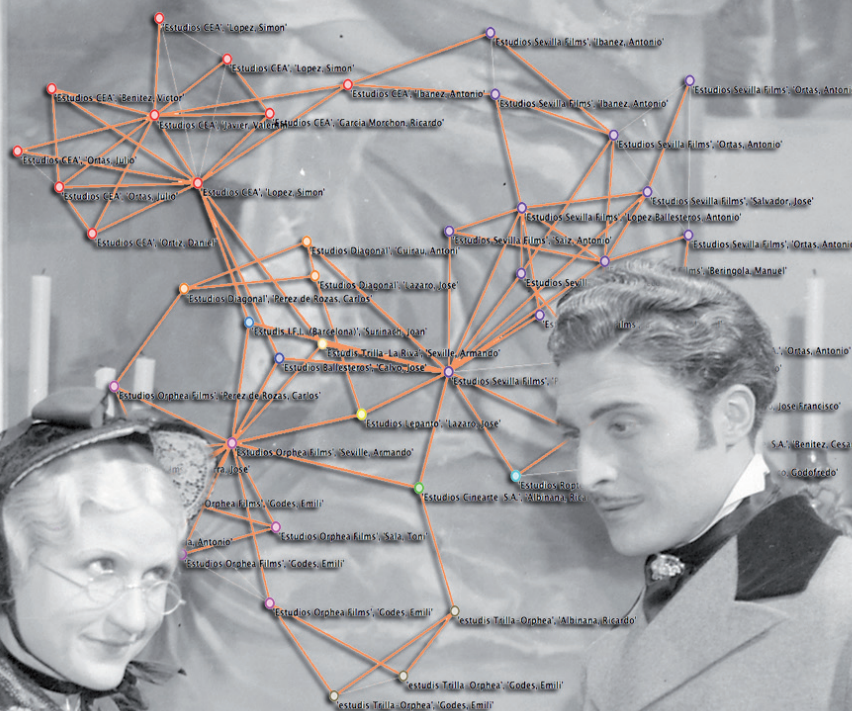
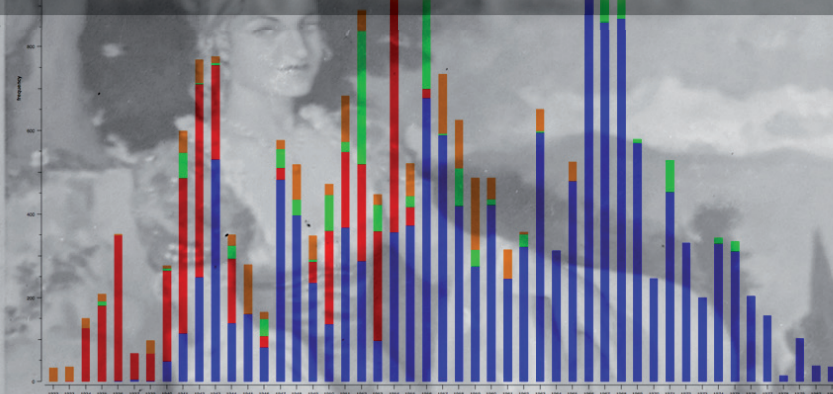


GESTIÓ DEL PATRIMONI CULTURAL ASSISTIDA PER MINERIA DE DADES. ANÀLISI DE LA CATALOGACIÓ D'UN FONS DE FOTOGRAFIES FIXES DE LA FILMOTECA DE CATALUNYA: EL CAS D'EMILI GODES



Miquel Àngel Pintanel Bassets

Treball final de Màster de Gestió Cultural (Orientació recerca) UOC-UdG – 2014
Directora: Dra. Muriel Gómez Pradas

A la Mireia

Pintanel Bassets, Miquel Àngel

Gestió del patrimoni cultural assistida per mineria de dades. Anàlisi de la catalogació d'un fons de fotografies fixes de la Filmoteca de Catalunya: el cas d'Emili Godes

Treball final de Màster de Gestió cultural (Orientació recerca) UOC – 2014

Directora: Dra. Muriel Gómez Pradas

Assessorament: Dr. Xavier Rubio Campillo

Fotografia de la portada: fotografia fixa d'Emili Godes de *La linda Beatriz* (Castellví, 1940)

Índex de continguts

Resum	7
1. Introducció	8
2. Objectius i preguntes de recerca	11
2.1. Objectius	11
2.2. Preguntes de recerca	11
3. Justificació	12
3.1. Catalogació del patrimoni com a primer pas de la gestió en un equipament patrimonial	12
3.2. Necessitat de realitzar un procés de Descobriment de Coneixement en Bases de dades	12
3.3. Tria del material per a realitzar el TFM	13
4. Marc teòric i conceptual	15
5. Metodologia prevista per a l'elaboració de la recerca	17
6. Anàlisi documental de la catalogació del fons estudiat	19
6.1. Origen del material de catalogació que ocupa aquest treball	19
6.1.1. Descripció dels fons de l'Arxiu gràfic de la Filmoteca de Catalunya	19
6.1.2. Descripció del fons Reproducciones Sabaté	21
6.2. Gènesi de les dades	22
6.2.1. Qui va catalogar el fons. Personal en formació <i>versus</i> personal professional	22
6.2.2. Descripció tècnica del procés de catalogació	23
7. Estudi de cas: el fotògraf Emili Godes	29
8. Aplicació del Descobriment de Coneixement en Bases de dades	34
8.1. Descripció de les dades	34
8.1.1. El sistema gestor de base de dades de la Filmoteca de Catalunya	34
8.2. Descripció i justificació de les eines utilitzades per al Descobriment de Coneixement en Bases de dades	36
8.2.1. Programari per a exportar i transformar les dades	36
8.2.2. Programari de mineria de dades	36
8.2.2. Tria de les eines utilitzades	37
8.3. Preparació i control de les dades per a la mineria de dades	37
8.3.1. Transformació de les dades	37
8.3.2. Anàlisi de les propietats de les dades i la seva qualitat	38
8.4. Coneixement adquirit mitjançant la mineria de dades	42
8.4.1. Amenaces, febleses i forteses del fons fotogràfic	43
8.4.2. Connexions entre les dades	47
8.5. Conclusions de l'aplicació del Descobriment de Coneixement en Bases de dades al cas d'Emili Godes	51
8.5.1. Material utilitzat per Emili Godes	51
8.5.2. Connexions d'Emili Godes	52
9. Conclusions	54
9.1. Febleses de la catalogació tradicional	54
9.2. Aportacions del procés de Descobriment de Coneixement en Bases de dades	54
9.3. Problemàtica del Descobriment de Coneixement en Bases de dades	55

9.4. Proposta d'aplicació del procés de Descobriment del Coneixement en Bases de dades en un equipament cultural.....	55
10. Bibliografia.....	57
11. Annexos.....	60
11.1 Guia d'entrada a la base de dades.....	61
11.2 Fitxa de descripció fotogràfica.....	66

Agraïments

La llista d'agraïments d'aquest treball és més llarga que els títols de crèdit d'una pel·lícula de gran pressupost, ja que inclou a tots els qui van fer possibles les dades que s'han utilitzat que són més de 400 estudiants, els col·laboradors que les van revisar i l'equip que va marcar-ne les directrius de catalogació. No cal oblidar a les companyies productores Filmoteca de Catalunya, sota el paraigües de la *major* Institut Català de les Empreses Culturals, i la Universitat de Barcelona. Com en un *flashback* hem d'anar enrere i he d'agrair la saviesa transmesa, la confiança i l'amistat de, per ordre d'aparició, Miquel Porter Moix i Palmira González López, sense ells no existiria aquest treball. Tampoc existiria, tornem al present, sense la Carme Belarte, que m'ha donat suport, confiança cega, ànim, i ha tingut la paciència necessària en els moments de crisi existencial. He abusat dels coneixements, de la seva predisposició a ajudar-me i escoltar-me de Josep Lluís, Laia Foix, Noemi Maya, Ariadna Elias i dels companys de la Filmoteca de Catalunya, els de l'associació Fotoconnexió i la gent de Les Escanes. He d'agrair el suport de la meua família, especialment del meu germà Xavi. Tampoc m'oblido de la Muriel Gómez Pradas, la Patricia Castellanos i en Xavier Rubio, que m'han aconsellat i guiat en tot moment pel bon camí. I amb la meua memòria de peix, segur que m'oblido d'un munt d'actors, tècnics, especialistes, dobles de llums i foto fixes que han fet possible aquest treball, espero que sàpiguen perdonar el meu oblit.

Resum

Aquest treball de recerca de gestió del patrimoni cultural té com a objectiu validar que es pot millorar la gestió d'un equipament patrimonial amb l'aplicació combinada a la catalogació dels seus fons d'una anàlisi documental, una recerca històrica i un procés de Descobriment de Coneixement en Bases de dades. Per tal de comprovar-ho, s'ha utilitzat un dels fons fotogràfics de la Filmoteca de Catalunya, el dels laboratoris Reproducciones Sabaté. Es realitza en tres fases diferents: en la primera, s'exporten les dades existents de la catalogació adquirides amb un mètode tradicional, es depuren, s'analitzen i es decideix quins són els camps i els registres que es poden utilitzar en els processos posteriors; en una segona fase es realitza una anàlisi documental i un estudi de cas sobre una part de les dades, les fotografies realitzades pel foto fixa Emili Godes; per últim, en una tercera fase s'analitzen aquestes dades existents de la catalogació tradicional utilitzant el Descobriment de Coneixement en Bases de dades, assistit per eines de mineria de dades i visualització. Com a resultat s'obté un millor coneixement de les característiques del fons que permet optimitzar-ne els protocols de catalogació, actuació i conservació.

Paraules clau: patrimoni, fotografia, cinema, Descobriment de Coneixement en Bases de dades, mineria de dades

1. Introducció

La recerca d'aquest treball final del Màster de Gestió Cultural (orientació recerca) s'ha fet a partir de les dades d'una catalogació existent dels fons fotogràfics de la Filmoteca de Catalunya. Des de l'any 1999 fins al 2011 a la Filmoteca de Catalunya es va dur a terme un projecte de col·laboració amb la Universitat de Barcelona en el marc del qual van participar estudiants d'Història de l'Art en la catalogació de fotografies fixes del seu arxiu. Com a resultat es van catalogar prop de 20.000 fotografies fixes, que corresponien a 472 pel·lícules produïdes a Espanya entre 1932 i 1984, i hi van intervenir prop de 400 estudiants. El procés tenia dues fases: descripció prèvia en una fitxa en paper pautada i introducció de les dades de forma tradicional en una base de dades relacional. L'autor d'aquest treball va participar en la formació dels estudiants i en la supervisió de la catalogació, com a responsable de la fototeca de la Filmoteca de Catalunya, i ha d'agrair la bona predisposició de la Filmoteca de Catalunya per a poder utilitzar les dades que han fet possible aquest treball.

S'han volgut combinar dues orientacions de recerca molt diferents per a poder comparar els resultats que retorna cada una: (a) un enfocament tradicional que consta d'una **anàlisi documental** del procés de catalogació i un **estudi de cas**, per al qual s'ha triat el foto fixa¹ Emili Godes, i (b) un procés de **Descobrimet de Coneixement en Bases de dades** (*Knowledge Discovery in Databases o KDD*) sobre el conjunt del fons assistit amb **minería de dades**² (*Data mining*) i eines de **visualització**.

Per a introduir l'estudi de cas és necessari explicar una mica en què consisteix la figura del foto fixa. És l'encarregat de realitzar les fotografies que s'utilitzaran per a la promoció de la pel·lícula. Un dels temes més complexos d'explicar és que no totes les fotografies que es realitzen durant un rodatge són fotografies fixes, ja que és una figura molt reglada i que té unes funcions molt concretes durant la filmació. La majoria de fotografies fixes, i en això les d'Emili Godes en són un clar exponent, són del que s'anomena fotografies d'escena, que representen una escena del film. El foto fixa ha de fer una mena de direcció d'actors en els descansos entre preses del rodatge per a

¹ S'ha triat aquesta forma d'anomenar a qui realitza les fotografies durant els rodatges seguint la normativa de la Neoloteca del TERMCAT (<http://www.termcat.cat/ca/Neoloteca/>). En la documentació d'època i en els títols de crèdit en castellà es fa servir el terme *Foto fija*, que tant serveix per a definir l'objecte com al seu realitzador, i aquesta forma d'anomenar-lo ha passat a la bibliografia especialitzada. Personalment opinem que a l'hora de fer la traducció al català s'hagués pogut triar un terme en que quedés més clara la diferència entre l'objecte, *Foto fixa*, de qui l'ha produït, ja que no hi havia una tradició per anomenar-lo. Tot i que existeix la possibilitat d'utilitzar *Fotògraf de foto fixa* no hem utilitzat aquest terme per trobar-lo redundant. En anglès s'utilitza el terme *Still photographer* i en francès *Photographe de plateau*.

² La minería de dades és l'aplicació sobre les dades d'algorismes que permeten classificar-les, agrupar-les i connectar-les entre elles. Es farà servir al text el terme minería de dades ja que s'ha convertit en la més comuna de les traduccions, tot i que alguns autors com Sangüesa *et alii* (2010) que utilitzen també el terme minería de dades, troben que el descriu millor la traducció prospecció de dades. En general s'ha optat en tot el treball per utilitzar la traducció en català d'una paraula tècnica si s'ha trobat la traducció en una font contrastada. Si hi ha diverses opcions de traducció s'especifica en nota els motius pels quals se n'ha triat una d'elles. Les fonts utilitzades principalment són TERMCAT i Sangüesa *et alii* (Ibidem).

realitzar les fotografies i ha de resoldre problemes com la disposició dels actors, ja que en els films moltes vegades no surten al mateix pla, però sí a la mateixa seqüència, com quan hi ha una conversa que es munta amb la solució de pla i contraplà. També ha de poder controlar la resta de l'equip per a que no interfereixin en la fotografia i per a que alguns, com els encarregats de la il·luminació, no deixin el seu lloc de treball. Del personal del rodatge és dels únics tècnics que està només a les ordres del productor del film i no, encara que sigui en part, a les del director.

El foto fixa és un personatge gairebé inexistent en la historiografia del cinema i té el que es pot anomenar una "autoria dèbil", ja que no és citat en la majoria d'anuaris i catàlegs sobre cinema i a vegades tampoc en els títols de crèdit, quan aquests són abreujats. El fet que no participa, aparentment, en el producte final cinematogràfic l'ha deixat apartat i, fins i tot en algunes obres bibliogràfiques especialitzades, les fotografies fixes són confoses amb fotogrames del film. Si es tingués una visió més global del fet cinematogràfic es podria veure com és d'important la figura del foto fixa, ja que en alguns casos les seves fotografies, com és el cas d'Emili Godes i altres autors, fan que la pel·lícula pugui semblar de més bona factura del que és.

Hi ha molts films del cinema espanyol que no s'han conservat i per tant no es poden utilitzar com a font primària d'informació per a documentar les fotografies. Per entendre com s'han perdut tantes pel·lícules es pot fer un paral·lel amb el fons estudiat en aquest treball. Les fotografies d'aquest fons s'han conservat gràcies a que la majoria dels productors no van recuperar els negatius dipositats al laboratori Reproducciones Sabaté, ja que aquestes havien deixat de tenir el seu valor promocional. Quan el laboratori se'n va voler desfer haguessin pogut anar a parar a un contenidor, a un drapaire o a una fàbrica de pintes, destins habituals del material cinematogràfic, però van poder ser adquirides per la Filmoteca de Catalunya.

En total hi ha 63 foto fixes identificats com a autors de les fotografies del fons i se n'ha triat Emili Godes per a fer l'estudi de cas, com s'ha comentat, i per controlar si es complien les hipòtesis que es plantejaven en la mineria de dades. Les seves fotografies destaquen per la gran qualitat i va tenir una llarga trajectòria com a foto fixa, de 1935 a 1961, representada al fons amb un alt volum de fotografies, el 7% de les 20.000 que el componen. És per tant un dels fotògrafs més representatius del fons catalogat.

En aquest treball hem aplicat el Descobriment de Coneixement en Bases de dades, que Fayyad *et alii* (1996b) defineixen com «(...) el procés no trivial d'identificar patrons en les dades que siguin vàlids, nous, potencialment útils i, en definitiva, comprensibles³», per a millorar la gestió del patrimoni cultural moble. Per a aquesta aplicació en la gestió del patrimoni cultural, sense oblidar que el procés ha de complir totes les condicions, ens hem de centrar en com **extreure patrons vàlids i potencialment útils** de les dades dels objectes per aplicar al dia a dia de la millora d'aquesta gestió. Per a millorar la difusió del seu coneixement i aportar-lo a la societat potenciem que aquests **patrons** siguin **nous i comprensibles**.

Aquest procés té varies capes i múltiples possibilitats, però se n'han volgut acotar els límits per aquest treball. La primera capa, imprescindible, és el tractament de les

³ Traducció de l'autor.

dades, la seva transformació i la seva tria, el que alguns autors han considerat la part més gran de la feina (Sangüesa *et alii*, 2010). La segona capa aplica algorismes de mineria de dades i a partir dels patrons creats per aquesta poder generar una tercera capa de visualització interactiva, que és la que permet mostrar conclusions útils de tot el procés. Totes aquestes capes depenen de dos factors inicials: el coneixement expert de les dades i, a partir d'aquest, saber realitzar les preguntes indicades i veure si les hipòtesis que es plantegen són correctes. La capa que es quedarà al tinter és com es poden aplicar les conclusions al dia a dia d'un equipament patrimonial amb eines d'aprenentatge automatitzat (*Machine learning*).

2. Objectius i preguntes de recerca

2.1. Objectius

Objectiu principal:

- Validar que es pot millorar la gestió d'un equipament patrimonial amb l'aplicació combinada a la catalogació dels seus fons d'una anàlisi documental, una recerca històrica i un procés de Descobriment de Coneixement en Bases de dades.

Objectius específics:

- Determinar si el volum, la qualitat i el coneixement previ sobre les dades de la catalogació són suficients per a un projecte de Descobriment de Coneixement en Bases de dades.
- Producte de les conclusions d'aquestes avaluacions, determinar quins són els algorismes aplicables per efectuar la mineria de dades.
- Fer una anàlisi documental assistida per mineria de dades de la catalogació d'un fons fotogràfic i valorar la utilitat dels resultats obtinguts.
- Proposar una forma d'aplicar les eines de Descobriment del Coneixement en Bases de dades en un equipament patrimonial.

2.2. Preguntes de recerca

Quines són les febleses d'un mètode tradicional de catalogació de fons fotogràfics?

Què pot aportar un procés de descobriment del coneixement en bases de dades (Knowledge Discovery in Databases o KDD) en la millora de la gestió d'un fons fotogràfic?

3. Justificació

3.1. Catalogació del patrimoni com a primer pas de la gestió en un equipament patrimonial

Partir d'un bon coneixement de l'equipament és crucial per a gestionar-lo correctament i en el cas d'un equipament patrimonial de béns mobles aquest coneixement correspon al dels objectes que custodia. La catalogació correcta és la base d'una bona gestió patrimonial. Moltes vegades, però, es perd la connexió necessària entre la gestió de l'equipament i la de les seves dades, amb la qual cosa l'única connexió entre elles acostuma a ser l'inventari, que només respon a les següents preguntes de gestió: ¿On tinc els objectes? ¿Què puc exposar? ¿Què puc explotar?

Pel camí es perd tot l'aprofitament que podrien tenir aquests centres com a fonts de coneixement, més enllà de la derivada pels estudis que es puguin fer per a publicacions científiques o divulgatives. Si la institució no fa una gestió del coneixement de les seves dades, essent qui en teoria les coneix millor, s'està perdent un valor important de retorn a la societat, com és aportar-li aquest nou coneixement. S'hauria d'anar un pas més enllà per aprofitar encara més la catalogació dels fons d'un equipament.

3.2. Necessitat de realitzar un procés de Descobriment de Coneixement en Bases de dades

Si es parteix de catalogacions ja realitzades i amb programes tradicionals es fa difícil descobrir coneixement de les nostres dades ja que el programari actual no està preparat per a aquesta funció. Per posar un exemple extrem, i que pot semblar trivial, es podria preguntar als responsables d'un equipament si el programari que gestiona la seva base de dades pot aportar un coneixement tan bàsic com la suma dels valors d'un camp. En molts casos la resposta seria negativa i si es pregunta com es fa normalment per sumar les dades d'un camp la resposta seria: exportant-les a un full de càlcul. Els gestors de bases de dades es limiten moltes vegades a treure llistes, però no saben res de sumes, percentatges, períodes, ni saben mostrar-ho de forma visual. Tampoc saben respondre a preguntes com quin és l'estat global de conservació de la col·lecció, ja que les seves dades no operen en conjunt, sinó que estan basades en objectes concrets.

La necessitat de realitzar aquest procés és doble, per una banda per a realitzar-ne una auditoria, i per l'altra per a extreure'n el coneixement.

Amb l'auditoria es podrà copsar com ha afectat el procediment de catalogació, el qual descriurem a l'**apartat 6.2.2**, a les dades. Les conclusions extreïdes podran anar més enllà de la intuïció o les impressions de les fitxes revisades perquè han estat objecte d'usos concrets. Això pot permetre prendre decisions futures més informades en el cas de voler seguir un procediment similar i respondre a preguntes sobre la fiabilitat de les dades generades.

Amb l'extracció de coneixement es pot posar en pràctica amb dades reals la teoria del

Descobriments de Coneixement en Bases de dades i poder veure si és d'utilitat en un equipament cultural. Aquesta teoria ja ha demostrat la seva validesa en altres disciplines, com en la medicina o el món dels negocis, però encara no ha estat prou posada en pràctica aplicada a dades d'objectes patrimonials. És una tecnologia prou madura, comptant que ja existeixen obres que en fixen els procediments l'any 1996 (Fayyad *et alii*, 1996a) i que encara són vàlides, així que no suposa una aventura intentar la seva aplicació. De fet, amb sentit de l'humor, però posant una mica el dit a la llaga s'ha arribat a dir «Els cínics, mirant amb ironia a l'explosió d'interès comercial (i desplegament publicitari) en aquesta àrea, equiparen la mineria de dades a la suma d'estadística més màrqueting⁴.» (Witten *et alii*, 2011).

La necessitat de fer una anàlisi amb un sistema assistit per mineria de dades i amb eines de visualització apareix quan el volum de dades és prou gran com per estar fora de l'abast d'un estudi individual de cada un dels registres de la base de dades i per no poder tenir-ne una visió global.

Les millores d'aplicar un procés de Descobriments de Coneixement en Bases de dades en la gestió patrimonial es podrien resumir en els següents punts:

- Identificació de patrons en les dades, la qual cosa n'augmenta el coneixement. Aquests patrons permeten crear:
 - a) Agrupacions de registres afins. Això permet:
 - Identificar registres amb necessitats afins d'actuació
 - Detectar les febleses i forteses del fons patrimonial
 - Descriure gràficament l'abast del fons patrimonial
 - b) Identificació de registres que no coincideixen amb la resta de dades. Això permet:
 - Identificar registres mal classificats i, per tant, erronis, amb molta probabilitat, o excepcionals
 - Classificació dels registres de la bases de dades
 - Augment del coneixement de les dades i possibilitat de crear arbres de classificació
 - c) Visualització del coneixement adquirit de les dades de forma gràfica

3.3. Tria del material per a realitzar el TFM

L'autor d'aquest treball final de màster va participar en tot el procés de catalogació del material escollit i va ser un dels que va establir-ne les pautes d'actuació, conjuntament amb Mariona Bruzzo, cap de l'Arxiu fílmic de la Filmoteca de Catalunya, Palmira González López, professora titular de la Universitat de Barcelona, José Luis Rubio Munt, fotògraf i historiador del cinema, i altre personal de la Filmoteca de Catalunya, especialment Iolanda Ribas. També hi va haver valuoses aportacions dels becaris que van donar suport, dels estudiants que van realitzar la catalogació i dels documentalistes que van col·laborar en la seva revisió.

