

# LA LONGEVITAT DE LES IMATGES DIGITALS RASTER

*René van Horik*

## Introducció

El problema de la creació d'imatges o retrats estables i permanents no es va solucionar amb la invenció de la fotografia l'any 1839. La cita següent ho il·lustra:<sup>1</sup>

...el daguerrotip era tan fràgil com l'ala d'una papallona, fugaç i molt més difícil de reproduir que un gravat. Hi havia un consens general sobre el fet que la fotografia esdevindria una força important només quan fos capaç de produir imatges permanents i que es poguessin repetir infinitament. [...] Aquest objectiu es va aconseguir parcialment al final del segle XIX, però no va assolir la plena maduresa comercial fins més tard.

Cap a les acaballes del segle XX, la captura i la reproducció d'imatges mitjançant aparells digitals va esdevenir accessible a gran escala. I, un cop més, la qüestió de l'estabilitat i la permanència de les imatges, ara en forma digital, es va posar damunt la taula, tal com ens ho il·lustra la cita següent:<sup>2</sup>

La digitalització dels objectes culturals hauria de proporcionar un testimoni electrònic durador per a l'accés acadèmic i universal, la preservació i l'estudi. Actualment, no obstant això, els projectes de digitalització s'estan duent a terme sense mètodes establerts d'enregistrament amb condicions precises de digitalització.

L'obsolescència del format d'imatge digital i el deteriorament del suport d'emmagatzematge de dades digitals són alguns dels factors que amenacen l'accés a llarg termini de les imatges digitals. Existeixen diversos enfocaments quant a la preservació digital per evitar que els objectes digitals deixin de ser accessibles en el futur. Partint de la tecnologia més avançada, aquest treball informa sobre allò que es pot fer per minimitzar el risc que les imatges digitals raster creades avui dia no es puguin reproduir en el futur.

En la primera part del treball s'estudia informació general de context sobre les imatges digitals. La secció següent descriu estratègies generals de preservació digital comunament acceptades actualment. L'objectiu d'aquesta secció és situar la preservació digital d'objectes digitals en un context més ampli.

La tercera part de l'estudi se centra en la durabilitat de les imatges digitals raster. Els temes principals que determinen l'accés a llarg termini de les imatges raster són el format de fitxer utilitzat per ajustar el fitxer digital que conté els píxels i la qualitat de la documentació o metadades.

## 1. Imatges digitals

El processament d'imatges per ordinador va començar als Estats Units al National Bureau of Standards l'any 1956.<sup>3</sup> Des d'aleshores, hi ha hagut una sèrie de decisions preses pels creadors del primer escàner que han influenciat els mètodes de l'enginyeria digital, per exemple l'ús de matrius rectangulars de píxels quadrats. *Píxel* és l'acrònim de *picture element*. I un píxel equival al punt més petit de resolució en una imatge digital raster. Un píxel també és la ubicació separada d'un fotosensor individual en una càmera digital. No es va fer cap intent per basar el protocol de digitalització en la naturalesa de la imatge, cosa que va derivar en imatges força grans. Kirsch il·lustra aquest fet mostrant un mosaic del segle IV que conté prop de 80 x 46 tesel·les acuradament pintades i formades. La digitalització d'aquest mosaic, fins i tot amb més píxels (100 x 58), dona una imatge inferior.

Es necessita un nombre molt més elevat de píxels per tal de reproduir els detalls del mosaic original.

Els fitxers gràfics es poden considerar fitxers que emmagatzemen qualsevol tipus de dades gràfiques persistents (en oposició a dades textuais, numèriques o fulls de càlcul, per exemple), i la finalitat última dels quals és ser reproduïts i exhibits. Aquesta definició es pot trobar a l'*Encyclopaedia of Graphics File Formats*, publicada el 1994.<sup>4</sup> En aquesta enciclopèdia hi apareixen gairebé cent formats.

El fet que hi hagi tants formats de fitxer diferents s'explica per diverses raons. La primera és que existeix un cert nombre de tipus diferents i fonamentals de dades gràfiques. Cada tipus necessita el seu propi format d'arxiu. Es poden distingir tres grans categories:<sup>5</sup>

- Dades raster: un grup de valors aleatoris, en espai tant bidimensional com tridimensional, que representa una imatge o que es pot processar com una imatge.<sup>6</sup>
- Dades geomètriques: descripció matemàtica d'espai, en espai tant bidimensional com tridimensional, que representa els components d'una imatge.<sup>7</sup>
- Dades d'imatges latents: dades no gràfiques que poden ser transformades en imatges útils mitjançant algun procés algorítmic.

El tipus d'imatge digital que es tracta en aquest treball és el fitxer d'imatge raster bidimensional o mapa de bits.

La segona raó de l'existència d'una gamma tan àmplia de formats de fitxer gràfic és que es van desenvolupar diversos formats de propietat per impedir-ne l'ús més enllà del control del creador original. El Kodak Photo CD és un exemple de format que en un principi era de propietat, però que, en una fase posterior, la Kodak (més o menys forçada per les lleis del mercat que demanaven més formats oberts) va permetre que fos utilitzat pels programadors segons les especificacions de format. Alguns formats de fitxer gràfic van ser destinats a l'ús en hardware específic dedicat, i per tant tenen un format específic.

El tercer motiu de tanta varietat de formats de fitxer gràfic és l'extens ventall de principis de disseny utilitzat pels programadors. Els dos punts fonamentals que van influenciar el desenvolupament de formats de fitxer gràfic específics van ser la velocitat de processament de la imatge i la memòria requerida per representar la imatge. La longevitat o durabilitat mai no es menciona com a objectiu de disseny específic del format de fitxer gràfic.

El disseny d'un format de fitxer gràfic es basa en els components de memòria, velocitat i sistema de circuits dels sistemes de hardware dirigits a usar les dades. La quantitat de memòria disponible pot afectar la velocitat de l'accés a les dades. Les aplicacions de gràfics són autèntiques devoradores de memòria. La independència de dispositius disminueix la velocitat de processament i augmenta les necessitats de memòria. De vegades els formats específics de hardware incrementen la velocitat i requereixen menys memòria. El fax n'és un exemple.

### 1.1. Imatges digitals raster

Un fitxer d'imatge raster o un fitxer de mapa de bits consisteix en una matriu de valors discrets de píxel creada pel CCD d'un dispositiu de captura d'imatge. El CCD (*Charged Coupled Device*) és un sensor sensible a la llum que es troba en un xip que serveix per convertir el senyal analògic en codis digitals discrets. L'interval de mostreig del dispositiu de captura d'imatge determina quants píxels hi ha emmagatzemats al fitxer raster. Aquests codis-píxels es poden escriure de diverses maneres. Els fitxers raster poden contenir dades addicionals com informació de la descripció de la imatge o una paleta de colors.

La capacitat de memòria en bytes necessària per emmagatzemar els píxels d'una imatge digital raster es pot expressar mitjançant la fórmula  $(PH \times PV \times PD) / 8$ . PH és el nombre de píxels en la dimensió horitzontal, PV és el nombre de píxels en la dimensió vertical i PD (*Pixel Depth*) és el nombre de bits necessaris per codificar el color d'un píxel individual. Segons la velocitat de la pel·lícula la resolució de mostreig de les pel·lícules fotogràfiques és l'equivalent de 2.000 a 5.000 píxels per polzada.<sup>8</sup>

A fi de reproduir un fitxer raster en una pantalla d'ordinador o d'imprimir-lo en paper s'ha de processar un mapa de bits emmagatzemat en la memòria de l'ordinador. Un mapa de bits d'alta resolució, per exemple, s'ha de reescalar per obtenir una vista en pantalla sencera en un monitor d'ordinador estàndard. Un altre exemple és l'algorisme de tramatge necessari per imprimir un fitxer raster en paper. Altres temes del processament d'imatges són la restauració i la millora de la imatge. Els algorismes de compressió de dades s'apliquen per tal de reduir la mida del fitxer. Atès que els fitxers de mapa de bits solen ser molt grans, sovint s'aplica la compressió. Els fitxers més petits necessiten menys capacitat d'emmagatzematge i menys amplada de banda. El mètode de compressió de fitxers JPEG s'utilitza a gran escala, principalment perquè els navegadors d'Internet poden comprimir les imatges JPEG automàticament. Les imatges comprimides JPEG ja no tenen l'estructura discreta raster i el mètode de compressió causa una pèrdua en la qualitat de la imatge. El fet que el JPEG sigui tant un algorisme de compressió com un format de fitxer pot produir confusions. És possible, per exemple, usar el mètode de compressió JPEG per a la creació de fitxers d'imatges TIFF. El format de fitxer d'imatge TIFF serà tractat més endavant en aquest treball. Quant a la durabilitat dels fitxers de mapa de bits, els algorismes de processament d'imatges es poden subdividir en dos grups:

- Algorismes de processament d'imatges que canvien els valors dels píxels en la seqüència de bits o *bitstream* del mapa de bits emmagatzemat. En són exemples: la compressió de la imatge i la modificació dels valors píxel a través de filtres digitals i ajustaments de contrast i lluminositat. La compressió de la imatge sovint també abandona l'estructura raster del mapa de bits.
- Algorismes de processament d'imatges que no canvien els píxels en el mapa de bits bàsic. En són exemples la visualització d'imatges i el software d'impressió d'imatges que manipula els píxels per optimitzar una sortida específica. Aquestes operacions es duen a terme en la memòria RAM de l'ordinador i no manipulen permanentment els valors píxel en el mapa de bits.

És millor no usar els algorismes de processament d'imatges per a un propòsit específic que canvien els valors píxel en el fitxer digital raster, perquè les característiques del fitxer s'alteren. Es desaconsella l'ús d'algorismes de compressió per diverses raons. Existeix el risc que el mètode de compressió esdevingui obsolet i si el mapa de bits es corromp és molt difícil reparar la imatge. Una imatge no comprimida es pot reparar molt més fàcilment. Finalment, l'aplicació de la compressió de la imatge sovint comporta una pèrdua de qualitat d'imatge.

