

# 数字信息资源长期保存技术策略分析\*

吴振新 张智雄 郭家义

(中国科学院文献情报中心 北京 100080)

**【摘要】** 技术策略作为数字信息资源长期保存中至关重要的组成部分,在很多项目中得到了广泛地研究。简略地介绍技术策略的发展概况,分析技术策略分析框架中的关键因素,详细描述几种独特的保存技术策略的理论和实践方法,并对这几种保存技术策略进行简略的总结对比。

**【关键词】** 数字信息资源 长期保存 技术策略 **【分类号】** G250.76

## Analyse of Technical Strategy of Long-term Digital Information Preservation

Wu Zhenxin Zhang Zhixiong Guo Jiayi

(Library of Chinese Academy of Sciences, Beijing 100080, China)

**【Abstract】** As the essential part of long-term digital information preservation, technical strategy has been widely studied. This paper gives an overview on the development of technical strategy. Specifically, the key factors of technical strategy analysis frame are analyzed, the theory and practice of several representative technical strategies are explained in detail, and a summary with comparisons is succinctly provided at the end of this paper.

**【Keywords】** Digital information Long-term preservation Technical strategy

近年来,随着数字技术的发展而产生的数字信息洪潮,使得数字信息资源长期保存成为科学研究所面临的一个重要挑战。技术策略作为数字信息资源长期保存活动中至关重要的组成部分,受到了许多项目和机构的广泛关注。

### 1 技术策略分析框架

数字资源保存首先要具有存储对象的能力,同时通过维护资源保持其生存能力、可呈现能力和可理解能力。数字资源的生存能力是指保持完整无缺的位流文件,保证从记录资源的媒体上的可读取性;可呈现能力是指可以把位流文件转换成人或机器可用的格式;可理解能力则是指被呈现的对象能够被用户说明和理解<sup>[1]</sup>。

被保存内容作为数字资源长期保存活动中的目标主体,是决定保存技术策略的根本因素。不同的保存内容因其自然属性不同,故对保存有不同的要求。通常要依据保存目标和具体的保存实践仔细分析数字资源的复杂组成,包括数字内容本身、用来呈现数字内容的显示软件

以及分发管理数字资源所需提供的各种系统功能等,从而确定被保存内容的深度。

同时,数字对象的生存能力、可呈现能力和可理解能力这3种不同层次的保持要求也极大地影响着技术策略的选择。比如,用于呈现的格式经常与底层内容格式不同,而呈现软件必然会随着时间而被替代,不同的格式对于技术变换的敏感度也不同,如果想保持最初显示的那种效果来保证其再次呈现的能力,技术上的要求是差别很大的。

长期保存所面临的一个很重要的挑战就是逾越随着时间的变迁而产生的技术障碍,数字资源需要通过特定的技术来实现数字资源的保存和呈现,因此,保存技术成为影响选择和制定技术策略的另一个主要因素。目前,经常为大家所熟知的是迁移、仿真和技术保存,这3种保存技术各有优劣,此处不再详述。

保存内容、保存能力和保存技术形成了技术策略分析框架中的3个关键要素。基于不同深度的保存内容、不同层次的保存能力和不同种类的保存技术的组配,在数字信息资源长期保存实践活动中形成了一些独具特色的保存技术策略,如图1所示。