Aquest coneixement de primera mà facilita en gran part la recerca ja que per a fer una anàlisi documental com la que és vol realitzar en aquest treball és necessària una

⁴ Traducció de l'autor.

preparació de les dades basada en els coneixements experts previs d'aquestes.

Així mateix, el fons triat és útil per a la finalitat de la recerca per diversos motius:

- a) És un fons tancat, ja que s'ha acabat la seva catalogació.
- b) És un fons homogeni i, per tant, vàlid per a treure'n fàcilment patrons útils.
- c) Té un volum de dades suficients i aquestes són consistents.
- d) És accessible per a realitzar l'estudi.

4. Marc teòric i conceptual

Les bases teòriques s'han organitzat en relació a diferents temàtiques que tractem en aquest treball de recerca.

A) El treball més complet, i l'antecedent bibliogràfic més clar, per a la gestió de fons i col·leccions fotogràfiques és l'obra de Boadas *et alii* (2001). Tot i que el seu enfocament és més clàssic del que tinc previst seguir, servirà com a marc teòric i conceptual en la part documental. Pel que fa a obres bàsiques en l'apartat documental, apart de la citada, han estat d'utilitat les obres de referència, normalment en forma de recomanacions o guies, d'institucions de referència. Entre aquestes hi ha la majoria d'obra del Centre de Documentació i Recerca de la Imatge de l'Ajuntament de Girona, les directrius i glossaris de la Federació Internacional d'Arxius Fílmics, les publicacions d'autors de L'Eastman House de Rochester i l'obra de DLM-Forum (1997).

B) Pel que fa al Descobriment de Coneixement en Bases de dades s'han fet servir les obres fundacionals de Fayyad *et alii* (1996a; 1996b). Sobre els processos d'aplicació de la mineria de dades s'ha emprat també la tesi de Gallardo Arancibia (2009) i Sangüesa *et alii* (2010). Apart d'aquestes obres cal apreciar que en altres àmbits, com en la medicina (Demšar, 2010), l'economia, la sociologia, les ciències aplicades, i en altres aplicacions dels processos de *descobriment de coneixement*, ja fa temps que s'utilitzen les tècniques que vull aplicar un cop realitzada l'anàlisi documental, així que és obligat estudiar com s'ha fet en aquests àmbits i comprovar com es poden aplicar al propi cas d'estudi. La meua aportació seria l'aplicació en la gestió de fons fotogràfics d'una metodologia que ja es fa servir en altres ciències o en l'àmbit empresarial. Per exemple, en el camp de la sociologia hi ha repositoris de coneixement com el de l'IPCRS⁵ (Inter-university Consortium for Political and Social Research), que es vol lligar (Vardigan & Whiteman, 2007) a objectes digitals utilitzant l'estàndard de consulta de dades documentals entre diferents sistemes de bases de dades OAIS⁶ (Open Archival Information System), fent el camí invers del que jo vull fer, ja que ells volen encaixar les dades sociològiques amb els documents per extreure'n coneixement.

C) En oposició al procés que jo plantejo en aquests moments el corrent dominant en els arxius fotogràfics és encomanar-se a les metadades⁷, ja que són, aparentment, una simplificació de la feina en contenir menys camps, apart del fet que permeten disseminar els fons amb molta facilitat. Molts arxius⁸ estan fent una preparació per a

⁵ Consultable a <https://www.icpsr.umich.edu>.

⁶ És pot consultar a <http://public.ccsds.org/publications/archive/650x0m2.pdf> la recomanació d'ús de OAIS.

⁷ Les metadades es defineixen normalment com «dades sobre dades» i poden formar part dels mateixos arxius digitals o anar associades en arxius XML, normalment amb una interfície web que les posen en comú amb l'objecte digital. És un estàndard interoperable, això vol dir que ha de ser independent del gestor de dades utilitzat, amb la qual cosa s'ha arribat a un consens en els camps que es poden utilitzar.

⁸ Se'n pot veure una mostra entre la llista dels participants de Memòria Digital de Catalunya, que es pot consultar a <http://mdc1.cbuc.cat/institutions.php>. A nivell europeu es pot consultar els participants a Europeana per a copsar aquesta activitat (<http://www.europeana.eu>).

afrontar l'ingent volum de dades que els està aportant la imatge digital i això ha provocat urgències que estan portant a simplificar les tasques, tot i que si s'apliquessin les eines d'aprenentatge automatitzat com més dades es tenen més senzill és entrar-ne de noves, ja que el sistema d'intel·ligència artificial és més efectiu ajudant al catalogador com més dades posseeix. Tots els formats de metadades i els protocols de consulta que els acompanyen, com OAIS, EADS, PREMIS, OAI-PMH o Dublin Core⁹ són eminentment descriptius (Iglésias Franch, 2012) i, per a ser compatibles entre arxius, una esquematització. El problema és quan les metadades no són conseqüència d'una catalogació i consolidació de les dades, sinó un intent de catalogació en elles mateixes i això ens allunya del coneixement ja que se salta un nivell de catalogació més intensiu. Es produeix una separació entre el coneixement i la descripció dels objectes, que es pot evitar si s'utilitzen eines com *linked data* o *open data*. No està clara, en els arxius fotogràfics, la forma de lligar les dades del coneixement amb els objectes corresponents i així, per exemple, el camp *creator* no és un URI (Universal Resource Identifier) a una fitxa linked data, sinó un parell *Cognoms, Nom* en la majoria de casos.

E) Cal saber clarament quina és l'estructura del coneixement que té el material catalogar per a fer un descobriment del coneixement. En el cas de les fotos fixes el marc de coneixement serà la producció cinematogràfica associada i els seus agents. Curiosament les fotos fixes tenen una càrrega d'autoria molt fluixa, ja que el foto fixa és dins d'un rodatge un personatge que depèn del productor i no del director, i genera un subproducte per a la promoció del film. Aquest fet fa que en moltes obres de literatura cinematogràfica no apareix-hi citat o que, fins i tot, en més d'un crèdit cinematogràfic no surti el seu nom, així que en la majoria de casos no es podrà aportar aquest coneixement. Alguns problemes d'autoria es podrien resoldre, tot i que queda molt lluny d'aquest estudi, aplicant eines de mineria de dades directament a les mateixes imatges a partir de les obres d'autor conegut, com fa Manovich (2009).

⁹ Les sigles esmentades es refereixen a estàndards de metadades, sobretot incideixen en el tipus d'arxiu, els termes que es poden emprar, la forma de comunicar la informació entre gestors de dades diferents, la localització única dels objectes digitals i la relació entre ells. Aquests estàndards són complementaris, tot i que Dublin Core acostuma a ser el que els uneix. Per a consultar més informació sobre tots aquests estàndards un bon punt d'entrada són els documents del Dublin Core Metadata Initiative (<http://dublincore.org>) o el World Wide Web Consortium (<http://www.w3.org>).

5. Metodologia prevista per a l'elaboració de la recerca

Per tal de desenvolupar la investigació s'ha considerat apropiat combinar diferents mètodes, que s'han aplicat per fases:

1) Anàlisi documental

En primer lloc s'ha realitzat una **anàlisi documental** del procés de catalogació del fons fotogràfic. Estudiant tota la documentació associada, per exemple les guies, les plantilles i els protocols establerts i la seva evolució en el temps en els 11 anys que va durar el projecte. Aquesta documentació s'ha estudiat també per a veure com es transmetia el coneixement previ dels que la van realitzar. En aquest apartat he aprofitat la meua experiència professional com a supervisor en l'entrada de les dades per part dels estudiants i per tant, coneixedor dels processos d'introducció de dades. Després se n'ha fet una tria per a contrastar com s'aplicaven aquestes directrius i si efectivament eren seguides pels catalogadors.

2) Estudi històric

S'ha realitzat un **estudi històric** en profunditat d'una part del fons utilitzant fonts documentals i bibliogràfiques, concretat en les fotografies del fons realitzades per **Emili Godes**. Aquesta part del fons servirà a les nostres intencions ja que és prou ampli, tant cronològicament com quantitativament, com per servir de referència en tot el procés.

3) Descobrimet de Coneixement en Bases de dades

Després d'aquesta primera fase d'anàlisi documental i estudi en profunditat, entrem en la segona part de la recerca, on s'aplica un procés de **Descobrimet de Coneixement en Bases de dades** utilitzant la **minería de dades**, necessària donat el volum d'informació, així com eines de **visualització** per a extreure de forma global el coneixement i contrastar-lo amb l'adquirit tradicionalment. En aquesta fase del treball he comptat amb l'assessorament de l'expert Xavier Rubio Campillo, investigador del Barcelona Supercomputing Center¹⁰, per a cobrir les possibles mancances i dubtes conceptuals i per a comprovar que la feina s'estava realitzant seguint uns estàndards científics.

Per a aquesta fase del treball he exportat la base de dades a documents en format XML que genera l'aplicació *Mobydoc Filmo*¹¹. He dissenyat, creat i implementat un programa, en el llenguatge de programació PHP, que realitza les següents funcions:

¹⁰ Es pot consultar a <http://www.bsc.es/about-bsc/staff-directory/rubio-campillo-xavier> o a <https://bsc-es.academia.edu/XavierRubio> la seva fitxa d'investigador

¹¹ Mobydoc (www.mobydoc.fr) és una empresa francesa que esta especialitzada en programari per a museus i arxius. La base de dades *Mobydoc Filmo*, que és la que utilitza la Filmoteca de Catalunya per a catalogar les seves col·leccions fílmiques, va ser creada a partir de l'experiència de la mateixa Filmoteca de Catalunya, es pot consultar la seva gènesi completa a Calle (2005) i la seva descripció a l'**apartat 8.1.1** d'aquest treball. Inicialment havia de ser utilitzada per més museus i arxius de la Generalitat de Catalunya, en versions adaptades a aquests equips, però finalment no es va realitzar aquesta implementació.

- a) Tria la sortida de dades, tant decidint quins camps s'exporten, com limitant els registres amb condicions de cerca. Aquestes condicions estan regulades per un document d'especificacions XML creat expressament.
- b) Transforma les fitxes en format tabulat, llegible pels programes de mineria de dades i visualització, per a poder fer cada estudi concret.
- c) Prepara les dades dels camps, ja que la majoria de vegades aquestes no es poden utilitzar directament donat que poden tenir buits d'informació, dades redundants o ser formalment incorrectes.

Per a poder fer un procés de mineria de dades, depenent dels algorismes que s'utilitzin, pot ser necessari realitzar conjunts de registres tant per comprovar, com per entrenar els algorismes. S'ha triat com un d'aquests conjunts de contrast les fitxes de les fotografies d'Emili Godes, ja que consta d'un volum suficient d'informació, aproximadament 1440 fotografies que correspondrien a un 7% de la col·lecció del fons Reproducciones Sabaté, i se'n té prou coneixement gràcies a l'estudi de cas. A això s'afegeix el fet que estudis recents han atribuït obres noves a l'autor que encara no consten a la base de dades en l'estat en què s'ha realitzat l'exportació per a l'estudi. Aquest fet serà utilitzat per a veure si el procés de mineria de dades aconseguirà trobar les connexions que han permès atribuir les noves fotografies, com pot ser la dedicació en exclusiva a un estudi cinematogràfic en un període de temps concret.

S'han utilitzat eines informàtiques de visualització per a poder aplicar la mineria de dades i interpretar el coneixement descobert. Els programes avaluats per al procés de mineria de dades i visualització són Weka (<http://www.cs.waikato.ac.nz/ml/weka/>) i Orange (<http://orange.biolab.si/>). Aquests són programes especialitzats en aquests processos que tenen interfícies visuals que faciliten l'aplicació dels algorismes i creen gràfiques interactives que serveixen per a poder fer amb facilitat el Descobriments de Coneixement en Bases de dades. Les característiques d'un treball escrit no permetran que es pugui mostrar aquesta interactivitat, la qual cosa s'intenta pal·liar amb les explicacions del text.

6. Anàlisi documental de la catalogació del fons estudiat

En els següents apartats farem primer una introducció als fons gràfics de la Filmoteca de Catalunya, i del fons que ocupa aquesta recerca, i una anàlisi documental de la catalogació del fons estudiat.

6.1. Origen del material de catalogació que ocupa aquest treball

6.1.1. Descripció dels fons de l'Arxiu gràfic de la Filmoteca de Catalunya

Els fons documental de la Filmoteca de Catalunya és compost, aproximadament, per 500.000 objectes. La paraula objecte no és triada a l'atzar, ja que la gran varietat de materials que hi són conservats va des d'aparells cinematogràfics, passant per llibres, fotografies, fins al marxandatge que promociona un film. Per a veure'n la complexitat només cal veure l'arbre de les **figures 6.1 i 6.2**, que es limiten al material que gestiona l'arxiu gràfic, una de les divisions de treball que té l'àrea de documentació. La diferència entre les **figures 6.1 i 6.2** és un exemple dels compromisos als quals s'ha d'arribar en els diferents aspectes de la gestió del patrimoni.

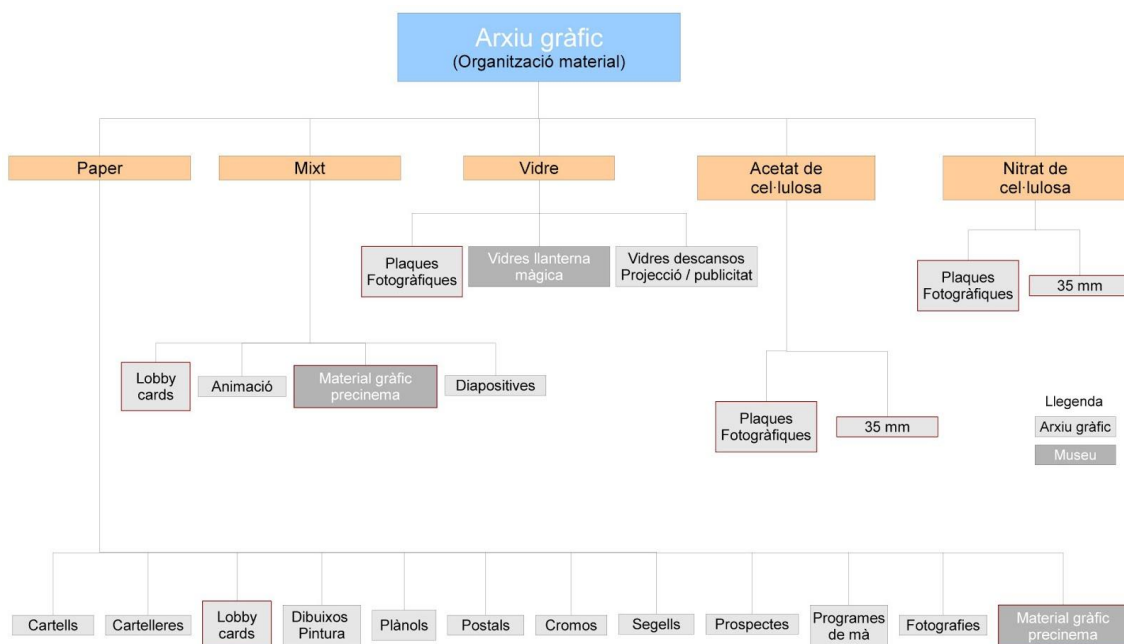


Figura 6.1. Classificació per tipus de materials dels objectes de l'arxiu gràfic del fons documental de la Filmoteca de Catalunya. Font: Filmoteca de Catalunya

A la **figura 6.1** veiem una classificació per materials, amb la qual cosa podem intuir les diferents decisions de conservació, preservació i restauració que s'han de prendre en un arxiu. Les decisions de conservació van des dels diferents llocs on s'han d'emmagatzemar els materials fins als contenidors o el material de conservació que s'ha d'utilitzar depenent dels materials que formen l'objecte, que en aquest cas poden ser múltiples, com acetat de cel·lulosa, paper, vidre, fusta o metall. La diferència entre la preservació i la restauració pot semblar subtil, però moltes vegades els problemes d'urgència o de pressupost fan que es decideixi preservar passant per davant, i a vegades impossibilitant, la restauració. La restauració és la restitució de l'objecte a un

estat el més semblant possible a l'original (Oprea, 1997) i la preservació és conservar activament l'estat actual de l'objecte, sobretot quan aquest és en perill de desaparèixer.

Quan compromet la preservació una restauració futura? Per exemple, quan el problema prové del contenidor original, la qual cosa fa que s'hagin de desmuntar àlbums amb coles que fan malbé les fotografies o treure anelles rovellades de *lobbycards* (fotografies encartonades que es col·loquen a les entrades dels cinemes). Aquestes accions dificulten la restitució a l'estat original de l'objecte i normalment les fem obligats per les circumstàncies. Per exemple, els contenidors originals de les fotografies d'aquest estudi estan conservats al magatzem del Departament de Cultura ja que ni la seva importància, ni les seves necessitats de conservació, no es consideren suficients per a ocupar espai a les zones de conservació de la Filmoteca.

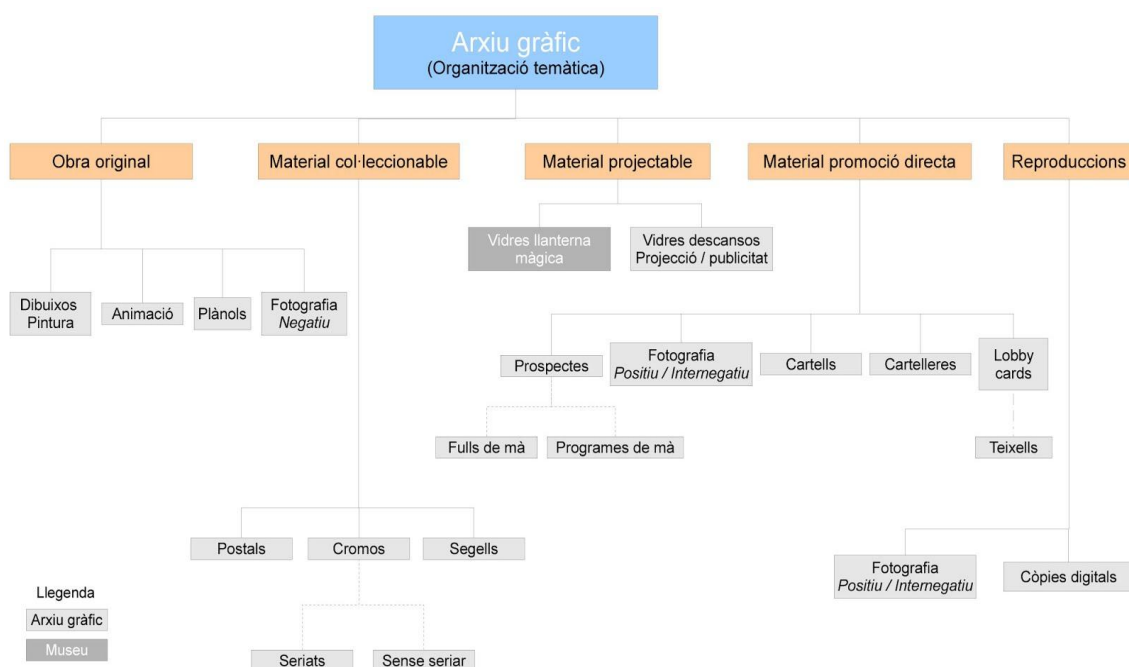


Figura 6.2. Classificació temàtica dels objectes de l'arxiu gràfic del fons documental de la Filmoteca de Catalunya. Font: Filmoteca de Catalunya.

Quin és el motiu de conservar aquests contenidors un cop documentats? El motiu és que el coneixement sobre la nostra col·lecció s'incrementa a mida que cataloguem cada cop més materials i a vegades un element amb el qual no havíem comptat pren importància. Un exemple d'això és que una convenció a la qual es va arribar en la descripció dels negatius fotogràfics va ser apuntar a la base de dades les mides estàndard, una aproximació en xifres arrodonides, i no les mides exactes en mil·límetres, per a millorar els resultats de la cerca de materials a la base de dades. Aquesta opció, que semblava innòcua, va resultar errònia quan es va demostrar que en molts casos era molt complicat distingir entre les plaques de 18 x 24 cm d'acetat de cel·lulosa i les de nitrat de cel·lulosa dels anys 50. Un estudi sobre un conjunt de materials va mostrar que el costat llarg dels nitrats era una mica més curt i que, per tant, trobar un material amb unes mides més grans de 236 mm suposava classificar-lo directament com acetat i estalviava una agressiva comprovació. Aquesta dada no estava a la literatura científica i per tant no es podia tenir en compte quan es va decidir abreviar aquesta informació.

A la **figura 6.2** podem veure com s'organitza el material des del punt de vista de les decisions de catalogació atenent al contingut del material. Les decisions són múltiples, ja que depenen per una banda de si formen part d'una sèrie, o són autònomes, i si depenen d'un film o no. Pot donar-se el cas que el que, aparentment, és el mateix tipus d'objecte pot variar depenent del que hi ha imprès al revers, ja que moltes vegades el distribuïdor del film feia un full de mà imprès per una banda amb una versió més petita del cartell del film i podia arribar a l'arxiu el full de mà sense revers, o imprès al revers amb la programació del cinema, o, fins i tot, amb la impressió d'un anunci comercial. Cada un d'aquests objectes té unes característiques molt concretes que afecten a la seva catalogació i gestió documental.

Acompanyant aquesta varietat i complexitat ens trobem que el fons sobre el qual gravita la Filmoteca de Catalunya és el fílmic, que és compostat per 180.000 llaunes de pel·lícula. Aquest centre gravitatori marca moltes de les decisions de la institució, ja que és la raó de ser de la Filmoteca. La Filmoteca forma part de la Federació Internacional d'Arxius Fílmics i, altra cop, el nom de la institució ens marca el punt principal d'acció, les pel·lícules.

6.1.2. Descripció del fons Reproducciones Sabaté

Tot i que ara depèn de l'àrea de documentació aquest fons va ser catalogat des de l'àrea de l'arxiu fílmic. Aquest context va marcar la decisió de catalogar aquest fons de forma intensiva i no un altre, i que les dades fossin compartides en la base de dades fílmica, anomenada Mobydoc, que ja va ser creada pensant en aquesta combinació.

Els motius de la tria van ser múltiples, algun d'ells basats en l'interès, entès en dues vessants: el fons més interessant des d'un punt de vista patrimonial, però també el que interessa més a la institució. Un altre dels motius de tria, aquest cop en el sentit d'acotar, és que la catalogació s'ha de poder fer en un temps raonable i aportant uns recursos raonables, amb la intensitat que s'ha decidit. També va pesar la decisió, que expliquem en el següent apartat del treball, de qui serien els encarregats de catalogar aquest material.

El fons escollit està format per negatius fotogràfics que eren guardats pel laboratori Reproducciones Sabaté i que van ser adquirits per la Filmoteca de Catalunya. Aquest laboratori conservava els negatius fotogràfics fets pels foto fixes durant els rodatges cinematogràfics per a fer-ne reproduccions, tant per a premsa com per a promoció del film.

