



**UNIVERSITAT DE
BARCELONA**

**Màster en Humanitats Digitals
Curs 2020-2021**

Projecte final de màster

**Creació de prototip de plataforma de difusió cultural per a
web estàtica i integració de visors d'imatge i notació musical
basats en esquemes IIF i MEI**

**Gerard Vidal Santos
Tutor: Miquel Centelles Velilla**

Agraïments

Als meus pares

A en Miquel, per la seva paciència i consells en el plantejament del treball

A Glen Robson, per les sessions del taller IIF i ajudar-me en la construcció del manifest

A Alex Gil i la resta dels membres del projecte Minicomp, per servir d'inspiració per aquest projecte

i a Paula (i especialment) a Irene, per les seves aportacions teòriques i la seva ajuda en la transcripció dels exemples

Resum executiu:

La democratització de l'accés a les tecnologies de la informació i dels requisits tècnics per a fer-ho possible ha permès a moltes institucions patrimonials incorporar la digitalització i la gestió digital de les seves col·leccions al seu nucli de treball diari i estendre, així, la carta de serveis a l'usuari en aquest mitjà, tant en el vessant més tècnic com en la difusió i interacció amb l'usuari en les xarxes socials (Raju, 2017; Grevtsova, 2020). El procés, destinat a facilitar l'accés als seus fons de manera no presencial i preservar digitalment els seus continguts a llarg termini s'ha caracteritzat per valorar molt positivament les xifres de documents accessibles i la proliferació de dipòsits digitals institucionals però, comparativament, els informes d'avaluació solen fer poc èmfasi en el retorn de la inversió produïda i altres supòsits que analitzin el seu efecte més enllà del merament quantitatiu. (Borrego, 2020; Shaw, 2016).

Aquest corrent, refermat per la popularitat de sistemes CMS i la fortalesa de sistemes de codi obert destinats a la comunitat investigadora i bibliotecària (Dspace, Omeka) ha propiciat la creació d'implementacions amb un disseny d'interacció amb poques variacions respecte a la seva matriu original, que assimilen les opcions per defecte de cerca i visualització de continguts i centren el tractament documental en models de descripció bàsics a l'espera que el suport del motor de cerca a text complet supleixi la falta de detall en la indexació i descripció dels objectes. Arguments en aquesta línia (Nualart, 2016) fan notar des d'una perspectiva pragmàtica que, com el nombre d'objectes digitals allotjats en aquestes plataformes creix exponencialment, la urgència per a oferir millors sistemes de recuperació per al seu accés augmenta a risc que, a mesura que la ingesta de materials tan diversos en la seva tipologia documental i disciplina no activi una recíproca inversió en interfícies que puguin gestionar tècnicament la complexitat de les seves col·leccions, augmenti el risc que un tractament documental homogeni malmeti l'accés a aquesta riquesa de continguts i faci encara més difícil el seu ús en recerca.

El tractament més habitual per a proveir d'accés telemàtic a col·leccions digitals no va més enllà de la publicació d'imatges o un tractament de reconeixement òptic de caràcters (OCR) per a facilitar la captura del text, fins i tot a risc d'oferir versions sense corregir. L'ús de tecnologies per a la geolocalització, l' anotació manual o automàtica, la implementació de relacions LOD o la visualització de materials complexos que no depenen exclusivament de text o imatge són poc comuns en aquesta mena de plataformes i solen presentar-se en plataformes específiques preparades per aquesta funció, amb una implementació tècnica a mesura que dificulta la seva integració en col·leccions digitals més generalistes o agregadors de contingut.

Seguint la descripció que planteja Lavoie (2020) per a la planificació de projectes digitals en el marc GLAM es pot observar com la necessitat d'innovació en els sistemes d'accés a les col·leccions i la sostenibilitat econòmica de cada projecte es veuen cada vegada més condicionats per la infraestructura tècnica que suporta la instal·lació, i com una mala planificació en aquest sentit pot obstruir nous desenvolupaments. Si aquestes estructures no estan preparades per a fer accessibles les col·leccions en un flux de treball modular, escalable i amb perspectives de suportar noves actualitzacions en un futur, els costos per a plantejar la logística necessària per a la millora/migració a un nou desenvolupament poden restringir en excés la inversió econòmica per a fer-ho possible.