## 1.2. Substituts digitals i imatges creades digitalment

Les biblioteques, els arxius i els museus normalment tracten amb dos tipus d'imatges raster: substituts digitals i imatges creades digitalment. Els substituts digitals són versions digitals d'imatges analògiques originals, com fotografies històriques o còpies fotomecàniques. Les imatges creades digitalment es creen directament amb un dispositiu de captura digital, i les imatges resultants no tenen una relació estreta amb els originals analògics. Pel que fa a la preservació digital dels dos tipus de fitxers d'imatge raster, hi ha molta similitud. En tots dos casos cal aplicar un format de fitxer d'imatge i un suport d'emmagatzematge digital

adequats, i la documentació o metadades rellevant per a la preservació digital ha de tenir en compte el caràcter i la sintaxi de la imatge.

En els substituïts digitals l'ús al qual està destinada la imatge digital i les característiques de l'original analògic influencien els paràmetres del dispositiu de captura digital i, per tant, les característiques del fitxer d'imatge raster. En moltes situacions els substituïts digitals es faran servir per a visualitzacions fidels en pantalla o per a la creació d'una reproducció en paper. Atès que en el futur els requeriments de sortida poden canviar, és evident que la qualitat i la longevitat dels objectes digitals és important. Això fa que la digitalització de recursos visuals sigui més aviat un procés circular i no pas lineal. L'enfocament cíclic comporta una política d'arxivament activa a fi de facilitar una utilització futura específica, per exemple la recerca acadèmica o la visualització en línia. El concepte *use-neutral* ('ús neutral') s'utilitza per expressar els requeriments d'objectes digitals d'alta qualitat creats en un projecte de conversió. Amb aquest enfocament, una imatge es digitalitza una vegada, al nivell de qualitat més alt que es pugui, i s'estableixen estàndards d'estudi, com la correspondència de colors i els nivells de contrast, de manera que la imatge és pugui emprar per a múltiples aplicacions.

En gran part, les característiques de les imatges digitals creades per dispositius de captura digital, per exemple una càmera fotogràfica digital, estan determinades per les propietats d'aquests dispositius. Alguns estàndards utilitzats per dispositius de captura digital s'haurien de tenir en compte per a la preservació digital d'imatges creades digitalment, com l'estàndard EXIF per emmagatzemar informació d'intercanvi en fitxers d'imatges, utilitzat per gairebé totes les càmeres digitals. Aquest tema es tractarà més endavant.

En general, es necessiten molts recursos per a la creació d'imatges digitals raster d'alta qualitat o imatges digitals originals. La pèrdua d'imatges digitals pot comportar una pèrdua de contingut valuós, cosa que justifica les inversions necessàries per garantir l'accés a llarg termini a les imatges digitals.

## 2. Preservació d'objectes digitals

Fins fa ben poc, la vulnerabilitat dels objectes digitals no es considerava un problema. Mentre valgui la pena mantenir vius els objectes digitals és només qüestió de traslladar-los d'un suport d'emmagatzematge a un altre de nou: de disquet a CD-ROM, de disc dur a cinta. De vegades també s'ha de convertir el format de dades a fi d'evitar que les dades no puguin tornar a ser processades.

L'article del *Scientific American* «Ensuring the Longevity of Digital Documents», de Jeff Rothenberg, publicat el 1995, és una publicació que se cita amb molta freqüència i que ha començat a cridar l'atenció sobre el fet que els documents digitals tenen una vida més aviat curta.<sup>9</sup> Rothenberg va escriure que els mitjans digitals «duraran per sempre —o cinc anys. Segons la prioritat». Però no és només el suport d'emmagatzematge que causa preocupació. El fet que les dades digitals siguin comprensibles en el futur també és important. Quin és el significat de la seqüència de bits en el suport d'emmagatzematge i com es pot interpretar en el futur?

### 2.1. Estratègies de preservació digital

Des de 1995 s'han dut a terme diversos projectes i estudis sobre preservació digital en un ampli ventall de temes. Aquests consistien en inventaris i valoracions de recursos digitals, eines i mètodes per conservar material digital i estàndards i directrius per a la preservació digital. La preservació digital fa referència a totes les accions necessàries per garantir l'accés als materials digitals més enllà dels límits de les avaries dels mitjans o dels canvis

tecnològics. L'any 2000 s'han desenvolupat tres estratègies principals per a la preservació digital.<sup>10</sup>

- L'estratègia de preservació tecnològica. La preservació del software i el hardware originals que van ser emprats per crear i donar accés a la informació. Inclou tant la preservació del sistema operatiu original com la del hardware per fer-lo funcionar.
- L'estratègia d'emulació tecnològica. Els sistemes informàtics futurs poden imitar plataformes informàtiques obsoletes. L'emulació és el procés d'imitar sistemes obsolets en generacions d'ordinadors futures.
- L'estratègia de migració d'informació digital. La informació digital es recodifica en nous formats abans que el format antic esdevingui obsolet. L'objectiu de la migració és preservar el contingut intel·lectual dels objectes digitals i assegurar que els usuaris puguin continuar recuperant-los, visualitzant-los i utilitzant-los malgrat els canvis tecnològics constants.

Tot i que existeix un acord sobre les estratègies disponibles per a la preservació digital, això encara no ha portat a uns punts comuns sobre com implementar aquestes estratègies i sobre quina estratègia de preservació s'adiu més amb els diferents tipus de material digital. Actualment s'estan duent a terme diversos experiments i estudis de viabilitat.

També s'ha de comentar que les tres estratègies de preservació digital s'apliquen a una gran varietat de materials digitals, entre els quals hi ha bases de dades, programes informàtics, imatges digitals, textos electrònics i pàgines web. La formació i l'opinió de les persones que implementen les estratègies de preservació digital determinen com s'entenen i classifiquen avui dia els materials digitals. Un lloc web és un tipus de text electrònic? Una base de dades està inextricablement connectada amb el seu sistema de gestió de bases de dades? N'hi ha prou amb preservar el resultat d'un càlcul per ordinador o també s'haurien de preservar els algorismes com a tals?

Com a conseqüència d'aquestes divergències d'opinió sobre els materials digitals, s'està desenvolupant una gran varietat de projectes i de línies d'investigació, de vegades amb aproximacions fonamentalment diferents tot i que la naturalesa del material digital és la mateixa.

La preservació digital és un camp de recerca relativament nou, i només les generacions futures podran jutjar si les estratègies de preservació digital implementades avui eren les correctes.

## 2.2. L'Open Archival Information System (OAIS)<sup>11</sup>

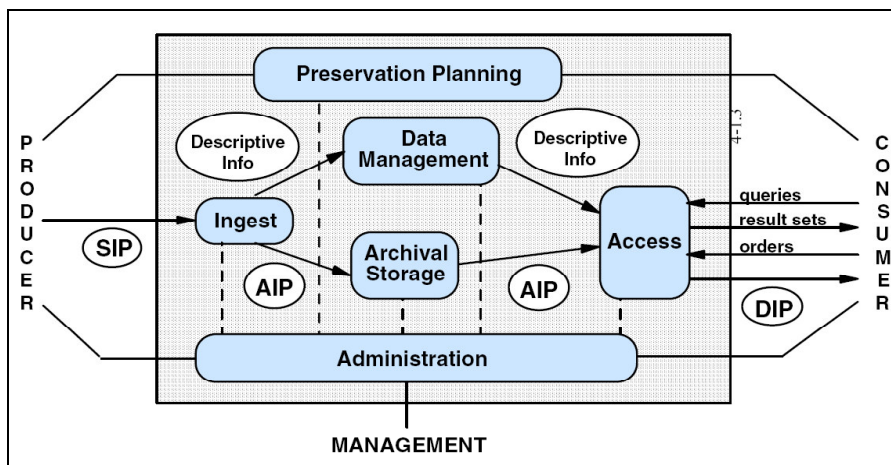
Recentment es va establir una norma internacional que està exercint una gran influència en l'àmbit de recerca de la preservació digital. Aquesta norma, la Reference Model for an Open Archival Information System (OAIS), s'està desenvolupant sota la direcció del Consultative Committee for Space Data Systems (CCSDS). El model de referència ha estat adoptat com a estàndard ISO 14721:2003 i estableix un marc comú de termes i conceptes rellevants per a l'arxivament a llarg termini de dades digitals. Un OAIS es defineix com «un arxiu, i consisteix en una organització de persones i sistemes que ha acceptat la responsabilitat de conservar informació i fer-la accessible a una comunitat designada». Una comunitat designada es defineix com «un grup de consumidors potencials identificat que hauria de ser capaç d'entendre un conjunt d'informació concret. La comunitat designada pot estar formada per moltes comunitats d'usuaris». El model OAIS ja s'empra àmpliament com a base per a una gran varietat d'iniciatives de preservació digital.

El model OAIS es pot considerar un marc conceptual que informa sobre el disseny d'arquitectures de sistema, però no assegura la coherència o la interoperabilitat entre implementacions. El model de referència OAIS conté tres conceptes clau d'alt nivell:

1. L'entorn d'un OAIS. Un OAIS o arxiu està envoltat de *Producers* ('productors', que proporcionen la informació que s'ha de conservar), *Consumers* ('consumidors', que interactuen amb els serveis OAIS per trobar i adquirir informació conservada d'interès), i *Management* ('gestió', que estableix la política general OAIS com un component en un domini de política més ampli).
2. Informació OAIS. El contingut digital és transportat del productor a l'arxiu, i de l'arxiu al consumidor, en forma de tipus d'objectes d'informació (*Information Packages*). Cal informació de representació (*Representation Information*) per tal d'entendre l'objecte de dades (*Data Object*) que hi ha arxivat. Aquest és bé un objecte físic bé un objecte digital. La informació de representació és aquella que relaciona un objecte de dades amb conceptes més significatius. Així, un objecte d'informació (*Information Object*) està format per dos components: l'objecte de dades i la informació de representació. Un OAIS consisteix en diversos tipus d'objectes d'informació, un contenidor conceptual de dos tipus d'informació: informació de contingut (*Content Information*) i informació descriptiva sobre la preservació (*Preservation Description Information*, PDI). La PDI es divideix en quatre tipus d'informació de preservació anomenats de procedència (*Provenance Information*), de context (*Context Information*), de referència (*Reference Information*) i de fixació (*Fixity Information*). Cal distingir entre un tipus d'objecte d'informació conservat per un OAIS i els tipus d'objecte d'informació que es lliuren a un OAIS i que es difonen des d'un OAIS. Aquestes variants s'anomenen *Submission Information Package* (SIP), *Archival Information Package* (AIP) i *Dissemination Information Package* (DIP).
3. Interaccions externes d'alt nivell. La interacció entre productor i consumidor amb l'OAIS es basa en tipus d'objectes d'informació específics. Un productor lliura un *Submission Information Package* (SIP) a l'OAIS perquè sigui utilitzat en la construcció d'un o més AIP. Un consumidor rep un *Dissemination Information Package* (DIP), procedent d'un o més AIP en resposta a una petició feta a l'OAIS.