El seu interès des del punt de vista patrimonial ve donat perquè es tracta de material original, dit així en el sentit que en la seva major part es tracta de material directe de càmera i revelat pel mateix autor de les imatges i que és, a més, material fet durant el mateix rodatge del film. Tot això, que ja de per si és un valor per a un arxiu fotogràfic, té un valor afegit per a un arxiu fílmic, ja que a vegades pot servir per documentar films que s'han perdut, o reforçar la catalogació dels conservats ja que són una mena de crònica gràfica del rodatge. Per exemple, en època de la postguerra espanyola s'aprofitava material fotogràfic provinent dels rodatges, ja que en els bastidors de les càmeres cinematogràfiques queda pel·lícula sense impressionar que és compatible amb les càmeres fotogràfiques. Aquest suport cinematogràfic té unes inscripcions que són úniques de cada exemplar fílmic negatiu i que poden permetre identificar gràcies a

les fotografies un material fílmic del qual no es coneix la procedència. Aquest material original també és molt útil per a veure la forma en com treballaven els fotògrafs de l'època, la qual cosa és de molta utilitat per als historiadors de la fotografia i els del cinema. Per tant, aquest fons complia el fet de ser interessant per ell mateix i per a les tasques de la institució.

Aquest fons té un nombre estimat de 200.000 negatius fotogràfics, la qual cosa el deixaria fora de l'abast del que es podria considerar un "temps raonable de catalogació intensiva". Per a reduir el nombre d'exemplars a catalogar es va decidir acotar la catalogació als films de producció espanyola, amb la qual cosa es va quedar el nombre en 20.000, els quals es van catalogar en 12 anys, des de l'any 1999 fins a 2011. Un cop feta aquesta tria es va donar prioritat a la catalogació dels materials més antics, ja que es va suposar que podien tenir unes condicions de conservació pitjors.

Un altre dels motius per utilitzar aquest fons és que era a priori molt homogeni i controlat, la qual cosa fa que sigui més fàcil establir procediments estandarditzats de catalogació i que la formació sigui més senzilla, sobretot a persones no especialitzades com en el cas que ens ocupa. Si pensem en la complexitat dels fons documentals que hem comentat anteriorment veurem que si s'ha de formar en la catalogació és més senzill partir de casuístiques com més senzilles millor. Tot i això en el material de postguerra s'ha pogut veure com un fotògraf pot fer servir diferents tipus de material, fins i tot utilitzant càmeres diferents, amb la qual cosa, com passa habitualment, les coses no són mai tant senzilles com un s'espera.

Algunes de les sorpreses van venir de les fonts bibliogràfiques existents, que provenen majoritàriament d'Estats Units i de França, ja que no coincidien algunes de les informacions publicades amb el que es va trobar en la catalogació. En aquests països el cinema és des dels inicis una activitat industrial, quan a Espanya té una part artesanal i una part d'aventures empresarials en fallida contínua, per la qual cosa no s'apliquen igual alguns processos industrials. Un cas concret que no coincideix amb aquestes fonts és el de la pervivència del material de nitrat de cel·lulosa, que hem documentat que s'utilitza fins l'any 1956, 4 anys més tard del que diu la literatura especialitzada. En els primers anys una part de la formació era deixar clares quines eren les dates d'ús de cada material i en els següents, després de constatar la seva poca fiabilitat, la màxima va ser dubtar de tot i comprovar cada material.

6.2. Gènesi de les dades

6.2.1. Qui va catalogar el fons. Personal en formació *versus* personal professional

Quan es va realitzar el projecte de catalogació es va decidir fer-ho en el marc de les pràctiques d'una assignatura d'Història del cinema. Un dels motius de fer-ho així va ser poder donar l'oportunitat a estudiants, tant de llicenciatura com de doctorat, de provar la investigació històrica, centrada en els documents disponibles en arxius, la qual era una de les possibles sortides a la carrera. En tot el procés van passar per l'arxiu més de 400 estudiants, la qual cosa donaria una mitjana de 35 per curs.

Una de les funcions de la Filmoteca de Catalunya és la formació en la cultura cinematogràfica (Llei del Cinema, 2011). Aquesta funció de formació s'ha contemplat sempre des del sentit ampli del terme i quan hi va haver la possibilitat de fer un

conveni amb el departament d'Història de l'Art de la Universitat de Barcelona, no es va desaprofitar la oportunitat.

També, com en qualsevol decisió d'un equipament cultural, es van tenir en compte dades més prosaiques com la quantitat de personal necessari per a realitzar la feina. Es va calcular que farien falta 4 persones fixes per a aconseguir el mateix volum de catalogació previst. Al conveni es va establir que hi hauria una persona encarregada des de la institució, un expert en fotografia de suport, un becari i la coordinació d'un professor de la universitat.

Un dels avantatges d'utilitzar personal no vinculat a la institució va ser que moltes de les tasques de documentació i recerca, com cercar una còpia editada de la pel·lícula en vídeo o obtenir les dades mercantils sobre l'empresa productora, es podien fer fora de l'arxiu. L'administració té, a excepció d'algunes proves recents de teletreball, molta rigidesa en la presència al lloc de treball, la qual cosa fa difícils les tasques de recerca cinematogràfica.

Quan es va arribar a la fase de repàs de les dades introduïdes es va comptar amb estudiants en pràctiques de la facultat de Documentació i Biblioteconomia. Això es va fer així ja que es considerava que les correccions a fer eren formals o relatives a la part documental de les fotografies. També documentalistes van realitzar el traspàs de dades des de documents de text, sistema amb el qual es va començar inicialment, ja que no hi havia base de dades operativa.

Entre els contres a la utilització de personal en formació hi ha varis factors, sobretot la provisionalitat de les dades entrades pels estudiants, la disparitat de qualitat entre elles i la varietat de criteris provocada pel mateix gran nombre de catalogadors. La tasca de repàs de les dades ha perdurat uns anys després del projecte, ja que les dades aportades superaven la capacitat de comprovació anual. Apart, sense que hi hagués hagut aquests factors, la tasca de catalogació és una tasca de revisió contínua, o hauria de ser-ho. El coneixement, com comentàvem, s'incrementa a mida que es cataloga i, per exemple, la base de dades de la Fílmoteca ha sobrepassat amb escreix el nombre d'actors reconeguts que es pot trobar a la bibliografia existent.

Un altre dels problemes d'utilitzar personal en formació és que en alguns ambients acadèmics aquest fet desprestigia les dades i el seu valor, tot i que aquestes siguin revisades per personal professional.

6.2.2. Descripció tècnica del procés de catalogació

a. Formació del personal catalogador

En els inicis s'impartia una classe magistral a la universitat explicant els processos i la fitxa amb els camps de la base de dades que s'havien d'omplir (veure **annex 1**). Aquesta classe es va demostrar innecessària, ja que els estudiants no retenien la informació, amb la qual cosa en els següents anys es va limitar a explicar magistralment els aspectes cinematogràfics com els tipus de pla, la funció del foto fixa o el funcionament dels rodatges.

La formació es va centrar en la pràctica i en els grups petits, quan era possible. Es va constatar com de diferent és el nivell de concentració quan el grup es prou petit per a provocar la consciència que el formador té contacte visual amb tots els membres.

També és molt important per evitar que sigui un dels estudiants que transmeti alguna informació a un altre estudiant que no ha sentit bé una explicació, enlloc del formador. Finalment, es repartien les caixes de fotografies a catalogar després d'una ràpida formació en la fitxa que s'omplia directament (veure **annex 2**). Per a repartir-les, ja que tenien volums i complexitats diferents, el formador comptava amb les impressions que havia tingut dels estudiants en aquest primer contacte depenent, sobretot, de l'interès observat o de la demostració de coneixements previs en cinema. Aquestes impressions no sempre eren fiables, ja que es feia de forma intuïtiva i no reglada anteriorment de forma qualitativa.

b. Material auxiliar

Els estudiants omplien una fitxa en paper, la qual es pot veure a l'**annex 2**, per les dues cares de forma manual amb un llapis. Aquesta fitxa es va anar afinant a mida que s'anaven trobant particularitats, sobretot introduint opcions que fos possible marcar amb creus, afegint en alguns casos espai per a observacions. Amb això per una part es controlava que no es possessin dades incorrectes i per l'altra es facilitava posteriorment la introducció de dades a l'ordinador, ja que seguia el mateix esquema.

Un dels canvis més grans de la fitxa va ser l'ordre en què es presentaven els camps. Inicialment aquest ordre no estava basat en la pràctica, sinó que estava plantejat per facilitar-ne la consulta. Es va comprovar que aquest ordre obligava a fer salts per la fitxa per omplir-la i, per tant, es va ajustar l'ordre de forma empírica al dels processos de catalogació i no als de representació de la fitxa. Un dels camps que va canviar totalment de posició va ser el dels **campus Estat físic i Estat Químic**, que en el primer moment estaven a dalt de l'**apartat Examen de conservació** ja que es consideren uns camps cabdals per a prendre les decisions de preservació i tenien un lloc preeminent en la consulta. L'observació va portar, però, a veure que és un camp que no es pot omplir fins que es coneixen tots els punts anteriors, la qual cosa porta a que fos finalment l'últim camp a omplir.

Aquesta estructura d'un foli per dues cares va ser utilitzada per a facilitar la tasca de descripció de les fotografies manualment sense ocupar molt de lloc al limitat espai de treball. També reflectia una part de l'estructura de la base de dades, ja que aquesta constava d'una jerarquia de tres taules, simplificant, que era la següent: Fitxa film -> Fitxa Fotografia -> Exemplar Fotografia i en aquest procés s'omplia la informació corresponent a les dues últimes taules.

També disposaven de material de conservació per a guardar definitivament la fotografia catalogada un cop omplien la part de descripció d'aquesta. En un sobre es posava el codi d'identificació i aquests sobres es col·locaven en caixes, separats per tipus de suport.

c. Processos de catalogació

Atès que el material fotogràfic que es catalogava era format negatius fotogràfics el procés de catalogació era seqüencial i el mateix per a tots els catalogadors, ja que normalment no es té l'habilitat d'interpretar els negatius i es fa necessària la seva digitalització i positivat per a descriure'ls. Aquesta característica seqüencial marcava molt la forma d'organitzar l'agenda de treball dels catalogadors, ja que aquests tenien un temps limitat, un quadrimestre en el cas dels estudiants de llicenciatura o grau, un

curs en el cas dels estudiants de màster o doctorat. En els següents punts descrivim els processos seguits.

c.1. Descripció del material

Va més enllà d'aquest treball descriure punt per punt la fitxa d'entrada de dades que es pot veure a l'**annex 2**, així que ens centrarem en el que eren les decisions que marcaven la catalogació i la posterior entrada de dades.

El primer punt de la fitxa, en el que sí ens estendrem, era entrar el número de registre, format per una combinació del número de registre que tenia el film a la base de dades comuna, amb el codi ".F" -que indicava que es tractava d'una fotografia-, i finalment un codi seqüencial alfanumèric de dos caràcters. Aquest número de registre, que es fa servir també en la catalogació d'exemplars fílmics, és una herència directa de la fitxa de paper que s'utilitzava abans dels ordinadors pel que fa al seu format i pel que fa a la seva llargada limitada, és una herència del primer programa utilitzat com a gestor de dades, anomenat DAC¹², que limitava aquest camp a 10 caràcters. El fet que sigui un codi descriptiu facilita que amb un cop d'ull es pugui veure de quin tipus de material es tracta, però suposa problemes conceptuals ja que en una base de dades és molt millor, sobretot per a inventariar, que el codi no sigui descriptiu sinó únic i incrementable.

Un dels problemes inesperats en la formació va ser la dificultat per a un nombre elevat d'estudiants d'entendre la necessitat de la transcripció literal de les inscripcions trobades, tant a contenidors com a materials, i per tant mantenir una coherència en aquest sentit en l'entrada de dades. Sobretot en les marques de fàbrica del material fotogràfic, un canvi en un espai o canviar una majúscula per una minúscula pot portar a errors en la identificació de materials. També era molt important la literalitat en el cas dels sobres originals ja que aquests se separaven del material quan es passaven als contenidors definitius per motius de conservació, ja que eren fets amb papers àcids i tenien fongs en la seva majoria. Els sobres originals duen escrits una sèrie de codis, molts dels quals aporten informacions com el destí de la fotografia o si és vàlida o no per a ser passada a positiu.

c.2. Pas a contenidors de conservació

El pas a contenidors definitius era senzill ja que els catalogadors l'única precaució que havien de tenir era la d'escriure de forma correcta amb llapis el número de registre en una camisa, un sobre obert per tres bandes, de paper barrera i després encaixar el material. Com a curiositat direm que es va convenir que es desarien les fotos amb el suport cap a davant per a facilitar la seva lectura, ja que els negatius es veuen així del dret sense treure'ls de la camisa i sense necessitat de girar-los ni tocar-los amb les mans. Aquesta opció que semblava la més pràctica es va demostra perjudicial per a la conservació, ja que el disseny de les caixes de conservació va fer que al final les fotografies quedessin guardades amb el suport a la part de dalt i, per tant, es balquessin més que si hagués estat de l'inrevés. L'observació posterior de com

¹² El programari DAC (Documentació Assistida de Col·leccions), va ser creat al 1992, gràcies a un acord entre la Generalitat de Catalunya i l'Ajuntament de Barcelona. Un programa que tenia com a objectiu sistematitzar i unificar la documentació dels objectes dels museus de Catalunya.

estaven col·locats els materials a les capsas originals va demostrar que originalment els tècnics de laboratori ja les desaven amb la part de l'emulsió cap a dalt.

c.3. Digitalització

La mida del fitxer generat en la digitalització va ser, en aquell moment, una de les decisions més polèmiques que es van prendre. Es va decidir realitzar escanejos de 2 Mb de mida, independentment de la mida del suport original, la qual cosa feia que diverses mides tinguin resolucions diferents, ja que els materials van des de fotografies de 35 mm, fins a plaques de 18 x 24 cm. La polèmica vindria per que aquesta mida està lluny dels anomenats estàndards de preservació digital. Es va considerar que la digitalització no era un substitució de l'original, sinó una estratègia per a que el material fos manipulat el menys possible i que la mida de 2 Mb, equivalent en els formats rectangulars a, aproximadament, 1600 x 1200 píxels, era suficient per a la seva consulta i, en realitat, per a la majoria d'usos editorials. Anys més tard de l'inici de la catalogació es considerava que la mida del costat llarg havia de ser de 3700 píxels per a preservació (Iglésias Franch, 2008) i actualment es realitzen digitalitzacions amb aquestes mides. També és cert que la relativa poca antiguitat del material permet no ser agressiu en aquest sentit, ja que no hi ha suports que tinguin degradacions actives i que necessiti ser preservats digitalment perquè estiguin en perill de destrucció.

De la mateixa forma que hem comentat que es va acotar el material a catalogar per a prendre una decisió de gestió, com ara que les tasques s'han de poder realitzar en un temps raonable, també la decisió de la mida del fitxer va anar en aquest sentit. Es va preferir que els catalogadors, que recordem estaven en formació, es prenguessin el seu temps en entendre com s'escaneja, aprendre què és un objecte digital i ajustar manualment el nivell de grisos de la fotografia, que en esperar un llarg temps de digitalització. Hem de comentar que qualsevol increment en la mida del fitxer augmenta exponencialment el temps de processat. Per tant, es van tornar a unir les necessitats de l'arxiu en el sentit de reforçar la formació (la major part dels estudiants no havien digitalitzat mai una imatge fins i tot entrada la dècada de 2010) i la de fer que la feina fos possible amb les eines disponibles, que al final del projecte eren 3 escàners. Actualment aquestes feines s'externalitzen, amb la qual cosa en alguns casos es perd el control de l'objecte digital final.

c.4. Documentació i visionat del film

Quan era possible es visionava el film durant el qual s'havien fet les fotografies. Això facilitava la descripció de les imatges i permetia tenir accés a una font primària que aporta més dades que molta bibliografia. Tot i que el cinema és una tasca col·lectiva la bibliografia, excepte en els catàlegs que són exhaustius, acostuma a incloure una fitxa tècnica abreujada. El problema amb la informació que aporten els títols de crèdit dels films és que moltes vegades contenen errors en els noms dels participants, o fins i tot, errors o transformacions voluntàries. Per a poder tenir informació fiable s'ha de consultar els anomenats cartons de rodatge¹³, que eren necessaris en temps de la

¹³ Els cartons de rodatge eren unes fitxes que s'havien d'enviar a la comissió de censura abans de fer un film i que contenien, entre altres coses, una relació dels participants en el rodatge que havien de ser acceptats abans de començar aquest. Les manipulacions dels productors podien venir donades perquè en alguns períodes hi havia un topall d'estrangers que es podien

dictadura per poder rodar, però que fins i tot a vegades són manipulats pels productors per a poder cobrar ajudes o realitzar el film.

Un dels aspectes en què s'insistia era que és imprescindible prendre notes durant el visionat i que el temps entre veure el film i catalogar les fotografies ha de ser el menor possible, ja que la retenció dels detalls de la trama es perd en molt poc temps.

c.5. Descripció del contingut

Les condicions per a la descripció del contingut de les imatges eren molt desiguals, i una de les condicions que canviava molt la forma d'enfrontar-se a les imatges era haver visionat el film. La consulta d'aquest ofereix dos avantatges, el primer que és té accés a una font primària d'informació com són els crèdits com hem comentat; el segon, que es té coneixement del context, el qual és molt útil a l'hora de descriure les imatges estàtiques. No obstant, aquest visionat dóna un problema, precisament producte d'haver adquirit informació del context, i és que es té molta més informació de la que es veu realment a la imatge i és té un record molt fort de l'acció que transcorre a la pel·lícula. Per a entendre perquè això suposa un problema hem d'explicar com es feien les fotografies fixes: la majoria es feien en els descansos del rodatge i no mentre es rodava, amb la qual cosa no són mai exactament el que es veu al film: els actors es col·locaven de forma semblant a l'escena rodada, però no exacta. A més, fins a els anys 60 les emulsions fotogràfiques són poc sensibles, i per tant les imatges es triguen més a realitzar i són molt més estàtiques, ja que qualsevol cosa que no estigui quieta el temps suficient sortirà moguda. Per si això no fos suficient a vegades es conserven fotografies d'escenes censurades o destinades a països estrangers, de preses modificades, o de modificacions com actors que ja porten el vestuari de l'escena següent. Per tant, tenim la força d'un record que pot no coincidir amb el que veiem realment i algunes vegades, a més, es tendeix a fer la descripció adaptant-la a aquest record i no tant al que veiem.

Per a solucionar els problemes que suposa el visionat, per exemple incloure un personatge que està fora de pla en la descripció de la imatge, es va decidir separar en dos paràgrafs aquesta descripció. En el primer paràgraf es descrivia el que es veia en la imatge, i en el segon es contextualitzava en l'acció del film, si era possible. La primera part es va decidir, si ho permetia la imatge, que es composaria d'una primera frase situant el lloc on transcorria aquesta, una següent amb l'acció principal i després tot el que servís per a aportar informació útil. La part de contextualització ha portat alguns problemes quan s'ha volgut utilitzar per a una finalitat diferent, com és la de divulgació. Com s'ha comentat, moltes vegades la descripció de les fotografies fixes serveix per a documentar films que no s'han conservat, amb la qual cosa es contextualitza de forma completa i, per tant, es donen detalls de la trama. Aquests detalls, que són útils per a la tasca de l'arxiu, en el cas de la divulgació es converteixen en el que s'anomena *spoilers* (parts de la trama que poden espatllar el visionat del film a qui no l'ha vist, com el final del film o el destí d'algun personatge).

Un punt apart va ser la introducció de les paraules clau, ja que es va prendre la decisió

contractar o perquè cada treballador havia d'estar donat d'alta en una professió concreta, però podia ser vàlid per a altres funcions. Això feia que a vegades es possessin noms en el cartó que no participaven en realitat en el rodatge.

de no limitar als introductors la creació de noves paraules clau per a no alentir el procés de catalogació. Quan es limita el tesaurus de paraules clau es crea l'obstacle de dependre de l'acció d'una altra persona quan no existeix un terme, tot i que els catalogadors estaven obligats a consultar cada canvi és diferent que esperar a que es creï un terme. Utilitzar un tesaurus tancat obliga a tenir coneixement expert d'aquest, i no hi havia temps per la formació en aquest aspecte. La revisió dels termes introduïts es feia feixuga, ja que no es podia buscar termes per data de creació, ni hi havia forma de saber quantes vegades era fet servir un terme, a més, els catalogadors tot i que estaven obligats a consultar l'entrada de nous termes a vegades no ho feien.

c.6. Entrada de dades al gestor de bases de dades

Els primers dos anys les dades es van entrar en una plantilla de document Word, ja que el programa de gestió de dades que s'utilitzava, anomenat DAC (Documentació Assistida de Col·leccions), va ser deixat d'utilitzar per la Filmoteca, ja que donava problemes continus de creuament de dades. Un dels problemes que té la Filmoteca amb els gestors de dades habituals en centres patrimonials és que cada fitxa d'una pel·lícula té moltes més dades de les habituals; recordem que el cinema és una tasca col·lectiva i molt tecnificada, amb la qual cosa genera un volum de dades en la seva descripció molt més gran que el que els objectes amb què tracten altres institucions. Aquest volum de dades, moltes d'elles enllaçades i relacionades, fa que moltes vegades hi hagi hagut dificultats amb el programari.

El fet d'entrar les dades en un document de text va obligar després, quan es va començar a utilitzar el gestor Mobydoc, a introduir aquestes dades de nou. Això va afegir el problema del temps transcorregut entre la introducció inicial i la posterior, i que aquesta era feta per una persona diferent a la inicial i amb un programa molt diferent, amb la qual cosa moltes vegades hi ha una reinterpretació de les dades.

7. Estudi de cas: el fotògraf Emili Godes

Emili Godes i Hurtado (Barcelona, 1895 - 1970) va ser un fotògraf conegut per diferents tasques, tot i que les més apreciades, fins al moment, són en la fotografia científica i macrofotografia, i com a membre del que s'anomena la *Nova Objectivitat* (IEFC). Era un fotògraf amb una obra d'una gran qualitat del qual es va fer una exposició l'any 1996 al Centre d'Art Santa Mònica¹⁴ i que forma part de la col·lecció permanent del MNAC¹⁵, així que les seves fotografies destaquen sobre la resta de foto fixa, juntament amb les d'altres fotògrafs reconeguts com Pérez de Rozas. Pel que fa a la foto fixa s'ha pogut constatar consultant els anuaris professionals¹⁶ que posseïa el carnet del *Sindicato nacional del espectáculo* amb la classificació de *Foto fija* per poder participar en rodatges des de l'any 1943 a l'any 1963. A la Filmoteca de Catalunya es conserven fotografies seves des de l'any 1935 fins a 1961 (veure relació dels títols de films als que corresponen a la **figura 7.1**).

Títol	Any	Fotos	Observacions
Farándula, La	1935	39	Coautor amb Armando Seville
Linda Beatriz, La	1940	62	
Vidas cruzadas	1942	56	
Huella de luz	1942	61	
Viaje sin destino	1942	38	Citat anuari 1943
Boda de Quinita Flores, La	1943	44	La Vanguardia 28/06/1943
Muralla feliz, La	1947	81	
Sirena negra, La	1947	73	
Parsifal	1951	201	Creiem que només una part són seves
Muchachas de Bagdad	1952	64	
Hay un camino a la derecha	1953	83	
Ases buscan la paz, Los	1954	40	
Cañas y barro	1954	67	
Padre pitillo, El	1955	74	
Vida es maravillosa, La	1955	34	
Zalacaín el aventurero	1955	39	
Carta a Sara	1956	45	Citat anuari 1963
Duende de Jerez, El	1957	70	
Tirana, La	1958	24	
Música de ayer	1958	40	
Ya tenemos coche	1958	15	Poc probable per l'estil, però citat anuari 1963
Casco blanco, El	1959	49	
Siega verde	1960	48	
Plácido	1961	61	Citat anuari 1963 i part del fons IEFC
		1440	

Figura 7.1. Films amb fotografies d'Emili Godes als fons de la Filmoteca de Catalunya. Font: Miquel Àngel Pintanel¹⁷

¹⁴ *Emili Godes, fotògraf de la nova objectivitat*. (1996). Barcelona: ACTAR, Generalitat de Catalunya.

