

# 数字资源特征与建设

郑州大学 臧国全

# 目 录

- 1 数字资源的来源
- 2 信息资源的属性
- 3 数字转换（信息资源数字化）
- 4 数字转换的原因
- 5 数字转换项目的实施者
- 6 数字资源的优势
- 7 数字资源的缺憾
- 8 数字资源的质量管理
- 9 数字资源的风险管理

# 1 数字资源的来源

- 原生型数字资源

如：现在CNKI的期刊全文数据库

- 数字转换的产品

如：LC的AMERICAN MEMORY

## 2 信息资源的属性

- 内容属性
- 内容标识属性（表征信息：如文字、代码、计算机存贮代码）
- 载体属性
- 阅读方式属性
- 传播属性

### 3 数字转换（信息资源数字化）

- 将非数字化载体转换为数字载体的过程
- 数字转换基本方法：数字扫描、数字拍照、模拟音频视频的  
数字转换系统
- 数字转换过程导致的信息资源属性的变化（内容属性、内容标识属性（文字、代码、计算机存贮代码）、载体属性、阅读方式属性、传播属性）

## 4 数字转换的原因

- 非数字化载体阅读的独占性占有
- 传统信息资源的保存问题（载体的脆弱性）等
- 网络的普及

## 5 资源转换项目的实施者

- 商业型企业。如GOOGLE
- 信息机构。如图书馆

## 6 数字资源的优势

(出版物——电子图书、电子期刊、电子报纸)

- 传播与更新速度快：网络传播、实时更新
- 信息量大、内容丰富：不受版面限制
- 范围广：受众分布在全球各地
- 检索方便：配备有数据库的检索功能
- 多媒体：内容形式有文本、图像、视频、音频等
- 超文本：超级链接、知识单元的多维表达
- 交互性：作者、读者、出版者之间实时交流
- 低成本、高效益：省去了制版、印刷、装订、投递等工序。

## 7 数字资源的缺憾

- 阅读问题
- 信息安全问题
- 长期保存问题
- 知识产权问题

# 缺憾 (1)

- 阅读问题：需阅读设备
- 信息安全问题：病毒、黑客等

## 缺憾 (2)

### ●保存问题：介质的刷新、数字迁移、数字仿真

**数字迁移：**当数字资源的文件格式出现过时危机时，就将其转化为新的文件格式。其主要优点在于数字资源总是以能够被普遍接受的格式为人们所使用，现行的硬件与软件能够毫无障碍地访问数字资源。迁移法的主要缺点有三：其一是在格式转化过程中，数字资源的一些样式（layout）甚至一些数据将会丢失。如果数字资源的原始样式需要重点保存，迁移法或许不是最好的选择。其二是迁移法的操作对象是电子出版物集合中的单件资源，如果数字资源集合庞大，转化过程需很长，工作量很大。其三是如果数字资源的文件格式已作废或所用的转换工具不再有效，那么采用这种方法进行格式转换就不可行。因此，采用数据迁移策略时，必须不断地开发并运行转换程序，以免数字资源的格式过时。

## 缺憾 (3)

**数字仿真：**不改变数字资源的文件格式，也不改变用于数字资源访问的软件或浏览器，但要提供采用原来浏览软件来浏览原始格式数字资源的一个工具，这个工具称为仿真工具。仿真工具在新的环境下工作，使原来的浏览器和原来的工作平台适应于新的环境。

仿真法的最大优点在于保持数字资源的外观（look-feel），比如，保存电子书籍时，其真实的外观就可长久被保存。但仿真法的一个最大缺点是开发和维护这种仿真工具是非常复杂的，也是很困难的，成本也是非常可观的。在实际应用中，要维护多个仿真工具，但即使这样也很难确保这些工具对未来计算机平台的有效性。

# 缺憾 (4)

## ● 产权问题：产权识别、产权许可与延伸许可、产权保护

产权（财产权益）有两种：有型产权、无形产权。此处为知识产权，知识的财产权益归谁所有？包括著作权、专利权、软件产权等等。数字资源的产权指该类信息资源中知识属性的权益的归属、授权许可（排他性许可、普通许可）、产权保护。

