

# LA DEFINICIÓ D'UN MAPA D'INFORMACIÓ CONCEPTUAL PER A LA GESTIÓ DE L'ARXIU FOTOGRÀFIC DIGITAL

David Iglésias Franch

Centre de Recerca i Difusió de la Imatge (CRDI)

Les imatges numèriques s'expliquen per si soles, ja que contenen tant la informació de cada píxel de la imatge com aquella que permet interpretar-ne la globalitat en diferents dispositius: una càmera, una pantalla o una impressora. No calen intermediaris més enllà d'aquells que imposa la naturalesa digital: hardware, software, especificacions tècniques, etc. Les regles que faciliten la comunicació i la interrelació de tots els actors que intervenen en aquest escenari són conegudes per tots ells i, per tant, constitueixen una realitat sòlida tot i que canviant. Aquest context tecnològic que permet visualitzar, editar o imprimir imatges no sembla que tingui límits definits. Qualsevol funcionalitat afegida pot ser integrada en un objecte digital, sense modificar-ne l'essència fotogràfica, però ampliant-ne les potencialitats en l'ús. Aquesta **extensibilitat funcional** representa un camp adobat per a moltes disciplines, entre elles l'arxivística i la documentació.

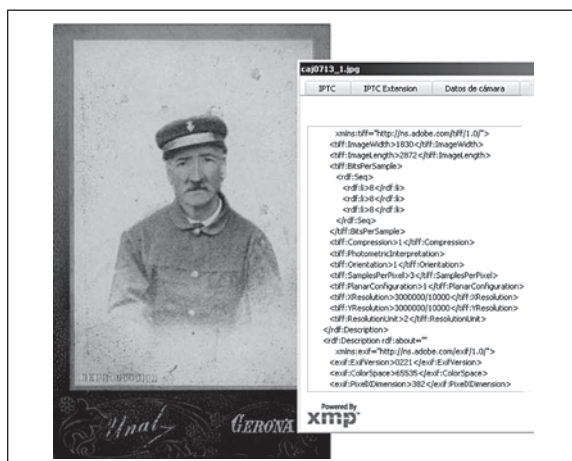
La gestió dels recursos d'informació ha experimentat un canvi significatiu per un motiu principal: la **disociació** entre l'objecte de custòdia i la informació referencial ja no és necessària. La naturalesa fisicoquímica de la fotografia no admet, per regla general, una concepció global de l'objecte informatiu, encara que sovint pot incloure un cert nivell d'informació referencial i de gestió. La imatge electrònica, en canvi, no es concep sense una informació referencial i de gestió, i no s'adapta fàcilment a la pràctica tradicional de separar la informació de la imatge d'aquella que en permet la seva gestió.

Aquesta circumstància ens porta a revisar determinades praxis pròpies de la professió per tal d'orientar-nos cap a la **sincronització de l'objecte**, amb l'objectiu d'adaptar-nos millor a l'entorn digital. És important tenir present que l'objecte de custòdia, la imatge digital, és similar en tots els arxius però que els objectius finals que es persegueixen no han de ser necessàriament coincidents en la seva totalitat. És important esclarir d'entrada quina informació tenim i reflexionar sobre quina informació necessitem per a dur a terme la nostra tasca.

Situats en aquest punt, cal procedir a una anàlisi més aprofundida per tal de conèixer la realitat del nostre entorn de treball. Aquestes són les qüestions principals a les quals caldria poder donar resposta:

- Quina informació forma part de l'objecte digital i com està estructurada?
- On es localitza aquesta informació?
- Es tracta de metadades intrínseques o extrínseques?
- La seva captura és automàtica?
- Quina dependència hi ha del software?
- Quina informació és del nostre interès i, per tant, volem que formi part de l'objecte digital?
- Quines funcionalitats aniran associades a l'objecte digital?

Per respondre a aquestes preguntes i per decidir quin entorn informatiu volem circumscriure és convenient conèixer com són les nostres imatges i tenir



Les metadades que inclou la imatge original són poques. La versió digital permet una expansió d'aquestes gairebé il·limitada. A més, són metadades normalitzades i codificades segons diferents estàndards.

Ajuntament de Girona. CRDI (Fotografia Unal)

clar quina realitat volem construir. En definitiva, cal projectar l'arxiu a partir d'uns fonaments sòlids, al marge de qualsevol circumstància temporal i sense dependències de tecnologies específiques. L'anàlisi inicial de la pròpia realitat és el primer pas per a la creació d'un **mapa conceptual d'informació**, erigit com a l'element fonamental i de base per al disseny de l'arxiu digital. És a partir d'aquest mapa que podem determinar de manera més consistent la concepció de l'objecte digital que aspirem a gestionar i, en funció dels recursos, com fer-ho. Per a la configuració d'aquest mapa caldrà centrar-se en els aspectes següents:

- Els blocs d'informació que formaran part del nostre mapa.
- Els contenidors de metadades estàndards existents.
- L'estructuració i combinació d'aquestes metadades tenint en compte qüestions sobre interoperabilitat i codificació.
- Les responsabilitats i funcions que tenim assignades que cal posar en relació als blocs de metadades i que han de determinar les necessitats de software.