¹⁵ Es pot consultar a <http://museunacional.cat/ca/advanced-piece-search>, utilitzant el formulari de cerca, les obres que formen part de la col·lecció del Museu Nacional d'Art de Catalunya.

¹⁶ Valero de Bernabé, A. (1944). *España cinematográfica: recopilación de cuanto concierne al arte, industria y comercio del cinema español: [anuario 1943]*. Madrid: Cinegrafos.

¹⁶ *Anuario español de cine: 1963*. Madrid: Sindicato Nacional del Espectáculo.

¹⁷ A partir d'ara, si no s'especifica el contrari, les figures són elaborades per l'autor.

Com hem comentat en la introducció hi ha molt poca bibliografia sobre els foto fixes, així que part de la investigació s'ha hagut de fer a partir de fonts bibliogràfiques d'època i documentació. En la franja cronològica en la qual treballa per al cinema Emili Godes, de 1935 a 1963, hi ha un problema afegit quan es consulten els anuaris cinematogràfics i les guies professionals, entre les fonts que aporten més dades. En aquestes obres podem obtenir molta informació d'un apartat que s'anomena «Películas en las que ha participado», quan hi ha la llista de foto fixa en actiu. Aquesta secció pateix una mena de buit històric, ja que ningú cita les produccions realitzades durant la República o la Guerra civil. Això pot ser degut a que aquest era un apartat publicitari i la informació depenia del que volgués destacar el fotògraf, però tampoc es pot assegurar que no hi hagués una censura sobre aquest tema. Tampoc es conserven anuaris dels anys de la guerra, així que no hi ha forma d'omplir aquest buit amb una font similar.

PERSONAL TECNICO			

1-B	Director realizador.	CASTELLVI	25.000.-
3	Primer ayudante.	Salvador	6.400.-
7	Segundo ayudante.	Castro Blanco	3.500.-
5	Operador fotografo.	MACASOLI	10.000.-
11	Primer ayudante de operador.	Foriscot ...	1.000.-
13	Segundo ayudante de operador.	Ripoll.....	800.-
14	Fotografo (incluido revelaje)	. Godes ...	1.200.-
48	Director dialogos .	Martinez de Rivera .	1.000.-
47	Director propaganda y ambiente.	Elias .	1.200.-
36	Director decorados.	Gosch	1.500.-
34	Maquillador y ayudane.	Carrasco y.Sra ..	1.600.-
1	Secretaria produccion.	Kety	1.600.-
41	Montadora.	Balansó	1.200.-
40	Resiseur.	Oejuela	1.600.-
	Ayudante resiseur.	Manolo	600.-
			58.200.-

Figura 7.2. Pressupost inicial del film *La linda Beatriz* (Josep Maria Castellví, 1940). Font: FdC.

Mirant els pressupostos i les factures associades als rodatges de la companyia productora S. Huguet Selecciones Capitolio¹⁸ s'han pogut confirmar, o conèixer en algun cas, moltes dades sobre el foto fixa i les fotografies que produïa. Els fets que

¹⁸ Informació consultada en els fons patrimonials de la Filmoteca de Catalunya (es pot consultar a http://catalegbeg.cultura.gencat.cat/iii/encore/record/C_Rb1517601 la fitxa NODAC). Aquesta empresa va estar activa com a productora i distribuïdora, tot i que passà per diferents propietaris, des dels anys 30 fins als 70 del segle XX i a la Filmoteca es conserva la documentació de l'etapa en que és dirigida per la família Mangrané, a partir dels anys 40. De la majoria de films només es conserva els pressupostos i els llibres de comptes, però de dos d'ells, *Parque de Madrid* i *El duende de Jerez*, es conserven les factures que es van generar durant el rodatge. Emili Godes va treballar per aquesta productora a l'esmentada *El duende de Jerez* i a *Parsifal*.

s'han pogut constatar sobre el foto fixa són els següents:

- a) Lluny de ser un element accessori en un rodatge es pot veure que cobrava un import similar a un càrrec de tercer nivell i formava part del personal que estava en nòmina del rodatge i que ja apareix a la pàgina inicial del pressupost, així que la seva feina es considerava fonamental (en podem veure un exemple a la **figura 7.2**).
- b) El productor del film pagava el material fotogràfic negatiu amb el qual es feien les fotografies al rodatge, així que es confirma que era el propietari de les fotografies negatives originals i no el foto fixa. Tot i que aquest fet s'intuïa (en algunes caixes buides del fons s'ha trobat un paper que deia que s'havia tornat al productor el seu contingut), no s'havia pogut confirmar fins ara ja que es desconeixia quin era aquest contingut.
- c) Que la seva tasca és tan especialitzada que en les factures es veu també reflectida la figura del *reporter*, que és qui fa les fotografies en actes públics o publicitaris durant el rodatge, però que no està en nòmina com el foto fixa.

Tot i que s'hauria d'ampliar la recerca a més fons d'empreses productores, creiem que no canviaria gaire aquestes conclusions, ja que s'ha consultat documentació que va de 1941 a 1958 i no hi ha canvis apreciables.

Com hem comentat en l'apartat del marc teòric és molt important el coneixement a priori per a realitzar el Descobriment de Coneixement en Bases de dades. En els següents paràgrafs farem una explicació dels fets que coneixem a partir de la catalogació, i de l'estudi posterior, que hem realitzat, de les seves fotografies.

Per quantitat de material, Emili Godes és un dels fotògrafs més representats en els fons estudiats de la Filmoteca de Catalunya amb 1440 objectes, el 7% del total, que corresponen a 24 films, el 5% dels films. El nombre de films en els quals participa és elevat en comparació amb altres fotògrafs i això és un reflex de la ingent producció realitzada als estudis Orphea, que eren els més importants de Barcelona fins que van ser destruïts per un incendi el 1962, i amb els quals Godes tenia una relació que explicarem més endavant. Una altra de les coses que augmenta la quantitat de material és el fet que si la mitjana de fotografies per film en el conjunt del fons és de 24, la mediana de Godes és de 49¹⁹.

No hi ha al fons estudiat fotografies signades per Emili Godes, així que aquest no ha pogut ser un sistema vàlid per atribuir obres a l'autor. Moltes fotografies en paper de temàtica cinematogràfica acostumen a tenir un segell a la part posterior amb l'autor, però aquest no és el cas de les fotografies del fons del laboratori Reproducciones Sabaté, i les de Godes no en són una excepció. Això és degut a que el que es conserva en aquest fons en paper no són còpies destinades a la promoció del film, sinó proves realitzades per a comprovar la qualitat de les fotografies en el mateix

¹⁹ Utilitzem la mediana dels valors per comparar ja que aquest càlcul en minimitzar l'impacte en el resultat dels valors més apartats de la norma compensa la distorsió de *Parsifal* que té 201 fotografies o de *Ya tenemos coche* que en té només 15. Si utilitzéssim la mitjana serien 57,6 fotografies per film, però creiem un valor més ajustat a la realitat per fer comparacions en aquest cas el de la mediana. Les diferències entre la mitjana i la mediana en el conjunt del fons serien poc apreciables ja que té un volum molt més alt de fotografies, 19.900, i de films, 472, tot i que no se n'ha pogut fer la comprovació empírica per a aquest treball.

laboratori. Sovint es troben fotografies en paper duplicades amb diferents virats o temps de revelat, probablement per mostrar al client les diferents possibilitats d'acabat. No és per tant un material que l'autor consideri necessari signar, comptant que en el cas de la fotografia fixa se signa no tant per protegir els drets associats a les fotografies, que són del productor del film, sinó per a promocionar-se.

Per tant, en ser un material que ha vingut sense intermediaris de la productora al laboratori, directament del material o els seus embolcalls originals no es pot treure cap conclusió d'autoria, ja que, menys les fotografies d'estudi i les d'algun autor concret, com Pérez de Rozas, no van signades, ni tenen l'atribució de l'autor. Si hi ha alguna autoria reflectida normalment és la de la companyia productora. Només es podrien agrupar per autors les que tenen informació escrita al sobre ja que els signes i codis són únics de cada fotògraf, i els van repetint. També es poden treure conclusions de les marques escrites als materials; per exemple, algun conjunt fotogràfic s'ha pogut corroborar que era d'Emili Godes, després de veure que el número de clixé, que en les plaques fotogràfiques de 180 x 240 és escrit a mà amb tinta, era característic d'ell. A partir d'aquest element, i altres indicis, s'ha pogut investigar més aquest conjunt en concret i confirmar la possible autoria en la majoria de casos.

Tot i que les factures demostren que el foto fixa cobra directament del productor, cosa que corroboraria el que diu la «tradició» que depèn d'ell, les dades dels films semblen demostrar que les fotografies de Godes corresponen a diferents productors, però estan realitzades majoritàriament per a dos únics estudis: als estudis Orphea Films de Barcelona, o als Trilla-Orphea dels mateixos propietaris.

Al que hem comentat en l'anterior paràgraf hauríem d'afegir que les fotografies realitzades en localitzacions exteriors dels films atribuïts a Godes són de qualitat inferior a les realitzades a l'estudi cinematogràfic, així que en algunes es podria dubtar de la seva autoria. Tot i que les condicions de treball en exteriors són més complicades que en interiors, ja que hi ha més pressa per rodar i les condicions de llum són menys controlades, paràmetres que s'han de tenir en compte per valorar-ne la qualitat, hi ha factors, com l'enquadrament i l'ús de la profunditat de camp, que són inherents a l'estil del fotògraf i que són pitjors a les fotografies d'exterior. S'hauria de comprovar si en el mateix moment s'estaven rodant varis films als estudis Orphea Films i això provoqués que Godes hagués de fotografiar a l'estudi les de d'un film i no es pogués desplaçar a l'altra localització. Si es confirmés que les fotografies d'exteriors no fossin de Godes es reforçaria la relació directa del fotògraf en estudi.

Abans de saber que era el mateix productor qui encarregava la pel·lícula fotogràfica es podia pensar que era factible buscar canvis en la pauta dels materials utilitzats per a distingir entre autors i resoldre alguns problemes d'autoria com l'esmentat al paràgraf anterior. També, tot i que no s'aplicarà en aquest treball, és possible crear eines informàtiques que permetin mesurar la possibilitat que les fotografies siguin d'un autor de forma automatitzada mitjançant l'anàlisi de les còpies digitals (Manovich *et alii*, 2009), tant pel que fa a l'enquadrament, balanç de grisos, com a les imperfeccions produïdes per la càmera fotogràfica o les feines de revelat.

D'una forma més manual es podria saber, precisament per les transformacions provocades pel revelat en el negatiu, qui ha revelat les fotografies, ja que aquestes feines deixen marques com les de les pinces per assecar els negatius en les plaques

fotogràfiques, ditades o retocs²⁰. Però tornariem a no poder tenir-ne la certesa perquè pot ser diferent la persona que fa la fotografia que la que revela o retoca el negatiu, tot i que com podem veure a la **figura 7.2** Godes cobrava pel revelat, pot haver variat factors, com que treballi amb ajudants, que facin que no realitzi directament el revelat. Amb aquesta observació, per tant, podríem confirmar un autor si se l'identifica positivament, però no descartar-lo.

Com hem pogut veure a l'**apartat 6.2.1**, les condicions de catalogació no permetien realitzar el treball de documentació que s'ha realitzat per aquest treball de recerca. El nivell de recerca es limitava a utilitzar catàlegs del cinema espanyol i fonts directes com el visionat del film.

Quan les fonts d'informació són indirectes el seu ús va més enllà del que es pot considerar catalogació i passa al nivell de recerca històrica, i més quan el temps i els recursos són limitats. S'ha de considerar que en un rodatge participen desenes de persones i, per tant, no es pot dedicar un esforç de temps a cada un dels participants. Normalment la recerca es realitza consultant les fitxes tècniques del film a anuaris o catàlegs, o a les notícies a la premsa especialitzada, però no a partir de buscar si un integrant concret del personal del rodatge té un anunci a un anuari. També s'ha fet servir elements com la publicitat dels estudis Orphea Films²¹, que es conserva a la Fílmoteca de Catalunya, un cop l'avanç del treball de recerca ha fet arribar a la conclusió que hi havia una relació estreta amb Godes. Tot i que hi ha altres fotògrafs que treballen pels mateixos estudis, la informació aportada servia per contrastar fonts.

Una altra recerca que s'ha realitzat per a aquest treball ha estat la de la visualització de la digitalització dels fons atribuïts a Godes en altres equipaments patrimonials, concretament els de l'IEFC i el MNAC. A partir dels actors presents a les fotografies de films i a l'eina «Credited with» de les fitxes de persones de la web IMDb²² s'ha pogut saber a quin film corresponien fotografies que fins ara restaven sense identificar.

Els conjunts de fotografies atribuïts a Godes després d'aquesta recerca, i que s'han pogut contrastar amb l'estudi que està realitzant Laia Foix de l'Institut d'Estudis Fotogràfics de Catalunya sobre el mateix autor, s'ha gairebé triplicat.

També s'ha de dir que sense la catalogació realitzada amb anterioritat moltes de les dades aconseguides amb aquest treball de recerca no s'haguessin trobat. Sense els indicis materials, com el número de clixé esmentat, les característiques de les fotografies o la presència en les fotografies de la tramoia dels estudis Orphea no s'haguessin pogut atribuir bona part de les fotografies a Godes. Tota aquesta informació estava a la base de dades i no a la historiografia o als documents d'època. Difícilment s'hagués pogut catalogar les fotografies necessàries per a obtenir totes les de Godes, ja que moltes, abans de realitzar el treball, no s'haguessin triat com a possibles obres seves.

²⁰ En l'època en què treballa Godes els negatius eren retocats directament sobre el suport o l'emulsió, la majoria de vegades per motius d'estètica dels actors, però també per a resoldre problemes com negatius poc exposats amb màscares translúcides. Alguns d'aquests retocs són autèntiques obres d'art i podrien ser útils per a identificar materials.

²¹ Fulletó publicitari Estudios Cinematográficos Orphea Films, S.A. (1960).

http://catalegbeg.cultura.gencat.cat/iii/encore/record/C_Rb1486543 [Consulta 6/7/2014]

²² <http://www.imdb.com>

8. Aplicació del Descobriment de Coneixement en Bases de dades

En aquest apartat comencem descrivint l'organització de la base de dades i del programari utilitzat en la catalogació realitzada a la Filmoteca de Catalunya. En els següents subapartats descriurem i aplicarem el procés de Descobriment de Coneixement en Bases de dades aplicant diferents tècniques de mineria de dades, estadística i visualització, triant un grup de camps que considerem que ens poden aportar conclusions vàlides i útils en cada subapartat. Hem de constatar que en aquest treball no és possible encabir totes les possibles anàlisis, ni tots els camps de la base de dades, i hem triat les variables que creiem més representatives en cada cas. Això és així perquè la intenció de la recerca no ha estat fer una anàlisi completa de les dades, sinó demostrar amb exemples la utilitat de l'aplicació d'aquest procés en la gestió patrimonial.

Per a fer tots els processos de mineria de dades, quan és possible, hem utilitzat la figura d'Emili Godes com a punt de referència per a comprovar la fiabilitat dels processos aplicats o per a validar hipòtesis, però les anàlisis s'han fet utilitzant tots els registres del fons Reproducciones Sabaté considerats vàlids per a cada estudi, excepció feta de l'**apartat 8.5**, dedicat en exclusiva a Emili Godes.

8.1. Descripció de les dades

8.1.1. El sistema gestor de base de dades de la Filmoteca de Catalunya

El sistema gestor de bases de dades utilitzat per la Filmoteca de Catalunya és *Mobydoc Filmo*, tal com hem comentat anteriorment. Integra, en una estructura jeràrquica, 3 taules principals per al cas que ens ocupa: **Film**, **Fotografia** i **Exemplar fotografia**. Per a la catalogació fílmica també existeix una taula anomenada **Exemplar film**, que no tractarem en aquest treball ja que l'hem centrat en els fons fotogràfics. La forma interna de gestionar les dades és relacional, tot i que té una marcada estructura jerarquitzada. Aquesta estructura ve bastant marcada per les característiques del material cinematogràfic i de com es va organitzar l'estructura de la base de dades. Es va crear una taula anomenada **Film** que no fa referència a cap objecte de l'arxiu, però que té les dades genèriques de la pel·lícula, ja que moltes vegades una pel·lícula està composta de diferents materials cinematogràfics, amb diferents versions i utilitats, i fotografies que tenen les seves pròpies dades. Per exemple, un registre **Film** té tots els actors participants a la pel·lícula i cada registre **Fotografia** corresponent només els actors que surten a la imatge concreta.

Com es pot llegir a Calle (2005) en un inici hi havia dues aplicacions, *Mobydoc Filmo* i *Mobydoc Photo*, que es van integrar en una de sola que va agafar el nom de la primera. En fer aquesta unificació es va provocar un problema de coherència a la base de dades que ha complicat el processat de les dades per aquest treball. Com es pot veure a l'esquema relacional de la **figura 8.1** les taules no segueixen una jerarquia de baix a dalt o de dalt a baix com faria preveure la jerarquització de les taules, ja que els registres **Exemplar fotografia** es relacionen amb els superiors en la jerarquia, però la

taula **Fotografia** no té relació directa amb la taula **Film**. És cada registre **Film** el que té un llistat dels registres **Fotografia** que en depenen i per tant aquests es converteixen en un atribut més i no en una dependència entre registres.

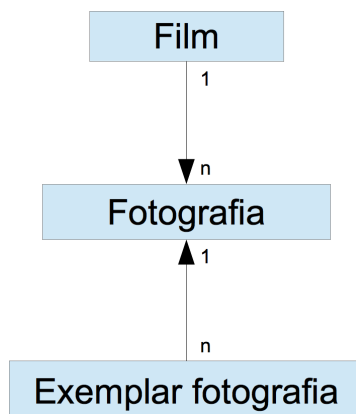


Figura 8.1. Estructura relacional de la base de dades amb la direcció de les relacions i la seva cardinalitat.

El gestor de base de dades Mobydoc, la interfície que utilitzen els usuaris, és un programa a mida i tancat, i el motor de bases de dades és una solució estàndard, ja que funciona amb un servidor SQL, però no deixa de ser una solució d'obertura parcial ja que no es permeten fer directament consultes SQL. El 2001, quan es va prendre la decisió de canviar de base de dades (Calle, 2005), encara no estava tan estesa la consciència de la necessitat que el programari de les institucions públiques tendeixi a ser lliure i obert²³. Ja que no es poden fer aquestes consultes SQL directament, ni el programa disposa de protocols d'interoperabilitat oberts com OAIS, per a poder fer l'exportació de les dades per aquest treball s'ha utilitzat l'exportació destinada al funcionament del servidor OPAC web²⁴.

La major part de l'entrada de dades és bastant controlada, amb diccionaris dels quals s'han de triar elements, tot i que la majoria d'apartats tenen un camp de precisions per a poder afegir informació textual. Els diccionaris són oberts, això vol dir que qualsevol usuari pot modificar-ne el contingut i crear nous elements. La decisió de deixar oberts els diccionaris té els seus pros i contres: per una banda deixar-los tancats provoca que els termes imprevistos moltes vegades no s'introdueixen, ja que reclamen una pausa en la feina, però deixar-los oberts obliga a una revisió contínua de les noves entrades. En el cas que ens ocupa els catalogadors tenien permís assistit d'inserció de nous camps, tot i que l'experiència ha demostrat que sovint no es demanava aquesta assistència, o no es rebia a temps, o es deixava a discreció del catalogador l'entrada de nous termes²⁵.

²³ En el cas de les bases de dades aquesta necessitat és més present, ja que les dades passen a ser controlades pel proveïdor i pot pressionar amb aquest fet a l'hora de negociar els serveis que presta.

²⁴ Correspon a les sigles *Online Public Acces Catalog* i facilita l'accés a la col·lecció de forma remota, tot i que no es fa directament des de l'aplicació Mobydoc, sinó que s'han de fer importacions periòdiques al servidor web que ho gestiona.

²⁵ Com es pot llegir a les memòries de l'Institut Català de les Empreses Culturals (<http://goo.gl/1ZinYu>) en algun any, com a 2003, es va arribar a tenir 97 estudiants en

El programa permetia duplicar un nombre de camps determinats de la fitxa anterior quan se'n creava una de nova. Això té, altra vegada, beneficis i perjudicis associats, ja que permet un estalvi de temps, però provoca errors si els objectes no són exactament iguals. Aquest estalvi de temps era considerable en un inici ja que la base de dades era extraordinàriament lenta, probablement per la complexitat de l'esquema de dades i pel fet que encara no tenia al darrera un motor de dades SQL potent. Només hi havia un camp que mai es va deixar duplicar i aquest era el de la ubicació del material, ja que un error en aquest camp fa que l'objecte resti perdut a l'arxiu.

Quan les dades s'entren en un camp de text o es creen de nou en un diccionari no hi ha cap mena de control, així que es poden introduir dades manifestament incorrectes en els camps, com ara dates de producció anteriors a 1895 (data de presentació pública del cinema), tot i que podria ser molt senzill evitar un error d'aquest tipus.

8.2. Descripció i justificació de les eines utilitzades per al Descobriment de Coneixement en Bases de dades

Passem a descriure les eines utilitzades per tal de dur a terme la recerca en el cas d'estudi que ens ocupa, així com a justificar la seva selecció i ús.

8.2.1. Programari per a exportar i transformar les dades

Les dades s'han tractat amb un programa a mida realitzat amb PHP i amb el suport de fitxers XML per a controlar la sortida i les opcions. Per a codificar s'ha utilitzat l'entorn de programació Eclipse²⁶ i per a fer anar el programa s'ha realitzat una senzilla interfície web, ja que PHP funcionava en un servidor web Apache.

8.2.2. Programari de mineria de dades

El programari que s'ha utilitzat per a la mineria de dades i la visualització ha estat Orange i és amb el que s'han realitzat totes les figures amb visualitzacions dels processos estadístics i de mineria de dades que es pot veure en el treball. S'ha triat aquest programa perquè és de codi obert i per la seva facilitat d'ús, ja que permet fer la connexió entre els diferents passos de cada procés de forma gràfica. Té molta varietat d'eines que es poden utilitzar i es configura molt fàcilment.

Tots els programes avaluats, tot i que es presenten com a conjunts d'eines de mineria de dades, contenen un alt percentatge d'eines dedicades a l'aprenentatge automatitzat, així que en realitat són programes d'intel·ligència artificial, que és la disciplina que les engloba. Aquesta barreja és fàcil, ja que són dues disciplines molt properes, si bé conceptualment molt diferents. L'aprenentatge automatitzat serveix per predir dades futures o per assignar valors desconeguts i la mineria de dades serveix per a descobrir-ne el coneixement ocult.

Si s'hagués d'aplicar el mateix procés en un entorn real s'hauria d'emprar alguna altra opció com R²⁷ que es pot integrar en altres programes de forma que els resultats de la

pràctiques catalogant fotografies amb una única persona de suport a temps complet.

²⁶ <http://www.eclipse.org>

²⁷ R (<http://www.r-project.org/>) és, de fet, un llenguatge de programació i una plataforma de software. Es pot integrar en altres aplicacions com un mòdul.

minería de dades es reflecteixin en els dades de l'equipament i el qual és bastant més potent que Orange.