A partir d'una revisió bibliogràfica de models d'optimització i millora d'infraestructures i interfícies per a la preservació i accés a col·leccions digitals aquest treball realitza un estudi per a identificar els patrons de construcció en la seva infraestructura tècnica, models de descripció dels objectes i enllaços a contingut relacionat en la seva infraestructura basat en la recol·lecció de les dades disponibles en línia per a la sistematització dels seus punts forts i les possibilitats d'interrelació amb desenvolupaments actuals en el camp de musicologia i la difusió de patrimoni musical a partir d'exemples ja consolidats.

A continuació s'ofereix el disseny i implementació d'un prototip basat en models JAMStack (Tramullas, 2020) capaç de gestionar objectes digitals multimèdia i oferir visualitzacions que permeten una major interacció amb una obra literària que conté text, imatge i notació musical. En aquesta secció es presenta la obra escollida, el tractat de Felip Pedrell *Cancionero musical popular español* de 1922 i s'infereixen els resultats de desenvolupaments similars en aquesta mena de plataformes (Gil, ; Saccomacco, 2020) per a establir una base de disseny per al tractament tècnic de cadascun dels seus components que pugui solucionar les mancances dels actuals models de publicació de contingut patrimonial per a la recerca assistida per tècniques computacionals amb els següents objectius:

- La transformació per mètodes de reconeixement òptic de caràcters (OCR) per a la part textual de l'obra a un símil digital que adapta l'experiència de lectura de l'obra a un entorn digital
- La disposició d'una edició facsímil que se sustenta en l'ús dels estàndards de la *International Image Interoperability Framework* (IIIF)
- La transformació de les partitures musicals a notació musical digital per mètodes de reconeixement òptic musical (OMR) i la seva transformació segons el model promogut per la Music Encoding Initiative (MEI)
- La implementació d'una plataforma d'infraestructura mínima per a l'accés en línia dels derivats d'aquestes transformacions

Finalment s'analitzen els resultats obtinguts, les possibilitats de l'ús d'aquests estàndards en els fluxos de treball dels centres documentals per a la promoció de la recerca assistida computacionalment i s'elabora una proposta d'integració del prototip en l'ecosistema de difusió del patrimoni documental català.

En el resultat final es valida que la construcció d'estàndards de descripció basats en els principis ODD permetin un tractament documental més ric que els estàndards usats en els repositoris actuals, que l'ús d'aquest sistema permet la creació de derivats digitals capaços de donar resposta a necessitats més específiques en el camp de recerca i la difusió i que la integració de sistemes JAMStack en un circuit de gestió de col·leccions digitals permet la creació de desenvolupaments tècnics més àgils, avançats i eficients tant en disseny d'interacció com en la preservació a llarg termini dels continguts que es disposen.

Sumari

1. Introducció.....	6
2. Revisió bibliogràfica.....	7
2.1 Situació de la codificació de textos digitals i música notada per a la recerca.....	7
2.2 Situació de la gestió de biblioteques digitals a institucions patrimonials i de recerca.....	9
2.3 Situació de plataformes digitals de recerca i difusió del patrimoni musical a Catalunya.....	12
2.4 Solucions modulars per a la interoperabilitat de plataformes digitals de recerca i difusió del patrimoni musical.....	14
2.4.1 Plataformes JAMStack.....	14
2.4.2 International Image Interoperability Framework (IIIF).....	16
2.4.3 Music Encoding Initiative.....	18
2.4.4 Exemples d'aplicació.....	20
3. Objectius del projecte.....	23
4. Consideracions metodològiques, tècniques i logístiques per a la creació del projecte.....	24
4.1 Anàlisi de la estructura formal de l'obra.....	24
4.2 Establiment de fases de treball i parametrització de processos.....	26
4.3 Criteris de selecció i altres consideracions logístiques.....	28
5.1 Digitalització, transformació de formats i edició de continguts.....	29
5.1.1 Procès de digitalització.....	29
5.1.2 Transcripció de textos.....	30
5.1.3 Transcripció de música notada.....	31
5.2 Creació de la edició facsímil.....	35
5.3 Disseny i prototipatge de la plataforma d'accés digital.....	38
6. Discussió del resultat.....	40
6.1 Limitacions del disseny i la parametrització de processos plantejats.....	40
6.2 Integració del prototip en altres sistemes de gestió de continguts.....	42
6.4 Pressupost estimat.....	44
6.3 Ús dels principis ODD en la consolidació d'iteracions en el tractament documental.....	46
7. Conclusions i futures actuacions.....	48
8. Bibliografia.....	49
9. Annexos.....	51

1. Introducció

En les últimes dues dècades la gestió de col·leccions digitals ha experimentat un creixement exponencial que ha posat en evidència els sistemes clàssics d'organització del coneixement en les disciplines que, al llarg dels anys, han vetllat per produir un cos teoricopràctic per al correcte funcionament de les unitats d'informació, la selecció i marc descriptiu dels objectes documentals que custodien i a la eficàcia en el disseny dels seus sistemes de recuperació per garantir-ne l'accés.