收稿日期:2005-12-27

收修改稿日期:2006-01-24

\*本文系国家科技图书文献中心(NSTL)“数字化科技信息资源长期保存体系与政策机制”项目(项目编号:NSTL2004-DP)的研究成果之一。

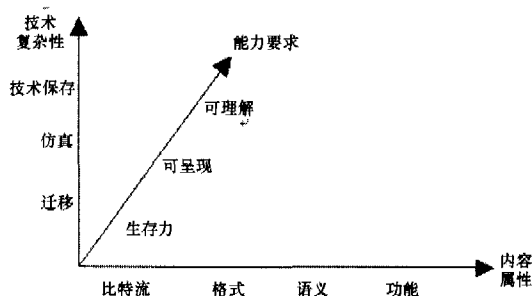


图1 技术策略分析框架的三维坐标模型

## 2 目前数字资源的保存层次划分<sup>[2,3]</sup>

### 2.1 位流保存

即保存纯粹的数据。只存储原始数字对象,维护原始位流的完整性和可读性,不考虑数字对象的外观、感知和相关功能。一般是利用复制技术,提供一种简单的数据备份存储能力,通常不能直接为用户提供访问服务,在需要将备份数据上载到一定系统上提供服务,这种保存有时也被称为阴暗存储(Dark Archive),如DAITSS项目<sup>[4]</sup>。比较流行的方式有多重备份、异地备份。另外,硬件迁移、载体迁移等也常用于位流文件保存。Dspace, NDAD<sup>[5]</sup>, Internet Archive<sup>[6]</sup>和 Kulturarw3<sup>[7]</sup>等项目都采用了备份和迁移的方式。

### 2.2 数据内容保存

保存数字对象的最基本内容,如文档类文件的文本方式保存,图像类文件的光栅方式保存等,数据内容不具备更多的特性,以最简略的格式进行保存和呈现,对于技术的要求较低,是一种介于位流和知识之间的相对简单的保存策略,通常作为一种辅助保存选择,如加利福尼亚数字图书馆的风干策略。

### 2.3 完整的知识保存

即在保存内容数据的同时还保存“与内容数据的存储和使用相关的元数据”。由于数字资源的利用依赖于对其格式和处理方式的理解,因此仅保留纯粹的内容数据不能够保证长期保存的可靠性。为了避免“死数据”的出现,存档者在提供内容数据备份存储能力的同时,也提供相关的元数据,并对内容数据进行定期上载、确认、使用检验机制,以保证数据及其相关功能的完整性和可使用性,实现对数字对象的数据外观、感觉、语义的完整保存。DAITSS项目除对数字对象的位流保存之外,还进行了完整保存实验。

### 2.4 服务保存

保存数字资源的一个主要目的就是数字资源的再利用,即重现原始的信息服务,所以数字资源与相应的服务

系统已经很难区分开来,特别是当前数字信息系统利用其知识组织机制、链接机制、交互处理机制等提供的丰富服务功能。只有“复原”原始系统提供的服务,才是真正意义上的数字资源可持续使用。因此,数字资源长期保存还需要保存数据的支撑运行环境。但如何析取“支撑运行环境”的成分还是一个难题。目前,对于技术元数据的研究意在进一步解决此问题,新兴的语义元数据的发展将会提供更多的帮助。加利福尼亚数字图书馆的Web保存正在就服务保存作进一步的研究。

## 3 现有的保存技术策略分析

### 3.1 按需迁移<sup>[8]</sup>的保存策略

保持数字对象的长期可用性是数字保存的重要内容,迁移是广泛使用的一种数字资源长期保存策略之一。当数据格式过时或发生其他情况时,利用迁移工具将原始的数字对象迁移到新的数据格式下,可保证数据在新的平台环境下的可用性。然而,传统的迁移方法存在着一些不足,即如果在迁移的某一步骤存在错误、遗漏或其他情况,就会影响以后的迁移,从而产生不同程度的失真。该方法还无法准确地保存和提供可信赖的还原机制,同时,需要迁移时就生产相应的迁移工具也会造成相关费用的提高。

Cedars项目认为在长期保存中,保存数字对象的原始字节流可以有效地解决保存中的相关问题<sup>[9]</sup>。如果保存的原始字节流并不随时间改变,则只需改变解释或还原原始格式的方式或工具。这种方法的明显问题是设计和使用的迁移工具的寿命比较短,一些研究项目为了解决这个问题,提出了一些设计工具的技巧。如CAMILEON项目<sup>[10]</sup>将这些技巧与软件工具设计结合起来,可以有效地延长软件使用的寿命,这些技巧也可以用于迁移工具的设计中。将软件设计的技巧与保存原始字节流的原则结合起来,可以产生新的迁移情况,如图2所示。该方法为执行迁移策略提供了更为实用的方法,按需迁移减轻了数字保存的负担。使用CAMILEON项目介绍的软件设计技巧的工具可以随时间还原所有数字对象,保存过程只需保存数字对象的原始格式。当以前支持的格式过时,只需在工具上添加可以解读新的数字格式的输出模块即可,同时也需要返回原始的数据格式,以保证按需迁移,从而减少在转换过程中产生错误的可能性。从原始版本向新版本转换永远只需一步,无论在何种情况下,需要迁移的是档案中的原始格式。