8.2.2. Tria de les eines utilitzades

El programa Orange té molts mòduls que es poden utilitzar però per aquest treball en concret ens hem limitat a aquells que es dediquen al descobriment de coneixement, a excepció del nomograma que permet experimentar i predir dades. D'aquesta manera hem utilitzat algunes eines molt més properes a l'estadística, que ja és una disciplina que es creua amb la minería de dades, com els diagrames de dispersió o les gràfiques de distribució. Les eines utilitzades, que explicarem en els apartats corresponents, són el nomograma, els arbres de classificació i l'anàlisi de xarxes.

8.3. Preparació i control de les dades per a la minería de dades

8.3.1. Transformació de les dades

L'exportació de dades de Mobydoc crea tres conjunts de fitxers que corresponen a les fitxes **Film**, **Fotografia** i **Exemplar fotografia**. Els fitxers XML²⁸ que crea quan exporta són en un format propietari del fabricant, així que no poden ser utilitzats per programes de minería de dades directament. El format XML és manipulable de forma automatitzada i es pot convertir a diferents formats de manera relativament simple.

Donades les peculiaritats de l'exportació, no hi havia cap altra forma per a transformar els fitxers d'origen que crear un programa a mida. El llenguatge de programació utilitzat per a fer-lo ha estat PHP, ja que és un llenguatge especialitzat en internet que té diverses funcions integrades per a manipular fitxers XML de forma senzilla. El programa no té una interfície amb la qual interactuar i es pot controlar canviant les dades d'uns fitxers XML que contenen tant les opcions d'exportació com les de transformació. Aquest programa ha estat realitzat per l'autor d'aquest treball.

El format de fitxer triat per utilitzar en la minería de dades ha estat el de text tabulat, ja que és l'opció més estàndard i que es pot utilitzar en més programes. Aquest format té, a més, l'avantatge afegit que Mobydoc no permet introduir tabuladors en els camps, així que no cal controlar-ne la possible presència quan es fa l'exportació.

Per a cada part de l'estudi s'han triat uns camps de sortida, que a vegades han combinat la informació de diferents fitxers. Un dels problemes que hi ha hagut per a aquesta combinació és que les dades podien no ser consistents com a conseqüència que l'estructura relacional de la base de dades no segueix una jerarquia unidireccional. S'ha hagut de crear una taula amb les relacions entre els fitxers per a poder relacionar els fitxers **Film** amb la resta i se n'ha hagut de comprovar la consistència ja que aquesta no estava assegurada. No es podia confiar en les dades de relació de la fitxa **Film**, ja que el programa no obliga a que existeixi aquesta dada, al contrari que amb la fitxa **Fotografia exemplar**, que no pot existir si no té una fitxa **Fotografia** amb la qual tingui relació.

²⁸ XML són les sigles de Extensible Markup Language. És un format de fitxers de text que utilitza uns codis per a estructurar la informació de forma que sigui fàcilment llegible pels ordinadors.

Apart de convertir els fitxers XML també s'ha hagut de convertir algunes de les dades dels camps per a que fossin útils per a les diverses anàlisis, per exemple:

- Convertir formats de dates, ja que en el programa d'origen tenen el format 00/00/1966 i s'havia d'eliminar la part inicial per a que només quedés l'any.
- Introduir un valor per defecte en els camps buits.
- Introduir el valor d'un segon camp, normalment el camp Precisions, quan el camp té un valor massa genèric, que es concreta en aquest camp Precisions.
- Convertir el text de les mides entrades manualment i separar la part numèrica per a possibles comprovacions.
- Unificar valors de camps que es referien al mateix concepte, però estan escrits amb petites diferències.
- Retornar el valor inferior quan s'havia agafat d'una llista jeràrquica, com la que es pot veure representada a la **figura 8.2**, el valor jeràrquic superior de forma incorrecta²⁹.
- Convertir els valor manifestament incorrectes.

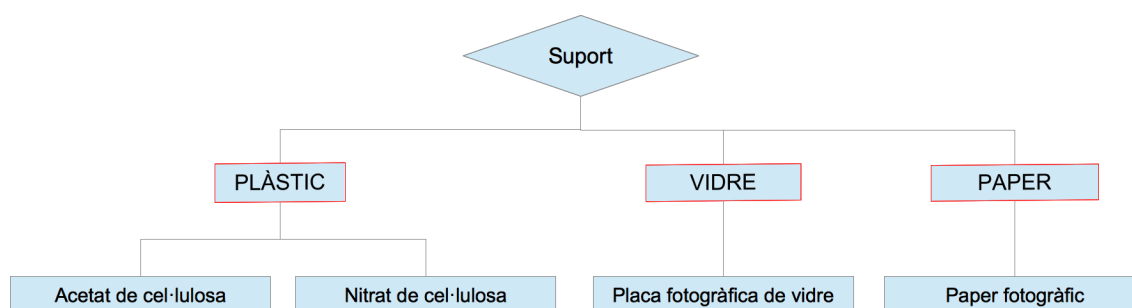


Figura 8.2. Jerarquia del camp Suport amb els termes que no s'havien d'utilitzar marcats en vermell.

Encara s'ha hagut de fer una transformació més, tot i que aquesta no ha estat producte de les dades, sinó dels programes utilitzats. Com els dos programes utilitzats són anglosaxons tenen molts problemes per a gestionar els caràcters accentuats o les lletres com la ñ. Finalment, ja que no afectava el resultat dels càlculs realitzats per la mineria de dades, s'ha optat per exportar convertint el text dels caràcters problemàtics a un equivalent que no donés problemes. Això es veu reflectit en el text d'alguna figura de les visualitzacions de resultats.

8.3.2. Anàlisi de les propietats de les dades i la seva qualitat

Per a poder fer el procés de mineria de dades s'ha de comprovar abans que les dades de la catalogació són suficients i que tenen la qualitat necessària per a fer el procés. Com no és viable, per la complexitat que suposa, comprovar tots els camps de la base de dades, és necessari triar-ne una part que compleixi aquestes condicions: (a) que la

²⁹ Això passava per la forma de mostrar els valors que s'havien de triar, ja que la llista no es veia en format jeràrquic i els termes superiors destacaven en majúscules, així que alguns introductors agafaven el terme superior. Per mantenir la coherència de nivells amb el fet que s'havia creat un nivell PLASTIC, del qual descendien Nitrat de cel·lulosa i Acetat de cel·lulosa, es van crear uns nivells VIDRE, que contenia Placa fotogràfica de vidre, i PAPER, que contenia Paper fotogràfic. Curiosament, en cap cas es va triar el terme PLASTIC, segurament perquè Nitrat i Acetat no són termes del llenguatge comú.

informació entrada sigui molt completa, preferentment que fos obligatori introduir-la; (b) que tinguem prou coneixement previ com per treure conclusions de la seva anàlisi; (c) que hi hagi una relació directa entre els camps i no depenguin d'un camp que no s'analitzi ja que podria distorsionar-ne les relacions. Ha estat de molta utilitat l'anàlisi documental per a triar els camps i les dades utilitzades.

a. Justificació de la tria dels camps emprats per valorar la qualitat de les dades

S'ha triat els camps **Mides**, **Tipus de procés** i **Suport**, que corresponen a la fitxa **Exemplar fotografia**, per avaluar la qualitat de les dades perquè, en principi, eren entre els més controlats quan s'omplien i uns dels que podien aportar relacions evidents amb les eines de visualització. Sense aplicar cap algorisme complex, la comparació d'aquests camps –**Mides**, **Tipus de procés** i **Suport**– amb un programa de mineria de dades dóna ja resultats molt interessants, més enllà de l'objectiu de verificar-ne la qualitat i integritat, els quals detallarem a l'**apartat 8.4.1**.

D'aquests tres camps, l'únic que podia portar sorpreses era el de **Mides**. Unes sorpreses o dificultats que podien venir del fet que normalment s'havia d'escollir una mida estàndard d'una llista, a no ser que la mida fos *prou diferent* d'aquest estàndard. Quan s'agafava de la llista el valor «Fora d'estàndard», s'introduïa el valor exacte en mil·límetres en el camp *Precisions* seguint un esquema preestablert. Aquest terme *prou diferent* va introduir una apreciació subjectiva a la introducció de les dades que ara hem vist que va donar massa iniciativa als catalogadors, agreujada pel fet que s'entraven les mides en un camp de text lliure. Això ha suposat una feina de filtratge i correcció prèvies de les dades manifestament incorrectes. Els errors que s'ha hagut de corregir abans de tractar les dades han estat, sobretot, de puntuació i separació entre paraules, ús incorrecte de les majúscules i caràcters estranys provocats per la base de dades original. La majoria d'errors són producte de no seguir l'esquema preestablert. Amb els canvis realitzats per eliminar aquests errors es va passar de 209 mides diferents a 144, i la majoria són prou semblants a un estàndard com per haver-lo pogut triar³⁰. Vist en perspectiva i després de l'anàlisi, potser hagués estat millor posar en el camp controlat la mida estàndard més propera. També hi havia el problema, en aquest cas de concepció de la base de dades, que les dades numèriques, com són les mides, haurien de ser dividides en dos camps numèrics controlats i no confiar que es compliran les mides estàndard d'una llista.

Què aporta aquesta primera aproximació? Veurem com les dades tendeixen a ser uniformes quan n'hi ha un gran volum, i encara més si corresponen a processos estandarditzats, com són els de creació i reproducció fotogràfica, així que la hipòtesi que els punts aïllats coincideixen amb els punts erronis o excepcionals es pot confirmar.

b. Anàlisi i visualització de la qualitat de les dades dels camps seleccionats

En les següents figures visualitzarem el resultat de comparar els valors dels camps **Mides**, **Tipus de procés** i **Suport** del total de 20.000 registres de la base de dades

³⁰ En l'apartat de mides hem deixat aquelles que es van considerar estàndard i figuraven a la llista, afegint amb cursiva les que predominen entre les entrades manualment pels catalogadors.

amb el programa de mineria de dades. La forma de representació seran diagrames de dispersió de les comparacions dels valors per parells de camps. El programa de representació permet mantenir el color dels diferents suports, amb la qual cosa encara aporta més informació en ser present sempre aquest valor. Aquesta forma de veure els valors comparats ens permet comprovar com aquestes s'agrupen i quines són les dades atípiques (*outliers*), que corresponen a les que no s'agrupen amb la resta. També ens permet veure si les agrupacions corresponen a les dades esperades pel coneixement previ que tenim. La gràfica que crea el programa utilitzat és interactiva i això ens permet saber amb un clic a quines dades es refereix cada punt i fer-ne la comprovació.

Els valors de configuració que hem triat per a representar les dades en les següents figures fa que destaquin més les dades atípiques, però es pot comprovar que la major part de dades estan agrupades si es canvien aquests ajustaments. S'ha deixat com a exemple l'opció que fa ressaltar més les dades atípiques ja que és la que es fa servir per a distingir entre els punts erronis i excepcionals.

Les parts més denses de la **figura 8.3** corresponen als materials estàndard que s'utilitzen en la fotografia professional en la cronologia del fons, amb la qual cosa segueixen la pauta prevista.

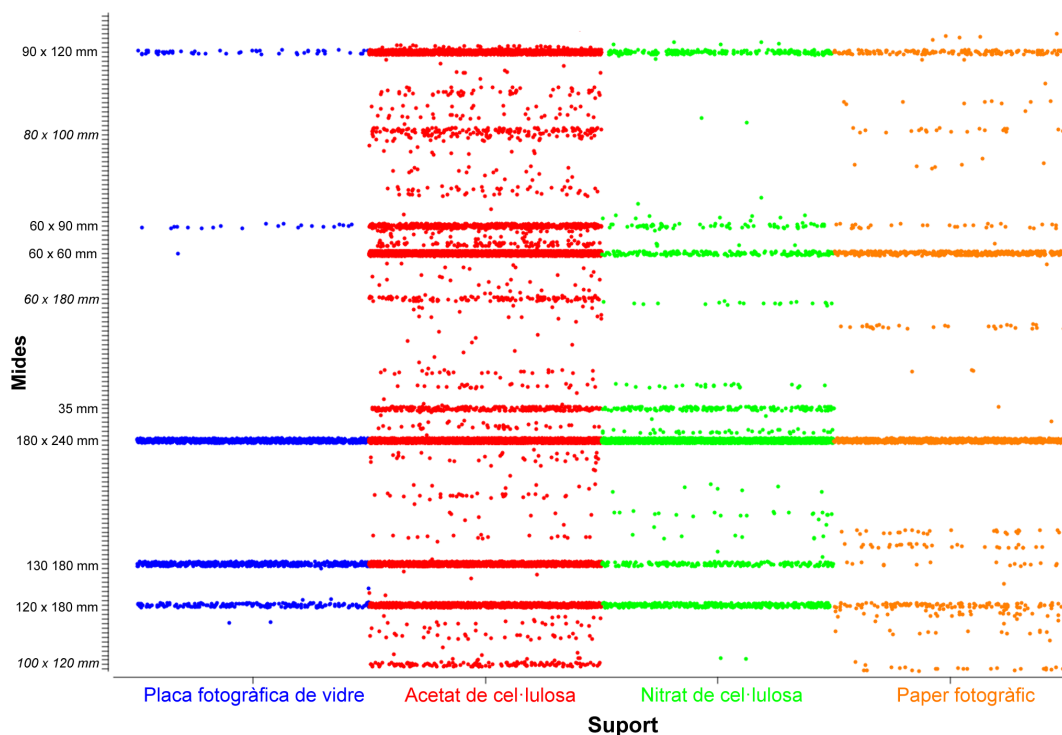


Figura 8.3. Gràfic amb la matriu de dades entre la mida i el suport del material fotogràfic

Un element de distorsió que podem veure a la **figura 8.3** és que l'acetat de cel·lulosa té una distribució diferent i molts més punts en totes les mides possibles. Això pot venir donat per dos fets: (1) que és el que ocupa una franja temporal més llarga, i (2) és utilitzat per a més usos que els altres materials. L'acetat de cel·lulosa, en ocupar una franja cronològica més llarga, pot haver sofert canvis en els processos de revelat i de duplicació, o l'aparició de nous estàndards i, a més, com es pot observar a la **figura 8.5**, és utilitzat per a més usos que els altres materials, a la qual cosa s'afegiria el fet

de ser el material més utilitzat. Tot això fa que sigui més difícil trobar els punts discordants, i per tant, per a poder trobar-los s'hauria de limitar la gràfica cronològicament o tipològicament si la finalitat de la gràfica fos corregir la base de dades.

A la **figura 8.3** sorprèn la concentració de punts en dues mides que no es consideraven estàndard: 80 x 100 mm i 100 x 120 mm. Per a comprovar aquest resultat, s'ha fet una verificació d'una mostra representativa i s'han mesurat aquests objectes per veure si realment els materials corresponen a aquestes mides³¹. La mida 60 x 180 no era a la llista, però s'explica comptant que hi ha objectes que eren formats per varies imatges de 60 x 60 mm o 60 x 90 mm que no se separaven i es catalogaven com un sol objecte.

A la **figura 8.4** podem veure com són d'evidents els elements amb les mides excepcionals o incorrectes amb el suport d'acetat de cel·lulosa i el procés positiu, ja que destaquen com a punts vermells entre els punts taronges. En aquest cas és necessària la utilitat d'ampliació del programari utilitzat per a identificar-los ja que la visió per defecte no permet fer-ho amb facilitat. El programa utilitzat permet també extreure en forma de taula aquestes dades per a poder avaluar-los amb més detall o realitzar altres processos a les mateixes.

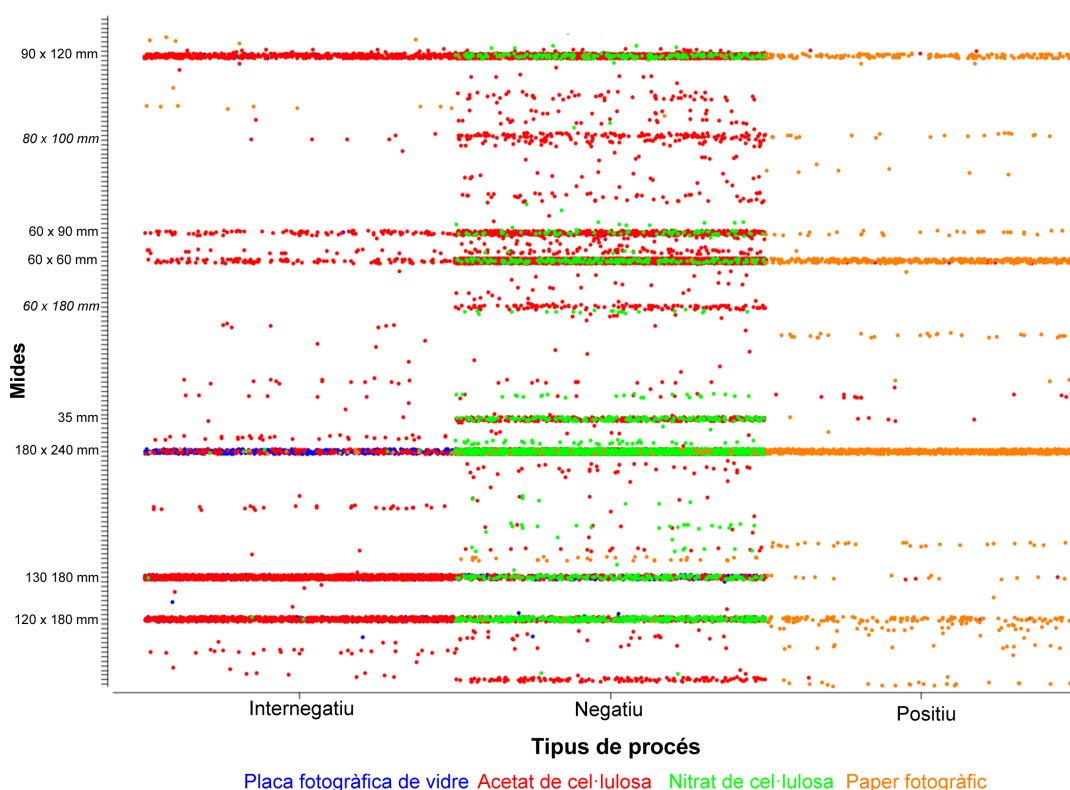


Figura 8.4. Gràfic amb la matriu de dades entre la mida i el tipus de procés del material fotogràfic, amb diferents colors depenent del suport.

Alguns punts del paper fotogràfic, però, de la gràfica de la **figura 8.5** són inconsistents amb els usos del material, com és la presència de punts que coincideixen amb

³¹ Realitzada aquest verificació s'ha pogut veure que corresponen en el primer cas a un material que per efecte d'una degradació havia encongit i en el segon a un error.

internegatius o negatius, i fa que sigui un cas apart per a considerar la validesa de les seves dades. Les còpies en paper són duplicats de negatius o internegatius existents en la col·lecció i eren tractats de forma més ràpida al final del procés de catalogació. A més, la base de dades permetia duplicar una part de les dades de la fitxa anterior, entre les quals el tipus de procés. Tots aquests elements, sumats al fet que tenen menys valor ja que són proves de baixa qualitat o part de fulls de contacte, versemblantment van produir una relaxació que va provocar més errors que a la resta de materials. Això fa evident que aquesta part de material requereix una revisió a fons per a evitar en la mesura del possible la distorsió que provoca.

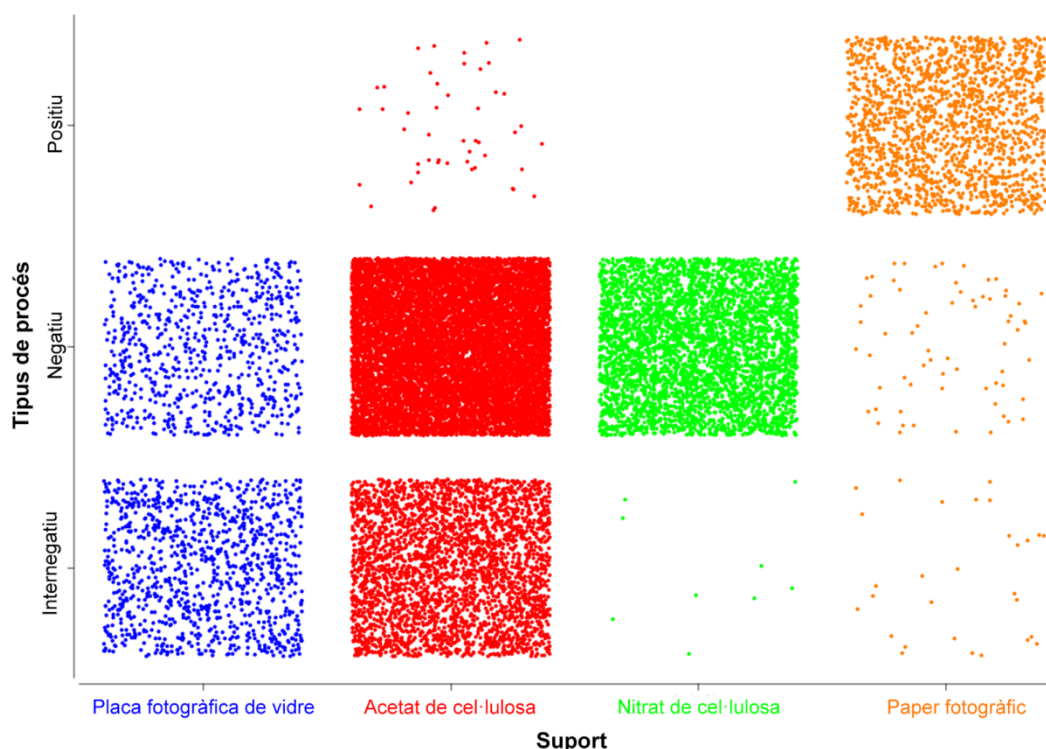


Figura 8.5. Diagrama de dispersió entre el tipus de procés i el suport del material fotogràfic

b. Conclusions sobre la qualitat de les dades

De l'anàlisi es pot concloure que és viable realitzar un procés de Descobriment del Coneixement en Bases de dades, ja que es correspon amb els valors esperats i la quantitat de registres utilitzables és suficient per aplicar la mineria de dades. Excepció feta del material en paper la majoria de les dades són correctes, tot i que s'han hagut de filtrar en alguns casos per problemes formals. Molts dels algorismes de mineria de dades fan una transformació de les dades a valors numèrics i per tant són sensibles a les diferències formals. Un dels passos que treu més temps, previ a l'anàlisi final, és la identificació d'aquests problemes formals i la seva correcció.

8.4. Coneixement adquirit mitjançant la mineria de dades

En aquest apartat s'analitzen tots els registres de la base de dades que es consideren vàlids per a cada anàlisi, comparant diferents camps en cada cas. L'**apartat 8.4.1** està més centrat en el coneixement aportat per la visualització i l'**apartat 8.4.2** en les diferents aplicacions de classificació que permet la mineria de dades.

En els casos en que és possible, a l'**apartat 8.4.2**, les hipòtesis a comprovar se centren en les fotografies de Godes, ja que és de l'autor de qui tenim més coneixement gràcies a l'estudi previ.

8.4.1. Amenaces, febleses i fortaleeses del fons fotogràfic

a. Tipus de materials conservats

Per a poder valorar els tipus de materials conservats utilitzarem les dades dels camps **Mides**, **Tipus de procés** i **Suport**, i les seves representacions gràfiques, que hem utilitzat a l'**apartat 8.3.2**.

Destaca la quantitat de materials diferents que formen el fons, com es pot veure a la **figura 8.3**, tant per la diversitat de mides estàndard com per la presència de tots els suports possibles per l'època. Això suposa un esforç extra en les condicions d'emmagatzematge, tant a nivell de condicions climàtiques com per les necessitats de conservació o la diversitat de contenidors necessaris.

També aquesta diversitat fa destacar el valor de la col·lecció, tant des del punt de vista d'aportar coneixement gràcies a la seva riquesa tipològica, com des del punt de vista de la seva composició. És un fons majoritàriament format per negatius i per tant serveix molt bé per explicar la forma de treballar dels fotògrafs i els laboratoris de fotografia industrial de 1932 a 1984.

