Challenges and Requirements in Preserving Cultural Heritage Materials

Qiang Xu

Shanghai Library & ISTIS

SHANGHAI LIBRARY
INSTITUTE OF SCIENTIFIC & TECHNICAL INFORMATION OF SHANGHAI

Introduction

- Software continually update
- The life of digital resources is limited
  - preserving methods
  - storage formats
- The long-term preservation of digital resources should be the key point of current digitalization work of ancient books

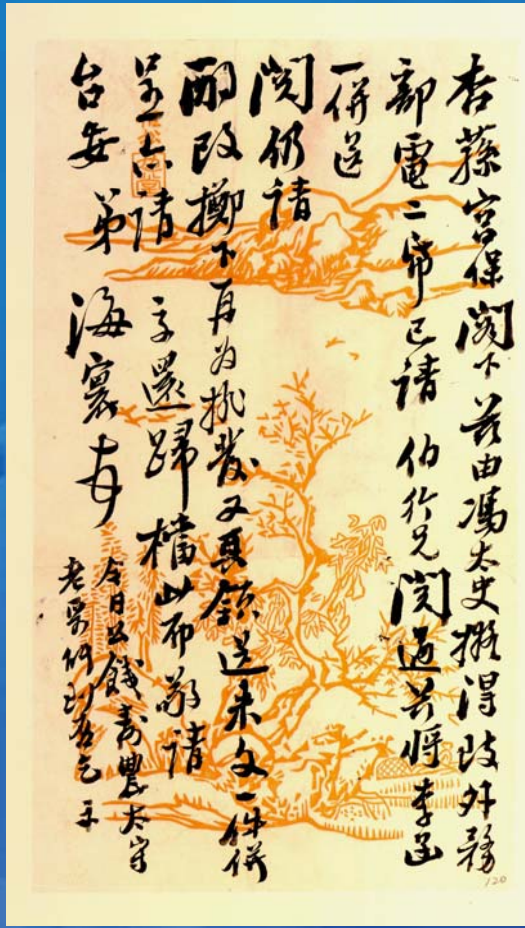
SHANGHAI LIBRARY
INSTITUTE OF SCIENTIFIC & TECHNICAL INFORMATION OF SHANGHAI

Digital Preservation of Cultural Heritage Materials

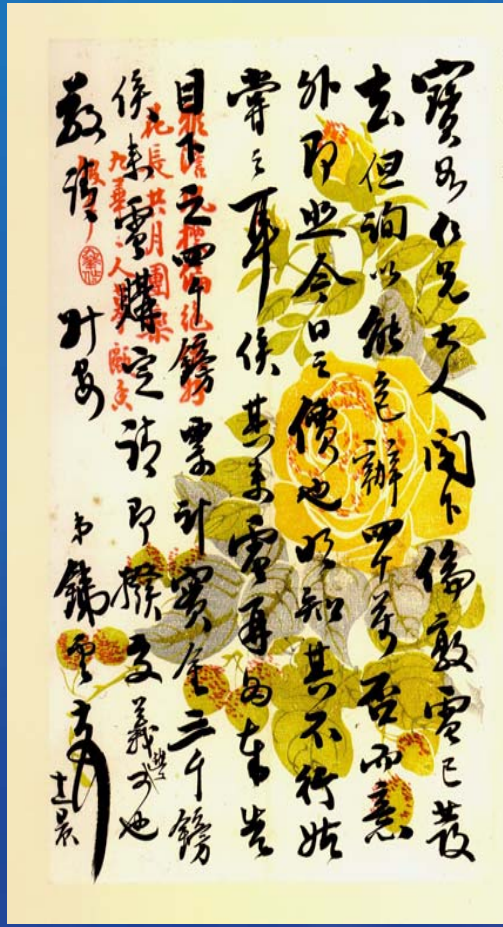
- **Digitizing Resources:**
  - Rare books: 3223 titles, 1.3 million pages
  - Genealogy: 34000 titles, 7.2 million pages
  - Archives of ShengXuanHuai: 178633 titles
  - MIN GUO books(1911-1949): 50000 titles
  - Manuscripts

