，并于2012年1月形成第二稿。2012年2月，在安特卫普的中期工作组会议上，第三稿出台。2013年8月，在新加坡会议上，工作组用两小时的公开会议时间介绍了指南草案，并征求听众意见。超过两百人出席了此次会议，并参加了45分钟活跃的问答讨论，由此专家组获取了许多宝贵的反馈意见。这项工作本应从来自其他地理区域的非工作组成员以及没有回应审查呼吁的人员中得到中肯意见。为了给数字化项目规划提供均衡的视角，工作组认为广泛听取全球各个地区的意见是很有必要的。我们积极鼓励来自世界其他地区的各种建议，我们将把它们纳入到以后的版本中。

工作组成员如下所列。（善本手稿专业委员会常委用星号（\*）表示。）

- 伊莎贝尔·加西亚-蒙赫 (Isabel Garcia-Monge), 主席, 西班牙\*
- 伊凡·博塞拉普 (Ivan Boserup), 丹麦
- 安妮·艾斯费尔特 (Anne Eidsfeldt), 挪威\*
- 皮拉·伊格斯科扎巴·卡拉斯科 (Pilar Egoscóbal Carrasco), 西班牙\*
- 克劳蒂亚·法比安 (Claudia Fabian), 德国\*
- 大卫·法恩斯 (David Farneth), 美国\*
- 瑟卡·哈武 (Sirikka Havu), 芬兰\*
- 沃尔夫冈-瓦伦丁·阿德 (Wolfgang-Valentin Ika), 德国
- 拉斐莱·穆朗 (Raphaële Mouren), 法国\*
- 安吉拉·诺沃 (Angela Nuovo), 意大利\*
- 克里斯特·奥斯特伦德 (Kristen Östlund), 瑞典\*
- 埃德文·C·施罗德 (Edwin C. Schroeder), 美国\*
- 加勒特·弗尔霍文 (Garrelt Verhoeven), 荷兰\*
- 玛丽娜·威尼亚 (Marina Venier), 意大利\*

其他参与、审阅或对草案提出建议的人员有:

- 简·博斯 (Jan Bos), 荷兰\*
- 马克·迪姆纳申 (Mark Dimunation), 美国\*
- 乔安娜·埃斯科贝多 (Joanna Escobedo), 西班牙
- 克里斯蒂娜·吉尔伦·贝尔梅霍 (Cristina Guillén Bermejo), 西班牙
- 皮拉·莫雷诺·加西亚 (Pilar Moreno García), 西班牙\*
- 亚历山大·萨马兰 (Alexander Samarin), 俄罗斯\*
- 卡罗琳·施莱伯 (Carolin Schreiber), 德国
- 温斯顿·塔布 (Winston Tabb), 美国\*
- 马尔塔·托雷斯·桑托·多明哥 (Marta Torres Santo Domingo), 西班牙
- 奥尔加·维嘉 (Olga Vega), 古巴

# 1 设计项目

首先仔细考虑一些至关重要的问题将有助于项目的成功设计和顺利实施。数字化项目可能产生不同的产品和服务。需要考虑可利用的资源。在项目开始时即明确目标和效果将有助于宣传和筹款，便于在项目结束时提供有意义的评估基准。因此，在项目伊始，厘清一些基本问题显得十分重要。

## 什么是项目的远景？什么是目标和任务？谁将会使用它？他们将如何使用它？

数字化项目的目的是为一般性研究服务还是为满足特定的研究需要？数字化是为了方便用户获取还是为了保存原始馆藏的需要，或者两者兼顾？如何将数字化项目融入到机构的整体目标之中？

图书馆员重视建立长期、全面的馆藏。学者和用户希望传递符合他们当前需求的文献。管理者则寻求能够增加本机构的知名度和声望的项目。最初的规划应该兼顾所有相关方面不同人士的需求。