El model funcional OAIS es compon de sis funcions arxivístiques:

1. Ingest ('ingestió'). Conté els serveis i les funcions que els *Submission Information Packages* (SIP) accepten dels productors, prepara els *Archival Information Packages* (AIP) per a l'emmagatzematge i garanteix que els AIP i la seva informació descriptiva de suport s'estableixin dintre de l'OAIS.
2. Archival Storage ('emmagatzematge arxivístic'). Conté els serveis i les funcions utilitzats per a l'emmagatzematge i la recuperació dels *Archival Information Packages* (AIP).
3. Data Management ('gestió de dades'). Conté els serveis i les funcions per proveir, mantenir i accedir a una àmplia varietat d'informació.
4. Administration ('administració'). Conté els serveis i les funcions necessàries per controlar el funcionament diari de la resta d'entitats funcionals OAIS.
5. Preservation Planning ('pla de preservació'). Conté els serveis i les funcions que controlen l'entorn de l'OAIS i proporcionen recomanacions per assegurar que la informació emmagatzemada a l'OAIS continuï essent accessible a llarg termini per a la comunitat designada, fins i tot si l'entorn informàtic original esdevé obsolet.
6. Access ('accés'). Conté els serveis i les funcions que fan que els documents arxivístics i els serveis que hi estan relacionats siguin visibles per als consumidors.



**Figura 1. Entitats funcionals de l'OAIS. (Font: OASIS/ISO14721:2003, p. 4-1.)**

La figura 1 conté tant els tres conceptes d'alt nivell com les sis entitats del model de referència OAIS. Es poden distingir tres àrees d'influència del model OAIS:

- Primer de tot, el model s'usa per a la compilació d'esquemes de metadades de preservació. Més endavant s'aportarà informació de fons sobre les metadades de preservació.
- En segon lloc, el model OAIS té un paper en l'arquitectura i el disseny dels sistemes de preservació digital d'informació. La normativa de l'OAIS declara: «Els implementadors utilitzaran aquest model de referència com una guia mentre desenvolupen una implementació específica per proporcionar serveis i continguts identificats».
- En tercer lloc, el model OAIS s'utilitza com una base per a la conformitat i molts sistemes de preservació digital d'informació reclamen el compliment de l'OAIS. Tot i així, encara no existeix un procés de certificació del model OAIS generalment acceptat. El Research Libraries Group (RLG) i la US National Archives and Records Administration (NARA) han creat un grup de treball per a la certificació de dipòsits digitals. El seu objectiu és elaborar uns requisits de certificació per establir i seleccionar dipòsits d'informació digital fiables.<sup>12</sup>

L'estàndard OAIS distingeix sis responsabilitats obligatòries que una organització ha de complir a fi d'operar un arxiu OAIS:

1. Negociar i acceptar informació apropiada dels productors d'informació.
2. Obtenir un control suficient de la informació aportada al nivell necessari per assegurar-ne la conservació a llarg termini.
3. Determinar quines comunitats haurien de formar part de la comunitat designada i, per tant, haurien de poder entendre la informació proporcionada.
4. Garantir que la informació que s'ha de preservar sigui comprensible de manera independent per a la comunitat designada.
5. Observar polítiques i procediments documentats que assegurin que la informació es conserva malgrat qualsevol contingència raonable, i que en permetin la difusió en còpies autèntiques de l'original, o de manera que es pugui localitzar l'original.
6. Posar la informació conservada a disposició de la comunitat designada.

Hi ha un reconeixement creixent del fet que els diversos tipus de dades digitals capturades de maneres diferents necessitaran diferents tipus de suport per preservar-se a llarg termini. Els materials digitals altament estructurats solen ser inherentment més fàcils de conservar i

d'accedir-hi al llarg del temps. Els materials menys estructurats acostumen a ser més difícils de tractar. Una altra manera de categoritzar la persistència inherent és veure si els materials són homogenis. Això vol dir, si són materials estretament relacionats amb normes conegudes i coherents pel que fa a estructura, paràmetres tècnics i metadades.

### 2.3. Metadades de preservació

Sense documentació un objecte digital només és una seqüència de díbits binaris. Si hi ha documentació disponible sobre els objectes digitals, la seva longevitat millorarà. Això és així perquè la informació sobre temes com el significat de la seqüència de bits que conforma l'objecte digital, la informació i les dades bibliogràfiques sobre les característiques formals de l'objecte informaran la gent i els sistemes sobre el contingut, el valor i el possible ús de l'objecte digital. D'aquesta manera, la documentació sobre l'objecte digital, o metadades, facilita moltíssim la longevitat de l'objecte.

El concepte *metadades* significa literalment 'dades sobre dades'. La bibliografia sobre biblioteconomia i ciències de la informació publicada en els darrers anys adjudica un paper significatiu a les metadades en relació amb la gestió d'objectes digitals.<sup>13</sup> Les metadades es poden considerar documentació que proporciona informació sobre les característiques de les «coses», tant analògiques com digitals. El fet que cada vegada es presti més atenció a la importància de les metadades es deu a l'ascens d'Internet. Internet no és una biblioteca estructurada i ben organitzada. Els objectes a Internet només es poden descobrir si hi ha metadades disponibles sobre ells. En una biblioteca és possible cercar entre les col·leccions. Les metadades poden tenir una àmplia gamma de funcions. Es poden utilitzar per identificar versions d'un objecte, per certificar-ne l'autenticitat, indicar-ne la categoria, controlar-ne els drets de propietat intel·lectual, marcar l'estructura del seu contingut, etcètera. Per a aquestes i altres funcions existeix un gran nombre d'iniciatives, projectes, estàndards i directrius.<sup>14</sup>

Les metadades no només són importants per descobrir recursos, sinó també per a la seva conservació. La informació sobre el context tecnològic i sobre altres contextos de la creació i l'ús d'un objecte digital es coneix com a «metadades de preservació». El model de referència OAIS es refereix a les metadades amb el nom d'«informació de representació», informació que correlaciona un objecte digital codificat amb informació més significativa. Aquesta operació inclou els aspectes sintàctics, estructurals i semàntics d'una codificació. La informació de representació no s'aplica tan sols als objectes digitals, sinó que també es pot aplicar als formats en què estan codificats els objectes.

No existeixen metadades de preservació generalment acceptades que facilitin la longevitat de les imatges digitals raster. Diverses comunitats van dissenyar conjunts d'elements de dades que es poden emprar per millorar la longevitat de les imatges digitals raster. En la propera secció d'aquest treball es presenten alguns conjunts d'elements de dades rellevants. A més, les característiques dels elements de dades utilitzats, com la formulació de la definició i del tipus de dades dels termes usats, difereix molt.

#### 2.3.1. METS: Metadata Encoding and Transmission Standard

Un mètode emergent per emmagatzemar i expressar metadades, especialment per a fonts audiovisuals i les seves representacions digitals, és l'ús d'un *wrapper* ('embolcall') d'acord amb les especificacions METS. El METS, Metadata Encoding and Transmission Standard,<sup>15</sup> s'ocupa de la proliferació dels conjunts d'elements de metadades en els darrers anys. En aquest sentit, es tracta d'una mena de meta-metadades. El METS es pot considerar una part integral del model de referència OAIS (vegeu la secció 2.2.) i serveix per facilitar



l'intercanvi d'objectes digitals d'un dipòsit a un altre.<sup>16</sup> Un document METS està format per set seccions principals:

1. Capçalera METS: la capçalera descriu el document METS mateix, per exemple dona informació sobre el creador.
2. Metadades descriptives: poden assenyalar metadades descriptives externes al document METS, per exemple un registre en un catàleg. Les metadades descriptives també poden estar incrustades internament.
3. Metadades administratives: proporcionen informació sobre com van ser creats i emmagatzemats els objectes digitals, sobre els drets de propietat intel·lectual, les metadades sobre l'objecte original del qual deriva l'objecte digital de la biblioteca i la informació sobre la procedència dels fitxers que componen l'objecte digital de la biblioteca. A l'igual de les metadades descriptives, les administratives poden ser bé externes al document METS, bé codificades internament.
4. Secció de fitxers: fa una relació de tots els fitxers amb contingut que conformen les versions electròniques de l'objecte digital.
5. Mapa estructural: és el cor d'un document METS. Traça una estructura jeràrquica per a l'objecte digital de la biblioteca i enllaça els elements de l'estructura amb els fitxers de contingut i metadades que pertanyen a cada element.
6. Enllaços estructurals: permet als creadors METS prendre nota de l'existència d'hiperenllaços entre nodes en la jerarquia esbossada al mapa estructural. Aquesta funció és especialment valuosa a l'hora d'utilitzar el METS per arxivar llocs web.
7. Comportament: es pot utilitzar per associar comportaments d'executables (programes informàtics) amb contingut a l'objecte METS. Aquesta secció no és rellevant per a imatges digitals raster.