**产权识别：**数字资源的产权状态有两类：其一是免产权许可，这些都属于公共领域（比如超过一定期限的图书），这类数字资源可以自由发布和使用；其二是产权归他人所有，这类数字资源在发布和使用之前必须得到产权拥有者的授权许可

**产权许可与延伸许可：**数字资源可以被收割到其他数据库中，这时，原来的许可条款可能就不足以支撑这种附加应用，就要考虑延伸许可，以确保第三方用户对数字资源的有效存取。

# 缺憾 (5)

## 产权保护方式

### 法律文件方式:

版权声明。在数字资源发布网站上发表产权声明，说明发布的数字资源受版权和相关产权保护。并应该展现在其主页上突出位置，或放置在主页的显著链接之处

不承诺声明。不承诺声明的目的在于避免针对数字资源发布网站中发布的数字资源将来可能导致的法律纠纷，比如：声明存在着网站发布的一些数字资源内容不准确的可能性，以避免未来可能产生的纠纷。对发布的数字资源内容不提供任何形式的担保，对使用发布内容所产生的任何后果概不负责

使用条件。比如，仅供用户个人研究之用，不能用于商业目的(法定许可)

### 技术方式:

分辨率技术、水印技术、数字产权管理系统

## 8 数字资源保存的质量管理

- 数字保存的质量保证标准
- 数字保存的质量认证

# 数字资源保存的质量标准

- 1、管理质量标准
- 2、性能质量标准
- 3、用户服务质量标准
- 4、技术支持质量标准

## （管理质量标准-1）

- 1、数字保存方针：系统职责声明、后续保存计划、数字资源获取方针、数字资源收录的内容覆盖性说明、数字保存方针的评审。
- 2、人力资源配置：员工数量要求、员工技能要求、员工技能更新要求。

## （管理质量标准-2）

- 3、系统运行日志管理：日志的产生、日志的内容、日志的分析、日志的维护
- 4、灾难性事件的防御与恢复：灾难性事件应对方案的制定、灾难性事件应对方案的测试、灾难性事件的恢复
- 5、系统风险管理：风险识别、风险评估、风险管理策略

## （管理质量标准-3）

6、财务管理： 财务计划、 财务运行、 财务分析报告

7、数字保存产权管理： 数字保存协议、 数字资源产权验证与跟踪、 数字资源产权纠纷的规避

## （性能质量标准-1）

- 1、**数字资源获取：** 数字资源保存内容的界定、数字资源来源的可信任性、SIP的正确性和完整性、数字资源获取过程的记录
- 2、**AIP的质量：** AIP的组成、SIP到AIP转换过程的记录、SIP的处理结果、AIP的唯一标识符、AIP的表征信息、AIP的正确性和完整性、AIP生成过程的记录

## （性能质量标准-2）

- 3、**AIP**的保存：保存策略、保存型元数据（数字对象识别元数据、数字对象生产元数据、数字对象修改元数据、产权管理元数据）
- 4、**数字资源检索**：检索语言、检索方式、检索功能、检索结果、检索效率

## （用户服务质量标准-1）

- 1、**目标用户群体：**目标用户群体的界定、描述、公开声明、目标用户群体与数字资源说明
- 2、**描述型元数据要求：**描述型元数据的设置、描述型元数据的创建、描述型元数据方案的选择、描述型元数据参照完整性的建立、描述型元数据参照完整性的维护

## （用户服务质量标准-2）

3、用户服务政策

4、用户访问管理：用户访问权限的限定、用户访问请求的反馈、用户访问结果、DIP的正确性和完整性、DIP的可追溯性、用户访问失败管理

5、数字资源的用户可用性与可理解性

## （技术支持质量标准-1）

- 1、**数据备份：** 数据备份的份数、存储位置和备份介质、数据备份的同步更新
- 2、**损坏和丢失数据的检测与恢复：** 检测、恢复、原因识别、记录
- 3、**关键过程的变更：** 关键过程的识别、关键过程变更的记录、关键过程变更的效果测试与变更复原

## （技术支持质量标准-2）

- 4、软件：软件性能原则、软件检测机制、软件技术环境监控、软件技术环境监控信息的评估、软件的更新
- 5、硬件：硬件的配置原则、硬件检测机制、硬件技术环境监控、硬件技术环境监控信息的评估、硬件的更新
- 6、系统安全：硬件设备安全、软件程序安全、物理环境安全、人力资源安全