## 谁应参与规划？

对一个有效的项目而言，团队合作至关重要。大多数成功的项目在规划过程中都将学者、用户、管理者、图书馆工作人员（编目人员、馆长、管理员）和技术员纳于其中。工作流程会涉及到图书馆的许多部门，在做决定时充分考虑大家的意见非常重要，因为这将影响他们正在进行的工作。

## 有没有外部融资的机会？

许多数字化项目是通过拨款资助，或是通过和其他机构合作的方式来实现的。外部资金提供方可能会有自己的准则来选择要对何种馆藏进行数字化。应确保对所有的合作伙伴一视同仁。

## 想达到何种水平？可达到何种水平？

一个雄心勃勃的项目，通常需要分阶段逐步完成。如果是这样，首先要集中建立优秀的、强大的，最重要的是，可用的资源。然后随着广大用户的参与，不断增强。

## 你想数字化何种馆藏，为什么？

选择哪些馆藏来进行数字化往往是整个项目中最困难的一个环节，尤其是当需求很高而资源又有限的时候（详见第二章）。

## 会不会存在文献的版权问题？

如果有，你就需要考虑这些问题（详见第二章）。

数字化过程应该由典藏机构本身还是外部服务供应商来完成？你是否具备场地、资金、设备和相关专业人士？外部供应商能做什么？



这些问题的答案取决于典藏机构的能力和该项目的性质。由典藏机构本身来进行数字化的优势有：掌握专业知识和建成相应的基础设施，以便于后续项目的顺利完成；减少破坏原始文献的风险（价值和保险问题都要考虑）；更好地控制图像质量和数据传送。在另一方面，外部服务供应商可以提供专业知识和更好的设备，降低成本，和（或）更快更好地完成项目，使得典藏机构的工作人员可以去完成其他任务。请记住，外包总是需要典藏机构的工作人员花费相当的时间来管理这个项目。

### **该项目的最终格式是什么？你能实现它吗？**

考虑该项目的最终格式是否能实现总体目标是非常重要的。要确定你是否拥有成功完成该项目所必备的设备、软件、设施、专业知识和资金。

### **可有预想到的社交网络，比如大量人群的转录或元数据增强？**

数字化项目具有广泛触及大众人群的潜力，在设计项目时应该考虑这些受众。想想在将来学者和其他用户的贡献可能会如何增强你的项目。

### **你将如何把质量管理纳入项目的所有阶段？**

质量管理不仅仅局限于数字化图像的创作。相反，质量管理要考虑到，并将之纳入项目的每一个阶段。整体质量参数包括：高质量的光学图像、完整性和元数据的可靠性（包括链接），忠实于原始馆藏和易用性。

## **2 选择原稿**

原稿的选择是发展数字馆藏的重要环节，主要依赖于专家馆员。要对照新的数字馆藏采集范围，研究并核对馆藏、著作、版本和副本。年表、地理、作者、主题、格式、所有者等，也要加以考虑。

下列问题可能有助于项目采集范围的确定和原稿的选择：

- 你想数字化专门性文献、综合性文献，比如“珍本”吗？
- 你想数字化现有的馆藏吗？
- 你想创建一个新的“虚拟”馆藏吗？例如，从不同的机构收集的文献具有某个共同点，比如共同出处。

在这一方面，数字馆藏采集的设计，取决于典藏机构的目标、职能和预期用户三方面。随着数字化馆藏和数字化项目与日俱增，考虑未来的发展，与相同或不同机构的数字馆藏互通有无，是很重要的。

手稿和大多数手工制作书籍都是独一无二的物品。同一版本的印刷书籍可能也会有显著的差异。那些有印章、批校题跋或藏书票的本子会别饶风致。数字化同一著作的哪些版本和副本，要视典藏机