Malgrat els clars avenços en l'ús d'eines més complexes i tecnològicament més adients per a les noves tipologies documentals que es van generant en l'àmbit digital i que la visió de fons en els estudis d'arxivística i biblioteconomia segueixen plenament vigents, algunes premisses de fons encara massa presents en algunes concepcions ancorades en el que podríem anomenar «espai de confort» --com el tractament de la descripció dirigida a l'objecte físic o l'ús d'estàndards de descripció poc permeables a visions complementàries provinents d'altres disciplines-- són peça mecànica que pot reverberar negativament en la evolució de la disciplina en el xoc que ha representat l'aposta per la digitalització en massa i l'obertura digital de les col·leccions dels centres.

La generalització simplista d'aquest fet i la presència d'exemples modèlics no ha de restar preocupació per l'evolució de les col·leccions digitals que massivament són allotjades i posades a disposició del gran públic i de les comunitats de recerca especialitzades a un ritme vertiginós i sincopat, que confia més en pràctiques agressives de SEO i la descoberta algorítmica poc controlada que a aquelles pràctiques orientades a dirigir esforços a ser part component del procés de recerca en totes les etapes en que són rellevants.

Per aquest motiu el punt de partida d'aquest treball és la rellevància en el procés de recerca i comunicació científica. En ell es parteix d'una orientació pràctica per a explicar el condicionants que poden crear obstacles a la generació de nou contingut a partir de materials digitalitzats en els repositoris patrimonials actuals, l'adopció de pautes de treball que faciliten la difusió de continguts i el treball computacional en altres disciplines i la formulació d'una proposta per a la creació d'una plataforma digital flexible per a la transliteració digital d'una obra física d'abast temàtic musical que pugui donar resposta a la adequada visualització dels materials que la componen i incrementar el catàleg de possibles usos després de la digitalització.

2. Revisió bibliogràfica

2.1 Situació de la codificació de textos digitals i música notada per a la recerca

La digitalització a les disciplines que componen el marc d'estudi de les humanitats ha estat un procés gens homogeni i ple de complexitats durant les dues últimes dècades. A partir de l'aparició de nous sistemes de tractament, anàlisi i difusió que han acompanyat l'aparició d'Internet, la comunitat ha viscut aquests processos de canvi de manera gens homogènia i les discussions als fòrums de debat han propiciat reaccions del tot oposades, segons el camp d'estudi i els agents implicats.

En cap cas es pot delimitar un marc general per aquest fenomen, els postulats més representatius els podem trobar en el ja clàssic *Apocalíptics i Integrats* d'Umberto Eco (1999) però, al llarg d'aquests últims anys es focalitza en aspectes més concrets com la preocupació en els plans docents dels ensenyaments (Grusin, 2014) i la representativitat dels valors associats a aquestes disciplines a la vida pública (Ordine, 2013). Des d'una perspectiva respectuosa amb els múltiples punts de vista d'aquestes comunitats i la evolució de la seva producció científica, es pot discernir que el xoc (potser disruptiu en alguns casos) està sent, més aviat, una «progressiva confluència» en el seu conjunt, de penetració fragmentada i molt diversa en cada disciplina segons els interessos de cada conjunt d'investigadors i les col·laboracions amb altres disciplines (Apollon et al., 2014; Zhang et al., 2015).

L'assimilació dels plantejaments originats en la gestió de projectes en l'àmbit de les ciències de la computació, especialment els referits a la facilitat d'establir repositoris de codi obert i el desenvolupament col·laboratiu a partir de sistemes *git* de control de versions, han fet possible un ràpid desplegament d'eines a l'abast de molts investigadors per accelerar la seva recerca i potenciar col·laboracions amb altres departaments, però segueixen havent contratemps en eleccions de base tècnica i metodològica per al desenvolupament d'eines per a la edició de textos crítics i música notada digitals.

El primer d'ells és la elecció de formats per al tractament i la reutilització de les dades de recerca. Com s'argumenta des de la iniciativa TEI¹, la complexitat d'integrar usos no previstos en les dades de recerca en un estàndard que pugui ser reutilitzat per altres investigadors recau sobretot en el consens dins de la mateixa disciplina, les afinitats amb altres comunitats de recerca i, en la gran majoria dels casos, la capacitat del programari que ha de tractar les dades per a poder fer conviure en un mateix fitxer aquestes particularitats.

