



Consorci de  
Biblioteques Universitàries  
de Catalunya

# Reconeixement de la digitalització com a un mètode de preservació per reformatatge

**Preparat per a l'ARL Preservation of Research Library Materials Committee per:**

**Kathleen Arthur, Head, Replacement & Reformatting, University of Chicago; Sherry Byrne, Preservation Librarian, University of Chicago; Elisabeth Long, Co-Director, Digital Library Development Center, University of Chicago; Carla Q. Montori, Head, Preservation Division, University of Michigan; Judith Nadler, Associate Director, University of Chicago**

Aquest article va ser publicat originalment el juny de 2004 a la pàgina web de l'*Association of Research Libraries (ARL) Preservation Program*, sota el títol "Recognizing digitization as a preservation reformatting method", [http://www.arl.org/preserv/digit\\_final.html](http://www.arl.org/preserv/digit_final.html)

La traducció catalana ha estat feta, amb permís de l'ARL, per Núria Ferran (Biblioteca Virtual de la Universitat Oberta de Catalunya).

L'article ha estat preparat per professionals de la Universitat de Chicago i de la Universitat de Michigan, amb aportacions d'altres experts en preservació així com de molts directors de l'ARL.

## RESUM

L'Association of Research Libraries (ARL) considera la digitalització com a una opció acceptada per a la preservació de diferents materials mitjançant el reformatatge ("reformatting"). Aquest document l'ha redactat un conjunt d'experts de diferents universitats, amb comentaris d'especialistes en preservació, amb l'objectiu de facilitar el desenvolupament i implementació de polítiques, estàndards, guies i millors pràctiques on encara no s'apliquen.

L'informe s'adreça principalment a la preservació del material en suport paper però també a com preservar, a través del reformatatge, materials que actualment es troben en diversos suports. Ofereix un estudi comparatiu de diferents tecnologies que poden ajudar a prendre una decisió sobre quina estratègia de preservació cal seguir, enumera quins beneficis comporta la digitalització per a la preservació, proporciona una introducció als estàndards i millors pràctiques de reformatatge digital i en descriu l'estat actual així com els compromisos que s'adopten per a la preservació a llarg termini. Finalment dona a conèixer els webs de projectes i iniciatives més rellevants en la matèria.

Les institucions culturals donen servei a la comunitat internacional a través de la creació, protecció, preservació i garantia d'accés a les col·leccions i als recursos. Els reptes de la preservació de les col·leccions han estat adreçats de maneres diferents al llarg del temps. Les biblioteques han utilitzat la conservació per preservar l'original mentre que les estratègies de reformatatge<sup>1</sup>, com ara la microfilmació i la creació de facsímils impresos, han servit per retenir el contingut, millorar l'accés i protegir l'original del desgast produït per l'ús. En els darrers anys, les biblioteques han anat escollint la digitalització com un mètode addicional per solucionar la fragilitat i el risc de pèrdua dels materials en paper tot preservant-los i proporcionant accés a les col·leccions bibliogràfiques.

L'Association of Research Libraries (ARL) està totalment d'acord amb què la digitalització és una opció acceptada per preservar una àmplia gamma de materials. L'ARL anima als seus membres així com a d'altres institucions que estan interessades en la digitalització o que ja l'utilitzen, a comprometre's a nivell econòmic i organitzatiu adherint-se als estàndards acceptats i a les millors pràctiques i a establir polítiques institucionals per mantenir els productes digitals a llarg termini. L'ARL sol·licita a les agències governamentals, a les fundacions privades i als assessors de subvencions que donin suport a les propostes que incorporin la digitalització amb la finalitat de preservació quan es donin les condicions exposades.

## **CONTEXT**

Les biblioteques necessiten utilitzar una varietat d'estratègies de reformatatge per satisfer les demandes de preservació dels diferents tipus de materials que formen part de les seves col·leccions. Cada mètode de canvi de format té els seus avantatges i inconvenients. L'elecció cal fer-la basant-se amb les característiques dels originals, les capacitats de cada procés de reformatatge, les necessitats actuals i futures dels usuaris i el cost. L'avaluació és la clau –no hi ha cap solució que satisfaci totes les necessitats. Els nostres programes de preservació han d'abordar múltiples facetes i ser efectius. Com més opcions de preservació tinguin les biblioteques, millor es podrà satisfer les necessitats de les col·leccions i les expectatives actuals i futures dels usuaris.

L'opció d'utilitzar la digitalització per a la preservació, així com qualsevol altre mètode de reformatatge, no és prescriptiva sinó que és una decisió local. Les institucions poden escollir la digitalització com a única opció o bé utilitzar un altre sistema de canvi de format o, fins i tot, utilitzar una combinació de sistemes. Molts enfocaments són possibles però la digitalització per reformatatge hauria de ser considerada una opció vàlida d'entre tots els mètodes de

---

<sup>1</sup> Nota del traductor: entenem per reformatatge l'acció de reformatar, és a dir, tal com defineix la Gran Enciclopèdia Catalana, adaptar un suport amb format a un altre format o al mateix format perquè el sistema el pugui llegir i escriure dades. En anglès "reformatting".

preservació de materials en paper. (Vegeu l'[Apèndix 1](#): "Comparació de tecnologies de reformatatge" per obtenir-ne més detalls).