构和项目的目标、读者和可利用资源的情况而定。至少，应该选择已有的品相最好的副本来进行数字化。

选择原稿时要对书籍的品相进行检查，因为原稿品相的好坏会影响数字化的流程。原稿选择者和其他步骤操作者之间的相互交流至关重要。

按需数字化是数字内容开发的另一个途径，但不能把它视为整个馆藏的基础。

数字化完整的知识整体总是较数字化其某一部分另为可取。因此，要数字化整本书或整个文件，而不是一个章节或页面。为一次性使用而只数字化部分内容，不是本指南的目标。

## 版权

馆藏数字化、传播和重复使用的程度要依具体作品的知识产权和隐私权而定。规制这些问题的法律条款，各个国家有很大差异。当典藏机构入藏新的实体藏品时，要理清哪些藏品不受版权限制，哪些受限制。在一定的裁决规定和协议下，受版权保护的资料也可以数字化。在通过和合作者以及其他第三方协议而开发的项目中，权利归属和“使用条款”也是非常重要的。在项目开始之前，必须预先商讨好并形成书面协议。另外，数字化产品可以看作是原始馆藏的一个新版本。因此，每一件数字物品和馆藏的可利用程度和使用条款应明确告知读者。

# 3 创建数字馆藏的工作流程

创建数字馆藏的过程可以分解成不同的阶段或步骤。各个典藏机构的操作情况或许稍有变化，但基本流程是一样的。要完整记录所有的过程。项目的总体设计和各参与人员的任务分工情况要有明确规定且要让人知道，这样错误和误解才可以避免。

## 第 1 步：检查和准备将要数字化之原件的相关材料：实体条件和现有的元数据

需要数字化的原件通常需要从其固定存储库房提走，要尽可能地紧密跟踪记录后继的移动情况。文献管理员应对脆弱易损原件进行评估，以尽量减少数字化过程中文献实体受损情况的发生。扫描技术人员应接受适当的培训，以确保能安全地扫描原件。并且，要告诉他们，如果某个项目有可能因数字化过程而受到损坏时，应立即向上汇报。在过程开始和结束时要对所有物件进行检查。

图书馆员应评估，当前的书目元数据是否能满足读者的需求。书目元数据应该是在同一个系统中，支持用户的检索，并能够提供指向数字化替代品的链接和从数字化替代品返回的链接。编目是创建馆藏的一个重要环节，它应该对相关协议，描述的水平 and 细节，以及用什么语言来进行编目作出决定。作为最低要求，要有与主目录语言相一致的描述。为了参与国家和国际项目，使更多的读者能获取这些数字馆藏，可以不断增补翻译成其他语言。其他多语言的工具或协议也可以提升访问量。描述的长度和深度要平衡一致。预期的受众人群和传播将会影响这些决定。如果没有足够的元数据，那么数字化之前就应该创建起来。（详见第 4 章的有关元数据的更详细的讨论。）

此外，足够的结构型元数据应该全方位记录数字化对象各实体组成部分，并且要对照元数据核对检查，以确保所有组成部分各就各位。必要时应为相关数字化对象编制卡片。手稿应对照卡片来核对。在归档时，应对照详细目录或索引检查文件盒中文件夹的顺序以及文件夹中各项目的顺序。数字化中省略的项目应加以标记。

## 第 2 步：数字化过程

### 设备的选择

要选择适合所需数字化对象并符合项目目标的数字捕获设备。例如，对于中世纪手稿和那些研究人员希望研究其细微之处的材料，推荐使用高分辨率的数码相机。平板扫描仪可能非常适合那些现代的图片类藏品，而专业书本扫描器能广泛用于印刷型书籍。

要尽可能做到对同一对象使用相同的处理流程，确保前后一致。由于有可能对数字化对象造成损毁，一些业已在大规模数字化项目中运用的自动化设备，可能并不适合数字化这些珍稀而质地脆弱的藏品。

### 图像质量

根据项目的一般设计，如何设置分辨率、色彩深度和布光，应考虑特定材料的任何特殊标准和普遍接受的建议、档案原件的需求以及展示和使用的需求。影响图像质量的其他因素也需要考虑，包括色彩饱和度，图像的亮度，图像完整性，以及光晕缺失（the absence of halos）和其他光缺陷（other optical flaws）。