## BLOCS D'INFORMACIÓ

La creació del mapa d'informació depèn del conjunt de dades que necessitem per tal de poder assumir les funcions derivades de les responsabilitats inherents a la custòdia i l'accés. Comptem d'entrada amb les dades de la imatge i les dades tècniques que en permeten la seva representació i la seva operativitat. A aquestes hi sumarem els diferents blocs de dades que determinem en funció dels nostres interessos i necessitats i principalment els blocs referents a: informació de captura, informació d'autor, informació d'arxiu, informació d'administració, informació històrica, informació d'entorn, i informació de preservació. Analitzem tot seguit els blocs d'informació principals.

**Informació primària.** La informació bàsica de la imatge és la informació a nivell de píxel, la unitat més petita en què es divideix la imatge numèrica. La codificació digital es produeix en un procés posterior a la captura en què el voltatge és enviat al codificador analògic-digital i s'assigna el valor del píxel. Aquest valor és el que determina el to de la imatge i, en funció de l'estadi de processament, el color. L'assignació d'aquest valor és clau per a la representativitat posterior de la imatge ja que una codificació a 16 bits, per exemple, ofereix més precisió i sobretot més opcions d'intervenció en el processat que una codificació a 8 bits. Com que estem parlant de mili-

ons de píxels, el codi és extens i poc intel·ligible en el seu conjunt. Per això calen un seguit de dades que el facin interpretable per a les màquines, aquestes són les metadades que formen el segon bloc.

**Informació de representació.** Aquest conjunt de metadades determinen les característiques tècniques de la imatge i en permeten la seva representació en diferents entorns. Les metadades principals són les que informen de la condició d'imatge del fitxer digital, de la seva codificació i també de la seva configuració, principalment la resolució espacial i la resolució de bit. També són imprescindibles les que permeten localitzar els diferents valors de píxel en el mapa de bits, la mida del fitxer i les que fan referència a la gestió del color, un aspecte fonamental i complex en la representació de la imatge fotogràfica. Moltes d'aquestes metadades venen condicionades per l'elecció del format de fitxer, ja que aquest determina i delimita en bona part les característiques tècniques de la imatge. Són de captura automàtica i formen part de l'encapçalament del fitxer.

**Informació de captura.** Aquesta informació ens permet sobretot conèixer les condicions de la captura. Es tracta principalment de les metadades referents a les condicions tècniques de la captura i també del processat, ja que a excepció dels fitxers RAW el processat és automatitzat segons la configuració de càmera. Les dades bàsiques de la captura són el temps d'exposició, l'obertura de diafragma, la velocitat ISO, la temperatura de color i l'ús del flaix. Entre les dades bàsiques del processat hi trobem el balanç de blancs, el contrast, la saturació i l'enfocament. També formen part d'aquest bloc les informacions sobre la data i l'hora de la captura, de gran valor per a l'arxiu. Addicionalment i segons les màquines, comptarem amb la informació geogràfica que aporten els sistemes GPS integrats. En el cas dels formats RAW, bona part d'aquesta informació es conserva en un fitxer adjunt, la qual cosa permet preservar la integritat de les dades de la captura sense cap tipus de processat. Totes aquestes metadades són de captura automàtica.

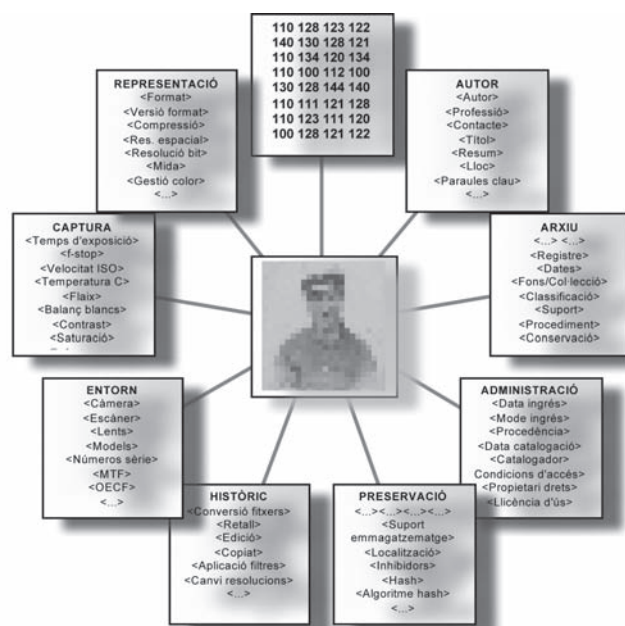
**Informació descriptiva de l'autor.** Aquestes metadades informen sobre l'autor però alhora poden incloure informació sobre l'escena aportada pel mateix autor. Encara que algunes d'aquestes metadades siguin coincidents amb les del bloc següent, cal contemplar-les inicialment de manera separada per les diferències existents en matisos semàntics i aspectes formals. Sobre l'autor, les metadades principals són el nom, la professió, l'adreça de contacte, etc. Sobre la descripció de la imatge feta per l'autor hem

de considerar principalment el títol, el resum, el lloc i les paraules clau, aquestes darreres d'escàs valor per a la descripció arxivística final. Aquestes metadades són d'inserció manual.