## **BENEFICIS DE LA DIGITALITZACIÓ COM UNA ESTRATÈGIA DE REFORMATATGE**

El desplegament de la digitalització com una estratègia de reformatatge proporciona uns resultats positius. La digitalització incrementa la capacitat de captura de molts tipus de materials en paper, com documents em grans formats i en color, pels quals no hi ha hagut una estratègia efectiva de reformatatge fins al moment. Funcionalitats com les que permet el *zoom*, permeten a l'usuari examinar de manera més profunda els detalls més petits i donar resposta a les necessitats més diverses. Els facsímils digitals reproduïen millor l'experiència de navegació d'un llibre que el format lineal d'un microfilm. Tot i que la preservació del material paper és el principal objecte d'aquest informe, la digitalització també té potencial per capturar informació actualment enregistrada en molts altres suports i, fins i tot, pot ser l'únic mètode per preservar aquest material.

Quan els facsímils dels materials en paper es fan accessibles a través de la xarxa, l'abast d'usuaris s'amplia ja que l'accés a les col·leccions es pot produir des de qualsevol lloc tant des d'on es conserva el document com des de fora. Un entorn virtual de fitxers digitals pot combinar continguts provinents de diferents tipus de recursos, incloent materials de fonts primàries i d'aquesta manera proporcionar oportunitats molt poderoses per integrar materials per a la instrucció i la gestió de cursos per a l'ensenyament i aprenentatge. La digitalització permet als usuaris crear col·leccions virtuals que donin suport a la recerca nova i creativa que es pot realitzar només en un entorn digital. (Vegeu l'[Apèndix 2](#): "Beneficis de la digitalització per a la preservació a través del reformatatge" per més detalls).

## **ESTÀNDARDS I MILLORS PRÀCTIQUES**

Assegurar la captura d'imatges d'alta qualitat i proporcionar la viabilitat a llarg termini dels objectes digitals és un repte, però les biblioteques ja tenen una llarga tradició desenvolupant estàndards i millors pràctiques que donin suport a operacions sostenibles i facilitin la col·laboració entre institucions. Aquesta tradició aporta confiança a l'hora de poder assolir els reptes que comporta la preservació digital.

### **Reformatatge**

Les institucions i les organitzacions que treballen cooperativament han identificat els aspectes i riscos relacionats amb la creació i manteniment d'objectes digitals i han fet progressos significatius per posar cada cosa al seu lloc. En la comunitat d'experts dedicats a la preservació ja hi ha un ampli consens i acceptació sobre la creació de fitxers d'arxiu<sup>2</sup> ("master files") digitals per a materials en paper. Existeixen pautes establertes sobre com cal capturar imatges i realitzar-ne el processament per assegurar que quedin d'alta qualitat i proporcionin representacions fidels de l'original. Es fan servir formats basats en estàndards i compressors d'arxiu.

### **Metadades**

Pel que fa a les metadades, hi ha encara reptes per solucionar. S'han establert estàndards per descriure registres bibliogràfics. S'està treballant per especificar el conjunt d'elements i formats de preservació i metadades internes ("administrative metadata"), per poder facilitar la disseminació i el manteniment dels facsímils digitals al llarg del temps. PREMIS, METS, i MODS són exemples en funcionament d'iniciatives centrades en les metadades. La realització d'informes regulars i discussions obertes mentre es desenvolupen aquests estàndards

---

<sup>2</sup> Nota del traductor: "master file" s'ha traduït per "fitxer d'arxiu", segons Tèrmens és el fitxer *mestre* o d'*arxiu* amb la màxima qualitat possible, que actua de còpia de seguretat i, alhora, serveix per generar les altres còpies o per a futurs reprocessaments si milloren els estàndards. <http://bd.ub.es/pub/termens/docs/formats.pdf>

permeten als programes de digitalització capturar i enregistrar informació que serà necessària un cop els estàndards estiguin acordats.

### **Preservació de l'objecte digital**

L'experiència fins al moment, mostra que els fitxers poder preservar-se i actualitzar-se ("refreshed") a curt termini. Amb tot, hi ha molts reptes a l'hora de gestionar els objectes digitals al llarg del temps, les institucions estan compromeses a desenvolupar solucions que assegurin la integritat i autenticitat, abordin l'obsolescència tecnològica i del suport i garanteixin l'accés a llarg termini. També s'estan realitzant esforços per identificar models de negoci i dissenys de sistema que siguin capaços de sostenir grans col·leccions d'objectes digitals. Mentrestant, ja s'estan realitzant pràctiques per assegurar que es captura i s'enregistra suficient informació i que es garanteix la permanència dels objectes digitals d'ara a través de la seva gestió. Les estratègies per garantir la seguretat dels fitxers d'arxiu a curt termini inclouen l'ús de mitjans d'alta qualitat i que assegurin l'emmagatzematge, múltiples sistemes de còpia de seguretat, testejos periòdics i una agenda per actualitzar dades. Aquestes estratègies són per a la preservació a curt termini i seran el pont cap a les solucions emergents que s'estan desenvolupant per assegurar l'accés i la disponibilitat a llarg termini. (Vegeu l'[Apèndix 3](#): "Estàndards i millors pràctiques de reformatatge digital" per més detalls).

### **COMPROMISOS NACIONALS I LOCALS**

A nivell nacional<sup>3</sup>, hi ha el compromís de finançar el desenvolupament d'infraestructures per donar suport a la preservació digital. La Library of Congress (LC) s'encarrega de desenvolupar el programa de preservació digital anomenat National Digital Information Infrastructure for Preservation (NDIIPP); el National Archives and Records Administration (NARA) ha creat amb el mateix objectiu l'Electronic Records Archive (ERA); i el Government Printing Office (GPO) ha establert el Legacy Digitization Project per escanejar les col·leccions del Federal Depository Program. La Biblioteca Nacional d'Austràlia (NLA) ha realitzat un ampli treball en la seva iniciativa Preserving Access to Digital Information (PADI), i el Joint Information Systems Committee (JISC) de Gran Bretanya ha donat suport a diverses iniciatives de preservació digital incloent els UK Data Archives. El Consultative Committee for Space Data Systems (CCSDS) ha desenvolupat un model de referència àmpliament acceptat per a l'Open Archival Information System (OAIS) el qual proporciona un marc conceptual i una terminologia comuna per descriure els elements i les funcions d'un arxiu digital. (Vegeu l'[Apèndix 4](#): "Estat actual dels compromisos de la preservació dels recursos electrònics a llarg termini" per més detalls).