在拍摄某些页面时如果采用了不同的设备，要按原稿顺序正确插入图像。

在设置拍摄分辨率时，一定要计算所需要的存储空间，要考虑研究人员所需要的下载时间。为了适应当前和未来的需求，那些希望一次性数字化全部馆藏的典构机构，应该考虑获取的图像质量为目前所需的最终格式的 1.5 倍以上。图像质量越高，文件在将来会越有用。

### 忠实于原件

在数字化善本和孤本时，尽可能保持和再现原件的外观和感觉是很重要的。应该摄取整个实体对象而不仅仅是知识内容。要摄取整个页面，包括正面和背面（包括外边缘），不可以在页面边缘内裁剪图像。装订成册的藏品应当从头至尾拍摄，包括扉页、空白页、封面裱衬纸和封面（前后封面及其内饰，书脊和书边）。不管采用何种格式，一定再现整个作品或工艺。另一方面，特地制作的一些特征（例如水印）常常是经由不一样的工序来实现数字化的，并且通常是加插在序列的末尾或者作为一个单独的文件。

为了能够将原稿的原始尺寸情况传递给研究者，图像中应当设置有标尺。处理同一档册时，不要随着图像处理的过程或者底本的翻转过程而随意改变页面的方向。

在获取图像时，通常扫描书籍的单个页面更为可取。因为这样做既与紧密的装订相适应，又能使得利用翻页软件展示页面时变得更为便利。有时，出于展示内容的需要，或者为了保持方向性和连贯性，而采用一次摄取迎面打开的两个对开页的方法。然而，这样做会导致以后的索引出现问题。

应在破损的页面后面垫上纸或者纸板。半透明的页面应该用米色或白色背景来衬托，以尽量减少“透光”；一些机构在摄取某些类型原稿的图像时会采用黑色作为背景，然而如此处理，通常会降低图像的对比度。在同一个项目中，大多数机构会尽量使用相同的背景。

试图再现外观和感觉，颜色是最重要和最复杂的问题之一。为了方便进行颜色校准，一份拍摄底稿中，至少要选定一页或一帧图像包含有目标色。每件设备都应校准到相同的色值标准（例如，CIR-实验室系统），同时要对设备进行周期性重新校准。显示设备同样也要进行校准调试。

### 涉及原件保护的问题

图像质量和文献原件保护之间的冲突是不可避免的，应该预先解决。许多特殊的原件，如手稿是独一无二的，往往是价值连城的文物。然而，如果数字化了，数字化复制件可以作为替代品提供使用，从而有利于原件的保护。对于那些材质脆弱的文献原件，数字文件也可以用来创建缩微胶片长期保存。

所有项目都要考虑原件保护的问题，特别是在项目外包的情况下，包括：

- 数字化之前和过程中训练有素的保护人员的配备
- 数字化和运输过程中环境和安全方面的控制
- 为尽量降低对原件的损伤，专用设备的配置（如加湿器、书籍支撑架等）
- 对扫描操作员的具体指导和培训，如翻开书页的适当角度和易碎对象的妥当操作

在获取数字图像时利用玻璃板来保持原件的平整，保护人员常常反对这样的做法。有些人认为这样做损坏书页或书脊的风险太大，其他人则接受这种惯常的做法。为了把玻璃板所施加的压力降低到最小，正在研制新的人工操作的书籍扫描装置。为了复制而拆开装订的方式不再认为是最好的做法。

在仔细检查完可能的损失或损害后，所有的原稿应尽快送回它们的固定存储库房。整个移动过程，都必须进行跟踪并记录存档。

### **第3步：所获取图像的处理流程和采集系统**

为确保最终产品的准确性和完整性（借助于校准监视器），已获取的图像应当经过一定的质量控制程序。不符合图像质量标准的图片，应重新摄制并予以更换。已丢失的图像应加以补充并按正确的顺序将其插入。必须重新审核结构型元数据并作必要的修订。