**Informació descriptiva d'arxiu.** Aquesta inclou informacions d'altres blocs, com les referents a continguts de la imatge (autor, títol i resum) i a la descripció de l'objecte (format, resolució, etc.), encara que en alguns casos caldrà adaptar-ne els valors per assolir-ne la normalització i incloure-hi els identificadors i descriptors temàtics. Però a més d'aquestes dades, hi ha un seguit d'informacions absolutament clau. És el cas de la gestió dels diferents registres existents que han de permetre identificar l'objecte digital, però també la imatge d'origen (quan es tracta d'una digitalització), o imatges relacionades. També són fonamentals les diferents dates rellevants per a l'accés i la gestió (com poden ser la data de captura i la del processat de la imatge original). Però el principal valor afegit en la descripció prové de la informació del context de producció que es concreta en metadades específiques com les de fons o col·lecció o les derivades d'una classificació funcional. Per últim cal considerar les metadades corresponents a la descripció de l'objecte físic, ja que en molts casos es tractarà d'una digitalització. Aquestes poden ser molt extenses i principalment n'hauríem de considerar el suport, el format, el procediment fotogràfic, el color i l'estat de conservació. Aquestes metadades són d'inserció manual.

**Informació d'administració.** Podem incloure en aquest bloc diferents tipus d'informacions i bàsicament aquelles corresponents a l'ingrés, al control de la catalogació, a l'accés i a l'ús de les imatges. Aquestes informacions es concreten en metadades com: data d'ingrés, mode d'ingrés, procedència, data de catalogació, nom del catalogador, restriccions d'accés o condicions d'accés. Però les metadades principals són aquelles que regulen la propietat intel·lectual, com poden ser: propietari dels drets, instruccions per a l'ús legal, data de domini públic, llicència d'ús, etc. Aquestes metadades són d'inserció manual.

**Informació històrica.** Es tracta del recull d'informacions referents a accions dutes a terme en el processat de la imatge i que diferenciem d'aquelles vinculades al processat original. Es tracta de registrar els canvis respecte a una imatge original plenament processada, la de captura, de la qual se'n crea un fitxer derivat per a l'arxiu. Entre les més habituals, hi trobem les accions referents a conversió de fitxers, retall, edició, copiat, aplicació de filtres, modificació



**Gràfic1.** Metadades agrupades per blocs d'informació. Aquest esquema constitueix la base del mapa conceptual, al qual cal associar-hi els diferents contenidors i formats estàndards.

de resolucions, etc. Aquestes metadades poden ser de captura automàtica.

**Informació d'entorn.** És important recollir la informació sobre els dispositius i el software utilitzat, ja que el resultat final de la captura està condicionat per les característiques d'aquests. Les metadades principals són les referents a la identificació dels fabricants de càmera, escàner i lents, i dels dispositius concrets, inclosos els números de sèrie. També el software utilitzat per a la captura i l'edició de les imatges, especificant-ne la versió. En el cas de la digitalització, podem incloure en aquest bloc totes les metadades derivades del text d'anàlisi fet sobre els dispositius i el software utilitzat per tal d'optimitzar el rendiment d'aquests. Es tracta principalment de la MTF i la OECF. Les metadades derivades de tests de rendiment són d'inserció manual, la resta poden ser de captura automàtica.

**Informació de preservació.** Podríem entendre que la major part de metadades desenvolupades en els diferents blocs són d'interès per a la preservació, sobretot si ens atenem a les indicacions de PREMIS quan esmenta que la funcionalitat de les metadades del seu diccionari responen a la necessitat de mantenir la viabilitat, la comprensió, la lectura, l'autenticitat i la identitat dels fitxers. Però podem considerar algunes metadades addicionals, tals com les referents al suport d'emmagatzematge, la localització en el repositori, els inhibidors (encriptacions), els resums numèrics del fitxer o de la cadena de bits (en els formats que és possible) i els algorismes aplicats per obtenir-los.

Cal dir que aquests blocs són purament orientatius i pensats per a la gestió de l'arxiu digital des de l'àmbit patrimonial. No obstant això són prou representatius de les informacions a tenir en compte a l'hora de crear el nostre mapa conceptual. Si fixéssim el nostre interès en altres àmbits professionals, com poden ser la premsa o el món editorial, la constitució d'aquests blocs diferiria parcialment i en alguns casos caldria incloure blocs de metadades addicionals.

## CONTENIDORS DE METADADES ESTÀNDARDS

La identificació dels blocs, el desenvolupament de metadades per a cada bloc i la selecció d'aquestes en base a les funcions que es volen assumir permeten dibuixar la base del mapa. No obstant això, per tal de treballar amb un mapa que sigui operatiu, cal estructurar aquesta informació d'acord amb els contenidors de metadades estàndards existents i ubicar-los en funció de les estructures jeràrquiques dels fitxers d'imatge en què treballem i segons els formats gràfics preferents. L'existència de diferents estàndards de metadades i les distintes normes de cadascun d'ells referents a l'emmagatzematge, l'organització i la codificació en el si del fitxer són qüestions que s'han de tenir presents a l'hora de desenvolupar un sistema de gestió, tenint en compte la incidència que les decisions preses sobre aquests temes poden tenir en la constitució final de l'objecte digital.

La sincronització de què parlàvem a l'inici d'aquest text ofereix unes potencialitats de gestió que no es poden desestimar i que ens permeten treballar de manera més optimitzada d'acord a l'entorn tecnològic vigent. No obstant això, perquè aquestes potencialitats siguin reals, no podem limitar-nos a les eines que ofereix el mercat, és a dir, estar limitats pel software que utilitzem. Cal conèixer primer quin és l'abast de la nostra responsabilitat en la gestió i quin treball d'estandardització s'ha fet al voltant de la imatge digital en concret i de la documentació electrònica en general, per després identificar aquells contenidors estàndards que ofereixen llistats de metadades plenament desenvolupats, amb estructures compatibles i amb codificacions universals.

