

# La preservació digital des de la perspectiva dels processos bibliotecaris

Alice Keefer Riva

P07/09061/02159



# Índex

<b>Introducció</b> .....	5
<b>1. L'adquisició de recursos digitals a les biblioteques: primer pas cap a la preservació</b> .....	7
1.1. La presència dels recursos digitals a les biblioteques .....	7
1.1.1. Els recursos en suports físics .....	7
1.1.2. Els recursos accessibles en web .....	8
1.1.3. El mètode de creació .....	9
1.2. L'adquisició de recursos externs .....	11
1.3. La incorporació al fons de materials interns .....	16
1.4. El compromís assumit i la manera de comunicar-lo a l'usuari ....	17
<b>2. Les metadades</b> .....	19
2.1. Metadades descriptives i el Dublin Core .....	20
2.2. Metadades de preservació .....	22
2.3. Emmagatzemament de metadades .....	27
<b>3. La gestió de la preservació</b> .....	28
3.1. Recursos econòmics .....	29
3.2. Recursos humans .....	33
3.2.1. Noves línies de comunicació .....	33
3.2.2. El nou perfil professional .....	35
3.3. Polítiques institucionals .....	37
<b>Bibliografia</b> .....	43



## **Introducció**

A més del tractament tècnic comentat en el mòdul anterior, la preservació digital inclou el tractament documental dels recursos custodiats i la gestió dels processos i dels recursos institucionals per a dur-los a terme. És un tema que urgeix molt als responsables de les biblioteques i arxius, ja que, a diferència dels problemes purament tècnics, la solució dels quals es troba a les mans dels informàtics, ells mateixos hauran de definir, acordar i desenvolupar les pautes i els procediments per a dur a terme la preservació digital de manera sostenible.

En aquest mòdul es fa un repàs de com les biblioteques, sigui com a col·lectiu o a títol individual, adapten certes operacions per a facilitar la integració dels recursos digitals al fons documental i gestionar els recursos institucionals necessaris per a poder-ne garantir els resultats –és a dir, la recuperació i utilització de la informació digital preservada– en el futur.



## **1. L'adquisició de recursos digitals a les biblioteques: primer pas cap a la preservació**

Els recursos digitals que s'incorporen al fons de la biblioteca poden procedir, igual que els documents tradicionals, de fonts internes o externes i s'adquireixen de diverses maneres: compra, dipòsit, distribució gratuïta o creació pròpia. Aquesta secció se centra en

- l'impacte dels recursos digitals en els processos tradicionals d'adquisició, i
- la importància de la política i dels procediments d'adquisició com un primer pas en la preservació d'aquests recursos.

La majoria d'institucions no prenen la decisió de preservar indefinidament els documents tradicionals en el moment d'adquirir-los. Aquesta decisió es pot ajornar llargament; així, si en un moment hi ha necessitat d'intervenir per a conservar un document, és en aquell moment quan es pot prendre la decisió de dedicar, o no, els esforços necessaris per a la seva preservació. En canvi, la preservació digital exigeix una gestió proactiva i continuada i, si determinades decisions es prenen des d'un principi, faciliten el procés de preservació quan s'ha assumit aquesta responsabilitat per al futur.

### **1.1. La presència dels recursos digitals a les biblioteques**

#### **1.1.1. Els recursos en suports físics**

Els primers recursos electrònics, adquirits en suports físics, és a dir CD, CD-ROM, disquets, etc., van començar a formar part dels fons de les biblioteques a partir de la dècada dels vuitanta. Alguns es compraven i d'altres eren gratuïts, com, per exemple:

##### **Recursos comprats:**

- CD amb enregistraments de música;
- CD-ROM amb bases de dades, enciclopèdies, etc., i
- disquets amb aplicacions de programari o jocs.

##### **Recursos adquirits gratuïtament:**

- material institucional;
- material de l'administració pública;
- material promocional, i
- material d'acompanyament de documents comprats.

El volum d'aquest tipus de fons creixia gradualment en els primers anys sense provocar grans desajustos en els procediments establerts. A diferència dels recursos de la web, els que arribaven en suport físic s'ajustaven –més o menys– a les regles i pautes existents, cosa que facilitava processos com la seva descripció, el seu préstec i el seu emmagatzemament.

Precisament perquè van ser tractats com qualsevol altre document, la seva gestió no va ser l'adequada per a facilitar-ne la preservació; igual que els altres documents, es guardaven a les prestatgeries però sense que es fes un seguiment de la seva llegibilitat; com els altres també, es descrivien segons les regles vigents; però aquestes no exigien les dades tècniques que ara es consideren com a imprescindibles per a la seva futura accessibilitat.

L'enquesta duta a terme el 1998 per encàrrec del Research Libraries Group sobre les necessitats de les biblioteques quant a la preservació digital (Hedstrom i Montgomery, 1998), va detectar la situació d'aquest tipus de material a les biblioteques participants i també la seva vulnerabilitat: gairebé la meitat dels enquestats van reconèixer que no podien accedir a una part del seu propi fons electrònic a causa de problemes com ara la falta d'equips de lectura originals o la impossibilitat d'interpretar les dades.

### 1.1.2. Els recursos accessibles en web

A partir de mitjan anys noranta i de l'aparició del World Wide Web, els canvis produïts per aquesta nova manera de distribuir informació electrònicament eren prou profunds per a requerir modificacions importants en la manera de tractar aquests recursos a les biblioteques.

Entre els primers recursos digitals adquirits per les biblioteques per aquest nou mitjà de distribució –és a dir, sense un suport físic– hi va haver les revistes científiques. La seva presència va introduir alguns canvis que van impactar sobre les maneres habituals de treballar i també van amenaçar alguns conceptes que constituïen els fonaments de les pràctiques tradicionals. Dos dels canvis de més impacte tenen a veure amb l'adquisició d'aquests recursos:

- el nou material ja no entrava físicament a la biblioteca sinó que s'hi accedia per mitjà de servidors remots;
- el model tradicional de **compra** se substituïa per un altre, basat en l'**accés** (*ownership* contra *access*), mitjançant la contractació del recurs durant un període determinat, per a uns usos i uns usuaris prèviament definits.

Aquest nou model d'adquisició va xocar frontalment amb el tradicional, caracteritzat per:

- la transferència física dels documents a la biblioteca i
- la conversió dels documents adquirits en propietat de la biblioteca.



L'adquisició de recursos digitals mitjançant el seu accés via web ha crescut espectacularment en l'última dècada. Aquest creixement continua a l'alça, com demostren les estadístiques de Rebiun: tan sols entre els anys 2003 i 2004 el nombre de revistes electròniques subscrietes per les biblioteques universitàries espanyoles va créixer un 78%, i el nombre mitjà de subscripcions per institució va passar de 4.069 a 7.225 (Rebiun, 2005).

Com es demostrarà en els paràgrafs següents, alguns dels canvis introduïts en la passada dècada tenen implicacions profundes per a la responsabilitat de la biblioteca i la seva capacitat d'actuar i han amenaçat en gran manera el seu paper en la preservació d'aquests nous recursos.

### **1.1.3. El mètode de creació**

Abans de centrar l'atenció en aspectes de la selecció i adquisició dels recursos digitals, és útil distingir-los segons el mètode en què es creen –els creats (o nascuts) originalment en format digital (*born digital*) i els digitalitzats, els que resulten de la conversió d'un document tradicional. El mètode de creació dels recursos pot repercutir en la seva gestió i en les possibilitats de preservar-los, per motius com ara:

- el format del document i la capacitat tècnica de la institució de gestionar-lo;
- el control que la institució pot exercir sobre les condicions de creació del document (normes seguides, criteris aplicats);
- l'existència, o no, d'un document original en suport tradicional que repercuteixi en la prioritat de la preservació de la versió digital.

#### **Recursos creats originalment en format digital**

Els documents que "neixen" en format digital no tenen versió "original" en suport tradicional. Hi ha milers de formats –molt variats i poc normalitzats– utilitzats per a generar-los i, a més, aquests experimenten una evolució constant. Els recursos poden ser creats per

- persones mitjançant equips informàtics (documents ofimàtics, pàgines web, blocs, fotografies, etc.);
- entitats (per exemple, pàgines web, directoris, bases de dades), i

- equips científics (així, les imatges de l'espai captades per telescopis o del cos humà captades per màquines de ressonància magnètica).

La quantitat d'agents creadors és enorme i creix incontrolablement: es deu principalment als nous dispositius que diàriament surten a la llum i a l'abaratiment de la tecnologia en general. Tot això ha servit per a obrir la porta a nous creadors institucionals, particulars i tecnològics. Un exemple del nou tipus de recurs digital són els vídeos realitzats amb telèfons mòbils i distribuïts mitjançant serveis com YouTube; alguns ja es reconeixen com a obres artístiques i d'altres, com és el cas de la filmació clandestina del penjament de Saddam Hussein<sup>1</sup>, són documents de valor polític i històric.

En fi, un dels grans reptes que plantegen els documents digitals és com respondre al creixement imparable de nous formats i, alhora, com evitar que els antics quedin obsolets i il·legibles.

### Recursos creats mitjançant la digitalització

El procés de digitalització converteix un contingut originalment produït en paper o un altre suport analògic en format digital. En el cas de documents impresos o gràfics, aquest procés es duu a terme amb la captura de la imatge del document per escàner o per càmeres digitals. La imatge digitalitzada es guarda en un format gràfic de mapa de bits, com el TIFF o JPEG. Del nombre reduït de tècniques de reproducció en resulta un nombre reduït de formats, un factor positiu per a la preservació dels recursos produïts. A més, com que es tracta d'un procés ja ben conegut i experimentat, els formats i els processos estan consolidats i normalitzats.

A diferència dels recursos nascuts en format digital que manquen d'una versió analògica original, els digitalitzats es generen com a còpia d'un document ja existent que, en molts casos, serà preservat. En aquests casos, la preservació de la còpia digital pot no tenir la mateixa urgència que per a aquells documents digitals de versió única. Per exemple, la prioritat assignada per la biblioteca de la Universitat de Cornell a determinats tipus de recursos per a la seva preservació varia segons l'existència o no d'un original en format tradicional (Cornell University Library, 2005).

Prioritat 1: materials "*born digital*"

Prioritat 2: materials digitalitzats sense versió analògica

Prioritat 3: materials digitalitzats amb versió analògica

A més de reconèixer els dos blocs de documents digitals segons el seu mètode de creació, hi ha una altra característica relacionada amb l'origen del recurs que pot condicionar la voluntat o la capacitat legal institucional de dur a terme accions de preservació. El fet que els recursos siguin de producció interna o externa pot repercutir en factors com ara:

#### Informe del NARA

Un recent estudi britànic esmenta un informe del NARA que calcula en uns 16.000 els formats gestionats per aquesta agència responsable de la preservació dels arxius oficials nord-americans (Jorum, pàg. 21).

<sup>(1)</sup>Hussein, expresident de l'Iraq, va ser penjat el 30.12.06.  
[http://news.bbc.co.uk/hi/spanish/international/newsid\\_6218000/6218539.stm](http://news.bbc.co.uk/hi/spanish/international/newsid_6218000/6218539.stm)

#### Recursos digitals i art

Per exemple, un museu d'art contemporani a Baltimore va organitzar una exposició anomenada "Cell Phone: art and the mobile phone", del gener a l'abril del 2007.  
<http://www.contemporary.org/exhibitions.html>

- el grau de responsabilitat que assumeix la biblioteca respecte a la custòdia i la preservació;
- l'autorització concedida a la biblioteca pels titulars perquè s'apliquin al recurs les tècniques necessàries per a la seva preservació, i
- la facilitat d'aconseguir l'autorització esmentada.