在此阶段，除了色彩校正外不进行其他任何图像处理。是否允许或者在什么时候进行颜色校正，相关典藏机构应该出台相关政策。该政策还要说明，进行图像颜色处理的操作方式该如何传递给用户。

应根据典藏机构的政策进行规范的文件命名。一些机构已经制定了文件命名的政策，把数字文件和该机构、与（或）实体对象相关联。每一个数字化的对象都应该有一个持久的标识符。

有些项目还会使用了一些其他的流程，以提高检索性能和可用性，如光学字符识别（OCR）处理，文本标记，与（或）添加地理空间的坐标。

机构水印有碍图像的使用。如果添加的话，不能形成对图像主体的干扰。

已获取的图片要转移到临时服务器以便系统摄取。在适当的时候，该档案的“主”图像应转移到一个永久的数字资源库，并将其从工作台和其他临时存储装置中删除。

该指南没有涵盖系统采集（将数字图像和元数据转移到管理发现系统的过程），因为这个过程要视当地技术基础设施和（或）特定的数字资源技术情况而定。

## 4 元数据

以往，图书馆员是利用书目元数据（编目）和某些类型的结构型和管理型元数据来管理和提供实体馆藏的存取。而今，图书馆员正在建设现代化的元数据模型，这些模型不仅能够提供对数字对象的访问，还能够提供所需长期保存的信息，并便于接入网络发现系统。下面，我们介绍四种适用于数字馆藏的元数据类型。

### 书目（或描述）型元数据

书目元数据描述了被数字化的实体对象，并包括知识内容方面的信息。所有被挑选出来要进行数字化的材料在数字化之前都要有某种书目元数据。在完成馆藏文献数字化后，应把指向数字版本的链接追加到相关书目记录上，该链接还要从该数字版本指回书目记录。

应采用公认的国际标准来创建这些记录，可以是最低水平的记录，也可以是完整的书目记录。正如前面提到的，相关物件的实体描述（详尽不等）也应该创建。

每个对象都应该在在线目录（可以是一个馆藏水平的记录）上进行编目登记，然后这些目录信息要能够被标示和（或）导入到其他系统。如果方便或有必要，相应的数字对象也可以在目录上有相关的说明。

### 结构型元数据

为了研究复杂的对象，比如中世纪的手稿、档案、信件或照片相册，研究者需要能够从一张张数字页面或图像中重新创建（或重组）实体对象。正在对中世纪手稿和其他类似复杂对象进行数字化的图书馆，应特别注意要进行优秀的校对，并提供其他形式的结构型元数据。至少，要保证研究者能够确定页面或图像的原始顺序。应该提供旧的和新的两种页码号（如果原来没有，在数字化之前应给这

本书编制页码)。应该给出总页数以及左页和右页的标识。其他重要要素还有编号方案、文本分段,重要的引用和插图。

### 图像(或技术)型的元数据

图像元数据(有时称为技术元数据)通常是由相机或扫描仪自动获得,并出现在文件的标头。它应包括以下要素:

- 以像素为单位的长度和宽度
- 采样
- 压缩方法
- 分辨率
- 以字节为单位的大小
- 设备信息(诸如相机或扫描仪的品牌和型号)
- 制作时间

如果图像是加工处理过的,与之相关的信息也要包含其中并且能够传递给用户。

### 管理型元数据

许多图书馆还需要附加管理元数据以协助数字文件的存取管理。管理元数据可能包括图片的拍摄者、图像或者原件的所有者、版权信息以及使用额度。管理元数据还可能包括如下信息:数字化过程中没有进行数字化的馆藏项目及其原因和形成的决议,还有常规的图像处理过程(整个馆藏的处理过程是一致的)的说明。