Algunes institucions i altres organitzacions han engegant també treballs importants per desenvolupar models, construir i testar arxius digitals, crear estàndards i desenvolupar sistemes i programari per donar suport a la preservació digital a llarg termini. El Global Digital Format Registry s'està dissenyant per capturar i compartir informació sobre una gran diversitat de formats de fitxers. Eines com JHOVE, estan disponibles per consolidar els formats de fitxers com a estàndards. Sistemes com ara LOCKSS, s'estan desenvolupant per gestionar arxius digitals i per intercanviar contingut digital i s'estan testant estratègies per migrar i comprovar dades digitals al llarg del temps. Les biblioteques continuen jugant un paper com a líders a l'hora de recomanar i testar estàndards i contribueixen àmpliament en trobar solucions adreçades a la preservació dels recursos digitals. (Vegeu per més detalls l'[Apèndix 3](#)).

L'ARL ha creat un grup de treball per a la digitalització de col·leccions de documents governamentals ("Working Group on Digitizing Government Document Collections") que s'encarrega de supervisar el *business plan* i està donant recomanacions per coordinar el paper de l'ARL i els propers passos. A més de l'objectiu del GPO de preservar contingut de manera permanent i millorar l'accés del públic a través fitxers derivats<sup>4</sup> ("derivative files"), el projecte establirà la línia de referència per a la creació de fitxers d'arxiu de qualitat per a la preservació digital, el lliandar mínim que han de complir totes les institucions participants. Alhora també

<sup>3</sup> Nota del traductor: es refereix als Estats Units.

<sup>4</sup> Nota del traductor: "derivated files" s'han traduït per "fitxers derivats" són aquells creats des del fitxer d'arxiu o master per a editar, millorar, comprimir i convertir a d'altres formats de fitxers.

desenvoluparà especificacions i guies per a la creació de metadades dirigides cap a la preservació de fitxers digitals i desenvoluparà una estructura de control. Aquestes accions comporten un esforç significatiu i a llarg termini que revertirà en la consolidació i documentació d'estàndards, pràctiques i altres guies útils per tots els professionals relacionats amb la digitalització per a la preservació.

### **TEMPS PER L'ACCIÓ**

Ara és un bon moment per adoptar la digitalització com un mètode de preservació per reformatatge. "Com més recursos neixin ja en format digital i una nova generació d'usuaris creixi amb el format digital com a sistema de lliurament per defecte, els recursos que no estan en format digital quedaran desprotegits amb el temps perquè estan en formats obsolets" (Abby Smith, Council on Library and Information Resources (CLIR), missatge de correu electrònic, 29 de març de 2004). Al mateix temps, en resposta a la creixent quantitat i dependència dels materials nascuts digitals ("born digital"), els sectors comercials i privats estan centrant-se en crear entorns on aquests materials es puguin mantenir a llarg termini.

Els aspectes tècnics que cal afrontar per a la preservació a llarg termini dels materials nascuts digitals és la mateixa que els que han estat convertits a un format digital. Per assegurar que els objectius de la preservació estan incorporats en solucions permanents a llarg termini, cal que siguem participants actius en el seu desenvolupament. Les biblioteques no poden esperar que aquestes solucions s'estableixin del tot sense haver-les provat. Per tant, cal que estem preparats pel canvi tecnològic persistent al llarg del temps.

L'ARL dona suport a la digitalització com una estratègia de preservació per reformatatge, actuarà com un catalitzador ajuntant comunitats i prendrà un rol de líder sent com una agència de recollida i distribució de la informació, promocionant l'ús dels estàndards i millors pràctiques i facilitant la implementació d'aquests estàndards per les institucions. L'ARL té el gran potencial per aconseguir aquest rol i no només amb el projecte GPO, sinó també amb la comunitat tant gran que treballa amb la digitalització per reformatatge d'un gran ventall de recursos.

"Les biblioteques vetllen pels recursos culturals i intel·lectuals de la societat. Per tal que les biblioteques puguin continuar complint aquesta missió, hauran d'aproximar-se a la preservació d'una nova manera. Caldrà que aquest rol s'integri en cada aspecte del treball bibliotecari. La preservació s'ha de considerar als més alts nivells de la institució i cal tornar-la a concebre per a l'entorn digital". (Extret del Prefaci de *The State of Preservation Programs in American College and Research Libraries: Building a Common Understanding and Action Agenda*, Deanna Marcum, President, Council on Library and Information Resources, desembre de 2002.)

## **APÈNDIX 1: COMPARACIÓ ENTRE TECNOLOGIES DE REFORMATATGE**

Les biblioteques necessiten una varietat d'estratègies de reformatatge per satisfer les demandes de preservació dels molts tipus de materials existents a les col·leccions. Cada estratègia de canvi de format té avantatges i inconvenients que fan que l'elecció es basi en les característiques del material original, en les capacitats de cada procés de canvi de format, en les necessitats dels usuaris i en el cost. La digitalització és ara una estratègia de reformatatge addicional per a la preservació. Com més opcions tinguin les biblioteques a la seva disposició poden ser més efectius els gestors de la col·lecció a l'hora de satisfer les necessitats actuals de les col·leccions i les expectatives dels usuaris actuals i futurs.