SHANGHAI LIBRARY
INSTITUTE OF SCIENTIFIC & TECHNICAL INFORMATION OF SHANGHAI

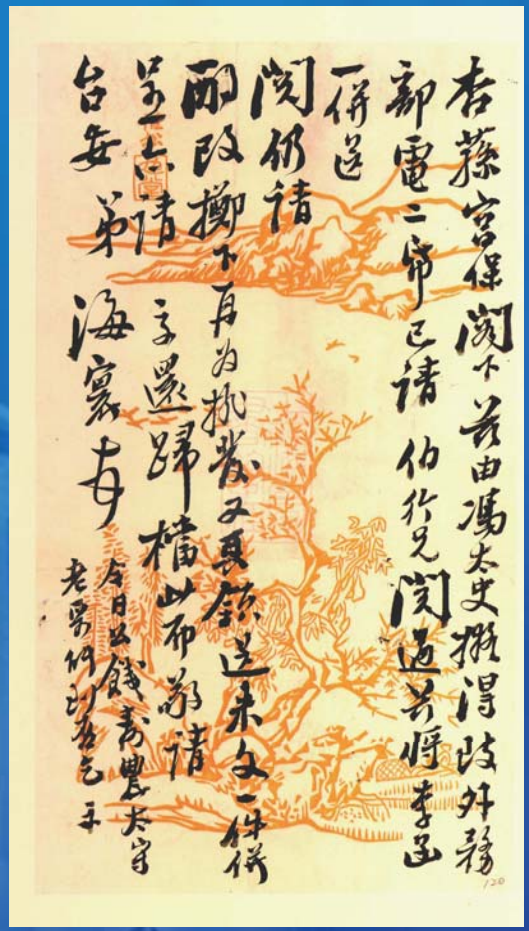
Digital Preservation of Cultural Heritage Materials