## 5 显示

用户通过常用的搜索引擎、简单的互动活动和展示网页来寻找开放免费的存取资源和易于获取的知识发现。展示网页是通过标准的网页浏览器和插件来呈显的,具备选项功能(双面显示和缩放功能),标记的功能(特别是为了以后的检索),个性化的注释功能,打印功能,下载、再利用及重组的功能。文稿的呈现格式要便于浏览和下载且容易操作。

演示文件通常是由“主”文件创建。该文件的显示和格式取决于项目的目标。例如,用户需求高倍放大时,适合使用 TIFF 或 JPEG2000 格式文件。原版高清 TIFF 格式文件往往用于存档而不用用于展示,

因为它们太大而不适宜于快速显示，也使得操作不方便，而 JPEG 和 PDF 格式都是流行的图像显示格式。（见第 8 章为了数字馆藏的长期保存。）

如果演示文件格式是经过处理的，不论是何种方式（例如裁剪），都应该告知用户。由于原本收藏不完整，数字化的二次产品，如果是编修或者改造过的版本，应明确标注。用户要能够很容易地确定所看到的内容出自文档或藏品何处，并且要能够轻松返回逻辑位置。至少要努力重建阅览室体验，而且如果可能的话，还因附加功能而超越这种体验。

要明确告知用户有何版权限制或其他的“使用条款”。可以考虑使用已有的规范来传递这一信息，诸如知识共享许可证。<sup>5</sup>

为确保对资源的持久存取，强烈建议使用持久标识符（PI），如 PURL（持久的统一资源标识符）、URN（统一资源名）、DOI（数字对象标识符）或句柄（Handle）。持久标识符（PI）不仅在唯一数字对象的本地 URL（统一资源标识符）发生改变时提供了便利的存取，还为引用和未来的验证提供了便利的方法。在今后任何关联数据环境中，它也是关键的。

## 6 传播、推广和再利用

数字化物件的书目描述和数字化文档（如果已创建）的书目记录，应包括在主目录中，并且要链接到相关文件。这些描述也要存储于数字图书馆系统，并且要建立返回主目录的链接。

大多数图书馆和赠款资助的项目有一个任务，就是要能够提供数字化馆藏的全球存取。这样的数字化馆藏通常将元数据和链接的分布也纳入其中，有时还有图像的分布。

增加访问和知名度的一种方式是利用门户网站和其他一些聚合网站，这些网站可能是国际的，国家的，还可能是特定主题的<sup>6</sup>。

许多图书馆还会利用博客和社交网站来推广宣传他们的数字馆藏，使得感兴趣的组织和协会得以了解它。在和外部服务机构和团体共享数字馆藏之前，图书馆应仔细查看“参与条款”并审查图像的法律权利，以确保它们不会和图书馆政策或其他现行的合作协议相冲突。

许多数字化馆藏是更大型研究项目的一部分。图书馆可以通过提供注释、评论以及其他相关信息来增加数字化馆藏的价值，使他们更为有用。新的分发系统提供了新技术，使得转录、注释和图像分析更为简易。

## 7 评估

大多数图书馆希望对数字化实施成效进行评估，也对他们的数字化项目和方案的传播和影响进行评估。一些组织已经开发了方便评估过程的表格。与产品和使用相关的统计数据为定量分析打下了良好基础，比如数字化书籍或对象的数量、门户页面的访问数量、数字对象的访问次数或下载次数以及它们被引用的次数或链接的次数等。

但是，更为重要且更为困难的是定性分析。因为定性分析往往需要从用户那里得到反馈。要考虑如下问题：

- 替代物忠实于原件的情况如何？图像是原件的有效替代品吗？研究者是不是至少需要查核原件一次才能达到研究目标？
- 该产品的可读性和可用性如何？
- 该技术满足研究需求的情况如何？
- 资源的使用情况如何？谁在使用？
- 资源的重复使用或用于其他目的的情况如何？
- 对实体馆藏的利用有什么冲击？