La gràfica de la **figura 8.5**, que compara el tipus de procés amb el suport, és una de les que dóna un coneixement més evident de la forma de treballar dels fotògrafs i els laboratoris. Com ja hem dit anteriorment, les dades tendeixen a ser uniformes i aquesta uniformització ens aporta coneixement sobre la realitat que reflecteixen. El primer que s'observa en aquest gràfic és el gran valor que tenia el nitrat per als fotògrafs, ja que dóna una qualitat molt superior d'imatge que els altres suports. Aquest valor es veu en el fet que és utilitzat gairebé exclusivament per a realitzar negatius originals i no per a tasques de reproducció com internegatius³². Aquesta és una de les informacions que es fa molt evident en una gràfica i que no és fàcil de veure en una llista de dades. També es pot observar a la gràfica els diferents usos que pot tenir l'acetat, ja que és present a tots els tipus de procés, i això correspon a les tasques del laboratori purament fotogràfic com les que fan referència a procediments fotomecànics, que correspondrien a feines d'impremta.

Si aquesta anàlisi s'hagués d'aplicar a la gestió de l'equipament cultural s'hauria d'ampliar amb eines d'aprenentatge automatitzat. Això serviria per a poder controlar de manera informada la compra de contenidors, ja que el sistema podria estimar quins serien més necessaris en un futur.

b. Estat de conservació de la col·lecció

Un punt molt important per a la gestió d'un fons patrimonial és conèixer l'estat de conservació dels objectes custodiats i en quines parts del fons es necessita actuar amb més urgència. Per poder fer això és molt útil tenir una estimació de quines parts

³² Són còpies d'altres materials positius destinades, sobretot, a processos fotomecànics.

tenen més possibilitats de necessitar una actuació en forma de preservació o restauració.

En la introducció de dades a Mobydoc es podien triar d'una llista 6 nivells d'estat de conservació: **Bo**, **Primer grau**, **Segon grau**, **Tercer grau**, **Inservible** i **Baixa**. Aquest nivells estan ordenats de millor a pitjor estat de conservació i tenien algunes particularitats: (a) no es recomanava l'ús del terme **Bo** ja que està reservat a reproduccions noves (b) el terme **Baixa** és un terme «històric», indica que el material ha estat destruït o enviat a un gestor de residus, i no afecta a la gestió de l'equipament. Així, de forma efectiva, la tria es feia en 3 nivells de conservació, ja que **Inservible** era un nivell d'últim recurs.

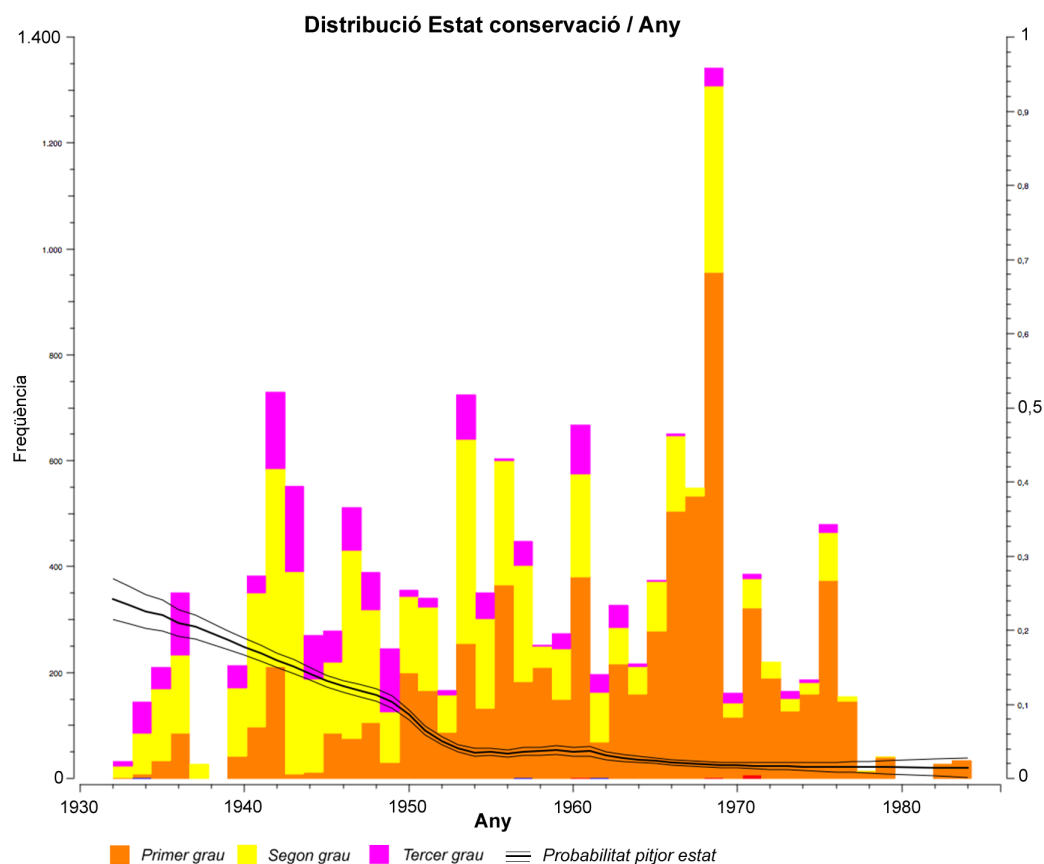


Figura 8.6. Distribució dels estats de conservació amb la probabilitat que estiguin en el pitjor estat possible / any.

Aquesta reducció a 3 nivells efectius va fer que aquests quedessin molt justos per a descriure l'estat de conservació i va fer que no quedés clara la frontera entre els nivells. Apart d'això, en fer la mineria de dades, ha donat problemes una decisió de catalogació: es va decidir canviar la forma de realitzar la valoració de l'estat de conservació uns anys després de començada la catalogació dividint-la en **Estat químic** i **Estat físic**. Quan es va fer aquest canvi no es va tornar a revisar el material per introduir aquest valor, ni es va posar un valor per defecte al valor que quedava buit, que va ser **Estat químic**. El fet que quedin valors buits distorsiona les gràfiques de distribució i per això en aquest estudi hem decidit descartar els valors de l'Estat químic, que són la minoria. El motiu de dividir els dos tipus d'estats va ser afavorir la tasca dels catalogadors i donar dades més precises sobre l'estat de conservació. Normalment s'ha de tenir més cura amb l'estat químic ja que aquest és reflex que el

material té algun tipus de degradació activa i pot estar amenaçat de destrucció. Dels problemes físics normalment no es dedueix que un material pugui estar en perill de canviar de grau a curt o mitjà termini, ja que la majoria dels defectes físics són ratllades, parts trencades i parts absents. Quan els defectes físics són reflex d'una degradació activa, suport balcat o arrugat, van acompanyats de defectes químics i normalment puja de la mateixa manera el nivell de degradació química.

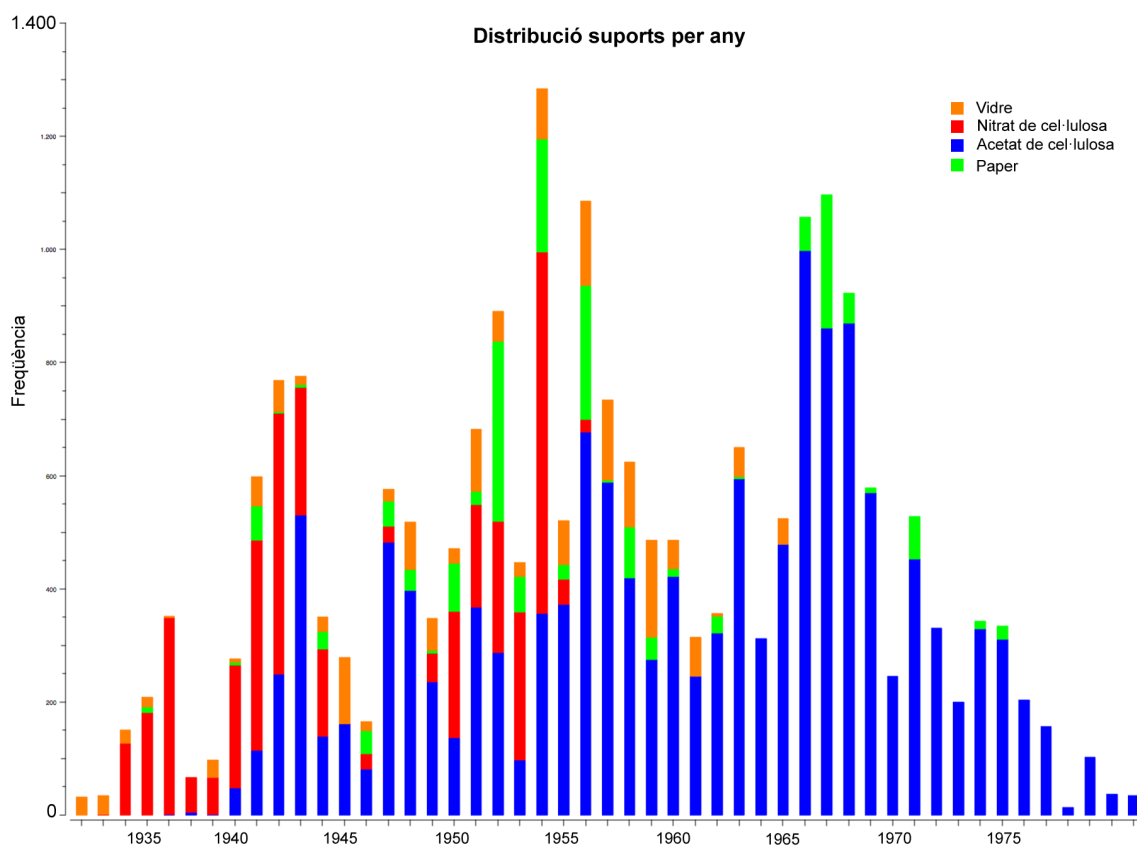
A la **figura 8.6** es pot veure la distribució d'estats de conservació, limitats en realitat a l'estat de conservació físic pels problemes esmentats anteriorment, ja que és el que té més dades representades. Es pot veure com a mida que el material és més modern es fa més preminent la presència d'objectes en **Primer grau** de conservació.

La línia més gruixuda indica una estimació de la probabilitat que un material estigui en el pitjor estat possible; en aquest cas s'ha triat el nivell **Tercer grau**. Les línies fines corresponen a la confidencialitat d'aquesta estimació, això és, el marge d'error que té l'estimació. Es comprova el que la lògica indica, això és, que com més antic és un material més possibilitats té d'estar malmès. Aquesta gràfica també reflecteix un problema que tenen una part dels materials de l'any 1958 a 1962 que s'han degradat més que la resta. Aquest objectes tenen les mateixes característiques físiques (osques, mides, color i la presència del número de clixé fotografiat), la qual cosa fa pensar que són una remesa, o un model concret, d'un fabricant de suport fotogràfic que tenia un defecte, o que va ser mal processat en el revelat.

b. Anàlisi cronològica del fons

Per conèixer el valor dels fons patrimonials un dels punts importants és saber si ocupa tota la franja cronològica de la qual s'ocupa l'equipament i en quin volum. Saber quin és l'estat de la qüestió cronològic en els fons ens servirà per a portar una política d'adquisicions que pal·liï els punts febles de la nostra col·lecció o en reforci els punts forts. Reforçar punts febles o forts participen de polítiques diferents i normalment s'ha de triar per un dels dos quan les possibilitats pressupostàries no són il·limitades. Si el que es vol és que l'equipament sigui especialitzat en un període mot concret reforçarem aquest període i si volem que en el nostre equipament es trobi un ampli espectre de materials reforçarem els punts febles.

A la gràfica de la **figura 8.7** veiem la distribució de materials per anys del fons estudiat, en ella podem veure l'evolució tant dels suports utilitzats com de la seva quantitat. És útil per veure com evoluciona l'ús dels materials durant els anys i per valorar l'activitat del laboratori. El laboratori Reproducciones Sabaté va funcionar després de l'any 1984, de fet encara és actiu. Que l'últim material sigui de l'any 1984 semblaria que obeeix a un canvi en la forma de tractar el material, que ja no devia ser dipositat pels clients a les instal·lacions del laboratori i no tant a una baixada tant gran de la producció. És en curs un treball d'investigació a càrrec de Núria F. Rius sobre la història de l'empresa Reproducciones Sabaté que aclarirà alguns d'aquests punts.



8.7. Distribució de nombre de material per anys del fons.

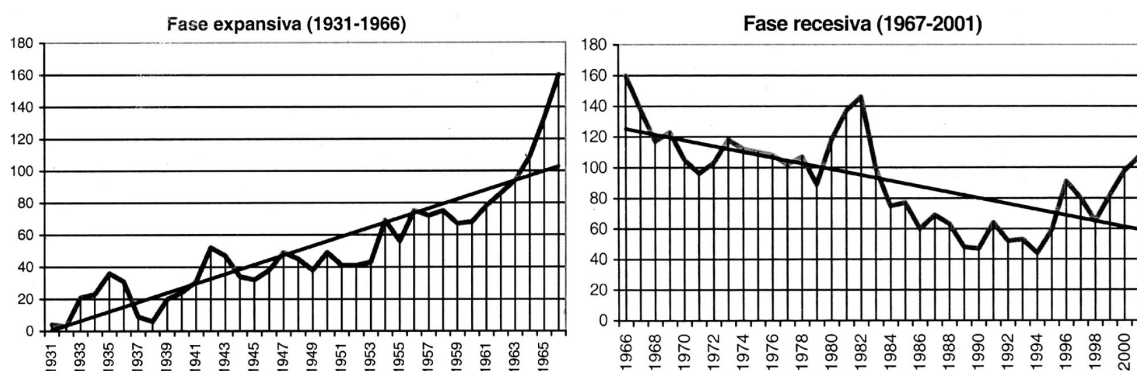


Figura 8.8. Gràfiques amb la producció cinematogràfica a Espanya de 1932 a 2000. Font: Diez Puertas (2003).

És un fons que està format majoritàriament de material que correspon a negatius i internegatius en tota la seva cronologia. Com és un fons tancat, ja que es va comprar tot l'inventari del laboratori i té unes característiques molt especials no es poden prendre decisions per ampliar-lo. El que sí que es pot dir és que donada la magnitud del fons i la presència majoritària de material original, molt difícil d'aconseguir en quantitat, és probable que no es pogués augmentar la fortalesa en aquest aspecte de l'equipament.

La gràfica de la **figura 8.7**, però, tal com està queda òrfena per a ajudar a prendre decisions o consideracions sobre el fons, sobretot per conèixer-ne les febleses o fortaleses, ja que no té relació amb la producció de films que corresponen a cada any. A la **figura 8.8** podem veure dues gràfiques amb els títols produïts per any a Espanya. Per exemple, la dada que pot ser més evident de comparar amb la **figura 8.7** és la de

l'any 1932, ja que en realitat els films que conserva el fons corresponen a un tant per cent més elevat del que podria semblar. La transició del mut al sonor va ser molt traumàtica a Espanya i gairebé no es van produir pel·lícules, amb la qual cosa seria molt difícil incrementar molt més la quantitat de materials d'aquests anys.

Una altra de les dades que es pot veure a la gràfica de la **figura 8.7** és la pervivència de material amb suport en nitrat de cel·lulosa fins a l'any 1956, quan aquest material teòricament deixa de fabricar-se i distribuir-se a 1952. Aquest és un dels casos en què la informació adquirida amb el descobriment de coneixement pot aportar un avís als historiadors, ja que s'hauria de buscar la raó d'aquesta presència, que a més, coincideix amb que durant l'any 1948 deixa d'utilitzar-se i el seu ús es reprèn amb força l'any 1950.

8.4.2. Connexions entre les dades

Per a poder posar un exemple de connexió entre les dades hem utilitzat els camps **Autor**, **Lloc de creació** i **Data de producció**. Hem triat aquest grup de camps perquè podien portar a conclusions sobre la forma de treballar dels fotògrafs i perquè hi ha relacions clares entre els valors dels camps. Amb això poden ser útils per a treure conclusions sobre el coneixement de la relació entre les dades, però també poden aportar exemples de com poden ajudar en l'entrada de dades.

Mostrarem diferents formes de visualitzar aquestes connexions, una d'elles interactiva, i que corresponen a diferents formes de mineria de dades. Les dues primeres parteixen del concepte de la classificació, així que ens mostren els resultats de les connexions, i la tercera aplica l'anàlisi de xarxes, així que és la que ens mostrarà més clarament les connexions entre les dades.

a. Nomograma

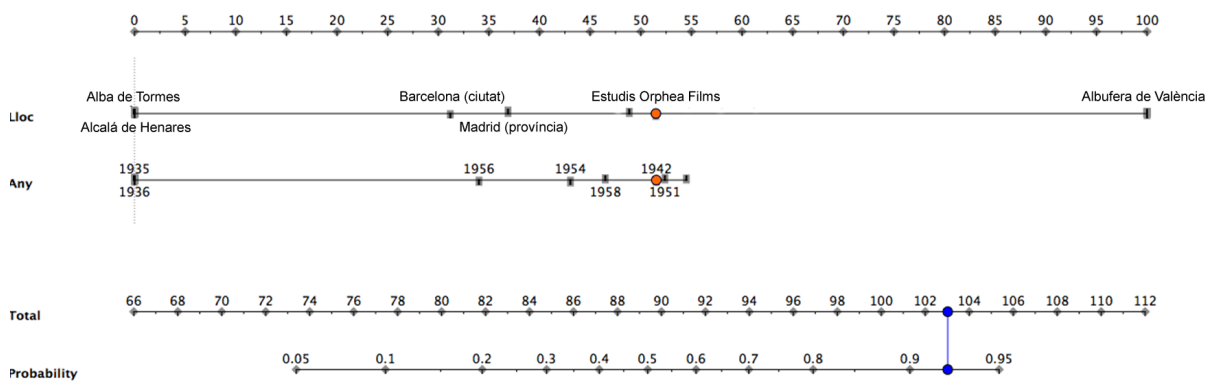


Figura 8.9. Nomograma amb la millor de les possibilitats que un film sigui d'Emili Godes segons l'any i el lloc on està rodat. En color taronja els punts que es poden moure.

La primera forma utilitzada per demostrar de forma visual la força de les relacions entre el lloc, l'any i l'autor és el nomograma. El nomograma és una eina interactiva que ens permet moure uns punts de forma semblant al funcionament d'un àbac. En aquest cas els punts mòbils corresponen al lloc i l'any, i com a resultat del canvi de posició mouen un altre parell de punts que ens donen la probabilitat que si es donen les dues condicions la combinació de partida sigui certa a partir de les connexions conegudes. En aquest cas que la fotografia sigui de Godes és la hipòtesi proposada.

Abans de poder utilitzar el nomograma s'ha hagut d'aplicar a les dades una classificació bayesiana ingènua (*Naive Bayes Classifier*) per a poder agrupar les dades i per a poder saber de forma numèrica quina influència té cada valor en el total en els exemples coneguts. A la representació visual la predicció de la probabilitat està entre els valors 0.05, improbable, i 0,95, molt probable.

Ja que el programa és interactiu presentem dues figures amb diferents estats per a poder copsar-ne el funcionament de forma més clara. La **figura 8.9** mostra la posició més favorable a que un film sigui d'Emili Godes que permet el nomograma generat pel programa³³. La relació de Godes amb aquest estudi ja era una de les hipòtesis que es confirmaven amb l'estudi de cas, però queda encara més reforçada amb els valors que dona el nomograma. De fet, la posició en la línia superior de percentatge del lloc dels estudis Orpheia Films ja ens deixa veure la influència en l'autor, doncs correspon a més d'un 50% de coincidències dels autors que hi treballen. Si es limita a períodes concrets com a la **figura 8.9** dona una probabilitat gairebé segura. A la **figura 8.10**, però, es pot veure la importància de combinar les dades del lloc amb les de l'any per a poder donar dades més fiables.

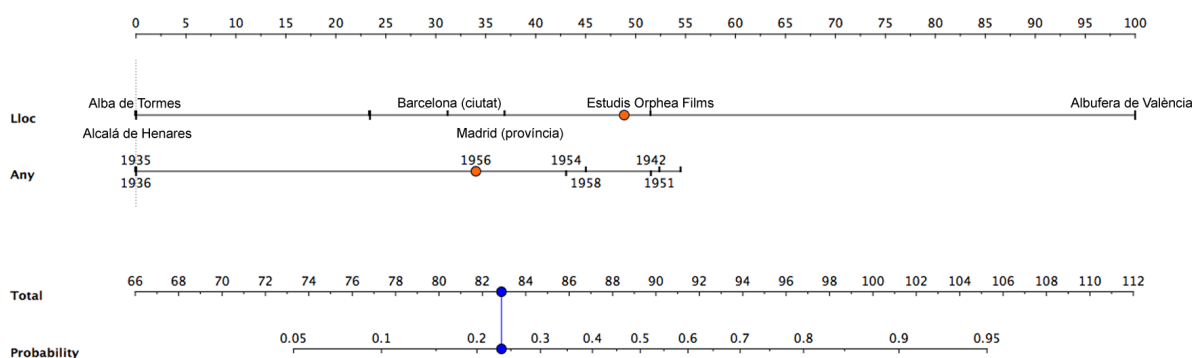


Figura 8.10. Nomograma amb menors possibilitats que un film sigui d'Emili Godes segons l'any i el lloc on està rodat.

Aquesta eina constata que es podria fer servir els camps **lloc**, **autor** i **any** per a ajudar a la introducció de noves dades: (a) per revisar si les fitxes amb autor desconegut realitzades a un lloc i període concrets són de l'autor més probable i (b) per afavorir l'entrada de noves fitxes si el gestor de bases de dades integrés un sistema d'aprenentatge automatitzat que presentés els autors més probables en primer lloc a l'introducció de les dades noves o avisés d'alguna forma si la dada que ha introduït és altament improbable.

³³ De fet, la millor posició es trobaria utilitzant com a lloc de rodatge l'Albufera de València, però aquest no és un element que es pugui comparar amb altres autors ja que només Emili Godes fa fotografies a aquest lloc, i per tant no hi ha correlacions possibles amb altres autors, i, a més, el nombre de fotografies que hi realitza és elevat. Aquesta eina és utilitzada per a diagnòsi mèdica comparant diversos símptomes (Demšar, 2010) i està pensada per a comparar més dades i amb relacions no tan marcades com l'esmentada. Cal dir que la finalitat d'aquestes eines de mineria de dades és mostrar la tendència a partir dels exemples coneguts de quines són les probabilitats i, per tant, si en tots els exemples de l'Albufera de València l'autor ha estat Godes és fàcil que es demostrés aquest fet des del punt de vista teòric.

b. Arbre de classificació

A la **figura 8.11** podem veure una part d'un arbre de classificació amb la distribució dels autors que treballen en un estudi cinematogràfic. S'ha limitat als estudis i no a totes les localitzacions per a fer més llegible el resultat i perquè les conclusions podien ser útils de cara a conèixer el funcionament d'aquests equipaments.

Apart de les relacions en aquesta visualització podem veure quin és l'autor més probable per un estudi i inclou un gràfic amb els percentatges d'altres autors. Si es col·loca el punter a sobre d'un dels gràfics de percentatges apareix un quadre amb el nom de la resta de fotògrafs i les probabilitats respectives. Amb aquesta eina veiem molt gràficament la distribució dels autors en la seva dependència amb l'estudi cinematogràfic. És una forma de tenir al dia la classificació del nostre fons en cadascuna de les seves àrees de coneixement amb l'afegit que ens pot ajudar a predir noves dades en mostrar la més probable. Un dels problemes d'aquesta visualització és que s'ha de limitar a alguna part del coneixement, ja que si no les dimensions de l'arbre fan molt difícil la seva lectura.