#### Exemple de diferents tipus de recursos segons el seu origen

	Extern	Intern
Nascut en format digital	DVD, revista científica en línia	Pàgina web, pautes, memòries, registres MARC
Digitalitzat	Col·lecció de mapes antics en CD-ROM	Fotos antigues Diaris històrics

Els paràgrafs anteriors han intentat demostrar com les característiques dels recursos digitals influeixen sobre la seva gestió a curt o llarg termini. La secció següent se centra en el procés d'adquisició: concretament en la seva alteració a causa de la transició cap als recursos digitals i en la importància de l'adquisició com a primer pas cap a l'eventual preservació d'aquests recursos a llarg termini. També quedarà patent que segurament la preservació eventual d'aquests recursos no la podran fer individualment les institucions "subscriptores", que havien aconseguit preservar les versions impreses dels mateixos títols. Però encara no s'ha trobat una solució definitiva: el col·lectiu de les biblioteques més afectades continua bolcat en la cerca de la millor manera d'assegurar l'accés a recursos tan importants com les revistes científiques en versió digital, entre d'altres.

### 1.2. L'adquisició de recursos externs

L'adquisició d'un recurs digital no suposa necessàriament que s'hagi de preservar a llarg termini, de la mateixa manera que poques biblioteques pretenen preservar tots els documents físics que adquireixen. Mentre que la decisió definitiva de preservar un document tradicional es pot ajornar durant molts anys, en l'entorn digital l'actuació ràpida és molt important per a facilitar les operacions de preservació en el futur. És per això que determinades decisions preses en el moment de la selecció o incorporació d'un recurs digital poden repercutir sobre la capacitat de la biblioteca d'assegurar el seu accés a curt o mitjà termini.

La selecció de recursos digitals per a la seva incorporació a la biblioteca sol seguir els mateixos criteris aplicats per les biblioteques per a seleccionar qual-sevol document nou: principalment, la seva qualitat, la seva pertinència i la demanda que n'hi hagi.

Però, per molt alta que sigui la valoració de qualitat, pertinència i demanda d'un determinat recurs digital, de vegades les biblioteques no l'adquireixen si no poden assegurar-ne l'accés durant el període necessari o de la manera exigida per les necessitats dels seus usuaris. És per això que, als criteris tradicionals de selecció ja esmentats, ara se n'hi sumen altres de nous, relacionats amb aspectes tècnics, institucionals i legals que poden tenir un impacte sobre l'accés i ús del recurs, ja que, evidentment, la preservació a llarg termini comença amb el manteniment de l'accés al recurs a curt termini. Si la biblioteca és incapaç de garantir aquest accés a curt termini, tindrà poques possibilitats d'assegurar-ho –o de pressionar per a la seva preservació– a llarg termini.

Els nous criteris inclouen:

- **Qüestions tècniques:** cal considerar si els requisits informàtics necessaris per a llegir o executar el recurs –programari, sistema operatiu, dispositius, increments de memòria, etc.– estan dins de les possibilitats de la biblioteca o si la seva instal·lació i implementació resultarien massa complexes o cares.
- **Qüestions institucionals:** cal considerar quin seria l'impacte de la incorporació del recurs sobre els serveis de la biblioteca, més enllà dels seus requisits tècnics. Alguns productes arriben a superar els recursos econòmics i humans disponibles; per exemple, alguns requereixen que els bibliotecaris o els usuaris rebin una formació; d'altres, equips informàtics especials.
- **Qüestions legals:** cal considerar les condicions imposades per la llicència d'ús i si permeten que la biblioteca compleixi les seves obligacions respecte als seus propis usuaris o els compromisos que pugui tenir en l'àmbit interinstitucional.

Fins i tot en aquells casos en els quals es decideix no preservar un recurs digital a llarg termini, és important assegurar-ne l'accés mentre en duri la utilitat per als usuaris. Si no, la inversió de diners i esforç en la seva adquisició i processament haurà estat debades.

#### La selecció i la desselecció

Les biblioteques no pretenen preservar tot el que adquireixen. El concepte de gestió de la col·lecció inclou tant la selecció com la retirada o expurgació de material perquè el fons es mantingui actualitzat segons com evolucionin els interessos i les necessitats de la institució i els seus usuaris. Els principals criteris per a la retirada de fons existents són semblants als de la selecció prèvia a la seva adquisició: 1) la qualitat de l'obra, 2) la seva pertinència per a la institució i els usuaris i 3) el nivell de demanda o ús.

Un dels principals riscos per a l'accés continuat als recursos web és la seva inestabilitat: el UK Web Archiving Consortium suggereix, a la seva pàgina web, que la vida mitjana d'un lloc web és comparable a la d'una mosca –uns quaranta-quatre dies. Però a més d'aquesta volatilitat natural en l'entorn web, també hi ha un altre motiu d'inestabilitat dels recursos comercials: la naturalesa volàtil del mercat editorial. Per exemple, en el cas de les revistes científiques, l'editor pot decidir deixar de publicar una revista o, fins i tot, en pot desaparèixer l'editorial –perquè ha deixat d'existir efectivament o perquè ha estat adquirida per una altra.

A part de la volatilitat del mercat editorial, hi ha altres factors que poden tenir un impacte sobre la gestió dels recursos digitals contractats per les biblioteques, com s'exposa en els paràgrafs següents.

Prenent com a exemple el cas de les revistes científiques, si hi ha dues versions del mateix títol –impresa i digital–, la tendència per part de les biblioteques al llarg de l'última dècada ha estat mantenir les dues: d'aquesta manera es millorava el servei per a l'usuari mitjançant la versió digital i, alhora, s'assegurava el futur accés al contingut mitjançant la conservació de la versió impresa. Però en els últims anys són cada vegada més freqüents els casos en els quals la versió digital desplaça la impresa als fons bibliotecaris. Sovint la desaparició de la versió en paper està motivada pel desig de les biblioteques d'estalviar –pressupost, espai o temps dedicat a la gestió i la manipulació del paper–, i és per això que anul·len la subscripció al paper una vegada que l'ús de la versió digital estigui consolidat.

Però de vegades el desplaçament definitiu cap a la versió digital és incentivat per l'editorial en el seu afany d'eliminar la publicació impresa, mitjançant la pujada del preu de la versió en paper o directament eliminant-la. Ara ningú no dubta que desapareixerà el paper. En un informe de la British Library es va fer la previsió que l'any 2016 la meitat de totes les publicacions en sèrie britàniques sortiran només en versió electrònica (British Library, 2004). Es preveu, a més, que les revistes científiques seran les primeres a fer aquesta transició.

De fet, el procés ja ha començat: a continuació es reproduïx un missatge emès el 2005 per l'American Geophysical Union (AGU) en el qual va anunciar la seva política de pujada progressiva de preus de la versió impresa de les seves revistes fins al punt que no sigui "viable" mantenir l'edició en paper; és a dir, quan no quedin biblioteques disposades a pagar els alts preus.

"... Fins que no es cobreixin tots els costos d'impressió, els subscriptors institucionals veuran que els preus de la versió impresa pugen més ràpidament que la inflació... El Comitè de Publicacions i el personal d'AGU preveu que arribarà un moment quan ja no serà viable econòmicament oferir material imprès."

"AGU Rebalances Pricing for Print and Electronic Versions" (6 de setembre de 2005). El missatge s'ha guardat en l'arxiu de la llista de distribució Liblicense: <http://www.library.yale.edu/~license/ListArchives/0509/msg00030.html>

### Exemple

En els últims anys el mercat ha experimentat compres i vendes que han portat algunes de les editorials més emblemàtiques a canviar de mans; és el cas d'Academic Press, Kluwer Academic Publishers, Pergamon, Springer Verlag. El cas més recent és del febrer del 2007: l'editorial John Wiley va comprar Blackwells Publishing. <http://www.publishersweekly.com/article/CA6413776.html?display=breaking>

### Previsió de la British Library

Una altra previsió de la BL (British Library, 2005) sobre les publicacions monogràfiques científiques va afirmar que l'any 2020, el 40% es publicaran només en versió digital i només el 10% es publicaran únicament en paper. El 50% restant mantindrà la publicació en les dues versions encara que la tendència sigui cap a la desaparició del paper.

Aquesta transformació del mercat, que respon tant a pressions econòmiques com decisions empresarials, té un impacte directe sobre la futura preservació d'aquests recursos, ja que una vegada desaparegui la versió impresa, no quedaràn documents físics per a preservar als fons bibliotecaris. Les biblioteques que sempre havien assumit el compromís de custodiar les revistes impreses s'hauran de replantejar el seu propi paper, que en molts casos –si no en la majoria– serà renunciar a aquesta funció.

Les biblioteques, que reconeixen la desaparició del paper dels seus fons en un futur proper i la seva incapacitat de continuar assumint la responsabilitat de la preservació com en el passat, reclamen una solució col·lectiva que garanteixi l'accés fiable al contingut de les revistes científiques subscrietes. Aquesta futura solució continua esperant l'elaboració de models i mecanismes estables i rendibles, com els que s'exploren actualment en els projectes com E-Depot, CLOCKSS o Portico.

A més de la incertesa sobre les garanties de l'accés a llarg termini a les revistes científiques digitals, les biblioteques també s'enfronten a riscos molt més immediats: les qüestions tècniques, legals i institucionals que poden posar en entredit l'accessibilitat als recursos subscrits a curt i a mitjà termini com ja s'ha comentat.

Justament es veu en la contractació de nous recursos una oportunitat per a pressionar els editors que es comprometin a col·laborar en la cerca de solucions per a la preservació del recurs a mitjà i a llarg termini, en interès de la comunitat científica actual i futura. Fa temps que s'instava les biblioteques que abans de contractar una subscripció s'informessin sobre la previsió de l'editor de garantir un accés perpetu. El compromís de l'editor pot conèixer-se mitjançant les seves respostes a preguntes com les següents:

- Quin mecanisme s'ofereix per a garantir l'accés continuat al material prèviament subscrit en cas de cancel·lació de la subscripció o de cessació de la revista?
- Quines tècniques s'apliquen per a assegurar que les dades es mantinguin llegibles?
- Quins sistemes de seguretat i de reproducció dels fitxers existeixen?
- Quines garanties d'accés hi ha en el cas de desaparició de l'editor?
- Quins acords té l'editor amb institucions "fiables", perquè assumeixin la responsabilitat de la preservació permanent del recurs?

Com a exemple d'un acord entre una editorial i institucions "fiables", s'ofereix el text següent, en el qual una editorial exposa els passos que ha pres per a assegurar l'accés a la seva revista digital en el cas que no pogués oferir-lo:

"Tenim un acord amb la Biblioteca del Congrés per establir un dipòsit de tot ... el material. A més, ...des de maig de 2001 s'ha hostatjat una còpia *mirror* a la biblioteca de la Universitat de Comell..."

Com es pot observar, la seva política ofereix dues opcions d'accés: una d'immediata i una altra, a llarg termini. La còpia *mirror* (una reproducció exacta i disponible per a l'accés actual) sota la gestió de la Universitat de Cornell ofereix accés al contingut immediatament en cas de problemes puntuals: per exemple, per congestió a la xarxa o per problemes tècnics que deixen inaccessible el servidor de l'editor. En canvi, la còpia d'arxiu dipositada a la Library of Congress serveix de garantia a llarg termini. Aquesta còpia no està destinada a remeiar problemes menors, sinó que només es posaria a disposició dels subscriptors en casos extrems, com la pèrdua catastròfica de les altres còpies o la desaparició de l'editorial.

### **El dilema dels "rolling years"**

El model de subscripció basat en "rotació anual" (*rolling years*) redueix la capacitat de la biblioteca de garantir accés al material contractat, fins i tot a curt termini. Sota aquest model, la subscripció vigent inclou el material publicat durant l'any en curs i un nombre fix d'anys anteriors, per exemple, els quatre últims. És a dir, la biblioteca contracta l'accés a material publicat durant un total de cinc anys. Aquest sistema d'"anys rodadors" o "rotació anual" comporta anualment la pèrdua de l'accés a material de l'any més antic.