要把数字化计划方面的统计数据纳入到图书馆活动和使用的所有定期报告中。

## 8 数字馆藏的长期保存

各典藏机构要制定数字馆藏长期保存的策略，尤其要根据数字化的成本、工作人员的投入情况、善本和孤本保护的紧迫程度来制定。数字馆藏的保存可以由典藏机构本身完成，也可以外包给服务供应商或服务组织，或通过分布式的、联营的模式来实现。可信任数字仓储（TDR）的标准业已存在，但大多数图书馆单凭自身的力量很难达到<sup>7</sup>。

至少，图书馆应确保他们定期备份的网络服务器中的数字馆藏是高分辨率的，并且有程序和系统会随时监测数字文件的完整性。在分散的不同地方存储多个副本也是一种公认的保存策略。定期评估馆藏迁移需求或进行软件功能更新的程序要准备就绪。



## 9 建议摘要

- 仔细规划，明确目的和目标，考虑版权限制、经费以及机构自身的能力。图书馆员、技术人员、文物保护工作者、研究者和管理人员多方面的人员配合。

- 要预测学者和数字研究方法的需求，其中可能包括大数据、资源聚合、详细图像分析、数据可视化，地理信息测绘、社交媒体等。

- 借助数字化项目提供那些过去一直深藏不露的藏品，并努力把它们与其他典藏机构的相关材料无形连接起来。

- 永远要数字化完整的知识整体，而不是它的某个部分。要保留完整原始对象的外观和感觉，并加强其作为历史文物的研究。既要保持原有的结构和内容，保护数字化复制件，还要坚持建立将数字化复制件连接回目录记录的链接。

- 数字化后永不丢弃原始馆藏。获取知识的完整程度，与数字替代物的互动永远不及与实体藏品的互动。

- 要提供高分辨率的彩色图像，它们包括图像的技术信息、版权的信息、标尺和色卡、缩放功能，并且在可能和适当的时候能够实现全文检索，全部或部分文本的转录，具备地理空间坐标。要使用数字资源的持久标识符。

- 要提供尽可能多的描述、结构、技术和管理元数据。提供一个能返回其固定存储库房中实体对象的相关描述链接。

- 如同新入藏实体藏品，对新创建的数字馆藏要进行文件记录，并要使研究人员能够获取这些信息。

- 向全球免费开放研究资源，并以一种容易聚集、使用和导入数字研究和出版平台的方式，为用户提供数字化对象、馆藏和元数据的下载服务。

- 通过外部门户网站和聚合服务机构展示数字馆藏。

- 要对项目成效进行评估和记录。

- 确保长期保存数字馆藏。

## 参考文献

除非有特别的说明，在线资源均于 2014 年 6 月 22 日访问所得。

1. Some standards that take a more comprehensive approach include:

Detusche Forschungspemeinschaft. (2013, February). *DFG Practical Guidelines on Digitisation*. Retrieved from [http://www.dfg.de/formulare/12\\_151/](http://www.dfg.de/formulare/12_151/) (English)

Bibliothèque nationale de France, Bayerische Staatsbibliothek, Universitat de València, Herzog August Bibliothek Wolfenbüttel, and Bibliothèque Royale de Belgique. *Project Europeana Regia*. The following three documents are available from: <http://www.europeana-regia.eu/en/project-europeana-regia/project-documentation>

- *Attractive Guidelines for Users*. (2009)
- *State of the Art in Image Processing* (2011)
- *Quality Management [in the field of medieval manuscripts and other prestigious objects]* (2011)

IFLA and UNESCO. (2002, March). *Guidelines for Digitization Projects: For Collections and Holdings in the Public Domain, Particularly Those Held by Libraries and Archives*. Retrieved from [http://portal.unesco.org/en/ev.php-URL\\_ID=7315&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/en/ev.php-URL_ID=7315&URL_DO=DO_TOPIC&URL_SECTION=201.html)