按需迁移可以减少错误,与传统的迁移方法相比,该方法具有以下优点:解释或读取特定文件格式的编码只

需执行一次;使用一步迁移策略可以保证迁移的准确性;保存数字对象的原始格式使得真实性问题变得简单;迁移工具模块的设计使得执行可逆的迁移试验变得简单、经济;迁移工具只是按需使用,因此在保存大量的数据时可以节省费用。

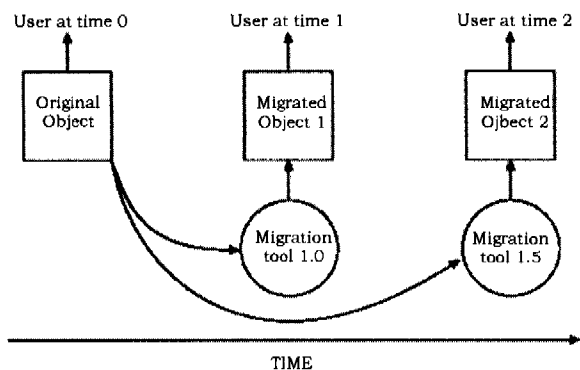


图2 按需迁移<sup>[11]</sup>

CAMILEON 项目<sup>[12]</sup>进行了矢量图像 Draw 格式和 WMF 文件格式的相关测试。DAITSS 项目<sup>[4]</sup>也开发了提供按需迁移功能的迁移器,通过修改迁移器配置文件来反映新格式的增加。

### 3.2 基于 UVC 的保存策略

仿真其实是生成一套软件,用于模拟保存、访问数据的硬件或软件,有时只是模拟硬件或软件的一部分功能,预期重现数字对象的原始操作环境,其优势在于与操作平台无关。

通用虚拟计算机(UVC)<sup>[13,14]</sup>是由 IBM 公司设计的一个方便的开发仿真或迁移工具的平台,可用于详细说明今天的过程,这些过程可能在将来的未知机器上运行。这种方法的唯一需求就是要有 UVC 仿真器。

实际上,UVC 是一个简单化的计算机虚拟表现,由于其简单,因此可以在任何可能的计算机系统上实现并工作,主要是在不断改变的软件和硬件上生成一个格外的层,为 UVC 程序提供一个稳定的平台。同时,由于其有详细说明,使得未来的开发者可以在任何时候重建一个 UVC。可以看出,UVC 实际上是为基于 UVC 编写的程序提供了一个经过仿真的平台,使得这些 UVC 程序同样具备了与平台无关的、技术独立的仿真特点。

在保存实践中,首先要编写一个基于 UVC 的格式解码程序,用于被保存内容的格式解码和呈现,该解码程序运行在仿真的 UVC 平台上,把保存内容转换成逻辑数据视图(Logical Data View,LDV)。LDV 是数字对象的结构化描述,通常依照一个特定的 Schema(Logical Data Schema,LDS)构建,如果未来有人想要浏览被保存内容,就可