### **Facsimils en microforma**

Avantatges:

- opcions adequades de captura per alguns tipus de materials en paper
  - existència d'estàndards, guies establertes i millors pràctiques per a preservació i que estan acceptats internacionalment.
- Els estàndards, guies i pautes asseguren:
- la llarga durada del suport, si s'ha processat i emmagatzemat adequadament
  - la còpia d'arxiu
  - assequible de realitzar, emmagatzemar, reproduir i distribuir
  - la producció i el préstec de títols amb drets d'autor estan acceptats pels drets d'explotació<sup>5</sup> ("copyright")

### **Facsimils en microforma**

Desavantatges:

- baixa satisfacció de l'usuari
- possibilitats limitades d'accés i distribució
- dificultats per incorporar-ho a un entorn d'escriptori
- funcionalitats limitades (per exemple, el text no és cercable, anàlisi del text)
- susceptible de rebre maltractes
- degradació del negatiu del fitxer d'arxiu amb duplicació i pèrdua de qualitat amb les successiva generació de còpies

### **Facsimils editats en paper**

Avantatges:

- capacitats adequades per a capturar diferents tipus de materials en paper incloent la producció de facsimils impressos de gran qualitat.
- estàndards establerts i internacionalment acceptats sobre la perdurabilitat del paper a través d'indicacions químiques i mecàniques que asseguren la permanència del mitjà (ANSI/NISO Z39.48-1992 (R1997))
- estàndards establerts i internacionalment acceptats sobre l'enquadernament de materials de la biblioteca, que asseguren la perdurabilitat dels volums (ANSI/NISO/LBI Z39.78-2000)
- la perdurabilitat del paper a través d'indicacions químiques i mecàniques que asseguren la permanència del mitjà (ANSI/NISO Z39.48-1992 (R1997))
- la producció i el préstec de títols amb drets d'autor estan permesos sota els drets d'explotació actuals
- alta satisfacció de l'usuari
- un cop a mà, les pàgines són fàcils d'accedir i explorar.

### **Facsimils editats en paper**

Desavantatges:

- no hi ha una còpia d'arxiu segura
- les possibilitats d'accés i distribució són limitades
- no és fàcil d'incorporar en un entorn d'escriptori
- funcionalitats limitades (per exemple, el text no és cercable, anàlisi del text)

---

<sup>5</sup> Nota del traductor: hem traduït "copyright" per "dret d'explotació", l'únic dret aplicable en la traducció anglosaxona que inclou la reproducció, distribució, comunicació pública i transformació.

- pèrdua de qualitat de les còpies

### **Facsimils digitals**

#### Avantatges:

- augmenta les capacitats de captura en diferents tipus de material de paper com ara documents en color en grans formats
- augmenta les capacitats de captura per molts tipus de materials
- no es degrada el fitxer d'arxiu amb la duplicació i no es perd la qualitat en les còpies.
- s'assegura la còpia dels fitxers d'arxiu amb l'adhesió als estàndards que estan apareixent, pautes i millors pràctiques en un repositori digital segur
- alta satisfacció dels usuaris amb alguns productes digitals
- augmenten les possibilitats de distribució i accés
- fàcilment incorporable a l'entorn d'escriptori
- augmenten les funcionalitats (per exemple de *zoom*, d'impressió, permet cercar en el text, anàlisi del text)

### **Facsimils digitals**

#### Desavantatges:

- els estàndards, les pautes i les millors pràctiques per produir i mantenir els facsimils digitals a llarg termini estan encara en fase de desenvolupament
- augmentar les capacitats de captura comporta un alt cost
- les funcionalitats limitades dels monitors d'ordinadors actuals
- repositori d'emmagatzematge i els costos de gestió pels fitxers d'arxiu digitals no són suficientment coneguts i els costos en alguns casos són previsiblement alts
- les lleis actuals de drets d'explotació afegeixen restriccions específiques a les còpies digitals de materials amb drets d'autor i requereix que les biblioteques prenguin passos addicionals per assegurar-ne el compliment –decidir com aquesta limitació estarà gestionada formarà part del programa de planificació de digitalització d'una institució
- els aspectes d'obsolescència tecnològica han de ser adreçats en la gestió dels objectes digitals al llarg del temps

## **APÈNDIX 2: BENEFICIS DE LA DIGITALITZACIÓ PER A LA PRESERVACIÓ A TRAVÉS DEL REFORMATATGE**

Els beneficis d'utilitzar la digitalització com a mètode de reformatatge per a materials en paper inclouen:

- Augment de l'accessibilitat
  - els documents són accessibles per tothom amb accés a Internet, per tant s'amplia el nombre d'usuaris que hi poden accedir.
  - múltiples usuaris es poden connectar al mateix temps independentment de la seva ubicació i en tot moment (7 dies a la setmana, 365 dies l'any).
  - els documents són fàcilment incorporables per aplicacions instruccionals i educatives.
  
- Augment de la funcionalitat
  - els usuaris poden comparar simultàniament dues o més pàgines seleccionades de l'original.
  - les imatges digitals es poden convertir en fitxers de text cercables per permetre la cerca i anàlisi del contingut a través dels documents independentment de la seva ubicació.
  - les còpies digitals es poden manipular (engrandir i canviar el contrast) per donar servei a alguns tipus d'estudis concrets.
  - els usuaris poden tenir opcions de retallar i enganxar per realitzar citacions de manera més fàcil, etc.
  