# 数字保存的质量认证

1、认证实施方案

2、《数字保存系统质量保证标准》

审核检查表

3、数字保存系统认证机构的要求

4、数字保存系统认证审核员的要求

## 9 数字保存的风险管理

- 1、数字保存的风险型元数据
- 2、数字保存的风险监控

# 数字保存的风险型元数据

- 1、数字对象的风险型元数据
- 2、保存事件的风险型元数据
- 3、行为主体的风险型元数据
- 4、保存政策的风险型元数据
- 5、知识产权的风险型元数据
- 6、保存系统的风险型元数据

# 数字保存的风险型元数据-1

## 1、数字对象的风险型元数据

数字对象的基本特征（内容方面，结构方面，行为方面，生产方面，访问方面，大小，格式，FIX校验、算法、校验值）；数字对象的存储（介质及其访问设备，介质有效期，介质加密，环境要求）；数字对象的使用环境（用途，软硬件的依赖）；数字对象的病毒检测（特征码，检验和，病毒密码）；数字对象的加密/解密；数字对象的数字签名；数字对象的数字水印

# 数字保存的风险型元数据-2

## 2、保存事件的风险型元数据

事件类型（导致数字对象之间发生关联的事件，如数字迁移导致迁移前后两个数字对象之间的来源关联；不导致数字对象之间发生关联的事件，如对数字对象的病毒检验）；事件执行结果（成功，部分成功，失败）；事件发生时间

# 数字保存的风险型元数据-3

## 3、行为主体的风险型元数据

行为主体类型（人，组织，软件等）；  
行为主体操作（拥有或许可产权，执行或授权执行事件，通过执行事件来创建或操作数字对象）

# 数字保存的风险型元数据-4

## 4、保存政策的风险型元数据

主要是数字保存操作的指标设置。元素示例：保存介质刷新的频率；保存介质的存储环境要求；数字对象的容错率；数据丢失的允许率；内容失真的允许率；软硬件的检测周期；灾难性事件应对方案的测试周期；数字迁移的准确率。

# 数字保存的风险型元数据-5

## 5、知识产权的风险型元数据

产权状态（保护状态，处于公共领域）； 产权管辖范围； 产权时间跨度； 产权内容识别（数字资源获取阶段的长期保存存档权，数字资源保存阶段的长期保存管理与处理权，数字资源服务阶段的服务提供权）； 产权许可的文件类型（纸介质协议，电子协议，口头协议）； 产权许可的法律基础（知识产权法，条例，判例）； 产权许可的内容（数字对象的修改、格式转换、复制、阅读、传播、删除）； 产权许可的类型（初始许可，延伸许可）； 与产权关联的数字对象和行为主体的识别

# 数字保存的风险型元数据-6

## 6、保存系统的风险型元数据

保存系统的防护（防火墙，入侵检测，入侵容忍，木马防御，可信计算，灾难备份与恢复）；保存系统的安全攻击（被动攻击，信息泄露，流量分析，主动攻击，黑客，病毒，安全漏洞，非法软件驻留，物理攻击）；保存系统的访问控制（身份认证模式，硬件认证，软件认证，访问控制机制，基于角色的访问控制，非法操作）

# 数字保存的风险监控

## 风险类型：

- ①**可获得性**。以确保数字对象能被用户长期使用。风险示例：数字对象的质量退化到保存技术无法恢复其价值和可用的程度。
- ②**可识别性**。以确保数字对象能被发现和检索。风险示例：用于描述数字对象的元数据元素与数字对象之间的链接断开。
- ③**持续完整性**。以确保存储介质中数字对象的二进制数据流不被破坏。风险示例：不适宜的存储环境导致介质退化。

# 数字保存的风险监控

- ④**可呈现性**。以确保数字对象的重要特征能被展现。风险示例：无法获得或操作用于展现数字对象重要特征所需的软硬件
- ⑤**可理解性**。以确保数字对象能被目标用户群体所理解。风险示例：用于支撑辅助信息理解的表征信息没有保存或遭到损坏。
- ⑥**真实性**。以确保数字对象是原始信息资源的等质量替代品。风险示例：数字对象来源信息、生产信息等真实性验证所需的信息缺失或没有随数字对象一起保存，或遭到恶意破坏，或描述不准确全面

# 数字保存的风险监控

风险监控的实现：

- 1、 风险检查
- 2、 风险统计
- 3、 风险监控系統

谢谢各位！