En molts casos els textos "perduts" passen a formar part d'un "arxiu històric" comercialitzat a part per l'editor. Amb aquest sistema, la biblioteca ha de tornar a comprar el contingut dels anys retirats per a continuar garantint l'accés al material prèviament subscrit, fins i tot a curt o mitjà termini. Si decideix no comprar l'arxiu històric, els seus usuaris no podran tornar a accedir a aquesta informació. Fins i tot les iniciatives que actualment cerquen una solució per a la preservació de revistes digitals –E-Depot, CLOCKSS, PORTICO– tampoc no permetrien a la biblioteca subscriptora l'accés als anys previs als últims cinc, ja que aquests dipòsits es comprometen a assegurar l'accés només a aquelles institucions amb dret d'accés mitjançant una llicència d'ús vigent.

Si, d'altra banda, la biblioteca accedeix a contractar l'arxiu històric per a garantir l'accés continuat a tot el material, el seu pressupost destinat a adquisicions noves disminueix, ja que una part haurà de dedicar-se a la compra dels mateixos textos adquirits prèviament.

La biblioteca de la Universitat de Yale va elaborar un document, *Expectations of Yale University regarding creating archives of and perpetual access to electronic resources*, que recull les seves expectatives; és a dir, les accions que espera que els proveïdors facin per a garantir l'accés perpetu als recursos digitals que la Universitat ha contractat:

- l'editor crearà un arxiu que es preservarà i es mantindrà sempre actualitzat;
- l'arxiu contindrà totes les publicacions contractades i també tot el material o programa informàtic necessari per a llegir-les;
- el proveïdor es comprometrà a no modificar la informació arxivada, cosa que assegurarà l'autenticitat i integritat del recurs arxivat.
- es faran les reproduccions i migracions periòdiques necessàries per a mantenir les dades llegibles i assegurar-ne la preservació.

En el paràgraf següent de les expectatives de Yale, es destaquen dos elements importants:

- l'ús de formats estàndard, no solament del text sinó d'altres elements incorporats al text, com poden ser les figures o les imatges;
- l'anticipació a possibles canvis i l'agilitat de respondre-hi.

"L'Arxiu de preservació es revisarà a intervals apropiats per actualitzar-lo o fer una migració a nous formats d'informació. El... material es trobarà en un estàndard actual predominant amb formats d'emmagatzematge actuals adequats i basats en els estàndards pel que fa a les xifres, imatges de pàgina i altres formes d'informació a mesura que es vagin desenvolupant. La informació es convertirà als nous formats abans que els formats actuals estiguin en perill de quedar obsolets o deixin de ser acceptats, o quan els nous formats ofereixin unes millores importants quant a les seves característiques sense cap pèrdua de contingut." (*Expectations if Yale...*)

El document de Yale serveix només de recomanació per als proveïdors, però altres experts, principalment els responsables de les grans biblioteques d'investigació, van més lluny, i arriben a proposar la cancel·lació de subscripcions a revistes com a represàlia per la falta de col·laboració de certs editors. En la seva declaració del 15 d'octubre de 2005, una dotzena d'aquests experts van instar les biblioteques a condicionar la contractació de subscripcions al compliment, per part dels editors, d'aquests dos compromisos:

- el lliurament dels arxius a un dipòsit digital fiable i
- la publicació de la seva política de preservació.

Aquesta declaració també reivindica la creació dels dipòsits esmentats, allunyats del control dels editors, ja que, segons aquests experts, serà l'única manera de garantir la gestió fiable, independent i estable d'aquests recursos. No obstant els projectes en marxa per a preservar les revistes científiques en format digital com Portico, CLOCKSS o E-Depot, els esforços que es fan encara estan lluny d'oferir una protecció global. La urgència explícita en la declaració demostra la preocupació de les biblioteques d'investigació per la situació de desemparament en la qual es troben aquestes publicacions (*Urgent action needed...*).

### **1.3. La incorporació al fons de materials interns**

Els recursos digitals interns que es troben sota la custòdia de la biblioteca inclouen, en primer lloc, els de la seva pròpia creació. També es consideren recursos interns aquells creats per membres de la mateixa organització que, per procediment, obligació o fortuïtament, s'incorporen al fons per a gestionar-los i –possiblement– preservar-los a llarg termini. Poden ser recursos tant "nascuts digitalment" com digitalitzats, en suport físic o emmagatzemat en un servidor local.

Almenys en teoria, la gestió dels recursos de procedència interna és més fàcil perquè hi ha més possibilitats d'influir sobre el cicle vital del recurs abans de la seva incorporació a la biblioteca: per exemple, es poden recomanar als autors les condicions de creació del recurs, seguint les normes establertes, cosa que difícilment pot ocórrer amb recursos de fonts externes.

Evidentment és molt més fàcil que la biblioteca exerceixi un control directe i des de l'inici sobre els recursos que ella mateixa crea, tant si s'elaboren dins de la biblioteca com si s'encarreguen a un servei extern. La contractació d'un servei extern es produeix sovint per a la realització de projectes concrets, com



la digitalització d'un fons històric. En aquests casos la biblioteca pot definir les normes i pautes que cal seguir, i així controlar des de l'inici aquells aspectes que facilitaran la gestió del recurs a llarg termini (aplicacions, formats, incorporació de metadades, etc.).

La imposició de criteris per a la creació de documents d'una mateixa organització però fora del control directe de la biblioteca és més difícil i, en tot cas, el seu èxit requereix un suport institucional explícit. A la pràctica, en lloc de la imposició de criteris, és més freqüent la formulació de recomanacions dirigides als creadors dels recursos –docents, investigadors, etc.– sobre la manera de crear i guardar un document digital abans de depositar-lo a la biblioteca, per a afavorir l'èxit de la seva posterior gestió i preservació. Per exemple, es pot informar sobre:

- els programes d'aplicació que tenen més possibilitat de mantenir-se viants durant un període més llarg o de tenir un relleu per a la transferència de les dades;
- els formats normalitzats recomanats per a documents de determinats tipus, usuaris o finalitat;
- les metadades o altres elements descriptius perquè determinades dades –tant bibliogràfiques com tècniques– conegudes o fàcilment captades en el moment de la creació es puguin usar en una descripció bàsica inicial, i
- els suports, ja que la qualitat del suport i la seva cura mentre està sota la custòdia del creador condicionarà la seva vida útil i l'estabilitat de la informació gravada en ell.

El seguiment de les recomanacions per part dels autors a l'hora de crear els recursos i durant el període en què quedin sota el seu control servirà per a facilitar la gestió i, d'aquesta manera, minimitzar el cost de la seva eventual preservació.

Algunes entitats –especialment les empreses privades o agències de l'administració pública– arriben a especificar el programari i els formats que cal utilitzar, però aquest nivell de normalització resulta molt difícil en altres entorns heterogenis o descentralitzats, com són les universitats, on cada facultat –o, fins i tot, cada autor– pot justificar l'ús d'un determinat programa o format, com el ja esmentat format LaTeX usat pels matemàtics per a la representació de símbols matemàtics.

#### **1.4. El compromís assumit i la manera de comunicar-lo a l'usuari**

És previsible que moltes biblioteques no assumeixin la responsabilitat de preservar i garantir l'accés continuat als recursos digitals ja incorporats al seu fons. La decisió definitiva pot basar-se en criteris ja comentats, com ara:

- la utilitat del contingut per a la institució i els usuaris actuals i futurs;
- la disponibilitat d'aquesta informació en altres servidors;

- la titularitat del material (per exemple, si pertany a la mateixa institució o a alguna que hi estigui relacionada), i
- la responsabilitat de preservació assumida per una altra institució.

A més d'aquests criteris, la decisió definitiva pot resultar senzillament de la capacitat, o no, de la biblioteca de dur a terme les operacions necessàries. Però és important reconèixer que la desaparició de l'accés a un recurs prèviament disponible pot confondre i, fins i tot, frustrar l'usuari. És per això que es recomana un mecanisme –dins del catàleg o de les guies temàtiques– per a informar els usuaris del grau de compromís assumit sobre els recursos descrits i accessibles des de la biblioteca. D'aquesta manera estaran avisats prèviament sobre la relativa estabilitat, o no, dels recursos que els interessin.

La National Library of Medicine (NLM) va elaborar un sistema per a qualificar el grau de permanència (*permanence ratings*) dels recursos digitals produïts per l'NLM. El seu propòsit és avisar els usuaris de la permanència previsible d'aquests recursos i també de la probabilitat que el contingut es modifiqui. Els quatre nivells corresponen a (Hodge i Frangakis, 2004, pàg. 51):

- Permanent: contingut sense canvis (*permanent: unchanging content*). Per exemple, una imatge digitalitzada.
- Permanent: contingut estable (*permanent: stable content*). Per exemple, un registre de MEDLINE.
- Permanent: contingut dinàmic (*permanent: dynamic content*). Per exemple, la pàgina web de l'NLM.
- Sense permanència garantida (*permanence not guaranteed*). Per exemple, calendaris de congressos.

## 2. Les metadades

La descripció bibliogràfica és una de les tasques fonamentals de les biblioteques i respon a la seva responsabilitat de representar els seus documents mitjançant l'aplicació de regles de catalogació. La descripció resultant –plasmada en un registre catalogràfic– conté informació sobre l'obra per si mateixa i sobre l'exemplar que posseeix la biblioteca, tal com reflecteixen els objectius originals de la catalogació definits per Charles A. Cutter<sup>2</sup>, el 1876:

<sup>(2)</sup> Charles Ammi Cutter (1876). *Rules for a printed dictionary catalogue*. Washington, DC: Government Printing Office.

- ajudar un lector a trobar un llibre concret, si coneix l'autor, el títol o la matèria;
- indicar els títols que té una biblioteca al seu fons, escrits per un determinat autor, amb un determinat títol o sobre un determinat tema;
- ajudar el lector en la selecció d'una determinada obra: per exemple, segons l'edició específica o segons la manera de presentar la matèria (llibre infantil o llibre científic).

Amb la introducció a les biblioteques dels fons digitals va ser evident que la unió intrínseca entre el contingut intel·lectual i el document (o suport) físic va deixar d'existir, ja que el contingut podia separar-se fàcilment del vehicle de transmissió. Es complicava encara més la descripció del recurs digital perquè l'aspecte d'un determinat recurs podia visualitzar-se de manera diferent segons factors particulars, com el navegador utilitzat per l'usuari, la configuració de la seva pantalla, la data en la qual accedia al recurs, etc.

Aproximadament al mateix temps que el col·lectiu bibliotecari començava a adonar-se de la necessitat d'adaptar les regles de catalogació, també es va detectar la necessitat de disposar d'un altre tipus de descripció no tan vinculada a la gestió bibliogràfica i que incorporés aspectes no requerits per aquesta. Les metadades són el vehicle elegit per a aquesta funció: són dades estandarditzades que acompanyen el document per a identificar-lo i per a facilitar-ne la recuperació i ús.

Quina relació hi ha entre la descripció i la preservació? La finalitat de la preservació digital, en cas d'assegurar l'accés continuat al recurs quan el requereixin els usuaris al llarg dels anys, exigeix, evidentment, un sistema de descripció uniforme, consistent i durador.

## 2.1. Metadades descriptives i el Dublin Core

Si una biblioteca descobreix un disquet o una cinta sense cap referència, es podran identificar els documents que hi estan arxivats?, es podrà identificar la informació necessària per a interpretar-los (el programari amb què van ser elaborats, incloses les seves versions i els formats originals)? Sense dades –internes o externes– que els identifiquin, la resposta probablement sigui negativa. Per desgràcia, els documents no identificats en molts casos tampoc no són identificables. Com a conseqüència, es pot impossibilitar la seva preservació per desconeixement dels elements tècnics necessaris per a les operacions i qualsevol intent d'allargar la vida del suport serà debades.