2. More minimal approaches with links to other resources include:

Kavlilil-lolil, A. (2001, August). *Selected Literature in the Field of Digitization*. Retrieved from [http://www.ictparliament.org/sites/default/files/2011\\_training\\_puerto\\_rico-alenka\\_kavcic-selected\\_literature\\_in\\_the\\_field\\_of\\_digitization.pdf](http://www.ictparliament.org/sites/default/files/2011_training_puerto_rico-alenka_kavcic-selected_literature_in_the_field_of_digitization.pdf)

UNESCO. (In progress). *Fundamental Principles of Digitization of Documentary Heritage*. [“The purpose of this text is to collate in one document, the basic information needed to understand the requirements of digitization.”] Retrieved from [http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/digitization\\_guidelines\\_for\\_web.pdf](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/digitization_guidelines_for_web.pdf)

3. Guidelines that focus on digital capture and specific media include:

American Library Association, Association for Library Collections and Technical Services, Preservation and Reformatting Section. (2013, June). *Minimum Digitization Capture Recommendations*. Retrieved from <http://www.ala.org/alcts/resources/preserv/minimum-digitization-capture-recommendations>

Federal Agencies Digitization Initiative Still Image Working Group. (2010, August.) *Technical Guidelines of Digitizing Cultural Heritage Materials: Creation of Raster Image*

*Master Files for the Following Originals: Manuscripts, Books, Graphic Illustrations, Artwork, Maps, Plans, Photographs, Aerial Photographs, and Objects and Artifacts.*

Retrieved from

[http://www.digitizationguidelines.gov/guidelines/FADGI\\_Still\\_Image\\_Tech\\_Guidelines\\_2010-08-24.pdf](http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image_Tech_Guidelines_2010-08-24.pdf)

4. Digital preservation remains a pressing concern, as evidenced by this UNESCO declaration:

UNESCO and University of British Columbia. (2012, September 26-28) .

*UNESCO/UBC Vancouver Declaration: The Memory of the World in the Digital Age: Digitization and Preservation*, 26-28. Retrieved from

[http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco\\_ubc\\_vancouver\\_declaration\\_en.pdf](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco_ubc_vancouver_declaration_en.pdf)

5. IFLA is now using Creative Commons licenses for its publications.

Creative Commons. (n.d.) . Retrieved June 27, 2014 from Creative Commons website,

<http://creativecommons.org/>

6. Aggregators and portals are proliferating. Two common international efforts include Europeana (<http://www.europeana.eu/>) and the Internet Archive (<https://archive.org/>) . Examples of national efforts include Gallica in France (<http://gallica.bnf.fr/>) , Trove in Australia (<http://trove.nla.gov.au/>) , and the Digital Public Library of America in the United States (<http://dp.la>) .

7. The Primary Trustworthy Digital Repository Authorisation Body (ISO-PTAB) plays a major role in training auditors and repository managers. There are three important ISO standards related to digital preservation:

International Organization for Standardization. (ISO 14721:2012) . *Space data and information transfer systems -- Open archival information system (OAIS) -- Reference model*. [A reference model for what is required for an archive to provide long-term preservation of digital information.] Retrieved from

[http://www.iso.org/iso/iso\\_catalogue/catalogue\\_ics/catalogue\\_detail\\_ics.htm?csnumber=57284](http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=57284)

International Organization for Standardization. (ISO 16363:2012) . *Space data and information transfer systems -- Audit and certification of trustworthy digital repositories* [Sets out comprehensive metrics for what an archive must do, based on OAIS.] Retrieved from

Retrieved from

[http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=56510](http://www.iso.org/iso/catalogue_detail.htm?csnumber=56510)

International Organization for Standardization. (ISO 16919, soon to be published) . *Space data and information transfer systems - Requirements for bodies providing audit and certification of candidate trustworthy digital repositories*. Retrieved from

[http://www.iso.org/iso/catalogue\\_detail.htm?csnumber](http://www.iso.org/iso/catalogue_detail.htm?csnumber)