El METS fa servir el llenguatge XML.<sup>17</sup> Un esquema XML defineix els continguts possibles d'un document XML. Amb el METS, una col·lecció d'objectes digitals relacionats, per exemple les pàgines digitalitzades d'un llibre o les fotografies digitalitzades d'un àlbum, es poden agrupar. El METS presenta una aproximació liberal pel que fa al format dels elements de dades que descriuen els objectes digitals: es pot utilitzar qualsevol format. Els sistemes que fan servir l'estàndard METS encara es troben en fase de prototipus. La figura 2 conté un exemple d'una petita part d'un document METS, concretament de la secció de fitxers. L'exemple fa evident que quatre imatges digitals pertanyen a un mateix grup. Les quatre imatges podrien ser imatges digitals raster d'una sèrie.

```
<METS:fileGrp>
<METS:file GROUPID="129131-1" MIMETYPE="image/tiff" ID="_79926" SEQ="1"></METS:file>
<METS:file GROUPID="129131-2" MIMETYPE="image/tiff" ID="_79927" SEQ="2"></METS:file>
<METS:file GROUPID="129131-3" MIMETYPE="image/tiff" ID="_79928" SEQ="3"></METS:file>
<METS:file GROUPID="129131-4" MIMETYPE="image/tiff" ID="_79929" SEQ="4"></METS:file>
</METS:fileGrp>
```

**Figura 2. Petita part de (de la secció de «Fitxers») d'un document METS, que estableix que quatre imatges digitals (en format TIFF) pertanyen al mateix grup.**

#### 2.4. Preservació de formats de fitxer digitals

La forma en què estan disposats els díbits binaris en un fitxer digital depèn del format del fitxer. La informació sobre la sintaxi i la semàntica internes del format del fitxer és important per tal d'entendre i processar el fitxer digital. Els registres de formats que contenen informació de representació sobre els formats digitals poden contribuir a assegurar l'accés a

llarg termini als fitxers digitals. Un registre de formats es pot utilitzar per identificar, validar, caracteritzar, transformar i lliurar objectes digitals, també a llarg termini.

El Global Digital Format Registry (GDFR) és un exemple d'iniciativa que investiga les possibilitats d'establir un registre de formats sostenible.<sup>18</sup> El model de dades provisional per al GDFR inclou propietats del registre mateix i propietats del format. Les propietats del format estan subdividides en propietats descriptives, propietats tècniques, propietats del sistema i propietats administratives. El disseny del model de dades es va regir per la pregunta: «Quina informació voldries tenir avui per manipular un objecte digital de fa cinquanta anys?» Actualment s'està desenvolupant un prototipus del GDFR, però encara estem lluny d'un registre de producció operacional.

Fa poc els Arxius Nacionals del Regne Unit van iniciar un registre de formats de fitxer amb el nom de PRONOM.<sup>19</sup> Tal com consta als seus llocs web: «PRONOM és una font en línia que aporta informació sobre formats de fitxer i productes de software. És un recurs per a qualsevol persona que necessiti informació tècnica imparcial i definitiva sobre els formats de fitxer emprats per emmagatzemar registres electrònics, i sobre els productes de software necessaris per crear, reproduir o migrar aquests formats.» Actualment el sistema PRONOM conté informació molt limitada sobre formats de dades.

A més dels registres de formats, s'estan desenvolupant eines per realitzar la identificació, la validació i la caracterització específica del format d'objectes digitals. La identificació és el procés de determinació del format específic d'un objecte digital. La validació és el procés que determina la conformitat d'un objecte digital a les especificacions del seu suposat format. La caracterització és el procés d'extreure informació de preservació (referència model OAIS) o metadades d'un objecte. Quan es trameten metadades externes a un dipòsit en associació amb objectes digitals, s'hauria de comprovar que fossin coherents.

El JHOVE (JSTOR/Harvard Object Validation Environment) és un marc extensible per a aquestes identificació, validació i caracterització específica del format d'objectes digitals.<sup>20</sup> El programa JHOVE disponible actualment conté mòduls per a diversos formats de fitxer d'imatges raster, com els populars GIF, JPEG, TIFF i PDF.

El registre de formats i la identificació, validació i caracterització d'objectes digitals encaixen molt bé en el model de referència OAIS (vegeu la secció 2.2.), principalment pel que fa a les entitats productor i arxiu.

## 2.5. Longevitat dels suports d'emmagatzematge dels objectes digitals

L'any 1995 el Departament de Defensa dels Estats Units va demanar al Laboratori Nacional de Mitjans que realitzés una investigació sobre l'expectativa de vida dels suports d'emmagatzematge per a les dades digitals. L'expectativa de vida real d'un suport d'emmagatzematge concret depèn de la qualitat dels suports manufacturats, la quantitat de vegades que s'hi accedeix durant la seva vida útil, la cura amb què es manipula, la temperatura i la humitat d'emmagatzematge, la netedat de l'entorn d'emmagatzematge i la qualitat de l'enregistrador usat per escriure al suport d'emmagatzematge.<sup>21</sup> La investigació va tenir en compte els tipus de suport següents: la cinta magnètica, el disc òptic, el paper i la pel·lícula. Els dos factors principals que afecten l'expectativa de vida dels suports són la temperatura d'emmagatzematge i la humitat relativa de l'aire. Una temperatura d'emmagatzematge de 10 °C i una humitat relativa del 25% garanteixen una expectativa de vida fiable d'almenys vint anys, tant per a Digital Linear Type (DLT) magnètic com per a CD-ROM com a disc òptic. Els millors venedors d'aquests productes poden oferir suports amb una expectativa de vida de cent anys com a mínim. S'assumeix que els nous suports són utilitzats, que s'hi accedeix amb poca freqüència, que estan ben emmagatzemats, segons

les condicions ambientals indicades, i que l'entorn d'emmagatzematge és net i no hi ha pols, fum, menjar, verdet, llum natural directa i contaminants gasosos.

Malgrat el fet que, en general, el paper i el microfilm tenen una expectativa de vida més llarga que el disc òptic i la cinta magnètica, la durabilitat de les dades digitals expressades en col·leccions de bits i bytes serà prou bona per a un emmagatzematge fiable al llarg d'un segle. La norma ISO 18921:2002 està disponible per estimar l'expectativa de vida dels CD-ROM partint dels efectes de la temperatura i la humitat relativa.<sup>22</sup> L'objectiu de la norma és establir una metodologia per calcular l'expectativa de vida de la informació emmagatzemada en CD-ROM. Aquesta metodologia proporciona un procediment sòlid, tant tècnicament com estadísticament, per obtenir i avaluar dades de proves accelerades. Una mesura important per determinar si un CD-ROM encara és accessible és la taxa d'error de blocs o BLER. Això és la ràtio de blocs erronis mesurats per segon a l'entrada al descodificador de dades.

Es pot concloure que existeixen suports fiables per emmagatzemar dades digitals durant llargs períodes de temps. És probable que el hardware per accedir a la seqüència de bits en els suports esdevingui obsolet ben aviat. Monitorar el hardware disponibles per llegir els suports és tan important com controlar els suports d'emmagatzematge. Un risc més gran de perdre les dades digitals el constitueix el fet que la interpretació i el processament de dades requereix aplicacions que poden esdevenir obsoletes. La durabilitat del format de dades té més importància que la durabilitat del suport d'emmagatzematge.

Recentment s'ha publicat l'informe *Care and Handling of CDs and DVDs: A Guide for Librarians and Archivists*. Aquest informe descriu en un llenguatge no tècnic els diversos tipus de CD i DVD usats actualment, com estan fets i com operen. L'informe també conté informació industrial actual sobre la longevitat dels suports, les condicions que influencien l'expectativa de vida i sobre com tenir cura dels suports òptics.<sup>23</sup>

### 3. Preservació digital d'imatges digitals raster

Aquesta secció tracta temes rellevants per a l'accés a llarg termini d'un objecte digital específic, concretament les imatges digitals raster, i consta de dues parts. La primera s'ocupa dels formats estàndard de fitxer d'imatge disponibles i intenta respondre la pregunta de quin format estàndard de fitxer d'imatge és el més durador. La segona part explica detalladament una sèrie de conjunts d'elements de metadades que faciliten la durabilitat d'una imatge digital raster.

#### 3.1. Formats estàndard de fitxer d'imatge

Una manera òbvia de crear objectes digitals duradors és utilitzar formats de dades estandarditzats. Un estàndard es considera un objecte ben dissenyat, àmpliament utilitzat i que gaudeix d'un gran suport. Empíricament els requisits per als formats de dades estàndard es donen seguits d'una avaluació dels formats de dades per a imatges digitals raster existents. Un format de dades estàndard per a objectes digitals ha de complir aquestes tres condicions:

1. Una comunitat gran ha de fer servir el format de dades durant un període de temps considerable. Convertir en obsolet un format de dades usat per una extensa comunitat tindrà una influència negativa en la reputació de l'organització que va crear el format de dades. L'organització probablement tindrà en compte la comunitat d'usuaris a l'hora de redissenyar el format de dades.

2. Les especificacions del format de dades s'han de trobar en el domini públic o han d'estar publicades i avalades per una organització creadora d'estàndards (SDO), com l'ISO.
3. Una àmplia gamma de sistemes importants ha d'utilitzar el format. Per exemple, un nombre significatiu de dispositius de captura d'imatges, així com de sistemes de processament d'imatges, ha de fer servir un format de dades estàndard d'imatges digitals. La funcionalitat entre plataformes del format de dades també és una característica d'aquest requisit.

Quant als tipus específics d'objectes digitals, concretament les imatges digitals raster, es poden formular tres requisits més per a un format de dades estàndard. Aquests requisits es basen en el principi que el format de dades ha de permetre la creació d'imatges digitals raster d'alta qualitat:

4. La compressió de dades no està permesa per dues raons. La compressió de dades provocarà una pèrdua de qualitat d'imatge, i una imatge digital comprimida té un risc més gran d'esdevenir il·legible que no pas una de no comprimida. Ara aclarirem aquestes dues qüestions. Les imatges raster solen ser molt grans, per això s'apliquen algorismes de compressió de manera generalitzada a fi de reduir l'emmagatzematge de les dades. La majoria d'algorismes de compressió per a imatges raster es basa en el principi que l'ull humà no és capaç de distingir els diferents colors que apareixen en una imatge. Si es dona el mateix codi a colors molt similars de l'espectre, es pot reduir el nombre de codis de dades requerit i, per tant, la mida del fitxer. En gairebé tots els casos la compressió d'imatges digitals de fotografies implicarà una pèrdua de qualitat. La compressió de dades eficient és una qüestió de disseny molt important pel que fa als formats de dades per a imatges de mapes de bits desenvolupats en els últims temps. L'estàndard JPEG 2000, que ha estat desenvolupat recentment, n'és un bon exemple.<sup>24</sup> Un bit corrupte en un fitxer d'imatge comprimit es tradueix en una «imatge morta»; en canvi, el més probable és que un bit corrupte en una imatge no comprimida només sigui un «píxel mort». Així doncs, una imatge raster no comprimida es considera més duradora que una de comprimida, ja que la primera segurament es podrà seguir interpretant encara que alguns bits s'hagin alterat amb el pas del temps.
5. Un format de dades durador per a imatges digitals raster hauria de contenir eines per emmagatzemar metadades de preservació. La qualitat i la granularitat de les metadades és un factor important per a l'ús futur del format de dades i de la seva longevitat.
6. Un format de dades d'imatges digitals ha de permetre la codificació de característiques específiques significatives, per exemple tots els colors, els detalls i el rang dinàmic d'un original o d'una escena capturada.