Aquesta necessitat de disposar de dades específiques relatives a l'entorn informàtic original per a poder accedir als recursos digitals va conduir, entre altres factors, a la recerca d'un consens sobre uns elements comuns d'identificació i descripció dels recursos digitals. Aquesta recerca va adquirir més urgència a causa de factors com:

- el creixement desorbitat de la quantitat i varietat de recursos generats,
- l'heterogeneïtat dels agents creadors,
- la dificultat de localitzar-los mitjançant recerques estructurades,
- la facilitat d'alteració, que posa en dubte la seva integritat i autenticitat i
- la necessitat d'identificar els titulars dels drets.

La solució proposada està determinada per metadades, elements descriptius que serveixen per a identificar i recuperar els documents digitals. De fet, la paraula *metadades* significa senzillament 'dades sobre dades', un concepte de llarga tradició a les biblioteques. Així, les fitxes de cartolina d'antany conteni- en les metadades dels documents del fons de la biblioteca. S'ofereixen a continuació altres definicions més concretes:

- **NISO:** informació estructurada que descriu, explica, localitza o, de qual- sevol altra manera, facilita la recuperació, l'ús o gestió d'un recurs d'informació.
- **Dublin Core Metadata Initiative (DCMI):** en general, "dades sobre les dades"; funcionalment, "dades estructurades sobre les dades". Les metada- des inclouen tant les dades associades amb un sistema d'informació, com les associades a un objecte d'informació amb la finalitat de descriure, ad- ministrar, assenyalar els requisits legals, la funcionalitat tècnica, l'ús i la utilització i la preservació.

Possiblement l'esquema de metadades més conegut i implementat en l'àmbit de les biblioteques sigui el Dublin Core<sup>3</sup>. Aquesta iniciativa –nascuda d'una reunió d'experts celebrada el 1995 a la ciutat nord-americana de Dublin, Ohio– va disposar des del seu inici de la participació d'experts de diferents camps i sectors: biblioteques, arxius, museus, empreses informàtiques, editorials i fotògrafs.

El 1998 es va publicar el *Dublin Core Metadata Element Set*, que consisteix en quinze elements considerats bàsics per a la descripció de tot tipus de material digital. El 2001 va ser aprovat com a norma nord-americana ANSI-NISO Z39.85 i, el 2003, es va convertir en norma internacional, ISO 15836:2003. A continuació s'enumeren els quinze elements<sup>4</sup> relacionats amb el contingut intel·lectual, la propietat intel·lectual o la instanciació<sup>5</sup>.

- 1) Títol (contingut)
- 2) Autor o creador (propietat intel·lectual)
- 3) Claus [o descriptors] (contingut)
- 4) Descripció (contingut)
- 5) Editor (propietat intel·lectual)
- 6) Altres col·laboradors (propietat intel·lectual)
- 7) Data (instanciació)
- 8) Tipus de recurs (instanciació)
- 9) Format (instanciació)
- 10) Identificador del recurs (instanciació)
- 11) Font (contingut)
- 12) Llengua (contingut)
- 13) Relació (contingut)
- 14) Cobertura (contingut)
- 15) Drets (propietat intel·lectual)

L'objectiu principal de les metadades Dublin Core és facilitar la identificació i localització o "descobriment" de recursos en el web. A partir d'aquesta llista bàsica, cada sector pot adaptar els elements als seus propis usos especialitzats. Per exemple, les biblioteques poden "mapar" aquestes dades als camps corresponents d'un registre MARC (*machine readable cataloging*), on es completarien amb altres dades estipulades per les regles catalogràfiques. La transferència inversa també és possible, la qual cosa permet la generació de metadades a partir de les dades bibliogràfiques d'un registre MARC. La Library of Congress ofereix un *crosswalk*<sup>6</sup>, o 'pont', entre els elements de Dublin Core i les dades bibliogràfiques de MARC21, com mostren a continuació els tres exemples:

- Element DC "Tipus" ⇒ Etiqueta MARC21 655 #7\$a
- Element DC "Font" ⇒ Etiqueta MARC21 786 0#\$n
- Element DC "Format" ⇒ Etiqueta MARC21 856 ##\$q

<sup>(3)</sup>Dublin Core Metadata Initiative (DCMI) <http://es.dublincore.org/> Seminari "Aplicación en España del Modelo de metadatos Dublin Core". Madrid, 2003. <http://es.dublincore.org/es/eventos/dc-mi-es1/>

<sup>(4)</sup>La traducció dels elements és: *Metainformació - Dublin Core: Elements del conjunt de metadades de Dublin Core: descripció de referència*. [http://www.rediris.es/metadada/dublin\\_core\\_elements.es.html](http://www.rediris.es/metadada/dublin_core_elements.es.html)

<sup>(5)</sup>La definició del vocable anglès *instantiation*, oferta en el glossari ja esmentat de la DCMI, és: "Exemplificació, il·lustració, concreció per mitjà d'exemples. Una ocurrència o esdeveniment d'una cosa identificable; en el cas del Dublin Core, una ocurrència específica d'un recurs d'informació".

<sup>(6)</sup>*Dublin Core/MARC/Global][Information ]L[ocator ]S[ervice]Crosswalk*. <http://www.loc.gov/marc/dccross.html>

En una nota sobre l'evolució de les metadades Dublin Core, Eva Méndez (Méndez, 2006) suggereix que el gran èxit de Dublin Core es deu a diversos factors, entre ells:

- Defineix una semàntica precisa, però és sintàctic independent.
- S'ha adoptat internacionalment [...]
- [...] no depèn de cap domini informatiu, però s'adapta a les diferents comunitats d'informació web i es porta bé amb altres esquemes de propòsit específic [...] servint de pedra Rosetta [...]
- [...] és simple, extensible i interoperable.

La necessitat d'una gestió activa dels recursos digitals al llarg de la seva vida, a diferència del fons documental tradicional, va conduir a l'expansió de les categories de metadades per a incloure informació més enllà de les dades descriptives bàsiques. Per exemple, la norma METS (*metadata encoding and transmission standards*) reconeix els tres grups funcionals següents:

- **descriptius:** dades sobre el contingut intel·lectual, que ajuden a identificar i localitzar un recurs;
- **administratius:** dades necessàries per a la gestió d'un recurs –el seu manteniment, emmagatzemament i lliurament–;
- **estructurals:** dades que es refereixen a l'estructura del recurs i als elements que el formen.

#### **La norma METS**

METS ofereix un marc per al manteniment i intercanvi de les metadades. Referent a això és semblant al format MARC, que no dicta el contingut però estableix la manera d'organitzar i codificar les dades. <http://www.loc.gov/standards/mets/>

## **2.2. Metadades de preservació**

També s'ha materialitzat en els últims anys una altra classe de metadades de gran importància: les metadades de preservació. Alguns experts les han inclòs com un grup més al costat dels tres identificats en el paràgraf anterior o d'altres. Per exemple, les directrius de la UNESCO, publicades el 2003, identifiquen tres classes de metadades, una de les quals són les metadades de preservació:

- **metadades de descobriment de recursos:** "[Permeten] trobar-los, avaluar-los i mantenir-los accessibles i comprensibles; [...] van des d'una simple llista de noms de fitxers fins a extenses descripcions que contenen una abundant informació contextual."
- **metadades de gestió de recursos:** "[Permeten] administrar els fluxos de treball: [...] com es crea, transfereix i utilitza el material, sobre els drets i les assignacions d'autoritacions, i sobre altres processos de gestió."

- **metadades de preservació:** "[Permeten] que els programes de preservació puguin comprendre com representar els materials digitals quan el seu accés sigui necessari; [...] Aquestes metadades són absolutament fonamentals per a qualsevol programa de preservació [...]." (*Directrices...*)

Però actualment se solen reconèixer com una categoria a part, per sobre de les altres, ja que les dades contingudes en els tres grups bàsics –descripció, estructura i administració– són necessàries per a dur a terme els processos de la preservació a llarg termini.

Lavoie i Gartner ofereixen la definició següent de metadades de preservació, en la qual es destaca aquest marc més ampli (Lavoie i Gartner, 2005):

"Són metadades de la preservació les metadades descriptives, estructurals i administratives que ajuden a la preservació a llarg termini dels recursos digitals [i la documenten]."

D'aquesta manera, les metadades de preservació permeten que un recurs digital "s'autodocumenti" durant qualsevol fase del seu cicle vital, sigui quin sigui l'entorn legal, tècnic o institucional. Les dades que constitueixen aquesta "autodocumentació" es refereixen a aspectes com ara:

- **les característiques i els entorns tècnics:** els requisits tecnològics necessaris per a la seva lectura i ús;
- **la procedència del recurs:** l'historial del recurs amb la relació dels responsables de la seva custòdia des de la seva creació;
- **les activitats de preservació:** les accions dutes a terme per a preservar el document, incloent-hi qualsevol alteració que hagin provocat i també les motivades per la preservació, ja que poden afectar la decisió sobre quines propietats s'han de preservar;
- **la validació de l'autenticitat:** la informació suficient per a verificar que el recurs no hagi estat alterat sense autorització; si s'ha alterat, cal deixar constància de les modificacions fetes;
- **els drets d'autor:** les condicions sota les quals es permet –o no– dur a terme accions necessàries per a allargar la vida dels recursos; per exemple,
  - la reproducció del recurs mitjançant el procés de "refrescament" en el qual les dades es transfereixen a nous suports,
  - la modificació del recurs mitjançant la migració de les dades a nous formats.

Les directrius sobre la preservació digital publicades per la UNESCO fan referència a les activitats facilitades per les metadades de la preservació (*Directri- ces...*):

- definir el material sota la responsabilitat d'un programa de preservació;
- indicar què es necessita per a mantenir i protegir les dades;
- indicar què es necessita per a representar l'objecte (o els seus elements essencials definits) a l'usuari quan aquest el sol·liciti, encara que les tecnologies d'emmagatzemament i accés hagin canviat;
- registrar la història de l'objecte i les conseqüències del que vagi succeint;
- documentar la identitat i la integritat de l'objecte per a certificar-ne l'autenticitat, i
- permetre a l'usuari i al programa de preservació comprendre el context de l'objecte emmagatzemat i en ús.

### Esquemes de metadades

Durant aquesta dècada hi ha hagut diverses propostes de metadades de preservació, amb cert grau de diferències –d'elements, de terminologia, de propòsit– entre elles. En el quadre següent, Candás compara l'aplicació dels principals elements de tres d'aquests esquemes: PREMIS, CEDARS i la Library of Congress (Candás, 2006).

PREMIS	CEDARS	LoC
'format' (identificació del format informàtic del DLO)		'associated_file_type' (format de l'arxiu informàtic)
'creatingApplication' (aplicació amb què es va crear el DLO)		
'preservationLevel' (identificació de les tasques de preservació que han d'aplicar-se)	preservation description information' (la informació necessària per a una preservació correcta)	'reformatting_guidelines' (polítiques i instruccions per al canvi de format)
'objectIdentifier' (la designació que identifica unívocament el DLO)	'reference information' (els identificadors del recurs)	'archive_ID' (identificador de la localització del DLO)
'relationship' (relacions existents entre el DLO i un altre o altres DLO)	'context information' (relació del DLO amb el seu context)	
'fixity' (informació sobre si el DLO ha estat alterat de manera no autoritzada o no registrada)	'provenance information' (la història del DLO)	'archive_history' (història de fets referents a l'arxiu del DLO)
'objectCharacteristics' (característiques tècniques del DLO)	'custody history' (dades de les organitzacions que han estat dipositàries)	'reformatted_original_information' (informació sobre les condicions del document original)
	'management history' (modificacions realitzades en el DLO per a la seva preservació)	'archive_date_time' (data i hora de creació, còpia de seguretat, etc.)
	'rights management' (informació sobre els drets de propietat intel·lectual relatius al DLO)	



Actualment, la iniciativa de metadades de preservació de més envergadura és el *Data dictionary for preservation metadata: final report of the PREMIS Working Group*, elaborat pel grup de treball PREMIS (**pres**ervation **meta**data **imple**mentation strategies) i publicat el 2005. Aquest document de finalitat doble –informe final i diccionari de dades– defineix un conjunt essencial de metadades de preservació. PREMIS defineix aquests conceptes de la manera següent:

- **Metadades de preservació:** "La informació que utilitza un dipòsit per a donar suport al procés de la preservació digital";
- **Conjunt essencial:** "Aquells aspectes que la majoria de dipòsits probablement necessitaran saber per a donar suport a la preservació digital".