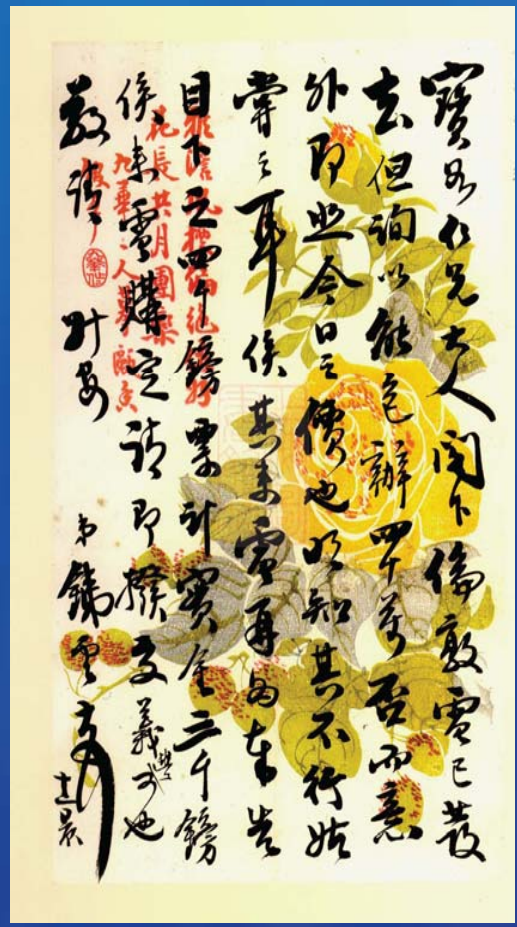
TIFF File



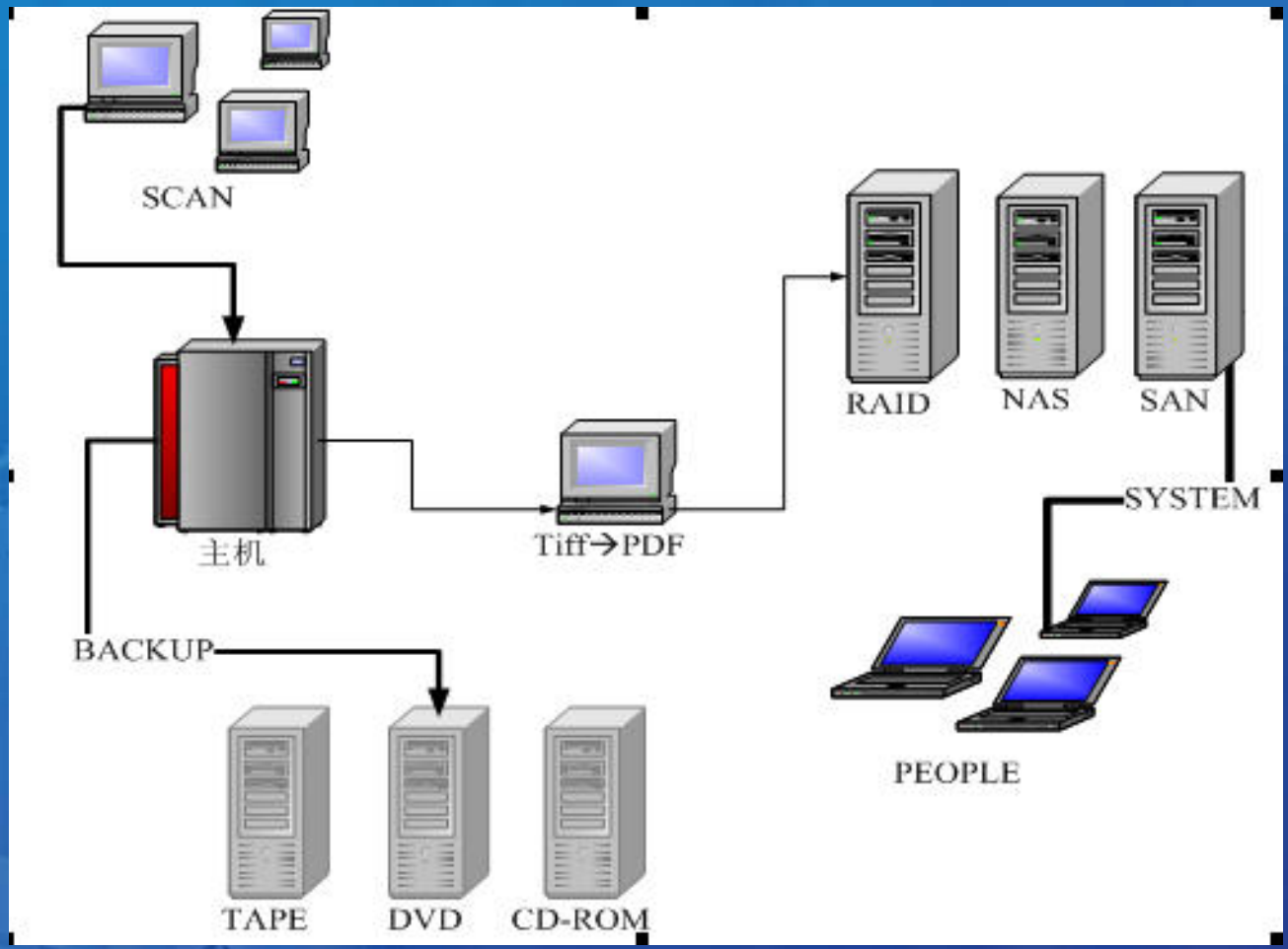
Digital Preservation of Cultural Heritage Materials



PDF



System Storage Architecture Figure



Digital Preservation Strategy

- **Establish a general metadata standard (DC)**
  - **Example:**

```
<?xml version="1.0" encoding="UTF-8" ?>
- <metadata xmlns:dc="http://purl.org/dc/elements/1.1/"
      xmlns:dcterms="http://purl.org/dc/terms/"
      xmlns:mods="http://www.loc.gov/mods">
  <mods:edition>稿本</mods:edition>
  <dc:title>霞房叢鈔十七種二十四卷</dc:title>
- <dc:creator>張紫琳</dc:creator>
-<dc:creator>黃力</dc:creator>
- <dc:date>1741</dc:date>
- <dc:description>有叢書附表伍紙</dc:description>
- <dc:relation>
  <dcterms:hasPart>T26023-37#1</dcterms:hasPart>
  <dcterms:hasPart>T26023-37#2</dcterms:hasPart>
</dc:relation>
  <dc:subject scheme="四库类名">集类,别集部,唐五代属</dc:subject>
  <identifier>788542-49</identifier>
</metadata>
```