1 Acrònim de *Text Encoding Initiative*

A diferència de les disciplines més tècniques, on el material primari és majoritàriament de patró numèric, els objectes digitalitzats que formen el nucli d'estudi al camp de les humanitats es componen de materials molt diversos, fortament entrelaçats a l'objecte físic que representen, difícils de ser reutilitzats sense modificar-los d'alguna manera i, aquest contratemps en penalitza la reutilització. (Schmidt, 2014; Botè-Tèrmens, 2019). Tant la edició crítica de textos literaris, de fonts històriques o musicals (les àrees on s'exploraran nous mecanismes de visualització en aquest treball) han mantingut tradicionalment cànons metodològics propis en la codificació dels textos per al seu anàlisi, i dins de la mateixa disciplina els criteris que emparen cada codificació poden variar segons «l'escola» que desenvolupa el projecte i els fins que pretén aconseguir. En aquest sentit, el desenvolupament de formats *ad-hoc* o els problemes d'obsolescència han estat a l'ordre de dia durant molt de temps i cada comunitat de recerca en musicologia ha prioritzat de manera autònoma les funcionalitats que creia més convenient per al seu projecte (Merchán Sánchez-Jara, 2019).

Malgrat la progressiva adopció d'iniciatives TEI en entorns digitals de recerca, la base d'interoperabilitat sota l'esquema *XML* es segueix resolent en la majoria de casos de manera pragmàtica a partir de transformacions i personalitzacions a l'esquema original o exportant a formats de text pla per a ser importats per a altres usos, evitant les molèsties tècniques que suposa explorar la possibilitat d'incorporar els resultats de noves investigacions en un mateix contenidor o entrelaçar-los per a facilitar-ne la reutilització.

Aquesta simplicitat logística frena la difícil translació de la proposta *ODD (One Document Does it All)* a escenaris reals, on el pes del cicle de publicació i els nínxols de cada disciplina segueixen tenint un pes molt important, i fa evident que el binomi necessitats dels investigadors/capacitat tecnològica de les plataformes de treball i difusió és part del problema. Les crítiques en aquest sentit denoten una situació encara en eferescència, plena de propostes i desenvolupaments per construir dreceres i ponts per connectar més fàcilment aquests usos però es remarcable com tots els escrits apunten al disseny de les plataformes digitals com part essencial per a la seva resolució (Davison, 2016; Schmidt, 2014).

La disponibilitat i la perspectiva en el de maquinari i sistemes de programari per gestionar els continguts digitalitzats afecten tant, o més, a l'accessibilitat, el descobriment, la usabilitat i la viabilitat a llarg termini de col·leccions digitalitzades que els recursos econòmics destinats en la digitalització de fons patrimonial (Shaw, 2016) i l'establiment d'estratègies digitals solvents i perdurables .

Si posem per cas els mètodes de treball basats en tècniques computacionals es fàcil preveure que la interrelació entre aquests i la custòdia de materials que són la font primària d'aquestes investigacions és tant forta que pot condicionar totalment les dinàmiques de creació de contingut en positiu, si el flux de treball dels dos agents implicats es coordina. En canvi, si aquestes particularitats no troben sintonia amb el treball intern dels centres de documentació, les possibilitats de col·laborar recíprocament disminueixen, evitant la

possibilitat d'incorporar nou material amb alt valor afegit en la recerca i el tractament de dades que les sustenta i la recíproca ampliació del catàleg de serveis a usuaris especialitzats.

Per aquest motiu l'establiment d'estratègies digitals solvents i perdurables a l'hora de planificar usos i la selecció de sistemes de gestió de continguts digitals en base a la interoperabilitat dels sistemes de recuperació i la reutilització dels materials dipositats afecta significativament la capacitat operativa de tots els agents que participen dels objectius dels centres que en són responsables i és un dels aspectes on es possible un ample marge de millora. .