A més d'identificar les principals metadades aplicables a la preservació, el document també ofereix una anàlisi de les opcions existents per a la seva codificació, l'emmagatzemament, la gestió i l'intercanvi.

El *Data dictionary* va ser l'últim d'una sèrie d'estudis sobre diferents aspectes de la preservació digital a llarg termini i coordinats per l'OCLC i l'RLG. La finalitat d'aquests estudis va ser fomentar que s'assolís un consens internacional sobre l'ús de les metadades per a facilitar-ne la preservació, basat en els coneixements i experiències ja existents. Per exemple, el 2004 es va publicar l'informe *Implementing preservation repositories for digital materials*, que va donar a conèixer els resultats d'una enquesta internacional feta a mig centenar d'organitzacions a tretze països de sectors variats: biblioteques, arxius, museus, agències de l'administració pública, centres d'investigació científica i empreses comercials. Entre els participants europeus, hi figuraven les biblioteques nacionals d'Alemanya, Àustria, Finlàndia, França, la Gran Bretanya, els Països Baixos, Portugal i Suïssa. Els comentaris d'aquestes organitzacions van contribuir a les conclusions i recomanacions del *Data dictionary*, publicat l'any següent. Algunes de les conclusions de l'estudi van ser:

- Més de la meitat de les organitzacions ja assignaven metadades descriptives, administratives, tècniques i sobre drets, procedència i estructura.
- El col·lectiu de patrimoni cultural –biblioteques, arxius i museus– tenia poca experiència amb la preservació digital.
- Encara no hi havia experiència suficient per a saber si les metadades que es recollien en el moment que es feia l'enquesta eren adequades per a la preservació.

Segons els resultats d'aquest estudi, l'origen de les metadades creades en les organitzacions entrevistades era el següent:

- autors (75%) –es limitaven majoritàriament a introduir les metadades descriptives;
- extracció automàtica dels recursos dipositats (75%);
- assignats per personal de la biblioteca, arxiu, etc. (63%).

#### Estudis anteriors al *Data dictionary*

Els estudis anteriors són:

- *Preservation metadata for digital objects: a review of the state of the art* (2001)
- *A metadata framework to support the preservation of digital objects* (2002)
- *Implementing preservation repositories for digital materials: current practice and emerging trends in the cultural heritage community* (2004)

### **Autors i metadades**

Hi ha moltes dades que l'autor possiblement desconeix, per exemple sobre aspectes tècnics o legals. Per exemple, dos estudis internacionals han detectat fins i tot una falta de coneixement de l'estat dels drets d'autor dels seus propis treballs: *Scholarly communication in the digital environment: what do authors want?* <http://www.ucl.ac.uk/ciber/ciber-pa-report.pdf>; *New journal publishing models: an international survey of senior researchers*. <http://www.slais.ucl.ac.uk/papers/dni-20050925.pdf>

Si les metadades les crea l'autor o es creen mitjançant extracció automàtica, se'n recomana una revisió en el moment de la incorporació del recurs al dipòsit per a assegurar que siguin correctes i completes.

El treball manual de creació o revisió de metadades és molt laboriós i per això molts experts opinen que l'automatització d'aquests processos és l'única manera de millorar-ne l'eficàcia i reduir els costos de la gestió. Com a exemple, durant les ja esmentades Jornades sobre la Preservació de Patrimoni Digital (Madrid, 2006), diversos ponents d'institucions europees pioneres en el tema van opinar que l'avenç de la preservació digital quedarà obstaculitzat fins que no es desenvolupin sistemes per a la creació automàtica de les metadades.

Precisament per això alguns centres exigeixen la presència de metadades com a condició prèvia a l'admissió del recurs a la seva custòdia, per a reduir els seus costos. N'és un exemple la recomanació del *The handbook* de la Digital Preservation Coalition d'esbrinar si el recurs té metadades o un altre sistema d'identificació abans d'assumir la responsabilitat de la seva preservació (Jones i Beagrie, 2001).

En resum, sense les metadades de preservació el dipòsit no podrà garantir que un recurs sigui recuperable i utilitzable en el futur, ja que:

- sense dades descriptives, no es podrà localitzar;
- sense dades tècniques, els documents localitzats no es podran visualitzar o utilitzar;
- sense dades administratives, no es tindrà la garantia de la integritat o l'autenticitat del document recuperat;
- sense dades legals, els responsables de la custòdia no sabran què podran fer amb el document per a prolongar-ne la vida o quines condicions existeixen per al seu ús per part de la comunitat d'usuaris.

L'alt percentatge de metadades generades pels autors (el 75%) detectat per l'estudi PREMIS, subratlla la importància de fomentar la col·laboració amb aquest col·lectiu, sobretot en el cas de recursos de creació interna en la institució. L'oferta de serveis d'ajuda i suport per part de la biblioteca o del dipòsit podria comportar una reducció de costos i d'esforç en el futur.

### 2.3. Emmagatzemament de metadades

Les metadades de preservació poden acompanyar el document –és a dir, quedar encapsulades o incrustades dins d'ell– o mantenir-se en un fitxer a part. Hi ha arguments a favor i en contra de cada opció:

- **La incrustació de les metadades:** les metadades incrustades al recurs serveixen perquè s'autoidentifiqui o s'autodocumenti. Qualsevol que trobi aquest recurs en el futur disposarà de la informació necessària per a identificar-lo i interpretar-lo. La tecnologia ja facilita l'encapsulació de les metadades: hi ha formats que ja les incorporen i hi ha dispositius de conversió, com escàners i càmeres digitals, que generen les metadades tècniques en el moment de fer la conversió i les deixen incorporades directament al recurs digital creat. Per contra, alguns experts desaconsellen la intervenció repetitiva en el recurs que representava l'actualització contínua de les metadades per a reflectir les operacions a què s'hagi sotmès el recurs i els canvis que hi hagi hagut.
- **Les metadades independents:** l'emmagatzemament de les metadades en un fitxer a part pot facilitar la localització, manipulació i ús dels recursos descrits, sobretot si el mateix recurs no és accessible a la xarxa pública (per exemple, si està arxivat en una intranet). Però l'inconvenient és que els fitxers que contenen les metadades també són recursos digitals que requereixen la mateixa gestió que qualsevol altre per a garantir-ne l'accessibilitat en el futur.

La tendència actual dels dipòsits sembla que sigui curar-se en salut: l'estudi ja citat dut a terme pel grup de treball de PREMIS (*Implementing...*) mostra que la majoria dels centres entrevistats preferien replicar aquesta informació, i afirmaven el seu interès a mantenir una còpia de les metadades en una base de dades, i una altra, dins del recurs corresponent. D'aquesta manera cada recurs portarà dins seu la informació necessària per a la seva identificació i utilització i, alhora, la base de dades agilitarà la localització dels recursos per navegadors i motors de cerca, sense que hagin d'entrar directament en el recurs mateix. Però aquesta solució dual, evidentment molt pràctica, corre el risc de la generació de versions diferents de metadades per a descriure el mateix recurs. Per a evitar-ho, és necessari un rigorós control de les actualitzacions per part dels gestors de la informació.

Potser aquest risc ha contribuït a la decisió presa per la majoria de les biblioteques nacionals més implicades en la preservació digital de només emmagatzemar les metadades de manera independent, i no incrustar-les dins del recurs digital. Algunes les mantenen fora del dipòsit (Alemanya, Austràlia, el Canadà, Nova Zelanda i els Països Baixos), mentre d'altres les guarden separadament però dins del dipòsit (Àustria, Dinamarca) (Verheul, 2006, pàg. 48).

### 3. La gestió de la preservació

Un dels principals canvis detectats en els últims anys és una menor atenció a la problemàtica tecnològica de la preservació. Si no es pot parlar encara d'una superació dels reptes tecnològics, almenys han estat transferits a un segon pla. Un exemple d'aquest canvi de perspectiva queda plasmat en els comentaris recollits en l'informe publicat per la Comunitat Europea i esmentats en el mòdul 2 (Geser i Pereira, 2004).

Cada vegada més, l'interès col·lectiu se centra en els aspectes institucionals, com la viabilitat de la institució per a dur a terme les operacions necessàries i la seva capacitat d'assegurar l'accés als recursos al llarg del temps. Aquesta transició és deguda, en part, als avenços fets fins ara; ja funcionen sistemes de preservació en nombroses institucions que aporten les seves experiències i coneixements perquè els altres puguin aprendre'n. Alhora, ja requereixen respostes als nous interrogants que els sorgeixen. És per això que la seva mirada –i la d'altres implicats– es fixa en gran manera en qüestions com la disponibilitat i l'administració dels recursos –humans, materials i econòmics– i les polítiques i els procediments que regeixen i guien les operacions ja en marxa.

Per molt progrés que hi hagi hagut, els experts afirmen que falta molt camí per arribar a tenir un entorn amb les característiques imprescindibles: "estructurat, extensible [*scalable*] i segur" (Verheul, 2006). Alguns fins i tot adverteixen que no es pot parlar de "millors pràctiques" –un terme molt utilitzat– per a referir-se a moltes de les operacions actualment en marxa, ja que encara es desconeix l'adequació, o no, dels seus resultats (OCLC i RLG PREMIS Working Group, 2004; Verheul, 2006). Molts també recomanen cautela a l'hora de treure conclusions concretes de les experiències d'altres, ja que les que hi ha hagut han resultat de la cerca per part de determinades institucions de solucions per als seus propis fons digitals i amb els seus propis recursos econòmics i humans: una combinació de factors difícilment comparable a les condicions existents en altres centres.

Aquesta secció exposa algunes qüestions centrals en qualsevol debat sobre la gestió de la preservació en l'àmbit institucional:

- els recursos econòmics i humans, i
- la política i planificació.

### 3.1. Recursos econòmics

Cada activitat relacionada amb la preservació genera despeses, directes o indirectes, i són gairebé inexistents les institucions que han trobat fonts estables de subvenció per a cobrir-les en el futur. El cost de les operacions varia molt segons factors com ara:

- el tipus i el format dels materials gestionats,
- l'origen dels documents,
- l'ús previst,
- la política del centre, i
- les tècniques aplicades.

Alhora hi haurà altres factors externs, encara desconeguts, que tindran un paper decisiu a l'hora d'alterar la base sobre la qual es calculen els costos actualment: en definitiva, els avenços tecnològics dels pròxims anys obriran pas a alternatives, difícils d'imaginar avui, que permetran fer front a una part de les despeses actuals.

Però mentrestant, s'ha observat com a tema recurrent en la documentació especialitzada del sector la preocupació pels costos: com conèixer-los?, com controlar-los?, com finançar-los?

Tal com s'ha afirmat en altres seccions d'aquest material docent, actualment l'anàlisi de la problemàtica de la preservació se centra en la sostenibilitat econòmica de les operacions necessàries per a garantir l'accés continuat als recursos preservats: és a dir, com es podrà assegurar la continuïtat dels sistemes, ja en funcionament, en nombroses institucions?

Per començar, la sostenibilitat comença amb la minimització dels costos actuals: és reconegut àmpliament que l'abaratiment dels costos actuals anirà estretament relacionat amb la capacitat de les institucions de fer el següent:

- **preveure els canvis externs** que tinguin un impacte sobre les operacions: per exemple, el desplaçament al mercat d'un format present en molts recursos gestionats pel dipòsit. Aquesta previsió permet una programació de respostes àgils, adequades i assequibles;
- **automatitzar els processos**, i així reduir les intervencions humanes que representen un alt percentatge dels costos actuals: els processos de la incorporació dels recursos al dipòsit, la generació de metadades i la migració de formats.