D'entrada, es poden diferenciar aquells estàndards que són específics, que no exclusius, de la imatge fixa, d'aquells que tenen un vocació més general. En el primer grup hi trobem: EXIF-TIF, IPTC i XMP; mentre que en el segon hi trobem EAD, DC i PREMIS. Les funcionalitats pròpies d'un arxiu fan que la consideració d'aquest segon grup de metadades si-

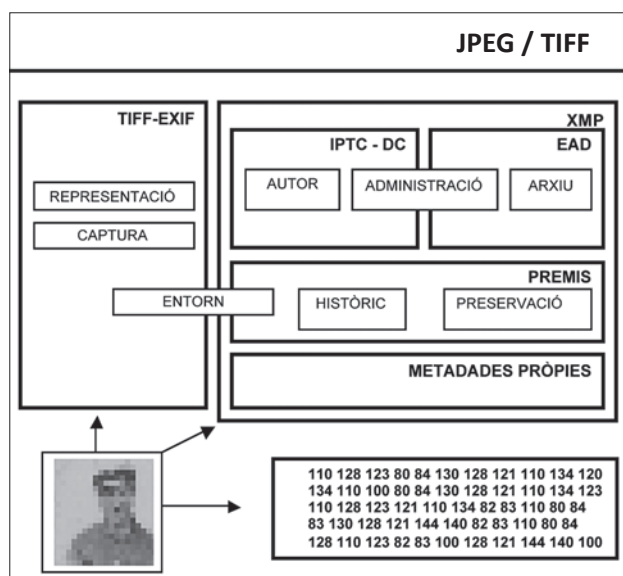
gui imprescindible i que l'esforç per integrar-les en el si de l'objecte digital sigui recomanable. No obstant això, en la gestió de la imatge digital podríem prioritzar els estàndards específics, perquè estan pensats per formar part del fitxer d'imatge i, per tant, poden ser utilitzats pel software. Podríem incloure també en aquest grup les metadades de Photoshop (PSIRs), per la generalització del seu ús, malgrat que es tracti d'un recurs d'un format propietari. La combinació dels estàndards específics resol bona part de les exigències derivades de l'accés i l'ús, encara que en l'àmbit patrimonial no es pot obviar el segon grup. La integració d'aquests dos grups en un sistema de gestió és un repte de futur que ha de permetre configurar un arxiu fotogràfic eficient en les funcions que tingui assignades, versàtil en l'adaptació tecnològica i sostenible per la seva condició d'estàndard.

**EXIF** (Exchangeable Image File Format) és un estàndard del JEITA (Japan Electronics and Information Technology Industries Association) i el CIPA (Camera and Imaging Products Association) desenvolupat a partir de la necessitat de poder comunicar les imatges creades per les càmeres digitals amb altres dispositius. Aquesta necessitat va portar a la creació de l'especificació DCF (Design Rule for Camera File System) que estableix les normes per a la gravació, la lectura i el suport de fitxers d'imatges i d'altres fitxers relacionats, i que defineix un subconjunt de l'EXIF on algunes propietats són opcionals en l'EXIF però obligatòries en el DCF. L'EXIF és de fet un format gràfic que presenta una estructura conforme al format JPEG per a imatges comprimides i conforme al format TIFF per a imatges sense compressió. En ambdós casos, l'EXIF descriu un conjunt d'etiquetes TIFF, segons el format descrit en la versió 6.0, i per aquelles informacions referents a càmera que no es contemplen en el TIFF, inclou metadades en un directori propi, diferenciat del directori de les metadades TIFF i també del directori de les metadades GPS. En correspondència als blocs d'informació descrits en aquest text, l'EXIF representa de manera codificada els de representació (metadades TIFF), captura (directori de metadades pròpies i directori de GPS) i entorn (inclou els valors OECF i SFR). També compta amb dues metadades addicionals per a la gestió dels drets: *Artist* i *Copyright*, de poca transcendència quan es combina amb altres estàndards.

**IPTC - IIM** (International Press Telecommunication Council). IIM (Information Interchange Model) és un format embolcall per a la transmissió de notícies en text i en imatge creat per un consorci d'agències de notícies. A partir de l'IIM es van crear els encapçalaments IPTC que van ser adoptats i inclosos per Ado-

be en el Photoshop amb la tecnologia Image Resource Block. Posteriorment es va substituir aquesta tecnologia per l'XMP i es van crear dos esquemes nous: primer l'IPTC Core i després l'IPTC Extension (ambdós exclusius de l'XMP). L'IPTC Core respecta l'estructura IIM i n'hereta moltes de les seves metadades, en substitució dels encapçalament originals de l'IIM. L'esquema IPTC associa les metadades amb els esquemes DC, Photoshop, Adobe Rights Management i IPTC IIM. De fet la relació amb l'esquema IPTC IIM fa que inclogui la majoria de metadades de l'original IIM. Cal tenir present que les metadades IIM estan generades en l'àmbit periodístic i que moltes d'elles no tenen valor per a l'arxiu. L'IPTC Extension és resultat dels requeriments d'incloure metadades addicionals a l'esquema inicial per part dels professionals de la fotografia, per la necessitat de gestionar determinades informacions del flux de negoci. Compta amb un nivell de compatibilitat més baix respecte al Core, encara que es treballa amb dades més especialitzades. En l'àmbit patrimonial les metadades principals d'interès són les de l'IPTC Core, moltes de les quals estan assimilades a l'esquema DC, com és el cas de l'autor, la descripció o el títol. Aquestes es corresponen principalment amb els blocs d'informació sobre dades descriptives d'autor i d'administració definits en aquest text. De tota manera, algunes metadades de l'Extension poden ser d'interès ja que fan referència al contingut de la imatge: identificació de persones, de llocs o descripció d'esdeveniments, mentre que d'altres estan clarament destinades a altres sectors professionals.