以编写一个 UVC 仿真器,然后运行解码程序,生成 LDV。同时,根据保存的 LDS,再开发一个浏览器,这样就实现了对被保存内容的重现,如图 3 所示。

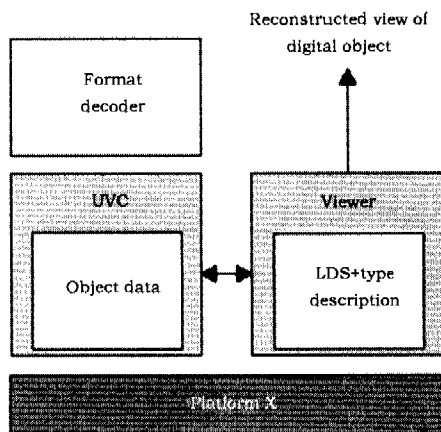


图3 UVC 保存方法说明<sup>[15]</sup>

该 UVC 方法结合了仿真和基于类 XML 的 LDV 说明的迁移这两种方法的优点。

IBM 在与荷兰国家图书馆(KB)的合作中,针对数据存档和程序存档的不同特性,提出了一个更为实用的 UVC 建议:对数据和原始程序分别进行存档的方案。

如图 4 所示,IBM 把存档和恢复数据时要做的工作分成两部分,信息以位流的方式从现在传递到未来。在目前的存档工作中,利用应用程序生成要保存的数据,同时还包括一个把数据呈现给用户的程序 P,数据和程序 P 被同时保存,程序 P 是一个可在 UVC 上执行的解码程序。在未来对数据进行恢复时,通过一个恢复应用程序的控制,读取保存的程序 P 并传递给 UVC 解释器(即 UVC 仿真器),使程序 P 在新环境中得以运行。然后恢复程序重复调用解释器,每次重复都返回一个新的带标识(这些标识同应用在 XML 中的标签很相似)的数据条目,这些标识使数据条目与语义建立关联,这样返回给用户的就是一种可理解的数据。

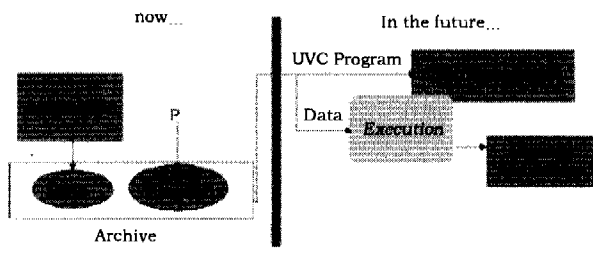


图4 基于 UVC 的保存系统的核心机制<sup>[16]</sup>

#### (1) 数据存档

数据存档不需要全面仿真,只需保存一个转换程序 P,它负责从原始格式的位流文件中抽取数据并以一种可理解的方

式返回给未来的用户。这样,数据便可以输出到新的系统。程序 P 是基于 UVC 编写的,是随着数据的生成而编写的,它可以在未来机器的仿真器上运行,未来所有要做的工作就是为执行程序 P 提供一个 UVC 的解释器(仿真器)。

### (2) 程序存档

程序存档则更为复杂,为了保存程序的行为和功能,仿真是不可避免的。IBM 建议采用同样的 UVC 来编写现在的 UVC 程序,这样,当将来需要运行现在的程序时,只要仿真现在的计算机就可以了。因此,在保存程序时需要保存程序代码和一个老机器的仿真器(用 UVC 语言编写)以及运行原始应用程序需要的数据文件。在未来,需要编写一个新机器的仿真器,新机器的仿真器通过解释 UVC 指令来仿真老指令,从而有效地生成老机器的等价物,然后运行原始应用程序代码,处理数据并生成同样的结果。针对该过程,元数据中必须简单地包括一个用户手册来指导如何运行该程序。

IBM 成功地对 PDF 文件格式进行了 UVC 实验。有些人士认为,UVC 是唯一能够实现真正长期保存的策略,能不受时间的限制,可被经济地和正确地维护。一些反对意见则认为,用 UVC 方法来保存和延长存档寿命是一个不实用而且很难实现的策略。已经有研究项目证明,UVC 方案无法有效地应用于交互式资源,该论点得到了 IBM 和其他观测人士的承认。