- Capacitat d'exportació a d'altres suports.
  - es poden crear facsímils impressos d'acord amb les especificacions de preservació per satisfer les necessitats dels usuaris.
  - a partir de la versió digital es pot generar un microfilm que compleixi amb les especificacions de preservació.
  - es poden generar múltiples còpies digitals sense pèrdua de qualitat.
  - els facsímils en paper es poden generar fàcilment a un cost assumible.
  
- Col·laboracions sistemàtiques
  - un cop una institució ha digitalitzat un treball o una col·lecció, diferents institucions poden fàcilment accedir i integrar en les seves pròpies col·leccions virtuals, tot respectant les restriccions de drets d'explotació.
  
- Identificació i selecció dels materials
  - Els avantatges dels materials digitals ha generat un major interès com a estratègia de preservació que no pas el procés de reformatatge en microfilms.
    - el professorat s'acosta activament a la biblioteca amb projectes i idees que combinen objectius de canvi de format i preservació, cosa que rarament passava amb el microfilm
    - s'estan forjant noves aliances i s'estan incorporant nous materials a les classes i a la recerca
  - els responsables de la selecció i el professorat estan pensant de manera més creativa sobre les perspectives de preservació de la col·lecció, incloent l'habilitat de crear col·leccions virtuals de materials digitalitzats per reformatatge.
  
- S'expandeix l'abast de les activitats de preservació
  - Serà més senzill de compartir les col·leccions digitals, de manera que els projectes o programes de preservació cooperatius o coordinats resultaran més exitosos.
  - La digitalització per reformatatge té l'habilitat de capturar i exposar un ventall molt ampli de materials amb característiques i atributs que amb altres tecnologies no són fàcils de reformatar (per exemple, documents en color, els materials més grans del normal i altres materials gràfics complexos).



- La digitalització pot adreçar-se a les necessitats de conversió dels usuaris d'altres tipus de suport més enllà dels materials impressos en paper ( per exemple, àudio, vídeo, films) i pot permetre que col·leccions que contenen una gran varietat de formats puguin presentar-se i ser accedits sense barreres des d'una única interfície.

### **APÈNDIX 3: ESTÀNDARDS I MILLORS PRÀCTIQUES DE REFORMATATGE DIGITAL**

Els experts en preservació han establert unes especificacions tècniques per capturar el contingut dels treballs impresos que continguin text i per a molts tipus d'imatges impreses. Les especificacions de captura per aquests materials estan àmpliament acceptades. I s'està treballant en desenvolupar les especificacions de captura per imatges en color. Els investigadors i editors han establert estàndards de codificació i nivells acceptables de precisió per al procés de cerca i recuperació a text complet. També s'ha treballat molt per establir metadades descriptives, d'estructura i administratives que assegurin la recuperació, estructura i gestió dels objectes digitals d'ara i del futur. S'està testejant per desenvolupar estratègies que s'adrecin a la preservació dels suports i a l'obsolescència tecnològica. En les àrees on els estàndards no estan acabats de definir, les millors pràctiques poden assegurar que els objectes digitals es gestionin mantenint-se segurs de moment, mentre s'espera que s'implementin estratègies ben definides per a la preservació a llarg termini.

#### **Marc per proporcionar accés continuat als objectes digitals**

Reformatar els objectes digitals per preservar-los es fa en un marc que assegurï una gestió dels objectes digitals a llarg termini seguint els estàndards, de manera sostenible i amb el suport institucional:

- OAIS Open Archival Information System Reference Model (ISO 14721:2003): un marc conceptual per a un sistema d'arxiu genèric i complet, amb termes i conceptes comuns per descriure les arquitectures dels repositoris i comparar implementacions.
- OCLC/RLG Preservation Metadata and the OAIS Information Model: un marc de metadades per donar suport a la preservació dels objectes digitals.

#### **Preparació per a la conversió a digital dels materials impresos**

La tasca de preparar els documents per reformatar es fonamenta en unes rutines establertes i consolidades per preparar el canvi de format dels microfilms. Els passos que cal prendre inclouran la preparació de l'original per assegurar que cada volum és complet i que encara no ha estat preservat.

- *RLG Preservation Microfilming Handbook*, Nancy Elkington, editor
- *RLG Archives Microfilming Handbook*, Nancy Elkington, editor
- *RLG Guidelines for Microfilming to Support Digitization*, Robin Dale, editor (<http://www.rlg.org/en/pdfs/microsuppl.pdf>)
- *Preservation Microfilming: a guide for librarians & archivists*, 2nd Ed. Lisa Fox, editor
- DLF/OCLC Registry of Digital Masters – per determinar si un títol ha estat digitalitzat (<http://www.diglib.org/collections/reg/reg.htm>)

#### **Captura d'imatges i control de qualitat de les imatges digitals**

Les guies establertes s'adrecen a la captura d'imatges i el seu processament per assegurar que les imatges siguin d'alta qualitat i siguin reproduccions fidels de l'original. La captura d'imatges digitals s'aconseguirà utilitzant el hardware i software que permeti seguir les guies per capturar el text en blanc i negre, l'escala de grisos o imatges en color. Els objectes digitals hauran d'orientar-se adequadament, ordenar-se i anomenar per reflectir la presentació del volum original i inclouran informació tècnica i d'altra informació rellevant sobre la creació de la imatge.

Tots els fitxers digitals creats per a la preservació per reformatatge seguiran estrictes inspeccions de control de qualitat que inclouran comprovacions per a la integritat bibliogràfica (que sigui complet, llegible i la col·locació adequada de les imatges) i compliran amb les especificacions tècniques.