Avui dia s'està desenvolupant una quantitat considerable de formats de fitxer gràfic, i en el futur també s'introduiran nous formats. A fi de determinar quin format de fitxer compleix millor les sis condicions presentades anteriorment, s'avaluaran i es compararan diversos formats de fitxer estàndard.

Un dels requisits d'un format de fitxer digital raster durador és que ha d'existir durant un període de temps considerable. La llista de formats de fitxer que apareix a l'*Encyclopedia of Graphics File Formats*,<sup>25</sup> publicada fa uns deu anys, s'utilitza com a referència per als formats estàndard de fitxer d'imatge raster potencialment rellevants. Hi ha quatre formats de fitxer raster esmentats al llibre que encara s'usen actualment. Aquests formats de fitxer són el TIFF (Tagged Image File Format), el JPEG (Joint Photographers Expert Group), el GIF (Graphic File Format) i el PNG (Portable Network Graphics).

Tres d'aquests formats de fitxer mencionats a l'enciclopèdia es fan servir amb força regularitat, concretament el TIFF, el JPEG i el GIF. El JPEG compta amb una comunitat d'usuaris molt gran, perquè l'utilitzen tots els navegadors web estàndard. Els navegadors web també fan servir el GIF, però aquest format de fitxer utilitza un algorisme de compressió patentat i no s'usa de manera tan generalitzada a la xarxa. El format PNG sembla molt adequat per funcionar com a format estàndard de fitxer per a imatges digitals duradores, principalment perquè va ser dissenyat per un grup independent. Però el PNG aplica per defecte un algorisme de compressió (sense pèrdua) i la comunitat d'usuaris de PNG no és gaire gran. A la taula 1 es mostra fins a quin punt els quatre formats de fitxer compleixen les condicions de durabilitat.<sup>26</sup>

	Requisits dels fitxers raster	TIFF	JPEG	GIF	PNG
1	És usat per una extensa comunitat durant un període de temps llarg	+	+	+	-
2	Hi ha publicada l'especificació del format de fitxer	+	+	+	+
3	És utilitzat per una gran varietat d'aplicacions	+	+	+	+
4	Utilitza imatges no comprimides / imatges d'una sola pàgina	+	-	-	-
5	Eines per a metadades de preservació	+	-	-	+
6	Permet la «captura total d'informació»	+	-	-	+

**Taula 1. Especificacions de durabilitat i formats de fitxer digital raster.**

El format de fitxer d'imatge TIFF sembla l'estàndard més durador per a la codificació d'imatges digitals que són substituïts digitals d'alta qualitat de fotografies històriques.

### 3.1.1. El format de fitxer d'imatge TIFF

El format de fitxer d'imatge raster TIFF (Tagged Image File Format) compleix els tres primers requisits per a objectes digitals duradors esmentats anteriorment. Aquest format de fitxer és independent de hardware i software i fa més de deu anys que funciona. Totes les iniciatives de conversió digital del sector de patrimoni cultural que aspiren a crear fitxers mestres digitals d'alta qualitat utilitzen aquest format per emmagatzemar imatges raster. L'especificació de la versió més recent del format de dades TIFF es troba disponible gratuïtament a la pàgina web d'Adobe.<sup>27</sup> La versió 6 de TIFF es va posar a disposició pública l'any 1992. L'especificació TIFF original va ser llançada al mercat el 1986 per l'Aldus Corporation, que més tard va ser adquirida per Adobe, com un mètode estàndard d'emmagatzematge d'imatges en blanc i negre creades per escàners i programes d'autoedició. La funcionalitat de les versions subsegüents del format de fitxer d'imatge raster TIFF va millorar enormement. La versió TIFF 6.0 fa servir la codificació del color, mètodes de compressió i metadades. No hi ha cap indici que s'hagi de publicar una nova versió del format de fitxer. Si es publicués, és previsible que el nou format fos compatible amb les versions anteriors. Segons rumors a Internet, en arxius de grups de discussió sobre gràfics, Adobe no va quedar satisfet amb el llegat de l'estàndard TIFF en adquirir Aldus, perquè és un competidor de l'estàndard PDF desenvolupat per Adobe. Però Adobe no pot abandonar l'ús de l'estàndard TIFF perquè la comunitat d'usuaris que l'empra és molt gran.

L'èxit del TIFF com un estàndard àmpliament utilitzat per a imatges digitals raster es deu a la seva naturalesa extensible i a l'ús de nombrosos esquemes de compressió de dades. D'aquesta manera, els programadors poden personalitzar el format per adaptar-lo a qualsevol necessitat concreta d'emmagatzematge de dades. Tant les imatges comprimides com les no comprimides es poden codificar amb l'estàndard TIFF i també és possible

emmagatzemar més d'una imatge en un format de fitxer formatat segons l'estàndard TIFF. L'estàndard TIFF permet incloure una quantitat il·limitada d'informació privada o per a propòsits especials, per exemple metadades de preservació. Això significa que els trets que escurcen la longevitat de les imatges digitals també formen part d'aquest format de fitxer estandarditzat.

Un inconvenient de l'estàndard TIFF és que els navegadors web no l'utilitzen. Cal una aplicació d'ajuda, un connector o la conversió a un altre format de fitxer (per exemple el JPEG) perquè el navegador web pugui processar una imatge. Un fitxer TIFF no pot tenir més de quatre gigabytes de dades raster.

## Base TIFF

L'especificació de la versió 6.0 està dividida en dues parts: la base TIFF i les extensions TIFF. La base TIFF és el nucli del format TIFF, els fonaments que tots els programadors TIFF haurien d'utilitzar en els seus productes. Les extensions TIFF són les característiques del format TIFF no necessàriament utilitzades per tots els lectors TIFF, cosa que pot dificultar un bon intercanvi i reduir la durabilitat de la imatge digital.

Un fitxer TIFF comença amb una capçalera de fitxer d'imatge de vuit bytes. Els dos primers bytes d'aquesta capçalera defineixen l'ordre de bytes utilitzat al fitxer. El segon i el tercer bytes contenen un número arbitrari en format acuradament escollit, també anomenat número màgic, que identifica el fitxer com a fitxer TIFF. Els darrers quatre bytes de la capçalera del fitxer contenen una referència a la ubicació del primer Image File Directory (IFD). Aquest «byte offset» sempre es refereix a una ubicació respecte del començament del fitxer TIFF. L'IFD és la segona secció d'un fitxer TIFF que conté camps d'informació o etiquetes que es descriuen a continuació. La tercera secció d'un fitxer TIFF conté les dades del mapa de bits. Un tret essencial de l'estàndard TIFF és que es compon de camps que contenen informació sobre les dades del mapa de bits. Els sistemes de processament d'imatges necessiten aquesta informació per reproduir la imatge. Altres camps es fan servir per emmagatzemar documentació textual sobre la imatge. La base TIFF té trenta-sis camps. A la taula 2 apareixen llistats en ordre ascendent per codi decimal.

Quant a les imatges digitals raster hi ha dos tipus d'imatges TIFF pertinents: escala de grisos i tot color. Els requisits de digitalització determinen si s'aplica una imatge a escala de grisos o una imatge a tot color. La base TIFF estableix una sèrie de camps obligatoris per a aquests tipus d'imatges que es poden trobar a la tercera columna de la taula 2.

Un fitxer base TIFF amb una imatge a escala de grisos necessita onze camps d'informació. Un fitxer base TIFF amb una imatge a tot color necessita dotze camps d'informació. El camp d'informació «SamplesPerPixel» és opcional per a imatges a escala de grisos, però és obligatori per a imatges a tot color. La ubicació real de les dades en un fitxer TIFF és força complexa, i tres camps d'informació obligatoris en una base TIFF gestionen la ubicació de les dades d'imatge. Aquests camps d'informació són «StripOffsets», «RowsPerStrip» i «StripByteCounts».

La base TIFF utilitza un petit nombre de mètodes de compressió de dades, codificats en el camp d'informació 259. Atès que les imatges duradores no haurien de ser comprimides, aquest camp d'informació ha de tenir el valor «1», que significa «no-compressió». La resta de camps d'informació de l'especificació de la base TIFF proporcionen les característiques dels píxels que conformen les dades raster. No es necessita un valor fixat per a la interpretació fotomètrica. El valor determina si «0» és reproduït com a blanc (valor = 0) o «0» és reproduït com a negre (valor = 1).

La cinquena columna de la taula 2 indica si les dades del camp d'informació es poden considerar metadades de preservació per a imatges digitals raster, bé en escala de grisos bé a tot color. Les metadades de preservació són documentació que ajuda els futurs usuaris de la imatge (persones i sistemes) a entendre i processar la imatge. Com més informació hi hagi millor, però els camps d'informació que no són rellevants per al tipus d'imatge o els camps d'informació que obstaculitzen la durabilitat digital, per exemple els camps d'informació que fan servir documents multipàgina, així com camps d'informació classificats com a «no recomanables per a l'intercanvi general» segons l'estàndard TIFF 6.0, estan exclosos d'aquesta llista.

La taula 2 presenta vint-i-quatre camps d'informació de baseline TIFF que observen la durabilitat digital. L'aplicació de cinc camps d'informació dificulta la durabilitat digital i set camps d'informació no són rellevants per a la durabilitat digital.