**XMP** (Adobe's Extensible Metadata Platform). És un estàndard per a la creació, processament i intercanvi de metadades. Ofereix una tecnologia d'etiquetatge que permet crear noves metadades i inserir-les en els mateixos fitxers. Es tracta de dades XML, emmagatzemades utilitzant un subconjunt del W3C Resource Description Framework (RDF). Això és especialment interessant per a la indústria informàtica ja que el software i els dispositius poden incloure informació pròpia als mateixos fitxers. Ho és també pel sector cultural per la possibilitat d'incloure metadades pròpies i degudament codificades en els contenidors XMP. En el cas dels arxius, suposa la possibilitat de plantejar la integració de les metadades EAD que d'aquesta manera passarien a ser metadades intrínseques pels formats d'imatge fixa TIFF, JPEG, JPEG2000, DNG, PSD i PNG. En els altres formats, es crea un fitxer de metadades independent. XMP defineix quatre blocs de metadades principals, el DC i tres de propis, més els blocs especialitzats: Adobe PDF, Photoshop, Camera Raw i Exif. La inclusió del DC en les propietats principals la converteix en una tecnologia de gran utilitat per a la



**Gràfic 2.** Blocs d'informació del mapa conceptual en relació als contenidors i esquemes de metadades estàndards segons la seva inclusió en l'estructura dels formats gràfics JPEG i TIFF.

comunicació entre diferents plataformes. Quant a les metadades exclusives de l'XMP, ens interessen principalment els tres primers blocs. El primer conté dades descriptives bàsiques del recurs digital, com la data de creació i la de modificació, o la data de modificació de les metadades. El segon bloc el formen dades per a la gestió dels drets, amb algunes metadades addicionals interessants com la que remet a un certificat de gestió de drets o la que remet a una declaració de propietat i drets d'ús del recurs. El tercer bloc està format per dades relacionades amb la identificació, composició i història, que permeten fer un seguiment de les transformacions experimentades per un recurs determinat. En definitiva, podem dir que les metadades principals que preveu l'XMP es corresponen parcialment amb els blocs de dades d'autor, dades d'arxiu, administració i històric dels blocs definits en aquest text. Des d'aquest punt de vista, l'aportació de l'XMP com a estàndard no seria massa interessant per a un arxiu, ja que es tracta de metadades que resolen altres estàndards. L'aportació principal de l'XMP és la seva extensibilitat i la possibilitat de codificar aquestes metadades addicionals en base a estàndards, i fins i tot metadades creades pel propi arxiu que no necessàriament han de formar part de cap estàndard codificat.

## ESTRUCTURACIÓ I INTEROPERABILITAT

La diversitat d'estàndards de metadades té la seva lògica complicació en l'organització i emmagatzematge d'aquestes, ja que es poden trobar en diferents estructures jeràrquiques. D'entrada, hem de tenir en

compte que els fitxers electrònics poden contenir diferents formats i que al mateix temps aquests poden contenir diferents cadenes de bits, amb metadades pròpies. Alhora els formats preveuen la integració de diferents contenidors de metadades que a més poden convergir en alguns dels seus objectius, fet que pot suposar la repetició d'algunes metadades en diferents contenidors. La intersecció entre formats de fitxer i contenidors de metadades provoca relacions complexes i confuses, amb metadades que són agrupades i emmagatzemades de diferents maneres segons els formats. En aquest text ens centrarem únicament en els formats TIFF i JPEG, per la generalització del seu ús, i en els contenidors EXIF, IPTC i XMP, contemplats tots ells en els formats esmentats i que tenen una presència destacada en la indústria fotogràfica. Altres estàndards com l'EAD o el DC quedarien integrats en els paquets XMP i, per tant, no complicarien l'estructuració, l'organització i l'emmagatzematge de metadades en el si dels fitxers. Partir d'aquest escenari simplifica la casuística derivada de les relacions d'aquests ítems alhora que ens apropa al context més habitual dels arxius.