Els estudis sobre els costos de la preservació, des dels primers intents d'abordar aquest tema a la darrerria dels anys noranta, han passat per diferents fases:

- la identificació de factors que influeixen sobre el cost de les operacions i la manera de minimitzar-ne l'efecte;
- l'elaboració de mecanismes per a calcular els costos d'operacions específiques, i
- el desenvolupament de models econòmics per a fer front als costos de manera sostenible.

A continuació es resumeixen alguns dels principals estudis sobre els aspectes econòmics de la preservació:

**CEDARS.** El projecte CEDARS, dut a terme per les universitats de Leeds, Oxford i Cambridge i amb finançament de JISC, va oferir una de les primeres aproximacions als elements de cost relacionats amb la preservació. Aquests elements, introduïts en el mòdul 2, són els tipus i formats de les dades, els drets sobre el material i el control exercit sobre ell a partir de la seva creació (per exemple, controls sobre la qualitat, sobre l'ús de formats estàndards o sobre les condicions d'emmagatzemament dels suports) (Granger, Russell i Weinberger, 2000).

La situació d'un recurs d'un o altre extrem del rang –des de menys a més complexitat– tindrà un impacte sobre el cost de gairebé totes les operacions relacionades amb la seva custòdia. En molts casos l'increment del cost per al material més problemàtic es deu a les intervencions individualitzades que requereix: com afirmen tots els experts, l'automatització dels processos és la principal manera de controlar els costos.

Segons l'anàlisi feta per CEDARS, les operacions que resultarien més costoses a causa de la complexitat dels recursos inclouen:

- la selecció,
- la negociació dels drets,
- les tècniques de preservació aplicada i d'accés,
- el procés de validació dels recursos incorporats, i
- la generació i actualització de les metadades.

Dos estudis posteriors al de CEDARS van revelar més dades sobre els costos eventuais de la preservació:

***Feasibility and requirements study on preservation of e-prints* (2003).** És un estudi encarregat per JISC l'objectiu del qual va ser analitzar els requeriments per a la preservació dels materials científics (*e-prints*) que començaven a dipositar-se en els flamants *repositories* institucionals de les universitats britàniques. Aquest document centra l'atenció, principalment, en dos aspectes que abaratirien els costos de la gestió d'aquests recursos: l'ús de formats estàndards per a la seva creació i la generació automàtica de les seves metadades.

**Erpa guidance cost orientation tool (2003).** Aquest document, fruit del projecte europeu ERPANET, té una visió molt àmplia que inclou elements de cost directe i indirecte. A més de la consideració de característiques tècniques dels recursos (format, etc.) i de les tècniques necessàries per a preservar-ne l'accessibilitat (migració, etc.), també s'hi inclouen costos més indirectes com els relacionats amb el personal, la formació i la comunicació, i la col·laboració amb altres entitats implicades en activitats relacionades amb la preservació.

Els estudis següents hi han contribuït mitjançant l'aportació de dades basades en experiències reals:

**LOCKSS.** Aquest sistema, dissenyat i coordinat per Stanford University Libraries, ofereix una mecanisme per a la preservació de les revistes científiques en format digital basat en una arquitectura descentralitzada i en la comprovació automàtica i contínua de la integritat de les dades mitjançant un programa d'igual a igual (*peer-to-peer*). L'esquema dels elements de cost es reparteix entre les tres grans àrees funcionals definides per l'OAIS: incorporació, preservació i accés/difusió. Cal destacar la importància donada en els càlculs als costos relacionats amb les metadades: no solament es refereixen a la seva generació, com en els dos estudis anteriors, sinó que també se'ls reconeixen tots els costos relacionats amb qualsevol altre objecte digital. És a dir, les metadades conservades en fitxers externs als objectes digitals que descriuen hauran de ser gestionades, actualitzades i preservades a llarg termini, igual que aquests (Rosenthal i altres, 2005).

**Amigos Library Services.** Basat en l'experiència d'aquest servei bibliogràfic regional (creat originalment com una xarxa regional dins d'OCLC), es va elaborar un model que permet comparar els costos relacionats amb la preservació: tant els operacionals directes i indirectes com els costos capitals. Un factor afegit a aquest esquema són les estadístiques relacionades amb la comunitat d'usuaris: el nombre i el nivell d'ús. A més, es distingeix entre dues finalitats, si el cost es relaciona amb:

- la preservació dels recursos o
- la difusió i l'ús dels recursos.

Aquest conjunt de factors permet el càlcul de costos relacionats amb la preservació i l'ús, segons les fórmules següents:

- $\text{cost per recurs preservat} = \frac{\text{costos capitals i operacionals de la preservació}}{\text{nombre de recursos gestionats}}$
- $\text{cost per ús} = \frac{(\text{costos capitals i operacionals de la preservació}) + (\text{costos capitals i operacionals de l'ús})}{\text{nombre d'usos efectuats}}$

Segons els responsables del servei, aquest esquema els permet anticipar les possibles despeses i, com a conseqüència, millorar el procés de programació operacional (Sanett, 2003).

Les iniciatives següents versen sobre el valor dels objectes d'informació digital i dels beneficis derivats de la seva preservació. Comparteixen un interès a trobar models econòmics sostenibles per a assegurar que les operacions de preservació es mantinguin viables en el futur:

***Incentives to preserve digital materials*** (Lavoie, 2003): en aquest estudi finançat per l'OCLC, l'autor postula que els objectius de la preservació han d'estar d'acord amb els interessos dels administradors que financen les operacions. És per això que cal buscar fórmules econòmiques que contrastin els costos amb els beneficis de la preservació. Les característiques següents, comunes a molts projectes de preservació, són identificades per aquest autor com a **obstacles** per a l'aprovació del finançament, per exemple:

- En molts casos, els titulars dels recursos no percebran els beneficis de la preservació: com a exemple, se sol parlar que la preservació va en benefici de les futures generacions, la qual cosa no aporta cap benefici als que han de subvencionar les operacions actuals.
- Si els recursos digitals no són documents únics, és més difícil que una institució s'ofereixi per a assumir els costos de la preservació, sabent que d'altres també es beneficiaran dels seus esforços. Per això són importants els projectes en què participin tots els implicats, com LOCKSS o PORTICO, en els quals tots els socis perceben un benefici i comparteixen els costos. A més, en aquests dos casos, no podran beneficiar-se dels materials altres entitats que no hagin contribuït a la seva preservació, econòmicament o materialment.
- L'heterogeneïtat de la demanda de serveis –segons el tipus d'institució i d'objecte digital, els usuaris previstos i el pressupost disponible– fa difícil que sorgeixin solucions generalitzades d'utilitat per a totes les necessitats. Això afecta el cost dels serveis desenvolupats: com més específics siguin els requeriments, més cara serà la solució, ja que no es poden aprofitar les economies d'escala.

**ESPIDA.** El segon recurs surt del projecte ESPIDA, dut a terme per la Universitat de Glasgow i finançat per JISC. Neix com a resposta a la dificultat dels administradors de les organitzacions –els que aproven els pressupostos i defineixen la política institucional– per a reconèixer el valor dels objectes o béns intangibles, com és la informació. El projecte destaca el fet que el valor varia al llarg del temps i, per això, és necessari considerar tant els beneficis actuals com els previsibles en el futur que puguin obtenir-se de la preservació.



Els possibles beneficis que la preservació de recursos digitals representa per a la institució són els següents:

- el valor per a aquelles persones o organitzacions externes a la institució que en reben un servei (clients, usuaris, col·laboradors, socis, etc.);
- el valor que representa per al personal i per a la mateixa organització quant a avenços d'innovació d'R+D;
- el valor que aporta al funcionament institucional;
- el valor econòmic que aporta a la institució.

El resultat del projecte ESPIDA ha estat l'elaboració d'un model que permet plasmar tant els costos com els beneficis de la preservació. D'aquesta manera es perfila com un instrument per a millorar la comunicació entre els responsables de la preservació i els administradors institucionals; alhora, serveix per a facilitar la presa de decisions relacionades amb la planificació de les operacions necessàries (Currall i McKinney, 2006).

### **3.2. Recursos humans**

Els aspectes relacionats amb els recursos humans són variats. En aquesta secció se centren principalment en les noves línies de comunicació –dins i fora de la institució– i les noves competències dels responsables de la preservació dels recursos digitals.

#### **3.2.1. Noves línies de comunicació**

Els gestors dels dipòsits hauran d'establir o reforçar les relacions amb els grups següents:

- els creadors dels recursos custodiats;
- els administradors institucionals, i
- altres especialistes els coneixements dels quals poden contribuir al bon funcionament de les operacions, com poden ser els informàtics o els assessors legals.

Tant els creadors com els especialistes podrien treballar dins de la mateixa institució, o bé ser externs a ella. En qualsevol dels casos, aquests contactes són imprescindibles perquè el dipòsit pugui dur a terme les seves funcions eficaçment.

#### **Contactes amb els creadors dels recursos**

Abans de comentar el tipus de comunicació entre els gestors dels recursos digitals i els seus creadors, és útil repassar un dels conceptes fonamentals del model OAIS: el de la comunitat designada d'usuaris. Segons l'OAIS, els recursos es

preserven perquè siguin utilitzables per un determinat perfil d'usuaris, la qual cosa permet que el dipòsit adopti les pràctiques més adequades per a servir els membres d'aquest grup, el perfil del qual condicionarà decisions sobre:

- els formats gestionats,
- els elements essencials preservats,
- les tècniques aplicades,
- el tipus d'accés ofert i
- els serveis elaborats per a facilitar la comprensió i usabilitat dels recursos preservats.

Qui millor pot assessorar sobre els usuaris i els usos potencials dels recursos són els mateixos creadors, ja que aquests, teòricament, dirigeixen les seves obres a determinats grups d'usuaris. En establir i mantenir els contactes amb els creadors, els gestors poden obtenir la informació necessària com, per exemple, la finalitat dels recursos, els usos previstos i les principals propietats la preservació de les quals s'haurà d'assegurar.

Si la relació amb els creadors ho permet, els contactes també poden aprofitar-se perquè els gestors del dipòsit els aconsellin sobre la millor manera de crear recursos nous, per a facilitar-ne la futura preservació: per exemple, l'ús de formats estàndard, la creació de metadades o el maneig físic dels suports.

### **Contactes amb els administradors institucionals**

El compliment de les responsabilitats del dipòsit requereix el suport institucional. La millor manera per a manifestar aquest suport és mitjançant l'elaboració d'una política formal que es comenta en la secció següent. Sense aquest suport el dipòsit no serà considerat "fiable" per altres del sector, ja que no hi ha una garantia explícita de la voluntat institucional de desenvolupar i mantenir la infraestructura i les operacions necessàries per a assegurar la preservació i l'accessibilitat dels recursos a llarg termini. Per a forjar una implicació institucional més gran, és necessària una comunicació fluïda entre els gestors del dipòsit i els administradors responsables de l'aprovació i assignació de partides pressupostàries i de la definició i implantació de polítiques institucionals.

Els gestors dels dipòsits tenen la responsabilitat d'informar els administradors de les necessitats per a mantenir les operacions i alhora exposar objectivament els costos de les accions proposades i els possibles riscos de no dur-les a terme. Les propostes de Lavoie (Lavoie, 2003) i del projecte ESPIDA (Currall i McKinney, 2006), exposades en la secció anterior, ajudaran els gestors a demostrar tant el valor dels recursos digitals custodiats com els beneficis –directes i indirectes– de la seva preservació.

Si els creadors dels recursos són de la mateixa institució, seria convenient que els responsables de la preservació digital impliquessin els administradors institucionals en l'elaboració de pautes i procediments per a la creació dels recursos d'acord amb les normes existents, la qual cosa n'afavoriria la futura preservació.