### 3.3 基于风干的保存策略<sup>[17,18,19]</sup>

在数字保存领域中,风干(Dessication)是指从复杂数字对象格式中提取有价值的内容,保存简单的、低技术含量、机器易于还原和容易被人理解的数字对象格式的过程。基于风干的策略,就是在保存数字对象的原始版本之外,还需要保存一个简单的、低技术含量的,经过干燥处理的数据版本。风干是数字保存策略的附加方案,也是一种以防万一的策略。

加利福尼亚数字图书馆(CDL)的 John Kunze 博士认为<sup>[17]</sup>,在数字保存系统中,加入更多的技术因素,反而会使得数字保存复杂化。如果今天采用最简单的技术来对数字信息进行保存,明天才有可能通过最简单的操作来读取和理解被保存的信息。虽然技术是数字保存的一个部分,但是如果对技术的依赖性越大,数字保存的风险也就越大。因此,他提出了除移植和仿真之外,还需要另外一种数字保存的方法——Dessication(风干)方法,作为补充保存策略。

CDL 认为,基于风干的方法对于类似文档的数字对象来说,在保存原始格式的同时,应该同时派生出一个低技术含量的版本进行保存。这种方法从高质量的、拥有保存的数字对象格式中生成一个只含少量数据的派生对象,这些派生的版本是原始文档的缩减呈现。在这个过

程中,必然丢失许多特征,如光栅图像的动画和热点链接,文本文件的字体、颜色和插入的图表等,但是它们保留了基本的文档属性,保存有大量高价值的信息。

这个策略的提出来自纸本资料的启示:纸作为记录和显示的介质,其寿命在正常条件下可持续 1000 年,而今天我们所使用的其它媒介不具备这样的寿命,除了物理实体的寿命限制外,最关键的地方在于人类理解和使用纸本记载的资料中间没有技术的需求,不存在技术的障碍。同样,对于数字资源的保存,今天如果我们采用最简单的技术对数字信息进行保存,明天才有可能通过最简单的操作来读取和理解我们保存的信息。

CDL 对 Web 格式文档数据的风干处理是过滤掉所要保存的复杂数字对象的字体、图像、色彩、读音符号等,只保存最基本的文本数据,整个处理过程是随着对原始数字对象的保存而进行的。以此类推,类似于纯文本格式的图像格式是基本的光栅文件。

这种风干数据在很多中长期存在的数字格式中得到了检验,IETF 的 RFCs 的仓储就是一个重要的例子。RFCs (Request for Comments)是一套关于 Internet 技术和组织的说明的仓储,从 1969 年开始成为 RFC 连续文档的备忘记录,长期保存着计算机网络各方面讨论的备忘录,包括协议、过程、程序、概念、会议记录、评价等等。它以纯文本格式保存文档,不具备丰富的特性,这种“干瘪”的数据成功地通过 Internet 上所有协议进行呈现。该仓储持续了 36 年,今天我们依然能很容易地阅读到几十年前的文档。

除了文本和图像,CDL 还没有明确地阐述关于音频、视频或多媒体数字资源的风干数据版本,同时,他们也提出了目前存在的问题,无法有效地提取有价值的隐藏的数据。

### 3.4 基于 UAF 的保存策略<sup>[20,21]</sup>

在长期保存中,一个值得强调的理念就是被长期保存的是数据本身,即位流文件(在实践中是字节流而不是位流),而不是存储着数据的媒体。数字资源的一个很特殊的也很重要的属性就是可复制性。所有的数据在某种意义上都是抽象的,无论数据被存储在何种媒体上,总有些方面只与媒体管理相关,而与信息内容及相关的操作无关。在数据采集的早期,决定如何把数据从媒体复制为长期保存的字节流的关键是识别保存数据的抽象模型。这个抽象模型应该包括数据的所有关键属性,该抽象模型被 Cedars 称为底层抽象格式(Underlying Abstract Form, UAF)。