<i>Nom d'etiqueta</i>	<i>Codi decimal</i>	<i>Obligatori per a imatges a escala de grisos / a tot color</i>	<i>Valor (si és aplicable per a imatges a escala de grisos i a tot color)</i>	<i>Metadades de preservació</i>	<i>L'ús n'obstaculitza la durabilitat</i>
NewSubfileType	254				X
SubfileType	255				X
Imagewith	256	X	Número de píxels	X	
Imagelength	257	X	Número de píxels	X	
BitsPerSample	258	X	8 per a escala de grisos i 8 8 8 per a tot color	X	
Compression	259	X	1 = «no comprimida»	X	
Photometric Interpretation	262	X	0 o 1 per a escala de grisos / 2 per a tot color	X	
Thresholding	263				
CellWidth	264				
CellLength	265				
FillOrder	266				
Image Description	270			X	
Make	271		Fabricant escàner	X	
Model	272		Model escàner	X	
StripOffsets	273	X		X	
Orientation	274		Base TIFF només usa valor «1»		
SamplesPer Pixel (required for full colour images)	277	X	«1» per a imatges a escala de grisos (opcional) «3» per a imatges a tot color (RGB)	X	
RowsPerStrip	278	X		X	
StripByteCounts	279	X		X	
MinSampleValue	280		Conté mín. / màx. valors píxel per a estadístiques	X	
MaxSampleValue	281			X	
XResolution	282	X	Número de píxels per unitat de resolució (= etiqueta 296)	X	
YResolution	283	X	Número de píxels per unitat de resolució (= etiqueta 296)	X	
Planar Configuration	284		Base TIFF només usa valor «1»		
FreeOffsets	288				X
FreeByteCounts	289				X
GrayResponse Unit	290			X	
GrayResponse Curve	291			X	
ResolutionUnit	296	X	1 (= cap) o 2 (= polzada) o 3 (= cm)	X	
Software	305			X	



DateTime	306		data / hora de creació de la imatge	X	
Colormap	320		Només rellevant per a imatges de paleta de color		
Artist	315			X	
HostComputer	316			X	
ExtraSamples	338				X
Copyright	33432		Informació sobre els drets d'autor	X	

**Taula 2. Avaluació dels camps d'informació Baseline TIFF 6.0.**

En principi, el format de fitxer base TIFF versió 6.0 compleix els sis requisits per a imatges raster duradores. Les extensions TIFF 6.0 s'analitzen seguidament a fi de determinar si existeixen millors mètodes per a la formulació de metadades de preservació i la codificació de totes les característiques significatives obligatòries.

#### Extensions TIFF

Les extensions TIFF són característiques TIFF que poden no ser utilitzades per tots els lectors TIFF. L'estàndard TIFF 6.0 oficial conté diverses extensions TIFF i n'hi ha unes quantes que estan publicades independentment de l'estàndard TIFF 6.0 publicat oficialment.

L'especificació TIFF 6.0 publicada oficialment conté quatre grups d'extensions. Un grup d'extensions fa referència als mètodes de compressió de dades. Un altre grup s'ocupa d'una alternativa per a l'organització de la imatge en mosaics en lloc de tires. El tercer tipus d'extensió millora la qualitat d'un tipus específic d'imatges, concretament les imatges a mitja tinta. L'única extensió rellevant per a la durabilitat de les imatges digitals és l'extensió TIFF 6.0 per a una millor gestió del color. L'espai de color CIELAB, utilitzat per l'extensió en l'especificació de la versió TIFF 6.0, té una excel·lent aplicabilitat per a la manipulació, independent de qualsevol dispositiu, d'imatges de tons continus. Per a substituïts digitals de fotografies en color, per exemple, aquest espai de color d'alta qualitat és molt important. Si s'usa l'espai de color CIELAB, l'etiqueta d'informació 262 (interpretació fotomètrica) hauria de tenir el valor «8». Aquest valor té el significat «1976 CIE L\*a\*b».

A més de les extensions que formen part de l'estàndard oficial TIFF també hi ha una sèrie d'extensions publicades de manera independent. N'esmentem alguns exemples: una especificació TIFF especial per a aplicacions GIS, una especificació TIFF que utilitza el mètode de compressió JPEG i una extensió TIFF com a format d'intercanvi de preimpresió.

En principi l'especificació TIFF compleix els criteris per funcionar com estàndard per a la creació d'imatges digitals duradores, però no s'haurien de fer servir les característiques que n'obstaculitzen la durabilitat, com la compressió i la inclusió d'etiquetes privades no estàndard. Això vol dir que l'estàndard TIFF és força relaxat i tolerant quan es tracta de complir amb la norma.<sup>28</sup>

Hi ha una extensió al format TIFF 6.0 que gaudeix de la categoria de norma internacional ISO i que es pot considerar un format de fitxer durador: el TIFF/EP (TIFF/Electronic Photography); el format de dades d'imatge TIFF/EP es desenvolupa com un estàndard per a la codificació d'imatges de càmeres electròniques i està basat en la versió 6.0 TIFF, però el TIFF/EP conté diverses etiquetes noves.<sup>29</sup> El TIFF/EP també es coneix amb el nom de

EXIF (Exchangeable Image File Format). Aquest format té un paper important en la formulació de metadades, en parlarem més endavant.

L'especificació TIFF 6.0 es pot considerar duradora, sempre que s'apliqui d'una manera específica. No s'ha d'utilitzar cap mètode de compressió i no s'han de crear imatges de múltiples pàgines. L'especificació TIFF 6.0 conté una sèrie d'etiquetes d'informació que permeten l'emmagatzematge de metadades de preservació. En canvi, a la base TIFF l'emmagatzematge de metadades de preservació és bastant limitat. Només es disposa d'una quantitat limitada de camps d'informació per a l'emmagatzematge de metadades de preservació. A més a més, l'abast i l'objectiu d'aquests camps d'informació es descriu de manera força vaga. Per exemple, els requisits del camp d'informació «ImageDescription» no s'especifiquen. La codificació de totes les característiques significatives també és problemàtica, perquè la base TIFF no aplica la codificació acurada de la informació de color. L'extensió TIFF 6.0 conté un espai de color d'alta qualitat (CIE LAB) que aplica la codificació acurada de colors.

### 3.2. Metadades de preservació rellevants per a la longevitat d'imatges digitals raster

Les metadades de preservació constitueixen una documentació que té un paper important quant a l'accés a llarg termini a objectes digitals. El model de referència OAIS (vegeu la secció 2.2.) es pot fer servir per a la formulació d'elements de dades que els usuaris del futur, la comunitat designada, necessitaran a fi d'entendre i processar l'objecte digital. Actualment existeix un cert nombre de conjunts d'elements de dades, o esquemes, que poden tenir un paper determinant en la longevitat de les imatges digitals raster. L'esquema SEPIADES,<sup>30</sup> orientat a la catalogació d'un gran nombre de característiques de col·leccions fotogràfiques, n'és un exemple. També es poden emprar estàndards per a la representació i la comunicació d'informació bibliogràfica, com el MARC,<sup>31</sup> per a la formulació de metadades de preservació de material «no-llibre». L'objectiu del projecte PREMIS és desenvolupar mètodes i recomanacions millors per a la implementació de metadades de preservació per a objectes digitals.<sup>32</sup> Les conclusions del projecte PREMIS apareixeran al final de 2004.

En general els elements de dades que formen part d'un esquema de metadades es poden emmagatzemar de tres maneres. En primer lloc, els elements de dades es poden emmagatzemar a la capçalera del fitxer d'imatge, com el camp d'informació en un fitxer TIFF (vegeu la secció 3.1.1.). Els elements de dades també es poden emmagatzemar en el sistema de fitxer mitjançant els noms dels directoris i els fitxers d'imatges. En tercer lloc, els elements de dades poden ser emmagatzemats en una base de dades independent.

Una prometedora construcció estandarditzada per expressar metadades de preservació és l'«embolcall» (*wrapper*) METS (vegeu la secció 2.3.1.). Un document METS és un document formatat en XML que conté totes les metadades rellevants d'un objecte digital o d'un conjunt d'objectes digitals relacionats, així com dels seus originals analògics.

En la propera secció d'aquest treball ens ocuparem amb més detall d'un tipus concret de metadades de preservació, les metadades tècniques per a imatges digitals raster. Les metadades tècniques només són un subgrup del conjunt d'eines complet de preservació necessari per a l'accés a llarg termini, però sovint se l'ha anomenat la primera línia de defensa contra la pèrdua d'accés. La documentació tècnica és rellevant pel que fa a dos camps estretament relacionats. En primer lloc, les metadades tècniques faciliten l'intercanvi fluid d'imatges digitals entre sistemes diferents. En segon lloc, un procés futur de migració mitjançant la còpia d'imatges a nous formats es beneficia de les metadades tècniques estandarditzades.

### 3.2.1. Metadades tècniques per a imatges digitals raster

El camp d'informació de l'estàndard TIFF conté diversos elements de dades que es poden considerar igual de rellevants per a l'accés a llarg termini (vegeu la taula 2). El conjunt de camps d'informació s'amplia gràcies a altres estàndards. Si s'utilitza un format de fitxer d'imatge no-TIFF es necessitarà una construcció alternativa per expressar les metadades de preservació. Hi ha dos estàndards relacionats amb l'estàndard TIFF que són importants per a l'expressió de metadades tècniques d'imatges digitals raster: l'estàndard EXIF i l'esborrany de l'estàndard NISO Z39.87.

L'estàndard EXIF i l'especificació relacionada DCF sorgeixen de la comunitat de fabricants de càmeres digitals.<sup>33</sup> Aquest estàndard és rellevant per a imatges creades digitalment. L'esborrany de l'estàndard NISO Z39.87 Technical Metadata for Digital Still Images l'ha iniciat la comunitat de patrimoni cultural i està dirigit principalment a la formulació de metadades tècniques de substituïts digitals d'originals analògics.<sup>34</sup>

#### EXIF i DCF

*EXIF* significa 'Exchangeable Image File Format', i és un estàndard per emmagatzemar informació d'intercanvi en fitxers d'imatge, especialment en les que utilitzen la compressió JPEG. L'especificació DCF (Design Rule for Camera File system) es va redactar a fi de simplificar l'intercanvi de fitxers d'imatges i de fitxers relacionats en càmeres digitals i altres aparells. El DCF formula els noms de fitxers i la disposició de directoris. EXIF emmagatzema metadades al començament dels fitxers i utilitza l'espai de color estàndard sRGB.<sup>35</sup> La majoria de càmeres digitals actuals fan servir el format EXIF i l'especificació DCF, per exemple les càmeres digitals fabricades per Canon, Kodak, Sony i Olympus. El format ha estat desenvolupat per l'Associació d'Indústries de Tecnologia de la Informació i Electrònica del Japó (JEITA). El format de fitxer d'imatge EXIF es va crear amb l'objectiu de convertir-se en un format comú per als fitxers d'imatge emprats amb càmeres fotogràfiques digitals i altres aparells relacionats, per tal de fer aquests productes més convenientes per als usuaris finals. La creixent popularitat de les càmeres digitals ha fet que augmentés la demanda de fitxers d'imatges intercanviables, que permetran que les imatges captades per una càmera es puguin veure en una altra, o que es puguin imprimir directament.