L'anàlisi de la relació de metadades a nivell estructural en els formats TIFF i JPEG posa de manifest la complexitat d'un encaix global. El format TIFF compta, a més de les pròpies metadades, amb 5 directors subsidiaris, els que corresponen a les metadades XMP, IPTC, PSIRs, EXIF i GPS. El format JPEG organitza les metadades en les diferents marques de secció de les quals es diferencien les corresponents al TIFF-EXIF, a l'XMP i la PSIRs (que inclou l'IPTC). Per tant, a diferència del TIFF, el JPEG no conté una secció específica per a l'IPTC, sinó que aquest forma part del bloc PSIR. Per contra, en el TIFF es pot donar una repetició de metadades IPTC, les natives i les del PSIR. La secció TIFF-EXIF del JPEG apunta a un directori TIFF que inclou les metadades pròpies TIFF i una etiqueta que apunta a les metadades EXIF. Cap dels dos formats gràfics contempla el bloc XMP dins el PSIR, ja que l'XMP té l'espai propi. En el cas excepcional que el paquet XMP superi els 64 Kb que preveu el JPEG, es preveu la possibilitat de partir el paquet en dos blocs: l'estàndard XMP i una extensió. Tant en el cas del TIFF com del JPEG hem vist que inclouen el PSIRs (Photoshop Image Resources), a més dels contenidors ja esmentats. Es tracta d'un recurs natiu del format PSD que és present en aquests dos formats i que inclou també recursos de metadades i, més concretament, els estàndards XMP, TIFF-EXIF i IPTC, encara que els continguts específics del PSIRs són diferents segons el format. Quant a la possibilitat d'incloure dades EXIF a XMP és especialment d'utilitat per aquells formats que no inclouen aques-

tes metadades en la seva estructura. No és el cas del TIFF ni del JPEG, en els quals no es recomana incloure les metadades EXIF a XMP, sinó que és preferible mantenir l'EXIF natiu. En tots els formats de fitxer les metadades natives (aquelles previstes en l'estructura del format) tenen preferència sobre l'XMP. Com a conclusió d'aquesta explicació sobre la relació entre formats i contenidors, podem dir que els TIFF i els JPEG integren els mateixos tipus de metadades encara que emmagatzemades de maneres lleugerament diferents.

La combinació de contenidors de metadades és complexa i cal ser-ne conscients, ja que la dificultat de la seva gestió deriva principalment de l'assignació correcta dels valors de les metadades, dels canvis en els diferents estàndards en relació amb l'evolució del software i dels requeriments de codificació. En aquest sentit, el comportament del software resulta essencial per mantenir la coherència, la compatibilitat i la consistència de les metadades. La dificultat de combinar els estàndards de metadades és deguda principalment doncs a qüestions estructurals dels fitxers, a l'emmagatzematge de dades, a la repetició parcial de contenidors, a la manera d'accedir a aquestes metadades i als problemes provocats per diferents softwares que no sempre operen en favor de la consistència i la compatibilitat de les metadades.

Els problemes d'interoperabilitat causats per la repetició de metadades en diferents contenidors ens resulten menys preocupants si tenim en compte l'estudi realitzat pel grup de treball sobre metadades en fitxers d'imatge (Metadata Working Group, 2010). Aquest estudi revela que tant sols existeixen quatre metadades que siguin coincidents en els estàndards dominants: EXIF, IPTC i XMP. Aquestes són:

Drets d'autor: Exif Copyright – IPTC CopyrightNotice – XMP (dc:rights)

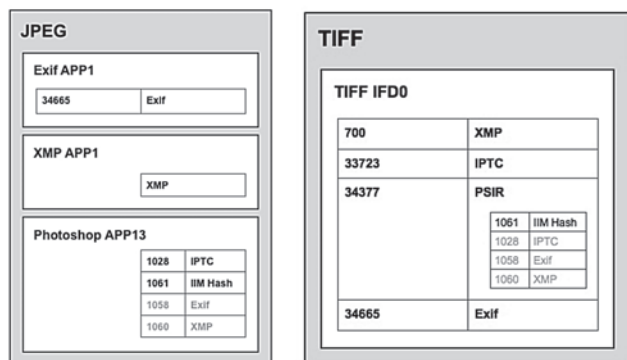
Descripció: ImageDescription – IPTC Caption – XMP (dc:description).

Data original: Exif DateTimeOriginal – IPTC DateCreated – XMP (photoshop:DateCreated)

Autor: Exif Artist – IPTC By-line – XMP (dc:creator)

El fet de tenir identificades les metadades coincidents i, sobretot, l'existència d'un grup de treball impulsat i sostingut per la indústria fotogràfica (Microsoft, Apple, Adobe, Canon, Nokia i Sony) que treballa per a resoldre de manera definitiva la integració de diferents tipus de metadades, ens fa pensar en un escenari futur en què la integració de metadades tècniques, descriptives i de gestió sigui una realitat.

De confirmar-se aquesta hipòtesi, la simplificació tindria conseqüències importants en el sector cultural, però difícilment podria complaure de manera satisfactòria el global de les seves necessitats.



**Gràfic 3.** Estructuració i organització de contenidors de metadades en els formats JPEG i TIFF. Gràfics publicats a: *Guidelines for Handling Image Metadata. Metadata Working Group, 2010.*