2.2 Situació de la gestió de biblioteques digitals a institucions patrimonials i de recerca²

La democratització de l'accés a les tecnologies i dels requisits tècnics per a desenvolupar plataformes digitals ha permès que moltes institucions hagin incorporat la gestió digital de les seves col·leccions al seu nucli de treball diari per fer extensible la carta de serveis a l'usuari en aquest mitjà, tant en la vessant més tècnica com en la difusió i interacció amb l'usuari a les xarxes socials (*Raju, 2017; Grevtsova, 2020*) malgrat que la definició de perfils i tasques associades no té prou establerts els criteris per establir les competències bàsiques de les persones encarregades de dur-les a terme (*Fuhr, 2022; Castro, 2021*) i que, a mesura que les necessitats de la plataforma augmenten en especificitat, els desenvolupaments són cada cop més costosos i requereixen de processos de transformació amb molta intervenció humana i un alt cost per als equips documentals d'aquestes institucions.

Qüestions com la preservació de contingut generat per programari amb dependències que poden veure-s afectades per obsolescències no desitjades, la persistència dels identificadors per a garantir l'accés a aquests materials i la demanda de nous desenvolupaments que aprofitin la tecnologia existent per donar millors condicions per a l'accés i la interacció amb els objectes digitals que es custodien han fet replantejar a molts agents les qüestions de base a l'hora d'engegar nous projectes de digitalització (*Newson, 2017; Diaz, 2018; Babcock et al, 2020*) davant de la perspectiva que «no hi ha, ni hi haurà, una forma canònica i estable per a la presentació i connexió de textos en el món digital»³, sinó un seguit de processos vagament interconnectats per a la creació d'aquests materials que depenen de les fortaleces com a grup de la comunitat de treball que en depèn.

2 Les conclusions d'aquest apartat s'han elaborat a partir d'exemples que poden ser representatius de les dificultats exposades en les seccions anteriors del text, que es citen a al final de la secció, però no pretenen establir-se com a una anàlisi exhaustiva de les solucions particulars que cada centre ha implementat. En la discussió del treball s'aborda de nou la implementació al territori prenent com a punt de partida la distribució de centres amb documentació musical a Catalunya

3 Cita atribuïda a Maria Morràs segons *Schmidt D. Towards an interoperable digital scholar edition*. En el text original s'aplica a la edició crítica de textos digitals però l'escenari presenta moltes similituds a la construcció de biblioteques digitals

A més, la proliferació de nou contingut ha fet més evident la necessitat del treball en xarxa en el camp de les humanitats i les institucions patrimonials veuen la necessitat d'establir sistemes d'informació que puguin plasmar en el seu disseny aquesta riquesa d'escenaris on poden ser d'utilitat i ser correspostos de força recíproca, ja sigui en la anàlisi de dades primàries que custodien, la seva visualització o altres maneres d'interactuar amb elles, segons intencionalitat (recerca, difusió) o el camp d'estudis que es poden (o volen) potenciar. Arguments en aquesta línia (Nualart, 2016) fan notar que, com el nombre d'objectes digitals allotjats en aquestes plataformes creix exponencialment, la urgència per oferir millors sistemes de recuperació per poder accedir-hi augmenta. A mesura que la periòdica ingesta de materials molt diversos en la seva tipologia documental i disciplina no activa una recíproca inversió en interfícies que puguin gestionar tècnicament la complexitat de les seves col·leccions augmenta el risc que un tractament homogeni malmeti l'accés a aquesta riquesa de continguts i faci encara més difícil el seu ús en recerca.

Malgrat que no es possible establir de manera sistemàtica les barreres per a la acceleració dels mecanismes per donar servei més especialitzats a comunitat de recerca i integrar les seves demandes en els mecanismes de tractament documental sense una avaluació específica centre a centre, hi ha certs elements que poden identificar patrons comuns i independents de la posició del centre en el organigrama intern de la seva institució o als factors personals adjacents a discursos de resistència al canvi i als rols de l'equip de treball en la mateixa unitat d'informació.

En primer lloc, la granularitat de les descripcions que acompanyen als objectes digitals és, sovint, poc exhaustiva i prioritza la homogeneïtat en convivència amb altres tipologies de materials o descripcions provinents d'altres centres a una catalogació més profunda per a millorar-ne l'accés o la interrelació amb altres objectes amb característiques comunes (Davison, 2016; Nualart, 2016).