OAIS 模型的实现和应用基于一个前提:数字对象必

须被转换为位流文件才能被保存,这个转换通过采集(Ingest)的2步过程完成:第一步:数据从保存其上的媒体中独立出来(即形成底层抽象格式);第二步:把从媒体上独立出来的数据映射成一个位流文件保存(使位流文件成为AIP的一部分)。随后,即是对位流文件的长期保存(在档案存储中持久保存AIP)。

Cedars认为,数据是一种实际的存在,而内容与其所依附的媒体无关,从媒体中剥离的UAF包括了数据的所有关键属性,并且独立于其所依附的媒体。而所谓的关键属性(Significant Properties),是一些能够重现数据对象原始目的的属性,每一个需要被保存的数字对象都需要被考虑哪些是关键属性,通常在数据提交协议中会详细说明数字对象的一系列重要属性。文件系统、CD镜像文件、带分隔符的数据文件以及来自于ASCII码的纯文本文件,都是可选的UAF。无论选择哪种格式,必须保证它能够重生一个文件系统,使之能够像原始对象那样用同样的方式操作,这样,文件系统才能通过使用该格式或其它软件把数据转换到字节流,如图5所示。

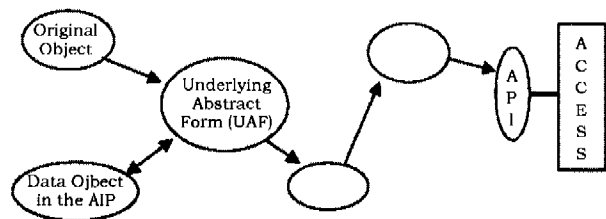


图5 对保存数字对象的映射和访问<sup>[22]</sup>

为了实现对一个特定数字对象的保存和访问,需要将具体的数字对象的底层抽象形式标识出来,然后转换为可以无限期保存的字节流存入AIP中。在对这一数字对象进行访问的时候,访问软件需要能够从AIP中将UAF提取出来,并且结合访问软件所处的环境,再对UAF进行多层不同路径映射的基础之上,还原出原始的数字对象。在Cedars系统的信息模型设计之中,希望通过将呈现信息(Representation Information)包含于AIP包的内容信息组件之中的方式来解决路径映射问题。

数据采集第一步的目的是把资源转换成单个的字节流文件。在某些情况下,资源已经以这种字节流格式存在,如PDF文件或带标记的文本文件,还有另外一些情况,如资源从CD中获得,或是使用已废弃技术的硬盘或磁带。第一步的重要性在于首先使数据从其存储的媒体中独立出来,这涉及到数据格式的选择和随后把这种格式复制到字节流文件中。在这个过程中,关键的概念就是UAF和重要属性。

在选择了UAF的格式之后,开始抽取、复制和保存

UAF(对于多媒体数据的UAF抽取目前还存在一定的困难)。随后保存系统将UAF打包成为AIP的一部分,同时抽取相关的元数据(包括RI),最后生成AIP包。

访问软件利用呈现信息所处的环境,决定在UAF向原始数字对象的转换过程中采用什么样的映射“路由”。例如,对于一个纯文本信息的存取,在Unix下使用Cat命令,而在Windows平台下使用Notepad。呈现信息帮助访问软件使被保存的数字对象以一种更加有意义的方式得到访问存取。

一个数字对象可能会有多个UAF,同一个UAF在不同的系统中也会需要不同的解析工具。呈现信息在UAF的生成和呈现过程中起了重要作用。

#### 4 结 语

按需迁移将迁移工具软件设计技巧与原始位流保存原则相结合,使用一步迁移策略,随时间改变实现原始数字对象还原;基于风干的保存策略是通过保存一个经过干燥处理的简单的、低技术含量的数据版本作为数字保存策略的附加方案;基于UAF的数据保存是在保存原始位流的基础上,以呈现信息元数据来支持原始数据的保存、呈现和使用,从而达到知识保存的目的;UVC策略则是对服务保存的一个尝试,从技术角度讲,它是一个转换和仿真的混合解决策略,是一种新的用于还原数字对象的方法,它不依赖现有的平台和格式,但它并没有对数字对象的所有功能和行为实现复原。