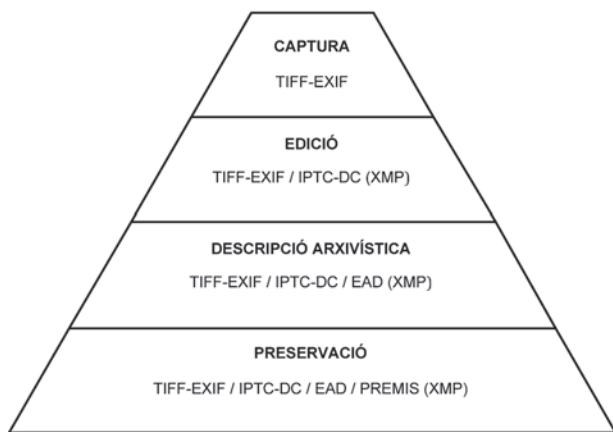
### METADADES AMB RELACIÓ AL FLUX DE TREBALL

Hem pogut observar com els estàndards de metadades han estat creats des de diferents sectors econòmics i professionals seguint objectius concrets: la indústria fotogràfica, a través del CIPA i el JEITA van desenvolupar l'EXIF per poder comunicar les imatges creades per les càmeres digitals amb altres dispositius; les agències de notícies, a través del seu consorci internacional, van desenvolupar l'IIM per a la transmissió de notícies en text i en imatge; el sector de la informació i la documentació, mitjançant un treball interdisciplinari, va desenvolupar el DC per a la descripció bàsica de recursos d'informació; els arxius van desenvolupar l'EAD per donar resposta a la descripció arxivística codificada; la indústria de software d'edició, representats per Adobe, van desenvolupar l'XMP per a la creació, processament i intercanvi de metadades; el sector de la documentació, principalment biblioteques, va desenvolupar el diccionari PREMIS per a la preservació de l'objecte digital; i podríem afegir altres estàndards que no hem esmentat en aquest text per no considerar-los tant rellevants per als arxius. És important tenir present els respectius objectius d'aquests estàndards per posar-los en relació amb les diferents funcions que deriven del flux de treball que s'inicia en la captura i que finalitza a l'arxiu. Les transformacions de la imatge al llarg d'aquest procés es produeixen sobre els continguts, però també sobre la informació que els acompanya, les metadades, ja que l'objectiu final que es persegueix és la creació d'un objecte digital accessible i sostenible.

En el moment de la captura es produeixen ja un seguit de metadades que variaran en funció del format gràfic però que, en tot cas, permeten sempre interpretar els continguts de la imatge de manera més o menys restringida, des de la màxima compatibilitat que ofereixen el JPEG o el TIFF a la limitació al software propietari dels RAW. A més de les metadades natives de cada format, es creen també de manera automatitzada un seguit de metadades de captura, les de l'EXIF, que passen a formar part del mateix fitxer en els TIFF i JPEG i que es mantenen normalment com a fitxers adjunts en els altres formats.

A partir d'aquí comença el treball d'edició de la imatge que serà diferent en funció del productor d'aquesta, si es tracta d'un fotògraf, d'un diari, d'una editorial, d'un particular, etc, però que en tots els casos derivarà en una imatge més o menys transformada i que compta amb metadades addicionals aportades pel software d'edició, com podria ser el Photoshop. Per tant, es tracta de metadades tècniques que poden quedar recollides a l'EXIF i que permeten fer una valoració tècnica del fitxer, analitzar-ne el format i la seva validesa i fins i tot determinar-ne l'autenticitat. Però aquest estadi no es limita a l'aportació de metadades tècniques, sinó que s'hi sol produir la inclusió de metadades descriptives per a la identificació de l'autor i dels continguts bàsics de la imatge i per a la gestió dels drets. Es tracta de les metadades que de manera exhaustiva recull l'IPTC i més concretament l'IPTC Core, si es treballa en la plataforma XMP. A partir d'aquí, si es pretén fer la inserció de metadades en el mateix fitxer és necessari treballar en XMP, ja que els contenidors nadius limiten considerablement l'extensibilitat de metadades. L'altra possibilitat, encara ara la més freqüent en els arxius, és fer ús de les base de dades, amb informació referencial sobre l'objecte.

Un cop acceptat l'ingrés i després d'haver fet la ingesta de les imatges al sistema, es duen a terme els treballs de descripció arxivística, amb especial atenció a les informacions de context del productor i a les informacions que han de donar suport als treballs d'administració. Per tant, s'introdueix en aquest moment l'EAD, que passa a conviure amb l'EXIF, l'IPTC i opcionalment amb l'XMP. Les responsabilitats sobre la preservació mereixen ser considerades apart, ja que les informacions necessàries dependran de les estratègies planificades que difícilment es poden limitar a l'àmbit de l'arxiu. El diccionari PREMIS ens dóna una idea de l'exhaustivitat d'aquestes informacions i de la necessitat d'articular un pla conjunt en el si d'una institució o d'una empresa amb la implicació de tècnics de diferents àrees.



**Gràfic 4.** Metadades utilitzades al llarg del flux de treball des del moment de la captura.

Per últim, cal considerar les funcions que deriven de l'accés i la difusió de les imatges de l'arxiu, que dependran en bona part de l'entorn en què es produeixi: un web, una intranet, una publicació, etc., i de les opcions que es plantegin: l'accés a la imatge final, l'accés a una imatge referencial, l'opció d'impressió, l'opció de compra, etc. En aquesta fase, la consideració d'un estàndard com el DC és molt recomanable, i en alguns casos imprescindible. Per tant, en els millor dels casos, comptarem amb els estàndards EXIF, DC, EAD, XMP i IPTC. No obstant això, cal tenir en compte que no podem ser exhaustius en la difusió d'aquestes metadades, ja que per una banda ens trobem amb dades personals que no poden fer-se públiques, com és el cas d'algunes metadades IPTC, o amb dades d'administració que han de limitar-se a l'ús de l'arxiu, com poden ser algunes metadades de l'EAD. A més, cal considerar la conveniència de ser més o menys exhaustius en determinats entorns, com pot ser el cas de portals compartits, on una descripció en els elements bàsics del DC pot ser suficient i sovint més adient. En tots els casos és important assegurar que el software que intervingui en les diferents fases i processos tingui la capacitat de mantenir i emmagatzemar degudament les metadades existents al marge de les modificacions i els canvis que es produeixen en el si de l'objecte digital.