La versió 2.2 EXIF, elaborada l'abril de 2002, especifica l'estructura dels fitxers de dades d'imatge i dels camps d'informació o etiquetes emprades per l'estàndard. L'estàndard EXIF amplia els camps d'informació TIFF obligatoris amb etiquetes EXIF addicionals. La figura 3 conté les metadades tècniques segons l'estàndard EXIF creades per una càmera digital corrent. Les metadades es mostren en forma de fitxer XML. L'estàndard EXIF conté moltes més etiquetes, però la càmera no les fa servir totes. D'acord amb l'estàndard EXIF, el codi «1» per a l'etiqueta <ColorSpace> es refereix a l'espai de color sRGB.

L'estàndard EXIF especifica les tres etiquetes següents per inserir dates; cadascuna té un significat específic:

- <DateTime> registra la data i l'hora de l'actualització del fitxer, com un segell de temps del fitxer.
- <DateTimeOriginal> registra la data i l'hora en què es va captar la imatge.
- <DateTimeDigitised> indica la data i l'hora en què es van crear les dades digitals.

Per a les càmeres digitals, en molts casos els continguts de les tres etiquetes són idèntics, com es pot veure en la figura 3.

```

<Exif>
  <CameraManufacturer>Canon</CameraManufacturer>
  <CameraModel>Canon PowerShot A70</CameraModel>
  <Orientation>top, left</Orientation>
  <XResolution>1/180</XResolution>
  <YResolution>1/180</YResolution>
  <ResolutionUnit>Inches</ResolutionUnit>
  <DateTime>2004:07:21 12:51:34</DateTime>
  <YCBCrPositioning>Centered</YCBCrPositioning>
  <ExposureTime>1/60 sec</ExposureTime>
  <FNumber>4.0</FNumber>
  <ExifVersion>0220</ExifVersion>
  <DateTimeOriginal>2004:07:21 12:51:34</DateTimeOriginal>
  <DateTimeDigitized>2004:07:21 12:51:34</DateTimeDigitized>
  <BitsperSample>2</BitsperSample>
  <ExposureBiasValue>0.0</ExposureBiasValue>
  <MaxApertureValue>4.0</MaxApertureValue>
  <MeteringMode>Multi Segment</MeteringMode>
  <Flash>Unknown</Flash>
  <FocalLength>11.10 mm</FocalLength>
  <FlashPixVersion>0100</FlashPixVersion>
  <ColorSpace>1</ColorSpace>
  <Width>1536 pixels</Width>
  <Height>2048 pixels</Height>
  <SensingMethod>One-chip color area sensor</SensingMethod>

```

**Figura 3. Metadades EXIF en format XML. (Nota: EXIF no conté gaires camps d'informació més a part dels que apareixen en la figura.)**

Els fitxers d'imatges comprimits s'enregistren com a JPEG<sup>36</sup> amb segments de marcador d'aplicacions inserits. Els fitxers no comprimits s'enregistren en format TIFF versió 6.0. La informació relacionada d'atributs tant per a fitxers comprimits com per a no comprimits s'emmagatzema en el format d'informació d'etiqueta definit segons la versió TIFF 6.0. La informació específica del sistema de la càmera, no definida en TIFF, s'emmagatzema en etiquetes privades enregistrades per a EXIF. El fet que EXIF utilitzi el format de fitxer d'imatge no comprimit TIFF no significa en absolut que els dispositius de captura digital que apliquen l'estàndard EXIF puguin crear imatges digitals no comprimides. Moltes càmeres digitals només poden processar fitxers d'imatges comprimides.

La norma DCF està orientada a la creació d'un entorn d'usuari en el qual els consumidors d'imatges digitals puguin combinar productes més lliurement i intercanviar suports sense problemes. La DCF especifica normes per enregistrar, llegir i manipular fitxers d'imatges i altres fitxers relacionats que s'utilitzen en càmeres digitals o en altres aparells. La DCF és aplicable a productes per escriure fitxers d'imatges en un mitjà d'emmagatzematge intercanviable. Segons la DCF el nom dels fitxers i directoris només pot contenir dígit, els vint-i-sis caràcters de l'alfabet llatí (sense distinció entre majúscules i minúscules) o el signe «\_» (guió baix).

La DCF consta de tres especificacions:

- *Media specification.* (Especifica l'estat de les dades en un mitjà d'emmagatzematge)
- *Writer specification.* (Especifica la funció d'enregistrament, per exemple mitjançant una càmera digital)
- *Reader specification.* (Especifica la funció de reproducció)

L'estàndard per a suports DCF defineix l'estructura del directori i els noms del directori als dispositius que emmagatzemen imatges digitals. El directori amb el nom DCIM (Digital Camera Images) directament a sota del directori arrel s'anomena directori arrel d'imatges DCF. Els directoris que emmagatzemen objectes DCF s'anomenen directoris DCF.

### ***NISO Z39.87 Diccionari de dades-Metadades tècniques per a imatges digitals***

```
<?xml version="1.0" encoding="UTF-8" ?>
<METS_Profile ...>
...
<extension_schema>
  <name>The MIX Technical Metadata for Still Images XML Schema</name>
  <URI>http://www.loc.gov/standards/mix/mix.xsd</URI>
</extension_schema>
...
<mix:mix>
  <mix:BasicImageParameters>
    <mix:Format>
      <mix:MIMEType>image/tiff</mix:MIMEType>
      <mix:ByteOrder>big-endian</mix:ByteOrder>
      <mix:Compression>
        <mix:CompressionScheme>1</mix:CompressionScheme>
      </mix:Compression>
      <mix:PhotometricInterpretation>
        <mix:ColorSpace>2</mix:ColorSpace>
      </mix:PhotometricInterpretation>
    </mix:Format>
  </mix:BasicImageParameters>
  ...
  <mix:ImageCreation>
  ...
  <mix:ImagingPerformanceAssessment>
  ...
</mix:mix>
...
</METS_Profile>
```

**Figura 4. Elements de dades de l'estàndard NISO Z39.87 expressats com a esquema XML «MIX» i inclosos en un embolcall METS.**

L'objectiu del diccionari de dades NISO Z39.87 és definir un estàndard d'elements de dades d'imatges digitals. El diccionari de dades ha estat dissenyat per facilitar la interoperabilitat entre sistemes, serveis i software, així com per permetre la gestió a llarg termini i l'accés continuat a col·leccions d'imatges digitals. Les normes NISO Z39.87 estan destinades a institucions culturals, editors, propietaris de drets d'autor i altres organitzacions implicades en la digitalització de materials visuals de col·leccions d'arxiu. Els elements de dades estan estructurats per incorporar mètodes associats a la còpia digital de fotografies com l'ús d'objectius tècnics i tècniques per a fotografies digitals directes d'escenes originals.

El diccionari de dades NISO Z39.87 inclou quatre categories de funcions:

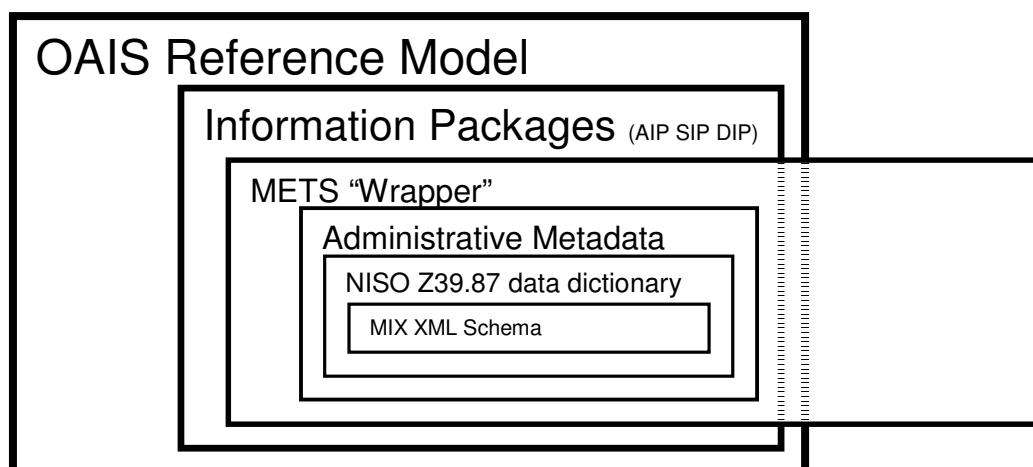
- Basic image parameters ('paràmetres d'imatge bàsics'), registren informació crucial per mostrar una imatge visible.

- Image creation data elements ('elements de dades per a la creació d'imatges'), registren informació important per entendre l'entorn tècnic en el qual va ser capturat un fitxer d'imatge digital.
- Imaging performance assessment ('valoració de la realització de la imatge'). Els elements de dades per a la valoració de la realització de la imatge registren informació que permet l'avaluació de la qualitat de la imatge digital, o la precisió de sortida.
- Change history ('canvi d'història'). Els elements de dades de canvi d'història registren informació sobre el procés aplicat a una imatge durant el seu cicle de vida.

Els elements de dades del diccionari de dades NISO Z39.87 afegixen i amplien les metadades tècniques disponibles en altres estàndards, per exemple el format de fitxer d'imatge TIFF versió 6.0.

També hi ha disponible un esquema XML que conté els elements de dades de l'estàndard NISO Z39.87. Aquest esquema XML, NISO Metadata for Images in XML Schema (MIX), proporciona un format per a l'intercanvi i/o emmagatzematge de metadades tècniques.<sup>37</sup> La figura 4 conté una petita part d'un document METS que fa referència a l'esquema MIX XML. L'element «Format», un dels elements de la secció «Basic Image Parameters», consisteix en una sèrie d'elements de dades. Segons el diccionari de dades NISO Z39.87, el valor «1» per a <CompressionScheme> significa «no comprimit» i el valor «2» per a <ColorSpace> representa «RGB».