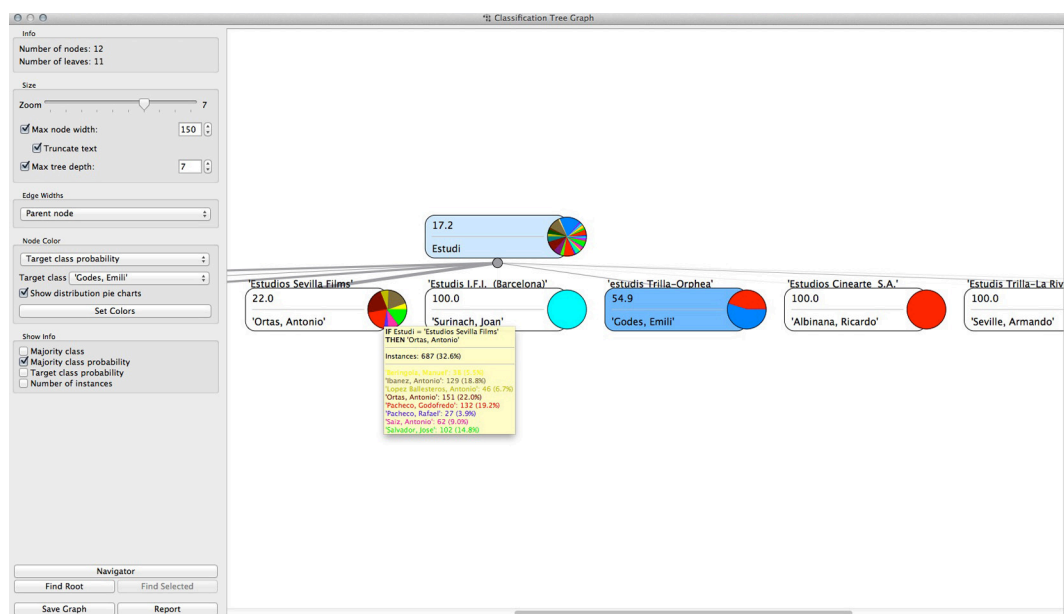


Figura 8.11. Captura de pantalla on es pot veure l'arbre de classificació amb els autors i els estudis.

En aquesta visualització hem utilitzat una opció que serveix per a estudiar un cas concret com és que marqui la possibilitat que una part de l'arbre sigui d'un autor. Hem triat Emili Godes com a la classe que s'ha de demostrar; així, a la part superior ens marca amb blau, color assignat a Godes, que si una fotografia és d'estudi hi ha el 17,2 de possibilitats que sigui d'ell. En el nivell inferior de l'arbre podem veure en blanc els casos que no corresponen a aquest autor i marcats en blau els sí ho fan.

L'anàlisi d'aquestes dades no permet concloure que en tots els casos hi hagi una relació tan forta com amb Godes i l'estudi cinematogràfic en el qual treballava més sovint. La conclusió és que és molt probable que en algun cas sigui així, però que no és una norma que es pugui aplicar a tots els casos. El cas més apartat d'aquesta possible norma serien els estudis Sevilla Films, que tenen 8 foto fixes diferents, com mostra el quadre d'informació ressaltat en groc a la **figura 8.11**.

Una altra de les conclusions que es poden extreure del gràfic és que Godes, a diferència de la majoria dels altres autors, és un autor de qui ja coneixíem més informació, tot i que les dades del fons són d'abans de realitzar l'estudi de cas. Això ho podem veure en el fet que les seves fotografies corresponen al 7% del global del fons, però en concret en els rodats en estudi hi ha una probabilitat del 17,2% que un film sigui seu. També pot ser que sigui un autor que treballi més en estudi que la resta d'autors, però per poder arribar a aquesta conclusió s'hauria de conèixer la localització de totes les fotografies, cosa que no passa.

c. Anàlisi de xarxes

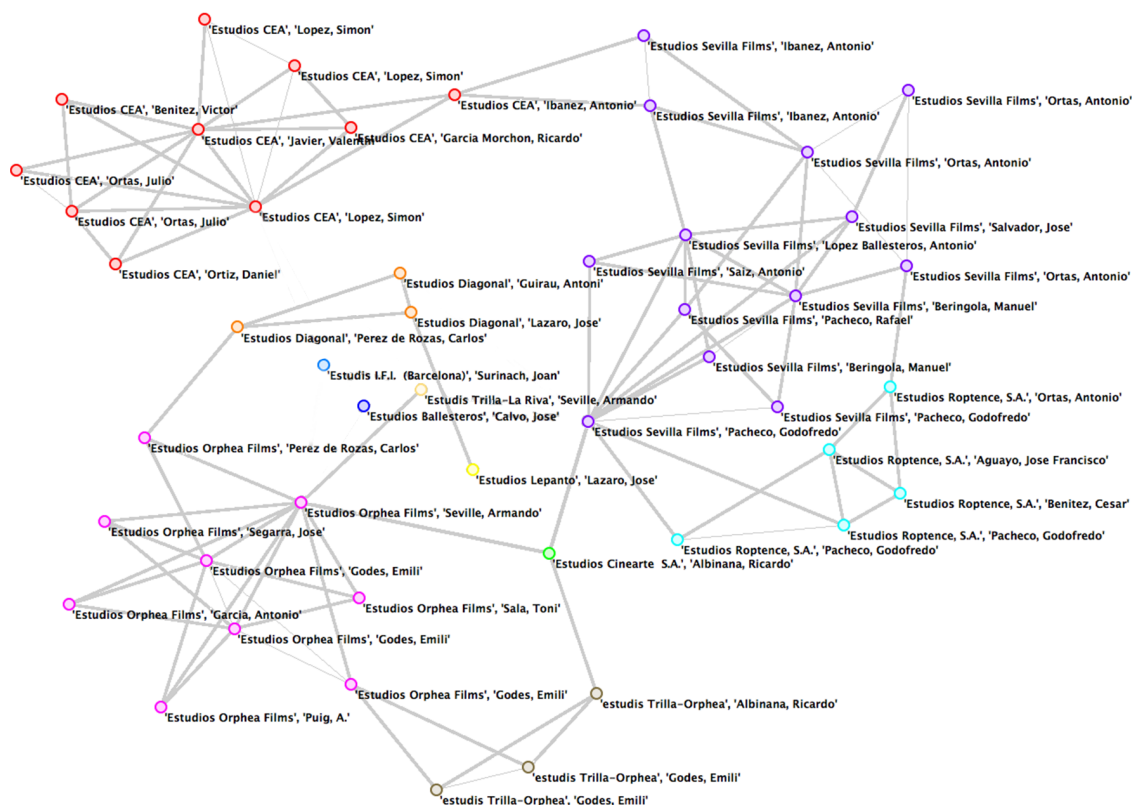


Figura 8.12. Xarxa amb les connexions entre els estudis cinematogràfics i els foto fixa.

A la xarxa de la **figura 8.12** podem veure una sèrie de punts que corresponen cadascun a un film del qual es coneix l'estudi cinematogràfic on va ser rodat i el foto fixa. S'ha marcat amb el mateix color cada estudi per a veure'n millor les agrupacions. Les línies mostren les connexions entre aquests punts i conformen la xarxa de relacions entre els fotògrafs i els estudis.

El primer que sorprèn en veure la xarxa és que només hi ha dos punts aïllats, els estudis I.F.I. i els Ballesteros, això vol dir que els foto fixes van passant d'un estudi a un altre i que hi ha una relació entre ells. Fins i tot alguns estudis que tenen un sol film utilitzen fotògrafs que han treballat a d'altres estudis. Això pot reflectir que la feina de foto fixa era molt especialitzada i controlada. De fet ja hem esmentat a l'**apartat 7** que existia un carnet amb la classificació de foto fixa del *Sindicato Nacional del Espectáculo*, sense el qual no es podia exercir aquesta professió. També podem veure en aquesta xarxa de forma molt clara com Godes no s'allunya de l'òrbita d'Orpheo.

Sempre hem de comptar, però, amb que les dades reflectides en aquestes gràfiques

són una mostra del cinema espanyol de l'època reduïda al material conservat al fons Reproducciones Sabaté i del qual s'ha pogut descobrir el lloc de rodatge. Això vol dir que la xarxa pot estar mostrant en realitat una relació entre els fotògrafs i el laboratori. Seria interessant poder comprovar com és la relació en el total de films produïts, i no només els conservats a la Fimoteca de Catalunya, i poder comparar-ho.

8.5. Conclusions de l'aplicació del Descobriment de Coneixement en Bases de dades al cas d'Emili Godes

El descobriment de coneixement global del fons realitzat a l'apartat anterior ha permès comparar Emili Godes amb la resta de foto fixes identificats. S'ha comprovat l'excepcionalitat que té respecte la resta d'autors en la seva relació amb un estudi cinematogràfic concret. Aquesta excepcionalitat no era coneguda prèviament a la realització d'aquest treball i no s'hagués pogut arribar a aquesta conclusió sense l'estudi del fons complet. En aquest apartat fem una recopilació de les dades observades sobre el material emprat per Godes.

Per a poder ampliar aquest coneixement comparat hem aplicat l'estudi del material utilitzat per Godes per a poder contrastar-lo amb les visualitzacions de les **figures 8.3 a 8.5** que corresponen a tota la col·lecció.

8.5.1. Material utilitzat per Emili Godes

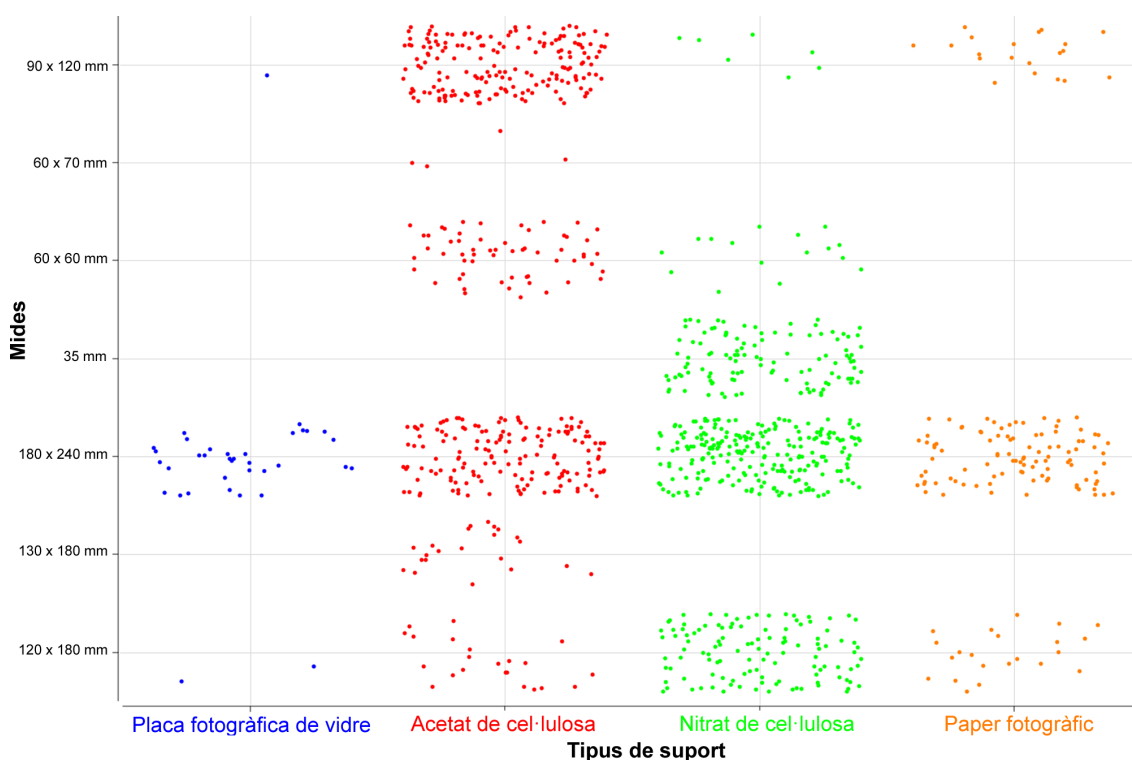


Figura 8.13. Relació entre mides i el suport de les fotografies d'Emili Godes.

Les gràfiques de dispersió del material utilitzat per Emili Godes que es poden veure a les **figures 8.13 a 8.14** ens donen una visió del tipus de fotògraf que és i de la seva relació de proximitat amb el laboratori. El reflex del fet de treballar a la mateixa ciutat on està situat el laboratori es pot veure en varis punts: (a) la majoria del seu material és negatiu i hi ha molt poca presència d'internegatius, la qual cosa indica que

dipositava directament el material al laboratori; (b) la presència també de material positiu en paper reforça la presència del fotògraf o el productor pròxims, ja que són proves destinades a ells. Si el fotògraf o la productora no tinguessin una relació directa amb el laboratori hi hauria una majoria de material internegatiu, ja que el material original estaria dipositat en un altre laboratori.

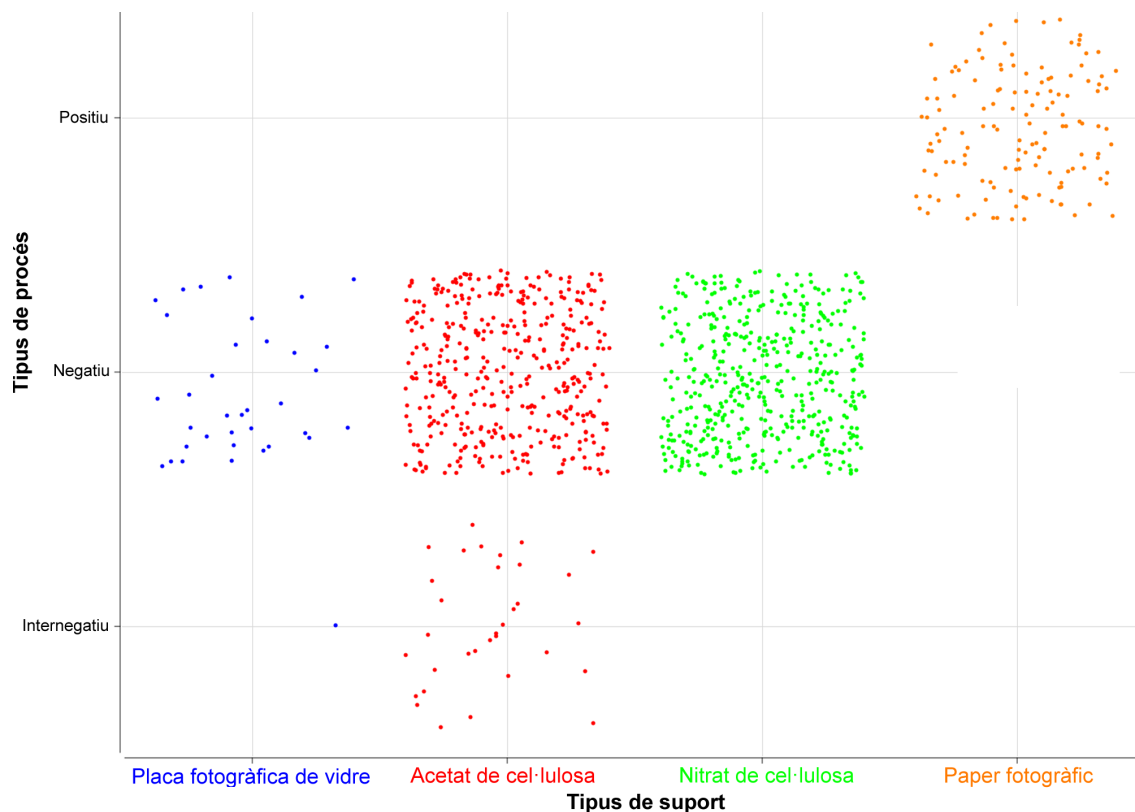


Figura 8.14. Relació entre el procés i el suport de les fotografies d'Emili Godes.

La diversitat de suports i mides utilitzades i la seva distribució són molt semblants a les de la tendència global, així que en aquest sentit Godes no difereix excessivament de la resta d'autors. Aquesta dada ja era l'esperada, ja que no hi ha moltes opcions tractant-se d'un professional que treballa de forma continuada i amb una àmplia franja cronològica. Les úniques característiques que no coincideixen amb la majoria de fotògrafs és que utilitza menys el material en rotlle, els formats 35 mm i 6 x 6. Aquesta seria, doncs, la seva marca en l'ADN fotogràfic que el diferencia del global dels autors.

8.5.2. Connexions d'Emili Godes

A l'anàlisi de xarxes de la **figura 8.12**, de la qual en podem veure una ampliació a la **figura 8.15** amb els nodes corresponents a Emili Godes, hem pogut confirmar que és un fotògraf que no s'allunya gaire de l'esfera dels estudis Orphea, a diferència d'altres autors que sí que estan repartits per nodes més allunyats de la xarxa. Per tant, confirmaria no només aquesta relació, que ja era coneguda gràcies a l'estudi de cas, sinó que fa patent la seva excepcionalitat.

També pot ser en gran part un reflex de la forma de funcionar dels estudis cinematogràfics Orphea Films, que en la seva publicitat s'anunciaven com a uns estudis que donaven serveis integrals i que fins i tot incorporaven un laboratori de revelat fotogràfic en els seves instal·lacions. Així, oferir un foto fixa de les qualitats

innegables d'Emili Godes podia ser un dels serveis afegits dels estudis. Seria necessari ampliar, per tant, la recerca sobre els estudis Orphea Films incidint en aquest aspecte concret.

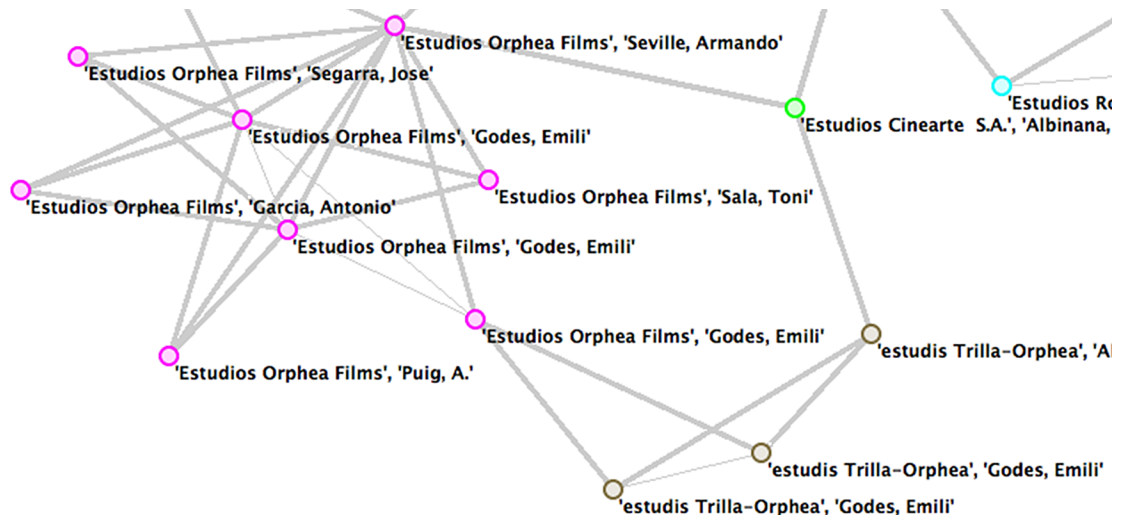


Figura 8.15. Ampliació de la part de la xarxa corresponent a Emili Godes.

9. Conclusions

9.1. Febleses de la catalogació tradicional

Tal i com ens plantejàvem en les preguntes de recerca, amb aquest treball hem pogut constatar algunes febleses del mètode tradicional de catalogació, que són resultat d'una manca de connexió entre els registres que no permet veure en perspectiva la seva organització. Aquesta manca de perspectiva permet duplicitats o valors inconsistents amb la resta de registres.

Els objectes d'un fons patrimonial, encara que siguin de tipus diversos com els de la Filmoteca de Catalunya, segueixen una tendència a unir-se en patrons determinats. Aquesta tendència a tenir patrons clars és més gran com major és la col·lecció, tot i que pugui semblar el contrari. Els objectes estan units per unes normes finites que marquen com podien manufacturar-se, com eren utilitzats i en quin context. Una de les febleses del mètode tradicional és que aquesta uniformitat no és utilitzada per a ajudar a entrar les noves dades ni per a afavorir-ne la correcció. Les eines informàtiques actuals no aprofiten el coneixement anterior i, per tant, no poden ajudar a aquells que introdueixen les noves dades.

Una altra de les febleses dels mètodes tradicionals és que tendeixen a dos extrems oposats: (a) confiar en la infal·libilitat dels introductors de les dades o (b) limitar-ne l'acció amb mètodes com llistes tancades. Aquests fets tenen els efectes respectius de provocar més errors o limitar el coneixement de la base de dades.

9.2. Aportacions del procés de Descobriment de Coneixement en Bases de dades

Responent a una de les nostres preguntes de recerca, l'aplicació del Descobriment de Coneixement en Bases de dades aporta les següents millores en la gestió d'un fons fotogràfic:

- **Qualitat.** Aporta una auditoria objectiva de les condicions i característiques dels objectes de la col·lecció i la seva catalogació que té diversos efectes:
 - La millora de les dades que aporta el procediment en ell mateix
 - Reflexió sobre la qualitat de les dades
 - Repensar els protocols de catalogació
 - Aprendre dels errors de planificació de la catalogació o de les decisions conceptuals i d'estructura de les dades
- **Presa de decisions.** Ajuda a prendre decisions informades, ja que aporta:
 - Dades objectives i comprensibles sobre les necessitats de conservació i actuacions sobre la col·lecció
- **Valor.** Augmentar el valor de la col·lecció a partir del seu coneixement:
 - Veure quin coneixement sobre el cinema i la gènesi de les fotografies ens proporciona la catalogació i aportar una eina de coneixement per a historiadors del cinema a partir de la relació entre els objectes i les dades de la seva creació.

- Veure de forma gràfica quines són les fortalezes de la col·lecció
- **Accés.** Crear noves formes d'accés a la col·lecció, més enllà de la cerca tradicional³⁴:
 - Donar ajudes a la cerca a partir de la mineria de dades sobre els textos i paraules claus que descriuen la col·lecció.
 - Donar accés al coneixement aportat per les visualitzacions generades amb la mineria de dades, anant un pas més enllà en la simple mostra dels fons de la col·lecció³⁵.

9.3. Problemàtica del Descobriment de Coneixement en Bases de dades

Un dels perills més evidents que poden aturar el descobriment de coneixement és el procés de preparació, apart de ser la part més gran de feina com hem esmentat anteriorment. En aquest moment és quan surten tots els errors i les petites trampes de la catalogació, ja que aquests es fan evidents per varis factors: (a) les dades es veuen connectades entre elles i en participar d'un mateix coneixement es fan més evidents els errors i (b) és necessari fer processos automatitzats que no funcionen si tots els valors no estan dins d'uns paràmetres. És en fer-se tan evidents els errors que poden desmoralitzar o evitar que es realitzi el procés, ja que suposarà aturar la catalogació per dedicar-se a la seva correcció.

Vistes les condicions de catalogació del fons estudiat hem de dir que la qualitat de les seves dades no és menyspreable i que en les revisions aleatòries s'ha pogut veure que els errors estan concentrats, sobretot, en alguns catalogadors concrets.

Un altre dels problemes que pot tenir el Descobriment de Coneixement en Bases de Dades és no realitzar l'imprescindible estudi previ de la informació emmagatzemada i fiar-ho tot a la tecnologia de mineria de dades.