### Contactes amb altres especialistes

Hi ha altres especialistes la col·laboració dels quals és imprescindible per al bon funcionament del dipòsit. És per això que els responsables de la preservació hauran d'establir i mantenir contactes actius amb ells, principalment amb els serveis informàtics i legals:

- Els serveis informàtics oferiran assessorament i faran les operacions necessàries per a la gestió continuada dels recursos; per exemple, la selecció, instal·lació i manteniment dels dispositius i el manteniment dels recursos custodiats: còpies de seguretat, controls d'accés, comprovació de la integritat de les dades, actualització i migració dels formats, renovació dels suports, etc.
- Els serveis legals assessoraran sobre els drets d'autor i la confidencialitat de les dades, entre altres qüestions, i també tramitaran els permisos i acords necessaris.

### Contactes amb agents externs

Són recomanables –i de vegades imprescindibles– els contactes amb altres agents, externs a la institució, directament o indirectament implicats en la preservació dels recursos digitals:

- **Editorials:** per a negociar els drets de gestionar –reproduir, modificar i donar accés a les publicacions dipositades.
- **Empreses de serveis:** per a subcontractar algunes operacions.
- **Institucions afins** (biblioteques, arxius, consorcis): per a compartir experiències i coneixements, desenvolupar aplicacions o pautes, fer operacions en col·laboració, etc.

### 3.2.2. El nou perfil professional

La preservació dels recursos digitals requerirà personal amb una nova combinació de coneixements i habilitats; també serà necessari que es mostrin flexibles per a poder adaptar-se a la nova realitat i als canvis constants que es presentaran en l'entorn. Aquests canvis afectaran la tecnologia, els estàndards, les pràctiques, la legislació vigent; totes aquelles àrees que poden impactar directament sobre les operacions del dipòsit. Si alguns han detectat la manca d'aquests requisits en el personal de moltes institucions, d'altres tenen una perspectiva més optimista i afirmen que els coneixements i aptituds dels pro-

professionals de la preservació de materials tradicionals són transferibles al nou entorn, com els identificats per les directrius de la UNESCO i esmentats en el mòdul 1 (*Directrices...*).

Segurament la majoria dels professionals actualment dedicats a la gestió i/o preservació dels recursos digitals veuran en els canvis l'oportunitat per a actualitzar els seus coneixements i situar-se bé davant dels nous reptes. Però d'altres se sentiran amenaçats i es resistiran als canvis.

Bárbara Muñoz de Solano y Palacios esmenta algunes causes de resistència als canvis per part del personal (Muñoz, 2006):

- La vivència d'experiències negatives anteriors.
- La gestió inadequada del procés en la seva part inicial.
- La ignorància sobre el mateix canvi.
- L'escassa informació relativa als beneficis derivats de la implantació del nou sistema.
- [El desconeixement de la implantació del nou sistema: en què consisteix?, per què s'efectua?, quin pot ser l'impacte directe sobre el professional?]
- La percepció de futures tasques poc assumibles o massa complicades.

La resistència serà menor si el personal està informat dels canvis i se sent preparat per a assumir-los. És per això que la institució pot facilitar la transició mitjançant la promoció de la formació, l'elaboració de noves descripcions de treball i l'assignació de responsabilitats clares.

Actualment es detecta una falta generalitzada de cursos formals sobre la gestió de la preservació digital (Harvey, 2005). En aquells casos en els quals les institucions no ofereixen la formació necessària, molts professionals prenen la iniciativa i renoven els seus coneixements mitjançant cursets aïllats –cercats i finançats per ells mateixos.

Els anuncis de feina ofereixen una idea de les noves competències. Al començament del 2007, una universitat nord-americana va informar d'una plaça disponible per al nou encarregat del seu dipòsit digital i va destacar les responsabilitats següents:

- el disseny i implantació d'un dipòsit digital;
- el disseny dels processos i fluxos de treball;
- la gestió de l'ús i adherència dels estàndards;
- el disseny i creació de metadades d'acord amb els estàndards;
- la migració de recursos digitals a formats més nous, i
- el seguiment de noves tecnologies, estàndards i aplicacions.

### 3.3. Polítiques institucionals

Un gran nombre –si no la majoria– de les iniciatives per a combatre la pèrdua d'informació digital es duen a terme com a accions aïllades: es financen amb un pressupost extraordinari i es fan amb personal assignat (o contractat) temporalment, seguint un calendari fix: per exemple, per al rescat de dades d'un format obsolet, la migració de dades d'un programari a un altre, la renovació de suports físics no llegibles. Malgrat la seva utilitat, aquestes iniciatives inconnexes no constitueixen un programa viable de preservació a causa de la falta de sistematització de les activitats i de continuïtat dels recursos necessaris per a sostenir-les.

Més que unes accions aïllades per contrarestar un perill específic, la preservació digital ha de ser abordada com un procés –programat, gestionat, continuat i modificable– perquè s'adapti als canvis interns i externs que es produeixin. Aquest procés ha de ser integrat, com qualsevol altre, a les operacions institucionals quant a pressupost, personal i planificació, i ser formalitzat mitjançant una política explícita. La política expressarà la voluntat institucional de mantenir la infraestructura i els recursos econòmics i humans necessaris per a garantir la preservació i l'accessibilitat del material custodiat; també definirà i assignarà les responsabilitats.

Les directrius de la UNESCO identifiquen així les característiques institucionals necessàries:

- **Responsabilitat:** un compromís fonamental amb la preservació dels objectes digitals considerats.
- **Viabilitat de l'organització,** incloent-hi la perspectiva d'un mandat permanent; un estatut jurídic que la defineixi com una organització que exercirà un paper permanent de preservació; i una capacitat demostrada per a reunir recursos, infraestructura i equips de treball que puguin fer-se càrrec de la complexitat de la preservació digital.
- **Durabilitat financera:** una perspectiva versemblant que l'organització serà capaç de continuar subministrant els recursos necessaris a llarg termini, amb un model de gestió sostenible per a dur a terme el seu encàrrec de preservació digital.
- **Adequació tecnològica i de procediment:** la utilització de sistemes i procediments apropiats per a dur a terme el que requereixen la gestió i la preservació dels recursos digitals.
- **Seguretat del sistema de qualitat molt alta.**

#### La política de la Universitat de Cornell

La política de la Universitat de Cornell, elaborada per la biblioteca universitària, exposa els objectius, responsabilitats, activitats i prioritats del programa de preservació digital.  
<http://commondepository.library.cornell.edu/cul-dp-framework.pdf>

- Transparència dels procediments, amb una clara atribució de les responsabilitats i mecanismes per a avaluar els resultats obtinguts i informar-ne.

Però el que s'ha exposat en els dos paràgrafs anteriors representa la teoria, tal com es planteja avui. A la pràctica, són pocs els casos de processos consolidats amb polítiques institucionals formalitzades. Per exemple, els resultats de l'estudi ja esmentat sobre la preservació digital en institucions públiques britàniques (Waller i Sharpe, 2006) mostren una realitat marcada per:

- la falta de compromís en el finançament per a la preservació a llarg termini,
- la falta d'assignació de responsables i de definició de responsabilitats i
- la falta d'una política institucional.

L'estat de la preservació digital queda reflectit en les paraules de diverses persones entrevistades en aquest estudi:

- "Molt poques organitzacions tenen una estratègia de preservació digital implementada."
- "Un obstacle per a la majoria de les organitzacions és la manca de línies clares de responsabilitat."
- "La preservació digital és una disciplina nova que encara està en via de desenvolupament i encara hi ha poques persones que en tinguin els coneixements necessaris."

Potser el més sorprenent d'aquestes afirmacions seria que descriu la situació actual a la Gran Bretanya, un país pioner a reconèixer els reptes de la preservació dels recursos digitals des de fa més d'una dècada, i que continua essent un líder en l'anàlisi dels problemes i la recerca de solucions.

Un altre cas que presenta la incongruència de la situació actual és el de la Universitat de Cornell. La biblioteca d'aquesta universitat és una de les primeres a involucrar-se activament en el tema de la preservació digital i en la recerca de solucions. Va prendre part en alguns dels estudis clau –per exemple el de *Trusted digital repositories*– i va ser de les primeres universitats nord-americanes que va crear un dipòsit per a la preservació dels recursos digitals generats per docents i investigadors. A més, ha pres la iniciativa per a compartir els seus propis coneixements i experiències mitjançant diversos canals:

- l'organització de tallers periòdicament, als quals assisteixen persones d'institucions nord-americanes i estrangeres, i
- l'elaboració de dos programes d'aprenentatge: un sobre la gestió de la digitalització de recursos analògics i l'altre sobre la gestió de la preservació digital. Estan disponibles gratuïtament en línia i han estat reconeguts degudament per la seva aportació al sector bibliotecari.

## Programes d'aprenentatge

Sobre la gestió de la digitalització de recursos analògics:

"Moving theory into practice: digital imaging tutorial". <http://www.library.cornell.edu/preservation/tutorial/contents.html>

Sobre la gestió de la preservació digital:

"Digital preservation management: implementing short-term strategies for long-term problems". Aquest programa d'aprenentatge va guanyar el premi de millor publicació atorgat per la Society of American Archivists. <http://www.library.cornell.edu/iris/tutorial/dpm/foundation/oais/index.html>

Malgrat aquesta gran experiència, els mateixos responsables reconeixen la lentitud del procés, fins i tot quan hi ha –com en el seu cas– el reconeixement dels problemes, la voluntat de resoldre'ls i els coneixements teòrics i tècnics per a fer-ho. Per exemple, una relació de les fases per les quals passen els centres cap a un servei ple de preservació digital –basades en l'experiència de Cornell– va ser elaborada per dos destacats membres de l'equip directiu de la biblioteca: Anne Kenney i Nancy McGovern. Les cinc fases van des del descobriment inicial del problema fins al funcionament d'un servei consolidat que assegurí la preservació i accessibilitat permanent dels recursos. Per a il·lustrar aquests passos, les autores van presentar-los mitjançant una original comparació amb les fases recomanades als alcohòlics per a combatre la seva malaltia (Kenney i McGovern, 2003; resumit a Keefer, 2005):

- 1) **Reconèixer el problema:** el reconeixement sovint ocorre quan hi ha hagut un problema específic. En el cas de Cornell, va ser un projecte per a recuperar dades d'un "jukebox" òptic el cost final del qual es va calcular en uns 150.000 dòlars.
- 2) **Actuar:** etapa caracteritzada per l'inici de projectes de preservació digital. Aquests solen ser aïllats i en resposta a amenaces específiques, la resolució de les quals té un vessant tecnològic.
- 3) **Consolidar:** reconèixer l'existència de normes i pautes externes i adonar-se que els projectes aïllats i independents no són eficaços ni rendibles. Els projectes es converteixen en activitats continuades i vinculades als objectius i les polítiques institucionals.
- 4) **Institucionalitzar:** s'apliquen esforços per a coordinar els projectes existents, aplicar les normes externes i actuar de manera coordinada en l'àmbit institucional. La institució es compromet a una dedicació pressupostària fixa a llarg termini.
- 5) **Externalitzar:** es busca la col·laboració d'altres institucions mitjançant relacions formals i informals.

En quin punt es troba la Universitat de Cornell? Doncs, durant una conferència pronunciada el juny del 2005, Anne Kenney va afirmar que la seva institució es trobava en aquell moment tan sols en el nivell 3, encara que havia emprès algunes iniciatives que l'apropava al nivell 4, gairebé deu anys després d'iniciar-se en el tema!

L'últim dels passos recomanats per Cornell –externalitzar– és molt important, ja que cap biblioteca no té suficients mitjans per a fer front sola a tots els reptes de la preservació digital. És per això que és important buscar col·laboracions amb altres institucions i col·lectius. De fet, la voluntat de col·laborar està generalitzada entre el col·lectiu bibliotecari i des de fa anys les biblioteques universitàries i les nacionals treballen conjuntament al costat d'altres del seu entorn.