- *DLF Benchmark for Faithful Reproduction of Monographs and Serials* (<http://www.diglib.org/standards/bmarkfin.htm>)
- *RLG/DLFGuides to Quality in Visual Resource Imaging* (<http://www.rlg.org/visguides>)
- *California Digital Library Best Practices for Image Capture*

- (<http://www.cdlib.org/news/pdf/BestPracticeImageCapture.pdf>)
- *Moving Theory into Practice: Digital Imaging for Libraries and Archives*, d'Anne Kenney i Oya Rieger  
(<http://www.library.cornell.edu/preservation/tutorial/contents.html>)
- *Guidelines for Image Capture*, d'Steven Chapman  
(<http://www.rlg.org/preserv/joint/chapman.html>)
- *Preservation Management of Digital Materials: A Handbook*, de Neil Beagrie i Maggie Jones  
(<http://www.dpconline.org/graphics/handbook/>)
- *Recommendations for the Evaluation of Digital Images Produced from Photographic, Microphotographic, and Various Paper Formats*, de James Reilly i Franziska Frey  
(<http://lcweb2.loc.gov/ammem/lpireprt.pdf>)
- *Introduction to Imaging*, by Howard Besser. Getty Research Institute, 2003.
- *NARA Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images* (revisió del document de 1998 – en fase de producció)
- Government Printing Office (GPO) – Projecte en desenvolupament per crear unes especificacions per a la digitalització  
(<http://www.gpoaccess.gov/about/reports/preservation.pdf>)

### **Formats dels fitxers per als fitxers d'arxiu**

Els fitxers digitals creats com a fitxers d'arxiu per a la preservació es desaran seguint els formats estàndards i es generaran dades per descriure aquests formats

- *California Digital Library Digital Image Format Standards*  
(<http://www.cdlib.org/news/pdf/CDLImageStd-2001.pdf>)
- Global Digital Format Registry (prèviament Registry for Digital Format Representation Information) – proporciona suficient informació estandaritzada sobre formats  
([http://www.ifla.org/IV/ifla69/papers/128e-Abrams\\_Seaman.pdf](http://www.ifla.org/IV/ifla69/papers/128e-Abrams_Seaman.pdf))

### **Producció d'imatges, objectes estructurats i fitxers de textos accessibles des del web**

Els formats i les interfícies específiques es basen en la naturalesa del material original i l'ús que es prevegi per als fitxers digitals. Els fitxers d'imatge que representen pàgines de text es poden acompanyar per metadades estructurals per tal de proporcionar una navegació bàsica a nivell de pàgina i d'alguns components especials com ara títols de les pàgines, taules de contingut i índexs. Els fitxers d'imatge poden convertir-se a text, mitjançant programari de reconeixement òptic de caràcters (Optical Character Recognition, OCR), o bé teclejant-ho manualment, i una vegada convertit a text es pot codificar usant esquemes estàndards per a possibilitar la cerca a la web i obtenció de text. L'accés online continuat als objectes digitals se suportarà amb l'ús d'identificadors persistents.

- *California Digital Library Digital Image Format Standards*
- *NARA Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images*  
([http://www.archives.gov/research\\_room/arc/arc\\_info/guidelines\\_for\\_digitizing\\_archival\\_materials.html](http://www.archives.gov/research_room/arc/arc_info/guidelines_for_digitizing_archival_materials.html))
- Global Digital Format Registry – proporciona informació sobre formats estàndards
- PDF/A – un estàndard d'arxiu emergent per guardar informació electrònicament, amb l'objectiu de ser utilitzat tant pel sector públic com pel privat
- EAD (Encoded Archival Description) – descripció arxivística codificada utilitzant XML
- TEI (Text Encoding Initiative) – per codificar text.
- SGML (Standard Generalized Mark-up Language) – per codificar text.
- XML (Extensible Mark-up Language) – per codificar i gestionar contingut i la seva visualització.
- METS (Metadata Transmission & Encoding Standard) – un esquema estandaritzat per codificar metadades descriptives, internes i estructurals usant XML.
- Persistent Identifier Schemes – per assegurar que els materials digitals es poden gestionar i localitzar en un futur [per exemple, URLs persistents (PURL) i URNs o Digital Object Identifier (DOI)]

## Metadades

Els estàndards basats en metadades proporcionen informació descriptiva per la recuperació, estructurar la informació sobre l'objecte digital, aportar informació contextual sobre la creació de l'objecte i informació administrativa per facilitar la gestió al llarg del temps.

- Useu els camps (530, 856) MARC per localitzar fitxers accessibles i relacionar-los amb d'altres versions i col·leccions
- Dublin Core - per a desenvolupar i promoure estàndards de metadades interoperables en línia que donin suport a un ampli ventall de propòsits i models de negoci
- OAI (Open Archives Initiative) – per desenvolupar i promoure estàndards interoperables que tinguin per objectiu facilitar la disseminació eficient del contingut
- MODS (Metadata Object Description Schema) – un esquema XML per un conjunt d'elements bibliogràfics que podrien ser utilitzats per una gran varietat de finalitats i particularment per als recursos electrònics
- PREMIS (Preservation Metadata Implementation Strategies) – un esforç realitzat conjuntament per l'IRL i OCLC per desenvolupar un conjunt bàsic de metadades i treballar en els aspectes pràctics de la implementació de metadades per a sistemes de preservació digital
- NISO (National Information Standards Organization) Technical Metadata for Digital Still Images (Z39.87-2002)
- METS (Metadata Transmission & Encoding Standard) – un esquema estandaritzat per codificar metadades utilitzant XML per a descriure, administrar i estructurar
- JHOVE (JSTOR/Harvard Object Validation Environment) – una eina per identificar, validar i caracteritzar objectes digitals
- National Library of Medicine: *Report of the Working Group on Permanence of NLM Electronic Publications* – una valoració per a la permanència dels recursos

## Preservació digital

Els fitxers creats amb estàndards amb les metadades apropiades cal que siguin gestionats en un entorn que n'asseguri la permanència i accessibilitat a llarg termini. És necessària una gestió activa dels fitxers digitals per fer front a l'obsolescència dels mitjans magnètics i òptics, els canvis ràpids de hardware així com de les configuracions de software.