通过上面多个保存技术策略的分析,不难发现,技术策略分析框架中的三要素是相互影响互相关联的,三者的不同组配则形成了不同的保存策略。因此,在现有的数字信息资源长期保存活动中,保存者采用了多种保存技术策略,他们从不同角度和深度对数字资源保存进行了探索和尝试;但是到目前为止,还没有哪一个项目把一种单独的技术方案作为完整的保存技术策略在数字保存系统中应用;保存技术策略的制定主要依赖资源的自然属性和保存者期望要保留的数字对象的能力(即保存的实现目标)、技术的选择以及经济的和组织的因素。因此,保存技术策略的制定决不仅仅是一个选择哪种技术的简单问题,而是需要分析数字资源保存所处的完整的上下文环境,识别影响保存技术策略的各种关键因素,根据既定目标,选择合适的保存技术,形成一个可行的、安全的混合保存技术策略,最终完成数字保存的使命。

#### 参考文献:

- 1 Web - Based Government Information: Evaluating Solutions for Capture, Curation, and Preservation, <http://www.cdlib.org/programs/>

- Web-based\_archiving\_mellon\_Final\_corrected.pdf (Accessed Nov. 23, 2005)
- 2 宛玲, 张晓林. 数字资源长期保存过程中的知识产权问题分析. 中国图书馆学报, 2005, 31(3): 65-69
- 3 Patricia Cruse. Environmental Scan; Preliminary Survey Results, [http://www.cdlib.org/inside/projects/preservation/govinfo/CDL\\_Mellon\\_environmental\\_scan\\_ver.3.pdf](http://www.cdlib.org/inside/projects/preservation/govinfo/CDL_Mellon_environmental_scan_ver.3.pdf) (Accessed Nov. 23, 2005)
- 4 DAITSS Overview. <http://www.fcla.edu/digitalArchive/pdfs/DAITSS.pdf> (Accessed Oct. 7, 2004)
- 5 NDAD. <http://www.ndad.nationalarchives.gov.uk/> (Accessed Nov. 23, 2005)
- 6 Internet Archive. <http://www.archive.org/> (Accessed Nov. 23, 2005)
- 7 Kulturarw3. <http://kulturarw3.kb.se/html/kulturarw3.eng.html> (Accessed Nov. 23, 2005)
- 8 Migration on Request. <http://www.si.umich.edu/CAMILEON/reports/mor/index.html> (Accessed Nov. 23, 2005)
- 9 Holdsworth, D and Sergeant, D M; A blueprint for Representation Information in the OAIS Model (1999) <http://www.personal.leeds.ac.uk/~eclhdh/cedars/iee00.html> (Accessed Nov. 23, 2005)
- 10 Holdsworth, D and Wheatley, P; Emulation, Preservation and Abstraction. RLG Diginews 5, 4 <http://www.rlg.org/preserv/diginews/diginews5-4.html#feature2> (Accessed Nov. 23, 2005)
- 11 同8
- 12 CAMILEON. <http://www.si.umich.edu/CAMILEON/> (Accessed Nov. 23, 2005)
- 13 Lorie, R, UVC "A Project on Preservation of Digital Data", <http://www.rlg.org/preserv/diginews/diginews5-3.html#feature2> (Accessed Nov. 23, 2005)
- 14 UVC. <http://www.kb.nl/kb/ict/dea/ltp/reports/4-uvc.pdf> (Accessed Nov. 23, 2005)
- 15 J. R. van der Hoeven, Development of a Universal Virtual Computer (UVC) for long-term preservation of digital objects, <http://www.alphaworks.ibm.com/tech/uvc> (Accessed Nov. 23, 2005)
- 16 同14
- 17 John Kunze, Future - Proofing The Web; What We Can Do Today, <http://rdd.sub.uni-goettingen.de/conferences/ipres/download/Future-ProofingTheWeb;WhatWeCanDoTodayJohnKunze.pdf> (Accessed Nov. 23, 2005)
- 18 Best Practices for Image Capture. <http://www.cdlib.org/news/pdf/BestPracticeImageCapture.pdf> (Accessed Nov. 23, 2005)
- 19 张智雄等. 国际数字对象保存会议 (iPres) 综述. 数字图书馆论坛, 2005(11): 56-58
- 20 Cedars Guide to: Digital Preservation Strategies, <http://www.leeds.ac.uk/cedars/guideto/dpstrategies/dpstrategies.html> (Accessed Nov. 23, 2005)
- 21 Cedars. <http://www.leeds.ac.uk/cedars/> (Accessed Nov. 23, 2005)
- 22 同18
- (作者 E-mail: wuzx@mail.las.ac.cn)