Laura Campbell, subdirectora de la Library of Congress i responsable del Programa nacional d'infraestructura nacional per a la preservació digital (NDIIPP) als Estats Units, va fer una sèrie de conferències durant la seva visita a Espanya el març del 2006. És una gran promotora de la col·laboració entre institucions, sectors i disciplines, ja que la considera com la millor manera d'identificar i resoldre els reptes de la preservació digital, tant tecnològics com legals i econòmics. També va reconèixer que durant la primera etapa de finançament de projectes dins del NDIIPP<sup>7</sup>, la LC va considerar més important la creació de sinergies entre institucions que la recopilació de contingut digital (Keefer, 2006).

<sup>(7)</sup>L'NDIIPP té un pressupost d'uns 175 milions de dòlars durant els deu anys que durarà el projecte.

Els resultats d'una enquesta duta a terme per la Comissió Europea el 2005, "i2010: bibliotecas digitales", revelen el desig, per part de molts implicats, que es coordinin activitats des de les altes esferes, per exemple: el desenvolupament i promulgació d'estàndards i bones pràctiques, el repartiment de responsabilitats entre les grans organitzacions o la creació de centres d'excel·lència especialitzats en determinats aspectes de la preservació o en la preservació de determinats tipus de recursos (European Commission, 2006).

Iniciatives a escala europea prometen facilitar aquest procés, entre les quals hi ha els projectes següents, que van començar el juny de 2006:

- **PLANETS** (<http://www.planets-project.eu/>). Aquest projecte europeu té la participació de biblioteques i arxius nacionals: la British Library (coordinadora), les biblioteques nacionals d'Holanda, Àustria i Dinamarca i els arxius nacionals de la Gran Bretanya, Holanda i Suïssa. Els seus objectius inclouen investigar, desenvolupar i promocionar solucions pràctiques per a la preservació digital. Espera tenir un impacte singular sobre el futur desenvolupament i gestió de dipòsits de preservació digital a Europa.
- **CASPAR** (<http://www.casparpreserves.eu/>). Es dedica a estudiar l'aplicació del protocol OAIS per a la preservació d'informació digital científica, artística i cultural. Té disset membres de diferents sectors.



Però l'eventual ajuda provinent d'entitats nacionals o supranacionals, per ben-vinguda que sigui, no deixa exempta de responsabilitat cada institució implicada en la creació o gestió de recursos digitals. Tard o d'hora, totes hauran d'enfrontar-se als reptes, dins dels límits de les seves possibilitats: informant-se de la problemàtica, analitzant els riscos que corre el seu propi fons i identificant els passos –per modestos que siguin– que puguin avançar.



## Bibliografia

**Beagrie, Neil** (2003). *National digital preservation initiatives: an overview of the developments in Australia, France, the Netherlands, and the United Kingdom and of related international activity*. <<http://www.clir.org/PUBS/reports/pub116/pub116.pdf>>

**British Library** (2005, 29 de juny). "British Library predicts «switch to digital by 2020»; British Library press release". <<http://www.bl.uk/news/2005/pressrelease20050629.html>>

**British Library** (2004). *Publishing output to 2020*. [Data de consulta: 29.1.2007]. <<http://www.bl.uk/about/articles/pdf/epsreport.pdf>>

**Candás Romero, Jorge** (2006, març-abril). "El papel de los metadatos en la preservación digital". *El Profesional de la Información* (vol. 15: 2, pàg. 126-136).

*Cedars guide to digital collection management* (2002).  
<[www.leeds.ac.uk/cedars/guideto/collmanagement/](http://www.leeds.ac.uk/cedars/guideto/collmanagement/)>

**Cornell University Library** (2005). *Cornell university library digital preservation policy framework*.

**Currall, James; McKinney, Peter** (2006). "Investing in value: a perspective on digital preservation". *D-Lib Magazine* (núm. 12, pàg. 4). <<http://www.dlib.org/dlib/april06/mckinney/04mckinney.html>>

*Data dictionary for preservation metadata: final report of the PREMIS Working Group* (2005) OCLC; RLG.  
<<http://www.oclc.org/research/projects/pmwg/premis-final.pdf>>

**Day, Michael** (2003). "Preservation metadata". A: *Metadata applications and management* (pàg. 253-273). Londres: Facet.

**Day, Michael** (2003). "Preservation metadata initiatives: practicality, sustainability and interoperability". A: *Metadata in preservation: selected papers from an ERPANET seminar*. <<http://www.ukoln.ac.uk/preservation/publications/erpanet-marburg/day-paper.pdf>>

*Digital lifecycles and file types: final report* (2006).  
<[http://rightsandrewards.lboro.ac.uk/files/resourcesmodule/@random43c8ae8b0d0ad/1148047621\\_DigitalLifecyclesV2.pdf](http://rightsandrewards.lboro.ac.uk/files/resourcesmodule/@random43c8ae8b0d0ad/1148047621_DigitalLifecyclesV2.pdf)>

*Directrices para la preservación del patrimonio digital* (2003). París: UNESCO.  
<<http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>>

*Erpa guidance: cost orientation tool* (2003). Erpanet.  
<<http://www.erpanet.org/guidance/docs/ERPANETCostingTool.pdf>>

**European Commission** (2006). "Results online consultation «i2010: digital libraries»". <[http://europa.eu.int/information\\_society/activities/digital\\_libraries/doc/results\\_online\\_consultation/en.pdf](http://europa.eu.int/information_society/activities/digital_libraries/doc/results_online_consultation/en.pdf)>

*Expectations of Yale University regarding creating archives of and perpetual access to electronic resources* (2004). Yale University (Development Council). Rev.  
<<http://www.library.yale.edu/ecollections/yalearchiving.pdf>>

*Feasibility and requirements study on preservation of e-prints* (2003). JISC.  
<[http://www.jisc.ac.uk/uploaded\\_documents/e-prints\\_report\\_final.pdf](http://www.jisc.ac.uk/uploaded_documents/e-prints_report_final.pdf)>

**Geser, Guntram; Pereira, John** (2004). *The future digital heritage space: an expedition report*. DigiCult Thematic issue 7.  
<[http://www.digicult.info/downloads/dc\\_thematic\\_issue7.pdf](http://www.digicult.info/downloads/dc_thematic_issue7.pdf)>

**Granger, Stewart; Russell, Kelly; Weinberger, Ellis** (2000). *Cost elements of digital preservation*. Ver. 4.  
<<http://www.leeds.ac.uk/cedars/colman/costElementsOfDP.doc>>

**Harvey, Ross** (2005). *Preserving digital materials*. Munic: K. G. Saur.

**Hedstrom, Margaret; Montgomery, Sheon** (1998). *Digital preservation needs and requirements in RLG member institutions*. Mountain View, Califòrnia: Research Library Group.  
<<http://www.rlg.org/preserv/digpres.pdf>>

**Hodge, Gail; Frangakis, Evelyn** (2004). *Digital preservation and permanent access to scientific information: the state of the practice*. CENDI; ICSTI.  
<[http://cendi.dtic.mil/publications/04-3dig\\_preserv.html](http://cendi.dtic.mil/publications/04-3dig_preserv.html)>

*Implementing preservation repositories for digital materials: current practice and emerging trends in the cultural heritage community; a report by the PREMIS Working Group* (2004). Dublín, Ohio.  
<<http://www.oclc.org/research/projects/pmwg/surveyreport.pdf>>

**Jones, Maggie; Beagrie, Neil** (2001). *Preservation management of digital materials: a handbook*. Londres: British Library.  
<<http://www.dpconline.org/graphics/intro/index.html>>

*Jorum preservation watch report* (2006). JISC. URL

**Keefer, Alice** (2006). "Laura Campbell, preservación digital en la LoC". *El Profesional de la Información* (vol. 3, núm. 15, pàg. 229).

**Keefer, Alice** (2005). "Preservación digital y depósitos institucionales". *El Profesional de la Información* (vol. 6, núm. 14, pàg. 404-406).

**Kenney, Anne R.; McGovern, Nancy Y** (2003). "The five organizational stages of digital preservation". *Digital Libraries: A Vision for the Twenty First Century, a festschrift to honor Wendy Lougee*.

**Lavoie, Brian F.** (2003). *The incentives to preserve digital materials: roles, scenarios, and economic decision-making*. Office of Research, OCLC.  
<<http://www.oclc.org/research/projects/digipres/incentives-dp.pdf>>

**Lavoie, Brian; Gartner, Richard** (2005). *Technology watch report: preservation metadata*. OCLC; Oxford University Library Services; Digital Preservation Coalition.  
<<http://www.dpconline.org/docs/reports/dpctw05-01.pdf>>

**Martínez Usero, José Angel** (2006, març-abril). "El uso de metadatos para mejorar la interoperabilidad del conocimiento en los servicios de administración electrónica". *El Profesional de la Información* (vol. 15: 2, pàg. 114-126).

**Méndez, Eva** (2002). *Metadatos y recuperación de información: estándares, problemas y aplicabilidad en bibliotecas digitales*. Gijón: Trea.

**Méndez, Eva** (2006, març-abril). "Dublin Core, metadatos y vocabularios". *El Profesional de la Información* (vol. 15: 2, pàg. 84-86).

**Méndez, Eva; Senso, José** (2004). *Introducción a los metadatos: aplicaciones y estándares*. Madrid: SEDIC. [Unitat d'autoformació en línia]  
<<http://www.sedic.es/autoformacion/metadatos/index.htm>>

**Muñoz de Solano y Palacios, Bárbara** (2006). "El profesional de la información: aspectos organizativos". *Anales de Documentación* (vol. 9, pàg. 143-163).  
<<http://www.um.es/fccd/anales/ad09/ad0910.pdf>>

**National Standards Organization** (2004). *Understanding metadata*. Bethesda, MD: NISO Press.  
<<http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>>

**OCLC/RLG PREMIS Working Group** (2004). *Implementing preservation repositories for digital materials: current practice and emerging trends in the cultural heritage community*.  
<<http://www.oclc.org/research/projects/pmwg/survey.doc>>

**Rebiun** (2005). *Anuario de las bibliotecas universitarias y científicas españolas: 2004*. [Data de consulta: 29.1.2007].  
<<http://bibliotecna.upc.es/Rebiun/nova/publicaciones/informeEstadistico04.pdf>>

**Rosenthal, David S. H. i altres** (2005). "Requirements for digital preservation". *D-Lib Magazine* (vol. 11: 11).  
<<http://www.dlib.org/dlib/november05/rosenthal/11rosenthal.html>>

**Salvador Benítez, Antonia; Ruíz Rodríguez, Antonio Ángel** (2005). "Metadatos para la preservación de colecciones digitales". *Cuadernos de Documentación Multimedia* (vol. 16).  
<<http://multidoc.rediris.es/cdm/viewarticle.php?id=35&layout=html>>

**Sanett, Shelby** (2003). "The cost to preserve authentic electronic records in perpetuity: comparing costs across cost models and cost frameworks". *RLG Diginews* (7: 4).

<<http://www.rlg.org/preserv/diginews/diginews7-4.html#feature2>>

*Seminario Aplicación en España del Modelo de Metadatos Dublin Core* (2003). Madrid.  
<<http://es.dublincore.org/es/eventos/dcmi-es1/>>

*Urgent Action Needed to Preserve Scholarly Electronic Journals* (15 d'octubre de 2005).  
<<http://www.diglib.org/pubs/waters051015.htm>>

**Verheul, Ingeborg** (2006). *Networking for digital preservation: current practice in 15 national libraries* (IFLA publications 119). Munic: K. G. Saur.  
<[www.ifla.org/V/pr/saur119.htm](http://www.ifla.org/V/pr/saur119.htm)>

**Waller, Martin; Sharpe, Robert** (2006). *Mind the gap: assessing digital preservation needs in the UK*. Digital Preservation Coalition.  
<<http://www.dpconline.org/docs/reports/uknamindthegap.pdf>>