D'acord amb la terminologia del model de referència OAIS, les metadades tècniques formen part de la informació de representació. En un dipòsit de preservació aquesta informació es convertirà en part d'un *Information Package* OAIS. L'especificació METS es compon d'un nombre d'objectes que es poden considerar una implementació dels *Information Packages* OAIS SIP, AIP i DIP (vegeu la figura 1). El METS és un esquema XML que es pot utilitzar per a la codificació de metadades descriptives, administratives i estructurals referides a objectes d'una biblioteca digital. El diccionari de dades NISO Z39.87 es pot considerar una part de les metadades administratives de l'especificació METS. Atès que no existeix una implementació de l'esquema XML del diccionari de dades NISO Z39.87 amb el nom «MIX» (vegeu la figura 4), els tres estàndards (OAIS, METS i NISO Z39.87), cadascun rellevant a un nivell específic per a la preservació digital d'imatges digitals raster, es poden combinar. Aquesta combinació s'il·lustra en la figura 5. Un esquema XML de l'especificació EXIF també encaixaria a la secció de metadades administratives d'un estàndard METS.



**Figura 5. Relació entre el model de referència OAIS, l'esquema METS i el diccionari de dades NISO Z39.87.**

## Conclusió

La preservació digital és un àmbit relativament nou. Existeix un conjunt de normes i principis a partir dels quals es poden elaborar solucions per possibilitar la durabilitat d'imatges digitals raster. Els dos factors principals que determinen l'accés a llarg termini a imatges digitals raster són l'aplicació d'un format de fitxer d'imatge estandarditzat i la creació de metadades de preservació.

A partir d'una anàlisi dels requisits necessaris per a un format estàndard de fitxer d'imatge i d'una revisió dels formats de fitxer d'imatge que s'han estat utilitzant durant els darrers deu anys, actualment el format base TIFF versió 6.0 sembla el format de fitxer d'imatge raster més durador.

Quant a les metadades de preservació, se'n poden distingir diverses categories. Per a la formulació i expressió de metadades tècniques hi ha unes quantes solucions disponibles. L'estàndard EXIF, aplicat per una gran quantitat de càmeres fotogràfiques digitals, conté una sèrie d'elements de dades rellevants per a l'accés en el futur a imatges digitals. Les institucions de patrimoni cultural que aspiren a oferir accés a llarg termini a imatges digitals van desenvolupar específicament el diccionari de dades NISO Z39.87 —metadades tècniques per a imatges fotogràfiques digitals. Els elements de dades d'aquest estàndard NISO provisional es poden emmagatzemar al format de dades XML mitjançant l'esquema MIX XML que pot formar part d'un document METS.

Actualment existeixen els contorns d'una solució per a la preservació digital d'imatges digitals raster. El model de referència OAIS és important com a marc conceptual per al disseny d'un arxiu digital. L'actuació proactiva i el compromís de mantenir les imatges digitals raster accessibles a llarg termini també és important. Aquest treball conté punts de partida pertinents per a la implementació d'una estratègia sòlida que minimitzi el risc que les imatges digitals raster deixin de ser accessibles en un futur pròxim.

- 
1. Cita de S. Aubenas, «The Photograph in Print. Multiplication and Stability of the Image», a M. Frizot (ed), *A New History of Photography*, Köln, Könemann Verlagsgesellschaft, 1998, p. 225.
  2. Cita de Ching-Chih and K. Kiernan (eds.), *Report of the DELOS-NSF Working Group on Digital Imagery for Significant Cultural and Historical Materials. Prepared for the National Science Foundation (NSF), Digital Library Initiative and the European Union under the Fifth Framework Programme by the Network of Excellence for Digital Libraries (DELOS)* 2003. Disponible en línia a <http://delos-noe.iei.pi.cnr.it/activities/internationalforum/Joint-WGs/digitalimaging/DigitalImaging.pdf>.
  3. Vegeu R. Kirsch, «SEAC and the Start of Image Processing at the National Bureau of Standards», a *Annals of the History of Computing*, IEEE, vol. 20, 1998, p. 7-13.
  4. J. D. Murray, J. D. i W. vanRyper, *Encyclopedia of Graphics File Formats*, Sebastopol, CA, O'Reilly & Associates, 1994.
  5. Ref. C. W. Browne i B.J. Sheperd, *Graphics File Formats. Reference and Guide*, Greenwich, Manning, 1995.
  6. Les dades raster també s'anomenen dades *bitmapped*.
  7. Les dades gràfiques vectorials són un exemple de dades geomètriques.
  8. Vegeu F. Frey i J. Reilly, *Digital Imaging for Photographic Collections. Foundations for Technical Standards*. Rochester, Image Permanence Institute, Rochester Institute of Technology, 1999. Disponible en línia a <http://www.rit.edu/ipi>, p. 21.
  9. J. Rothenberg, «Ensuring the Longevity of Digital Documents», a *Scientific American*, gener 1995, p. 42-47.
  10. Per a més informació sobre preservació digital, vegeu M. Jones i N. Beagrie, *Preservation Management of Digital Materials*, London, The British Library, 2001.
  11. *Reference Model for an Open Archival Information System (OAIS)*, publicat pel Consultative Committee for Space Data Systems, CCSDS 650.0-B-1, Blue Book, gener 2002. Disponible com a norma ISO 14721:2003. Disponible en línia a <http://www.classic.ccsds.org/documents/pdf/CCSDS-650.0-B-1.pdf>.
  12. La pàgina web del grup de treball RLG/NARA sobre la certificació de dipòsits digitals es pot trobar a <http://www.rlg.org/longterm/certification.html>.
  13. Vegeu, per exemple, M. Day, «Metadata for Digital Preservation: A Review of Recent Development», a *Proceedings of the 5th European Conference for Digital Libraries*, Darmstadt, Springer, 2001, p. 161-172.

- 
14. Per a una funció específica de les metadades, descobriment de recursos, s'usa un estàndard a gran escala, el «Dublin Core Metadata Element Set» (DCMES), vegeu <http://www.dublincore.org>. Els quinze elements de dades del DCMES s'apliquen a un gran nombre de projectes i iniciatives a fi de permetre el descobriment d'objectes a Internet.
  15. Per a més informació sobre el METS, vegeu <http://www.loc.gov/standards/mets>.
  16. Depenent del seu ús, un document METS es podria fer servir com a *Submission Information Package* (SIP), *Archival Information Package* (AIP) o *Dissemination Information Package* (DIP), tal com queden explicats en el model de referència OAIS.
  17. Vegeu <http://www.w3c.org/XML/Schema>.
  18. Es pot trobar més informació sobre el GDFR i enllaços a referències a <http://hul.harvard.edu/gdfr>.
  19. El sistema PRONOM és accessible a través de <http://www.nationalarchives.gov.uk/pronom/>.
  20. El software JHOVE es troba disponible públicament sota la GNU General Public License (GPL) de la pàgina web del projecte <http://hul.harvard.edu/jhove>.
  21. Vegeu C. Dollar, *Authentic Electronic Records: Strategies for Long-term Access*, Chicago, Cohasset Associates, 2000, p. 215.
  22. ISO 18921:2002 *Imaging Materials-Compact Discs (CD-ROM)-Method for Estimating the Life Exptency Based on the Effects of Temperature and Relative Humidity*, International Organisation for Standardisation.
  23. Vegeu F. Byers, *Care and Handling of CDs and DVDs: A Guide for Librarians and Archivists*, Washington, Council on Library and Information Resources, 2004. Disponible en línia a <http://www.clir.org/pubs/reports/pub121/pub121.pdf>.
  24. El lloc web de l'estàndard JPEG 2000 es troba a <http://www.jpeg.org/JPEG2000.html>.
  25. Vegeu la nota 4. S'utilitza tant la primera com la segona edició de l'enciclopèdia. La primera edició va ser publicada el 1994, la segona, el 1996. Fins ara no se n'ha publicat cap actualització.
  26. Els formats de fitxer gràfic que no estan disponibles en diverses plataformes, com el Microsoft Windows Bitmap (BMP), i els formats de fitxer gràfic que formen part d'un sistema d'imatgeria, com el Kodak Photo CD (PCD), estan exclosos de la selecció.
  27. Les especificacions de la versió 6.0 del format de fitxer TIFF es poden trobar a <http://partners.adobe.com/asn/tech/tiff/>.
  28. El format de fitxer d'imatge raster PNG (Portable Network Graphics), una recomanació de la W3C es considera un format superior, ben dissenyat i de fàcil accés, però després d'un començament prometedor, aquest format no va aconseguir una gran comunitat d'usuaris. Vegeu <http://www.w3c.org/Graphics/PNG>.
  29. Vegeu ISO 12234-2:2001 *Electronic Still-picture Imaging-Removable Memory-Part 2: TIFF/EP 2Image Data Format*, International Organisation for Standardisation.
  30. Per a més informació sobre SEPIADES, vegeu E. Klijn (ed.), *SEPIADES. Recommendations for Cataloguing Photographic Collections*, Amsterdam, European Commission on Preservation and Access, 2003.
  31. Vegeu <http://www.loc.gov/marc>.
  32. PREMIS significa 'PREservation Metadata: Implementation Strategies'. Vegeu <http://www.oclc.org/research/projects/pmwg/>. El projecte es basa en el model de referència OAIS.
  33. Es pot trobar més informació sobre EXIF i DCF a <http://www.exif.org>.
  34. El NISO Z39.87-2002 Data Dictionary-Technical Metadata for Digital Still Images es pot descarregar des de [http://www.niso.org/standards/resources/Z39\\_87\\_trial\\_use.pdf](http://www.niso.org/standards/resources/Z39_87_trial_use.pdf).
  35. Es pot trobar més informació sobre l'espai de color sRGB a <http://www.w3.org/Graphics/Color/sRGB.html>.
  36. El mètode de compressió JPEG es defineix com a estàndard ISO/IEC 10918-1. Vegeu <http://www.jpeg.org/jpeg/index.html>.
  37. Aquest esquema XML s'anomena NISO Metadata for Images in XML, i s'abreua com a MIX. Vegeu <http://www.loc.gov/standards/mix>.