Les estratègies inclourien:

- Refrescament (“refreshing”) (moure els fitxers d'un suport a un altre periòdicament sense alterar el seu format o contingut)
- Comprovacions puntals per avaluar la integritat de l'objecte digital (autenticitat i integritat), utilitzat per exemple una *checksum value*
- Redundància (mantenir moltes còpies de fitxers digitals i comparar-les unes contra les altres per assegurar que no hi ha pèrdua de dades o no estan corruptes
  - LOCKSS – “Lots of Copies Keeps Stuff Safe” (un programa per construir eines que donen suport a les biblioteques a crear, preservar i arxivar col·leccions electròniques locals)
- Migració (transformació periòdica de fitxers cap a nous formats digitals per assegurar la continuïtat de la compatibilitat entre formats de fitxers i aplicacions)
- Emulació (habilitar sistemes obsolets per a què funcionin en un futur amb sistemes desconeguts actualment, fent possible la recuperació, visualització i l'ús dels documents digitals amb el seu software original.)
  - CAMiLEON – un projecte conjunt de les universitats de Michigan i Leeds per desenvolupar i avaluar un ventall d'estratègies tècniques de preservació a llarg termini de materials digitals.

#### **APÈNDIX 4: ESTAT ACTUAL I COMPROMISOS DE PRESERVACIÓ DELS RECURSOS ELECTRÒNICS A LLARG TERMINI**

L'esforç de desenvolupar estàndards, guies i millors pràctiques per utilitzar la conversió digital com una opció de reformatatge està tenint lloc en un context molt més ampli. Aquest fet proporciona seguretat addicional en el treball que s'està portant a terme a l'hora de crear arxius digitals. Hi ha hagut un gran compromís i finançament en aquesta àrea, incloent:

- National Digital Information Infrastructure and Preservation Program (NDIIPP)
- National Science Foundation (NSF) Digital Government Research Program
- National Archives Electronic Records Archive (ERA)
- Government Printing Office Legacy Digitization Project (GPO) – Un projecte per a la reconversió digital (“scanning”) de la col·lecció del Federal Depository Library Program
- Digital Preservation Coalition (DPC) al Regne Unit
- JISC Digital Curation Centre –Joint Information Systems Committee de Gran Bretanya dóna suport a diverses iniciatives de preservació digital com ara els UK Data Archives
- PADI - The National Library of Australia's Preserving Access to Digital Information Initiative
- Elsevier Science Policy for Archiving Electronic Journals with Koninklijke Bibliotheek (KB), la Biblioteca Nacional dels Països Baixos
- Desenvolupament de repositoris digitals
  - CDL - The California Digital Library
  - DNEP – Depot van Nederlandse Elektronische Publicaties projecte per establir una primera versió d'una biblioteca electrònica de dipòsit.
  - D-Space – Un sistema de biblioteca digital per a capturar, emmagatzemar, indexar, preservar i redistribuir els resultats de recerca dels professors en formats digitals
  - Eprints – Eines per a la gestió d'arxius en línia
  - FEDORA – Un repositori d'objectes digitals que pot ser utilitzat per donar suport a repositoris institucionals, biblioteques digitals i per a la preservació digital.
  - JSTOR's Electronic-Archiving Initiative – Desenvolupament de la infraestructura organitzacional i tecnològica per assegurar la preservació a llarg termini i l'accés als recursos electrònics acadèmics
  - LOCKSS – (Lots of Copies Keeps Stuff Safe) – Un programa per construir eines que donin suport a les biblioteques per crear, preservar i arxivar les col·leccions electròniques locals
  - OCLC's Digital Archive – un arxiu basat en estàndards

## **APÈNDIX 5: WEBSITES RELLEVANTS**

ANSI/NISO Z39.48-1992 (R1997) Permanence of Paper for Publications and Documents in Libraries and Archives ([http://www.techstreet.com/cgi-bin/detail?product\\_id=36497](http://www.techstreet.com/cgi-bin/detail?product_id=36497))

ANSI/NISO/LBI Z39.78-2000 Library Binding ([http://www.techstreet.com/cgi-bin/detail?product\\_id=229053](http://www.techstreet.com/cgi-bin/detail?product_id=229053))

ARL, Association of Research Libraries (<http://www.arl.org/preserv/>)

CAMiLEON, Creative Archiving at Michigan and Leeds: Emulating the Old On the New (<http://www.si.umich.edu/CAMiLEON/>; <http://129.11.152.25/CAMiLEON/dh/ep5.html>)

CCSDS, Consultative Committee for Space Data Systems (<http://www.ccsds.org/>)

CDL, California Digital Library (<http://www.cdlib.org/>; <http://www.cdlib.org/inside/diglib/>)

CLIR, Council on Library and Information Resources (<http://www.clir.org/>)

Digital Evans Edition (Early American Imprints) ([http://www.readex.com/scholarl/eai\\_digi.html](http://www.readex.com/scholarl/eai_digi.html))

DLF, Digital Library Federation (<http://www.diglib.org/>; Benchmark for Faithful Digital Reproductions of Monographs and Serials <http://www.diglib.org/standards/bmarkfin.htm>)