(上接第7页)

- ence Proceedings of The State of Digital Preservation: An International Perspective, Washington D. C. <http://www.clir.org/pubs/reports/pub107/thibodeau.html> (Accessed Dec. 18, 2005)
- 5 Kenneth Thibodeau, NARA's Electronic Records Archives Program, Presentation to the National Digital Strategy Advisory Board, Library of Congress, [http://www.digitalpreservation.gov/data\\_presentations/Thibodeau.pps](http://www.digitalpreservation.gov/data_presentations/Thibodeau.pps) (Accessed Dec. 18, 2005)
- 6 Priscilla Caplan. FCLA Digital Archive. 2005 Joint Meeting of the CSUL Committees. <http://www.fcla.edu/digitalArchive/presents/FDAJointMeeting2005.ppt> (Accessed Dec. 18, 2005)
- 7 Migration on Request. <http://www.si.umich.edu/CAMILEON/reports/mor/index.html> (Accessed Dec. 18, 2005)
- 8 Raymond Lorie. The UVC: a Method of Preserving Digital Documents - Proof of Concept. 2002-11. [http://www.kb.nl/hrd/dd/dd\\_onderzoek/reports/4-uvc.pdf](http://www.kb.nl/hrd/dd/dd_onderzoek/reports/4-uvc.pdf) (Accessed Dec. 18, 2005)
- 9 John Kunze. Future - Proofing the Web; What We Can Do Today. International Conference on Preservation of Digital Objects 15 - 16 September 2005 in Göttingen, Germany. <http://rdd.sub.uni-goettingen.de/conferences/ipres/download/Future-ProofingTheWeb;WhatWeCanDoTodayJohnKunze.pdf> (Accessed Dec. 18, 2005)
- 10 Cedars Guide to: Digital Preservation Strategies, <http://www.leeds.ac.uk/cedars/guideto/dpstrategies/dpstrategies.html> (Accessed Dec. 18, 2005)
- 11 Raymond J. van Diessen and Ing. Ben J. van Rijnsoever Managing Media Migration in a Deposit System, [http://www.kb.nl/hrd/dd/dd\\_onderzoek/reports/5-medi migration.pdf](http://www.kb.nl/hrd/dd/dd_onderzoek/reports/5-medi migration.pdf) (Accessed Dec. 18, 2005)
- 12 Rachel Heery, Sheila Anderson. Digital Repositories Review. 2005-02-19 [http://www.jisc.ac.uk/uploaded\\_documents/digital-repositories-review-2005.pdf](http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf) (Accessed Dec. 18, 2005)
- 13 Handle System. <http://www.handle.net/introduction.html> (Accessed Dec. 18, 2005)
- 14 Persistent URL. [http://purl.oclc.org/docs/new\\_purl\\_summary.html](http://purl.oclc.org/docs/new_purl_summary.html) (Accessed Dec. 18, 2005)
- 15 Archival Resource Key (ARK). <http://www.cdlib.org/inside/diglib/ark/> (Accessed Dec. 18, 2005)
- 16 Storage Resources Broker. <http://www.sdsc.edu/srb/> (Accessed Dec. 18, 2005)
- 17 Shibboleth. <http://shibboleth.internet2.edu/> (Accessed Dec. 18, 2005)
- (作者 E-mail: zhangzhx@mail.las.ac.cn)