9.4. Proposta d'aplicació del procés de Descobriment del Coneixement en Bases de dades en un equipament cultural

Un dels objectius d'aquest treball és proposar la forma d'aplicar el procés de Descobriment del Coneixement en Bases de dades en un equipament de forma que en millori la gestió. Per a realitzar-ho proposem el següent pla:

- a) Realitzar una anàlisi documental i un estudi sobre el coneixement associat del fons complet
- b) Integrar, o connectar, el gestor de bases de dades amb les eines d'intel·ligència artificial per a que la tasca de revisió i preparació de les dades quedi reflectida a la base de dades
- c) Fer varies iteracions de revisió de la informació fins que considerem que els patrons són vàlids de forma acceptable
- d) Aplicar un procés de mineria de dades per a extreure'n el coneixement

³⁴ En aquest treball aquesta part només serà una instantània del que es podria aconseguir en una segona fase, la de l'aplicació a una situació real.

³⁵ Es poden veure exemples d'aquest ús del valor de la visualització com a eina de difusió de coneixement a la pàgina web <http://www.visualcomplexity.com/vc/index.cfm?domain=Art>.

- e) Redactar un informe de les conclusions del procés, que inclogui les actuacions que s'haurien de fer, que poden afectar a:
- Modificar l'estructura de les dades
 - Preveure nous protocols de catalogació
 - Realitzar actuacions de preservació i conservació activa en els fons patrimonials
 - Prendre decisions d'adquisició raonades basades en el coneixement adquirit
 - Valorar el coneixement nou, comprensible i valuós que es pot transmetre a la societat
- f) Preveure que es puguin integrar eines d'aprenentatge automatitzat en un futur programari gestor de bases de dades

Abans de començar a aplicar aquest pla s'ha de comptar amb que la seva aplicació hauria de tenir les conseqüències plantejades a l'informe final en la institució en que es realitzi, ja que és un procés costós. Si el procés es queda només en una extracció de coneixement puntual s'ha de comptar amb aquests costos per decidir-se a realitzar-lo.

Una altra proposta que pretenia oferir aquest treball era fer un esbós de com es podien aplicar les eines d'aprenentatge automatitzat en un gestor de bases de dades per a assistir a la introducció de noves dades. Una conclusió a què hem arribat realitzant el treball és que és molt difícil que es puguin crear eines universals d'intel·ligència artificial aplicades a tot tipus d'equipaments patrimonials. Cada equipament té les seves característiques concretes de coneixement i, si ja moltes vegades costa que es pugui utilitzar un programari comú, o adaptar-ne un d'existent³⁶, encara ho és més un de tant complex com aquest. En aquests moments, per exemple, s'utilitza a Catalunya un gestor de col·leccions, MuseumPlus, i a la resta de l'estat espanyol un altre, Domus, així que si ja es difícil compartir eines tradicionals encara ho seria més unes que canvien alguns paradigmes de funcionament.

És tècnicament possible integrar les eines d'intel·ligència artificial en un gestor de bases de dades existent, però el problema és ajustar-lo a cada casuística i aquesta integració l'han de realitzar experts informàtics treballant en col·laboració amb experts en el coneixement de l'equipament cultural. Ens hem de fixar en la característica de no ser trivials dels patrons que es descobreixin per veure com no es pot deixar a l'atzar l'ús d'aquestes eines. A més, aquest procés té una dificultat doble d'adaptació ja que en realitat parlem de dos processos diferents que comparteixen les mateixes eines: (a) unes eines de correcció prèvia, d'adaptació de la informació i de descobriment del coneixement, que aprofitarien per a mantenir i difondre les dades, i (b) unes ajudes a la introducció futura de dades. També seria necessària una formació en quin són els límits de l'aprenentatge automatitzat, ja que, com hem vist demostrat en el cas del nomograma a l'**apartat 8.4.2**, no són eines que donin un coneixement absolut, sinó l'opció més probable a partir del coneixement anterior, així que no substitueixen la necessitat de coneixement expert del catalogador.

³⁶ Ara mateix a Catalunya s'està adaptant el programa Museumplus a l'especificitat dels arxius fotogràfics, o als arxius i museus que contenen fotografies (Dahl, 2014).

10. Bibliografia

- Abreu, C. (2004). El análisis cualitativo de la foto de prensa. *Revista Latina de Comunicación Social*, (57). Retrieved from <http://www.ull.es/publicaciones/latina/20040757abreu.htm>
- Aluja, T. (2001). La minería de datos, entre la estadística y la inteligencia artificial. *Questiió (Quaderns d'Estadística I Investigació Operativa)*, 25(3), 479–498.
- Anuario español de cine: 1963*. Madrid: Sindicato Nacional del Espectáculo.
- Baeza-Yates, R., & Ribeiro, B. de A. N. (1999). *Modern information retrieval* (p. 513). Reading, Mass.: Addison-Wesley Longman. Retrieved from <http://biblioteca.uoc.edu/lilibres/16423.htm>
- Bates, M. J. (2006). Fundamental forms of information. *Journal of the American Society for Information Science and Technology*, 57(8), 1033–1045.
- Belkin, N. J., & Robertson, S. E. (1976). Information science and the phenomenon of information. *Journal of the American Society for Information Science*, 27(4), 197–204.
- Bessièrè, I., & Gili, J. A. (Eds.). (2004). *Histoire du cinéma : problématique des sources*. Paris: Institut National d'Histoire de l'Art.
- Blake, C. (2011). Text mining. *Annual Review of Information Science and Technology*, 45(1), 121–155.
- Boadas i Raset, J., Casellas, L.-E., & Suguet, M. À. (2001). *Manual para la gestión de fondos y colecciones fotográficas. Biblioteca de la imagen* (Vol. 3, p. 425). Girona: CCG : Centre de Recerca i Difusió de la Imatge CRDI.
- Brecht, B. (1963). Las cinco dificultades para decir la verdad (1934). *Boletín Del Seminario de Derecho Político*, (29-30). Retrieved from <http://labyrinth.uma.es>
- Buckland, M. K. (1991). Information as thing. *Journal of the American Society for Information Science*, 42(5), 351–360.
- Calle Buendía, J. (2005). La Catalogación en una Filmoteca : Arxiu de la Filmoteca de Catalunya. *Cuadernos de Documentación Multimedia*, 16, 64–76.
- Campagnoni, D. P. (2006). The Preservation, Care and Exploitation of Documentation Related to the Cinema: An Unresolved Issue. *Film History*, 18(3), 306–318.
- Capurro, R., & Hj rland, B. (2003). The concept of information. *Annual Review of Information Science and Technology*, 37(1), 343–411.
- Carreras Monfort, C. (2009). *Evaluación TIC en el patrimonio cultural metodologías y estudio de casos*. Barcelona: Editorial UOC.
- Cordón García, J. A., López Lucas, J., & Vaquero Pulido, J. R. (2001). *Manual de investigación bibliográfica y documental : teoría y práctica*. Madrid :: Pirámide. Retrieved from http://catalog.uoc.edu/record=b1010316~S1*cat
- Dahl, S. (2014). Estat de la questió: MuseumPlus foto. Retrieved May 14, 2014, from <http://arxiufotograficmmb.blogspot.com.es/2014/05/estat-de-la-questio-museumplus-foto.html>

- Demšar, J. (2010). Material for Data Mining Lectures, Kyoto University, July 2010. Retrieved June 10, 2014, from <http://www.birolab.si/janez/kyoto/>
- Diez Puertas, E. (2003). *Historia social del cine en España*. Madrid: Editorial Fundamentos.
- DLM-Forum. (1997). *Guidelines on best practices for using electronic information*. Luxemburg: Office for Official Publications of the European Communities.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996a). From data mining to knowledge discovery in databases. *AI Magazine*, 37–54.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996b). Knowledge Discovery and Data Mining: Towards a Unifying Framework. In *Second International Conference on Knowledge Discovery and Data Mining*.
- Gallardo Arancibia, J. A. (2009). *Metodología para la definición de requisitos en proyectos de data mining (ER-DM)*. (Ó. Marbán Gallego, Ed.). Universidad Politécnica de Madrid Departamento de Lenguajes y Sistemas Informáticos e Ingeniería de Software, Madrid. Retrieved from http://oa.upm.es/1946/1/JOSE_ALBERTO_GALLARDO_ARANCIBIA.pdf
- Iglésias Franch, D. (2008). *La fotografía digital en los archivos. Qué es y cómo se trata*. Gijón: Ediciones Trea.
- Iglésias Franch, D. (2012). La definició d'un mapa d'informació conceptual per a la gestió de l'arxiu fotogràfic digital. In *Imatge i Recerca: Jornades Antoni Varés*. Girona: CRDI. Retrieved from <http://www.girona.cat/sgdap/docs/vu23efriglesias.pdf>
- López-de-Quintana-Sáenz, E. (2014). Rasgos y trayectorias de la documentación audiovisual: logros, retos y quimeras. *El Profesional de La Información*, 23(1), 5–12.
- Lund, N. W. (2009). Document theory. *Annual Review of Information Science and Technology*, 43(1), 1–55.
- Lyotard, J.-F. J.-F. (1979). *The Postmodern Condition: A Report on Knowledge*. (B. Massumi, Ed.) *Poetics Today* (Vol. 10, p. 886). University of Minnesota Press.
- Manovich, L. (2011). What is Visualization? *Visual Studies*.
- Manovich, L. (2012). Media Visualization: Visual Techniques for Exploring Large Media Collections. *Media Studies Futures*.
- Manovich, L., & Douglass, J. (2009). Visualizing temporal patterns in visual media.
- Mariscal, G., Marbán, Ó., & Fernández, C. (2010). A survey of data mining and knowledge discovery process models and methodologies. *The Knowledge Engineering Review*, 25:2, 137–166.
- Martín Carmona, R., & Repiso Caballero, R. (n.d.). IMDB y su utilidad para la investigación cinematográfica: ejemplos de uso de datos desde la metodología de análisis de redes sociales. In *IV Congreso Internacional sobre análisis fílmico*.
- Oliva i Barceló, N. (2010). *Metodología qualitativa d'anàlisi d'imatges*. Universitat Oberta de Catalunya. Retrieved from <http://hdl.handle.net/10609/1265>
- Oprea, F. (1997). Els principis de la base de la restauració de documents històrics. *Lligall*, (11), 211–218.

- Oyarce, S. (2012). *In pursuit of image: How we think about photographs we seek*.
- Pérez-Montoro Gutiérrez, M. (2003). Gestión del conocimiento y documentación digital: un estudio de caso. *Hipertext.net*.
- Pintanel, M. À. (2008). Foto fixa. Més enllà dels marges. In *10es Jornades Imatge i Recerca*.
- Pintanel, M. À. (2010). 11 anys de catalogació de fotos fixes a la Filmoteca de Catalunya. In *11es Jornades Imatge i Recerca*.
- Riambau, E., & Torreiro, C. (2008). *Productores en el cine español: estado, dependencias y mercado*. Madrid: Filmoteca Española.
- Sangüesa Solé, R. (2010). El procés de descobriment de coneixement a partir de dades. Objectius, fases i problemàtica. In *Mineria de dades*. Barcelona: Universitat Oberta de Catalunya.
- Sangüesa Solé, R., Mor i Pera, E., & Molina Félix, L. C. (2010). *Mineria de dades*. Barcelona: Universitat Oberta de Catalunya.
- Sonnenburg, S., Braun, M., Ong, C. S., Bengio, S., Bottou, L., Holmes, G., Williamson, R. (2007). The Need for Open Source Software in Machine Learning. *Journal of Machine Learning Research*, 8, 2443–2466. Retrieved from <http://eprints.pascal-network.org/archive/00003273/>
- Trybula, W. (1997). Data Mining and Knowledge Discovery. *Annual Review of Information Science and Technology*, 32.
- Tufte, E. R. (1997). *Visual explanations :images and quantities, evidence and narrative* (p. 156). Cheshire, Conn.: Graphics Press. Retrieved from <http://biblioteca.uoc.edu/l1ibres/9019.htm>
- Valero de Bernabé, A. (1944). *España cinematográfica : recopilación de cuanto concierne al arte, industria y comercio del cinema español : [anuario 1943]*. Madrid: Cinegrafos.
- Vardigan, M., & Whiteman, C. (2007). ICPSR meets OAIS: applying the OAIS reference model to the social science archive context. *Archival Science*, 7(1), 73–87.
- Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining. Practical Machine Learning Tools and Techniques* (3rd ed.). Burlington: Morgan Kaufman Publishers.
- Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., Steinberg, D. (2007). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1), 1–37.

11. Annexos

11.1 Guia d'entrada a la base de dades

Identificació

Núm registre objecte : [Número registre].F/[Número fotografia]
Número d'ordre : [Número registre]
Organització museu : Fotografia
Nom de l'objecte : [Tipus d'emulsió. Normalment una d'aquestes
opcions: positiu, negatiu,
negatiu/internegatiu]
Status : [normalment: complet]
Nombre d'exemplars : [només cal omplir-ho en el cas de tenir còpies
idèntiques. Normalment positius paper]
Altres números : [posar el número d'inventari Sabaté, és el
número escrit a mà a les caixes de cartró]

Dades títol/responsabilitat

Títol / nom propi : [Títol original tot en majúscules]
Títol paral·lel : [altres títols tot en majúscules]
Precisions al títol : [només omplir si s'ha de precisar alguna cosa
sobre els títols]
Autor del document : [seguits, separats per comes, autors del
document en el format: nom cognom (funció).
Els autors que s'han de posar són: foto fixa,
director, director de fotografia, decorador.]
Precisions autor : [només omplir si s'ha de precisar alguna cosa
sobre els autors del document]
Intèrprets : [seguits, separats per comes, **només els
intèrprets que surten a la fotografia** en el
format: nom cognom (paper). Si hi ha actors
que no coneixem posar "desconegut" o
"desconeguda", els desconeguts que apareixen
en més d'una foto posar un identificador, per
exemple "desconegut (Home 1)".]
Personatge : [no omplir si no són personatges de dibuixos
animats]
Preci intèr/personat : [només omplir si s'ha de precisar alguna cosa
sobre els intèrprets]
Nom de la sèrie : [no omplir]
Número dins la sèrie : [no omplir]
Nombre capítol sèrie : [no omplir]

Dades de producció i edició

Nacionalitat document : [tot en majúscules, país de producció. Si és
una coproducció separar els països amb barres
"/"]
Entitat productora : [empresa productora. Si n'hi ha més d'una
separar per barres "/"]
Lloc.entitat product : [ciutat seu de les productores]
Serveis de producció : Reproducciones Sabaté (Barcelona)
Col.laboradors prod : [entitats que han col·laborat en la producció,
però no són empreses productores]
Preci entitat product : [només omplir si s'ha de precisar alguna cosa
sobre la producció del film]
Distrib document : [empresa distribuïdora]
Lloc distribució : [país de distribució]
Preci distr document : [només omplir si s'ha de precisar alguna cosa
sobre la distribució del film]
Data inici producció : [només omplir si coneixem la data d'inici del
film, posar en el format dia/mes/any]

Data final producció : 00/00/0000 [dia/mes/any si no tenim la data exacta deixem a 00 el que no sabem]
Justificació datació : [Opcions: contextualització, publicació, font, inscripció, referència, datat]
Precisions datació : [si hem omplert el camp anterior, precisar la justificació de la datació, si són fonts o publicacions han d'estar citades al camp bibliografia]
Lloc d'execució : [omplir si el lloc real de rodatge és diferent del lloc que representa la imatge.
Nivells: si és a Catalunya en aquest camp podem posar fins a pobles, si és Espanya fins a províncies, si és estranger posar el país]
Lloc precís : [lloc més concret de rodatge, que es pugui situar en un mapa o plànol. Pot ser el nom de l'estudi, si el coneixem i la imatge és un interior]

Descripció del contingut

Gènere audiovisual : [opcions: fotografia de rodatge, fotografia d'estudi, fotografia de reportatge]
Sinopsi : [descripció de l'escena representada a la fotografia. Si hem pogut veure el film posar-la en relació a ell, si no l'hem pogut veure podem posar una descripció més abstracta, per exemple "Parella besant-se apassionadament en uns escales"]
Descripció imatge/so : [seguits, separats per comes, termes que descriu la imatge]
Persones : [no omplir si no surten a la fotografia persones que al film s'interpretin a si mateixes o que apareguin en imatges documentals]
Lloc imatge / so : [Lloc que representa la imatge.
Nivells: si és a Catalunya en aquest camp podem posar fins a pobles, si és Espanya fins a províncies, si és estranger posar el país]
Lloc pre imatg mòbil : [lloc més concret de rodatge, que es pugui situar en un mapa o plànol]
Data imatge mòbil : [dia/mes/any, si es pot precisar]
Accident geogràfic : [quan apareix a la foto un accident geogràfic conegut. Exemples: Muntanya de Montserrat, Llac de Sant Maurici, Illa de Capri, etc]
Autor imatge/so repr : [quan a la foto es veu una obra d'art de la qual puguem donar el nom i l'autor]
Preci imatge/so : [omplir si hi ha alguna cosa a precisar sobre el grup de camps *Descripció del contingut*]
Carpeta : [no omplir]

Bibliografia

Bibliografia : [bibliografia, videografia i informació electrònica (internet, cd-rom) consultada per fer les fitxes. Cal seguir les normes de citació de l'arxiu]
Precisions bibliogra : [omplir si hi ha alguna cosa a precisar sobre el camp de bibliografia]
Objectes relacionats : [no omplir]

Descripció material

Material : [material de la fotografia, opcions: acetat de cel·lulosa, Nitrat de cel·lulosa, paper, vidre]

Precisions mat/tec : [omplir si hi ha alguna cosa a precisar sobre el camp de material. Per exemple retocs (indicant tipus: amb llapis, per fregat, per incisió), marques d'enquadrament, màscares (translúcides, opaques)]

Dimensions : [Si és gairebé la mida estàndard, posar-la, si no, posar la mida en mil·límetres. Mides estàndard: en negatiu són 180 x 240 / 120 x 180 / 90 x 120 / 60 x 90 / 60 x 70 / 60 x 60 / 45 x 60 / 35 mm; en paper són 180 x 240 / 120 x 180 / 90 x 120 / 70 x 100]

Escala : [no omplir]

Tipus de llum : [opcions: llum artificial, llum mixta, llum natural]

Tècnica fotogràfica : Gelatino-bromur

Enquadrament / pla : [tipus de pla. Primer pla, pla general, etc.]

Color : [opcions: blanc i negre, color]

Pas : [no omplir]

Format : [no omplir]

Tipus forma : [no omplir si no són formats especials: ovalat, escairat, etc.]

Status altres suport : [opcions: positiu, internegatiu]

Tipus alt. sup. : [no omplir]

Localització alt. sup.: [no omplir]

Referència alt. sup. : [número de registre del positiu relacionat]

Autor alt. sup. : [no omplir]

Data alt. sup. : [no omplir]

Preci alt. sup. : [no omplir]

Associació : [número de registre del negatiu original]

Inscripcions

Tipus d'inscripció : [opcions: inscripció, marca de fàbrica]

Text inscripció : [còpia literal de la inscripció o marca de fàbrica. Si alguna cosa està mal escrita posar-la tal com està i (sic) al darrera, si falta text posar [?] i si no s'entén el que està escrit posar [illegible]]

Preci inscripcions : [precisions a la inscripció. Descripció: incisa, escrita amb tinta, impresa amb tinta, fotografiada. Situació: indicar si està al suport o a l'emulsió. Posició: amb el suport cap a nosaltres el lloc de la fotografia on estan]

Perforacions : [Descripció de les osques de la fotografia]

Presentació : [Descriure primer la capsa apuntant el text de forma literal, si hi ha alguna targeta o paper indicar-ho. Després descriure el sobre, si té algun text apuntar-lo textualment]

Condicció / restauració

Estat conservació : [opcions: bo, primer grau, segon grau, tercer grau]

Data est conservació : [data del dia que heu mirat la fotografia,
format dia/mes/any]
Memòria estat conser : [consultar]
Percentatge pèrdua : [no omplir]
Parts que manquen : [no omplir]
Número de les parts : [no omplir]
Alteració/desperfect : [no omplir]
Precisions condició : [no omplir]
Restauracions : [no omplir]

Inventari

Ubicació : [no omplir]
Precisions ubicació : [quan un suport contingui dues imatges les
relacionarem en aquest camp, exemple: "en el
mateix suport que 13680.F/12"]
Memòria ubicació : [no omplir]
Sèrie Fotogràfica : [no omplir]
Número de clixè : [no omplir]

Ingrés al museu

Forma d'ingrés : Compra
Data d'ingrés : 13/06/1990
Font d'ingrés : [no omplir]
Col.lec procedència : Reproducciones Sabaté
Precisions ingrés : [no omplir]
Expedient : [no omplir]

Drets explotació

Copyright : [no omplir]
Drets d'utilització : [no omplir]
Precisions als drets : [no omplir]
Expedient contracte : [no omplir]

Història de l'objecte

Premis : [no omplir]
Història objecte : [no omplir]
Propietaris anterior : [no omplir]
Emplaçament d'origen : [no omplir]
Emplaçament precís : [no omplir]
Emplaçaments succe : [no omplir]

Control dades generals

Registre fet per : [nom de la persona que ha fet la fitxa]
Data de redacció : 00/00/0000 [data en que es fa la fitxa en el
format dia/mes/any]
Registre revisat per : [no omplir]
Data de revisió : [no omplir]

11.2 Fitxa de descripció fotogràfica

Núm. registreF/

Condicions de conservació

Presentació caps:

embolcall: sobre solapa de punta sobre solapa lateral sobre dossier altres

material

color

text

Descripció

Support imatge acetat nitrat paper vidre **Qualitat** negatiu negatiu/internegatiu positiu

Dimensions x mm **Color** blanc i negre color

Inscripcions / marques

retocs: incisió fregat llapis màscara translúcida
altres/precisions

marques d'enquadrament: tinta incisió màscara opaca
altres/precisions

Tipus d'inscripció / marca d'origen

inscripció marca de fàbrica segell
 tinta incisa fotografiada

Transcripció

Ubicació de la inscripció suport emulsió

Descripció

Tipus d'inscripció / marca d'origen

inscripció marca de fàbrica segell
 tinta incisa fotografiada

Transcripció

Ubicació de la inscripció suport emulsió

Descripció

Tipus d'inscripció / marca d'origen

inscripció marca de fàbrica segell
 tinta incisa fotografiada

Transcripció

Ubicació de la inscripció suport emulsió

Descripció

Tipus d'inscripció / marca d'origen

inscripció marca de fàbrica segell
 tinta incisa fotografiada

Transcripció

Ubicació de la inscripció suport emulsió

Descripció

Perforacions

Examen de conservació

Data de l'examen/...../.....

bruta fongs suport emulsió

mirall de plata desprendiment emulsió

ratlles emulsió: fines pronunciades taques

suport: fines pronunciades degradació

picada arrugada

ditades emulsió suport balcada

bosses d'aire

Precisions

Estat físic Bo primer grau segon grau tercer grau inservible

Estat químic Bo primer grau segon grau tercer grau descomposició

Núm. registre .FF/

Fotògraf

Fotògraf:

Data de creació

Lloc de creació

Classificació genèrica**Fotografia:** d'escena de rodatge d'estudi cartell de reportatge collage altres**Descripció analítica****Descripció analítica:**

.....

.....

.....

.....

.....

.....

Indexació**Assumpte / tema:**

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Condicions de la presa de vista**Il·luminació:** llum natural llum mixta llum artificial**Enquadrament:** pla general pla general curt pla americà pla mitjà pla mitjà llarg pla mitjà curt primer pla