DLF/OCLC Registry of Digital Masters (<http://www.diglib.org/collections/reg/reg.htm>)

DNEP, Deposit of Netherlands Electronic Publications ([http://www.kb.nl/kb/resources/frameset\\_kb.html?/kb/menu/ken-arch-en.html](http://www.kb.nl/kb/resources/frameset_kb.html?/kb/menu/ken-arch-en.html)) [web inactiva, suggerim la consulta del següent recurs alternatiu: <http://webdoc.gwdg.de/edoc/aw/liber/lq-2-01/09Sijtsma.pdf>]

DOI, Digital Object Identifier System (<http://www.doi.org/>)

DPC, Digital Preservation Coalition (<http://www.dpconline.org/>; <http://www.dpconline.org/graphics/intro/index.html>)

D-Space, (<http://www.dspace.org/>)

Dublin Core Metadata Initiative (<http://www.dublincore.org>)

EAD, Encoded Archival Description (<http://www.loc.gov/ead/>)

EEBO, Early English Books Online [suggerim l'enllaç a <http://eebo.chadwyck.com/marketing/about.htm>]

EPrints (<http://www.eprints.org/>)

Elsevier Science Policy for Archiving Electronic Journals (<http://www.elsevier.com/inca/publications/misc/ni2164.pdf>)

FEDORA, Flexible Extensible Digital Object and Repository Architecture (<http://www.fedora.info/>)

GDFR, Global Digital Format Registry ([http://www.ifla.org/IV/ifla69/papers/128e-Abrams\\_Seaman.pdf](http://www.ifla.org/IV/ifla69/papers/128e-Abrams_Seaman.pdf))

Getty Research Institute ([http://www.getty.edu/research/conducting\\_research/standards/introimages/homepage.html](http://www.getty.edu/research/conducting_research/standards/introimages/homepage.html)) [suggerim també [http://www.getty.edu/research/conducting\\_research/standards/introimages/Tbl.html](http://www.getty.edu/research/conducting_research/standards/introimages/Tbl.html)]

GPO Legacy Digitization Project (<http://www.gpoaccess.gov/about/reports/preservation.pdf>)

ISO 14721:2003 Space Data and Information Transfer Systems -- Open Archival Information System – Reference Model (<http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=24683&ICS1=49&ICS2=140&ICS3>)

JISC, Joint Information Systems Committee (<http://www.jisc.ac.uk/>)

JISC Digital Curation Centre ([http://www.jisc.ac.uk/digcentre\\_townmeeting.html](http://www.jisc.ac.uk/digcentre_townmeeting.html))

JHOVE, Format-Specific Digital Object Validation (<http://hul.harvard.edu/jhove/>)

JSTOR, the Scholarly Journal Archive (<http://www.jstor.org/>; <http://www.jstor.org/about/earchive.html>)

LOCKSS, "Lots of Copies Keep Stuff Safe" (<http://lockss.stanford.edu/>)

MARC, Machine-Readable Cataloging (<http://www.loc.gov/marc/>)

METS, Metadata Encoding & Transmission Standard (<http://www.loc.gov/standards/mets/>)

MODS, Metadata Object Description Schema (<http://www.loc.gov/standards/mods/>)

NARA, U.S. National Archives & Records Administration, Technical Guidelines (*in progress*) and ERA, Electronic Records Archive ([http://www.archives.gov/electronic\\_records\\_archives/index.html](http://www.archives.gov/electronic_records_archives/index.html))

NDIIPP, National Digital Information Infrastructure and Preservation Program (<http://www.digitalpreservation.gov/index.php?nav=1>)

NISO, National Information Standards Organization ([http://www.niso.org/standards/resources/Z39\\_87\\_trial\\_use.pdf](http://www.niso.org/standards/resources/Z39_87_trial_use.pdf))

NLM, National Library of Medicine, Permanence Rating (<http://www.nlm.nih.gov/pubs/reports/permanence.pdf>)

NSF, National Science Foundation's Digital Government Research Program (<http://www.digitalgovernment.org/>; <http://www.rlg.org/preserv/diginews/diginews6-5.html>; <http://www.iso.ch/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=24683&ICS1=49&ICS2=140&ICS3=>)

OAI, Open Archives Initiative (<http://www.openarchives.org/>)

OAIS, Open Archive Information System ([http://www.rlg.org/longterm/oais\\_schematics.html](http://www.rlg.org/longterm/oais_schematics.html))

OCLC Preservation & Digital Services; Digital Archive (<http://www.oclc.org/services/preservation/default.htm>; <http://www.oclc.org/digitalarchive/>)

PADI, Preservation Access to Digital Information (<http://www.nla.gov.au/padi/>)

PDF/A, PDF-Archive ([http://www.aiim.org/pdf\\_a/](http://www.aiim.org/pdf_a/))

PREMIS, PREservation Metadata: Implementation Strategies (<http://www.oclc.org/research/pmwg/>)

PURLS, Persistent Uniform Resource Locator (<http://purl.oclc.org/>)

RLG, Research Libraries Group (<http://www.rlg.org>; <http://www.rlg.org/longterm/repositories.pdf>; [http://www.oclc.org/research/projects/pmwg/pm\\_framework.pdf](http://www.oclc.org/research/projects/pmwg/pm_framework.pdf))

SGML -Standard Generalized Markup Language (<http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=16387&ICS1=35&ICS2=240&ICS3=30>)

TEI, Text Encoding Initiative (<http://www.tei-c.org/>)

UK Data Archive (<http://www.data-archive.ac.uk/>)

XML, Extensible Markup Language (<http://www.w3.org/TR/REC-xml/>)