SHANGHAI LIBRARY
INSTITUTE OF SCIENTIFIC & TECHNICAL INFORMATION OF SHANGHAI

Digital Preservation Strategy

- Manage the content of ancient book
  - Content Management
  - adopt each ancient book with DOI

Syntax of DOI:

<DOI> ::= <DIR>“.”<REG>“/”<DSS>

<DIR> ::= Directory Code
<REG> ::= Registrant's Code
<DSS> ::= DOI Suffix String
“.” ::= Separator
“/” ::= Separator

SHANGHAI LIBRARY
INSTITUTE OF SCIENTIFIC & TECHNICAL INFORMATION OF SHANGHAI

Digital Preservation Strategy

- Combined with the need for DOI, we put forward some principles for making DOI Suffix String:
  - Uniqueness
  - Stability
  - Feasibility

SHANGHAI LIBRARY
INSTITUTE OF SCIENTIFIC & TECHNICAL INFORMATION OF SHANGHAI

Digital Preservation Strategy

- **DOI Suffix String：self-definition**
  - **<DSS> ::= “SL”<Type Domain><Format Domain><Application Domain><Index Domain>**
    - **“SL” ::=Local Description。**
    - **<Type Domain> ::= Resource Media Type**
    - **<Format Domain> ::= Resource Storage Format**
    - **<Application Domain> ::= Application Resource Identify**
    - **<Index Domain> ::= Local Resource Index Identify，non-repetitive**

Digital Preservation Strategy

- Use XML to packaging digital object with metadata
- METS as the packaging format
- Platform-Independent

SHANGHAI LIBRARY
INSTITUTE OF SCIENTIFIC & TECHNICAL INFORMATION OF SHANGHAI

```
<?xml version="1.0" encoding="GBK" standalone="no"?>
<mets LABEL="古籍善本" xmlns="http://www.loc.gov/METS/"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http:// purl.org
/dc/terms/" xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.loc.gov/METS/
http://www.loc.gov/standards/mets/mets.xsd">
    <metsHdr CREATEDATE="2005-1-25T00:00:00">
        <agent ROLE="CREATOR">
            <name>shanghai library</name>
        </agent>
    </metsHdr>
    <dmdSec>
        <mdRef MDTYPE="DC">
        </mdRef>
    </dmdSec>
    <fileSec>
        <fileGrp>
            <file ID="FILE1" MIMETYPE="image/tif" SEQ="1">
                <FLocat LOCTYPE="URN" xlink:href="$directory$\000001.tif"/>
            </file>
        ......
        </fileGrp>
            </fileSec>
            <structMap TYPE="LOGICAL">
                <div TYPE="BOOK" LABEL="">
                    <div TYPE="CHAPTER" LABEL="目录">
                        <div TYPE="PAGE" ORDER="1">
                            <fptr FILEID="FILE1"/>
                        </div>
                        <div TYPE="PAGE" ORDER="2">
```

Practice in the Digital Preservation of Culture Heritage Materials

- Media preservation
- Technical update

SHANGHAI LIBRARY
INSTITUTE OF SCIENTIFIC & TECHNICAL INFORMATION OF SHANGHAI

Difficulties in the Long-term Preservation

- Lack of efficient standards
- Effect by outdated technology
- Security issues

SHANGHAI LIBRARY
INSTITUTE OF SCIENTIFIC & TECHNICAL INFORMATION OF SHANGHAI

Q&A

SHANGHAI LIBRARY
INSTITUTE OF SCIENTIFIC & TECHNICAL INFORMATION OF SHANGHAI

Practice and Requirements in the Digital Preservation of Cultural Heritage Materials

Qiang Xu

Shanghai Library & ISTIS

qxu@libnet.sh.cn

SHANGHAI LIBRARY
INSTITUTE OF SCIENTIFIC & TECHNICAL INFORMATION OF SHANGHAI