## CONCLUSIONS

El mapa conceptual per a la gestió de la imatge digital ha de donar resposta a les dues preguntes principals formulades a l'inici d'aquest text sobre quina informació tenim i quina informació necessitem per a dur a terme la nostra tasca. La seva concepció depèn d'un treball que de manera individualitzada pot elaborar cada arxiu i que ha de tenir en compte les següents fases:

- Identificar les necessitats informatives de l'arxiu i elaborar un llistat de metadades degudament estructurades.
- Valorar i considerar els blocs de metadades existents, ja siguin específics de formats gràfics (metadades natives) com transversals a diferents formats (els contenidors), com en forma de diccionari o esquemes de referència.
- Seleccionar aquells estàndards que millor responen a les nostres necessitats i posar-los en relació amb l'estructura inicial que hàgim acordat.
- Determinar les funcions específiques de les diferents fases del flux de treball i posar-les en relació amb els blocs de metadades.

El mapa conceptual ha de ser concebut per a ser operatiu i per tant caldrà donar resposta també a totes aquelles qüestions més concretes que permetran prendre decisions sobre la manera de treballar i sobre el software necessari per a l'execució del nostre treball. És imprescindible saber quines metadades formen part de l'objecte digital i com estan estructurades, fet que requereix d'un bon coneixement dels formats en què treballem i d'un software que ens permeti valorar la realitat dels nostres objectes. Conegudes aquestes metadades i la manera com estan estructurades, cal veure on es localitzen: en el mateix fitxer, en fitxers adjunts, en bases de dades o, com és habitual, repartides entre fitxers i bases de dades. També cal diferenciar quan provenen d'una captura automàtica, i en quin moment es produeix aquesta captura, de quan són d'inserció manual. Per aquestes segones, cal automatitzar tot allò que sigui possible per fer que els diferents processos documentals siguin assumibles. I encara caldrà determinar si hi ha dependència de software, quelcom que condiciona la nostra adaptació en l'evolució contínua de la tecnologia. Per últim, cal preveure quines funcionalitats aniran associades a l'objecte original i, per tant, ser conscients de si disposem de la informació necessària per dur-les a terme. En funció dels resultats d'aquesta anàlisi podrem desenvolupar o adquirir el software que ens permetrà dur a terme les funcions que tenim assignades de manera que el mapa conceptual sigui operatiu en base a la realitat informativa i tecnològica del nostre entorn.

## BIBLIOGRAFIA

- ANSI/NISO Z39.87. Data Dictionary – Technical Metadata for Digital Still Images. NISO, 2006.
- CIPA DC- 009; JEITA CP-3461B. Design rule for Camera File system: DCF version 2.0. 2010.
- CIPA DC-010. Exif 2.3 metadata for XMP. Camera & Imaging Products Association. 2012.



Exchangeable image file format for digital still cameras: Exif version 2.3. JEITA; CIPA. 2010.

Extensible Metadata Platform (XMP) Specification: Part 1, Data Model, Serialitization and Core Properties. Adobe, 2012.

Extensible Metadata Platform (XMP) Specification: Part 2, Additional Properties. Adobe, 2012.

Extensible Metadata Platform (XMP) Specification: Part 3, Storage in Files. Adobe, 2012.

Guidelines for Handling Image Metadata. Version 2.0. Metadata Working Group, 2010.

IPTC Standard. Information Interchange Model (IIM). IIM Schema for XMP. Specifications 1.0. Document revision 1. IPTC, 2008.

IPTC Standard. Photo Metadata: IPTC Core Specification Version 1.1 / IPTC Extension Specification Version 1.1. Doc Rev 1. 2010.

JPEG File interchange format, version 1.02. 1992.

Photoshop Image Resources. Adobe, 2008

TIFF 6.0 specification. Adobe, 1992.

## RESUMEN

El cambio tecnológico en la creación de las imágenes fotográficas ha tenido consecuencias lógicas para los archivos. La imagen digital requiere de planteamientos específicos a nivel de preservación, acceso, etc. Sin embargo, hay un hecho fundamental a la hora de diseñar el archivo digital: la equiparación de la

naturaleza informativa de los datos que configuran la imagen y de aquellos que permiten su gestión en un sentido global. Esta circunstancia nos permite subrayar la necesidad de crear un mapa de información general que nos permita tomar decisiones tanto en actuaciones básicas, la preservación y el acceso, como en las actuaciones más concretas, como pueden ser la decisión acerca del software de trabajo o la manera en que se procesan y se estructuran los datos.

## SUMMARY

The technological change in the creation of photographic images has had logical consequences for archives. The digital image requires specific approaches for preservation, access, etc. However, there is something fundamental when designing the digital archive: data that are part of the image and data that allows the management of those images in a global sense have the same informative nature. This circumstance let us to highlight the need to create a general information map, that allow us to make decisions concerning basic actions, related to preservation and access, or concerning more specific actions, as the decision on the software or the way that we process and structure data.