Els motius d'aquest mínim comú denominador varien segons l'origen de les col·leccions, el tractament previ documental i la capacitat d'avaluació iterativa dels centres per corregir i actualitzar inconsistències a l'hora de fer-lo accessible a la xarxa però, sense entrar en detalls en cada un dels casos observats, es poden generalitzar en la sistematització d'errors que Hannah Tarver i Steven Gentry (2021) identifiquen per operacions de manteniment a les col·leccions digitals o els problemes identificats per Gil-Leiva et al. (2020) en el seu anàlisi sobre l'abandonament de l'assignació d'encapçalaments de matèria y codis de classificació en les biblioteques universitàries degut a la incorporació massiva del paquets de subscripció de llibres electrònics, assimilable a la introducció massiva de registres bibliogràfics a OPACs en la compra consorciada de revistes electròniques o la creació de sistemes compartits de dipòsit i publicació (consorciats o per agregació) i els seus efectes en la gestió dels sistemes de descoberta (Sälgö, 2022).

Per una altra banda, els esforços en la digitalització de patrimoni cultural al territori s'han centrat en valorar molt positivament les xifres de documents accessibles i la proliferació de dipòsits digitals institucionals a

nivell quantitatiu, prioritzant el contingut per sobre del contingut i tractant de manera unitària les col·leccions textuais, d'imatges, mapes o música notada. Aquesta situació sovint predisposa a generar visors i *sites* externs a la plataforma i una càrrega de treball elevada en la gestió de la col·lecció digital que facilita la dispersió de recursos en petits regnes de taifes (efecte *silo*) o conjunts consorciats en un dispositiu col·lectiu que homogeneïtza la disposició de les col·leccions allotja.

En aquest sentit l'ús intensiu de plataformes CMS⁴ dissenyades a la publicació de continguts culturals i recerca a partir de la infraestructura LAMP⁵, que ha facilitat la creació de moltes plataformes digitals en els últims anys, es veu compromès en l'augment de la complexitat de gestió d'aquests sistemes i la seva necessària actualització, la implantació de prestacions *ad-hoc* imprescindibles per a les funcions del centre amb dependències que no es possible gestionar a partir de la base instal·lada i, de manera creixent, la coordinació del sistema amb fonts d'informacions externes a través d'interfícies de programació d'aplicacions o APIs⁶ (Tramullas, 2020) o la seva integració amb altres sistemes de gestió del mateix centre (OPACs, DAMs).

Aquests condicionants prenen forma en exemples propers al nostre territori com la construcció de *sites* per a projectes a la xarxa CORE de patrimoni cultural de la UAB⁷, les solucions proposades des del *Servicio de Biblioteca* de la *Universidad Carlos III de Madrid*⁸ per l'increment de demandes IT a la unitat per crear portals de difusió dels projectes que duen a terme els grups de recerca de la universitat, o els problemes d'interoperabilitat en l

La càrrega de treball necessària per aprofundir en cada un dels supòsits anteriors pot ser considerablement elevada i poc assumible per moltes institucions. En tots dos casos, les característiques del sistema de publicació escollit, ja sigui en la creació de projectes basats en la coordinació MySQL+PHP o CMS individuals amb els dipòsits institucionals del centre o l'adopció d'OMEKA S, una variant del CMS que permet la publicació de múltiples instàncies des de la mateixa BBDD del servidor, són indicatius dels problemes d'integració en el conjunt de serveis de la institució que les gestiona quan el volum sobrepasa la capacitat de resposta.

4 Acrònim de Content Management Systems

5 Acrònim de *Linux, Apache, Mysql* i *PHP/Perl*

6 Acrònim de *Application Programming Interface*

7 Vicente O. Transferència i Innovació del Patrimoni Digital: de la universitat a la comunitat. Barcelona: Centre de Recerca en Informació, Comunicació i Cultura (CRICC); Consorci de Serveis Universitaris de Catalunya (CSUC); Facultat d'Informació i Mitjans Audiovisuals (UB), 2019 [citad 22 març 2021]. Disponible a: <https://www.slideshare.net/FBDUB/transfencia-i-innovaci-del-patrimoni-digital-de-la-universitat-a-la-comunitat>

8 Malo de Molina T. La biblioteca como soporte imprescindible para la investigación en la Universidad Carlos III de Madrid. A Vitoria-Gasteiz: Artium, Centro-Museo Vasco de Arte Contemporáneo; 2020 [citad 22 març 2021]. Disponible a: <https://www.youtube.com/watch?v=0Doa6ErplXU&list=PLt9md-9a8f65HCCyVT4kJ3ok3Bwve8EWu&index=8>

2.3 Situació de plataformes digitals de recerca i difusió del patrimoni musical a Catalunya

Aquestes situacions es repliquen en el conjunt de plataformes que posen a disposició el patrimoni musical dipositat a les unitats d'informació catalanes. Segons el Mapa de patrimoni musical desenvolupat pel *Centro de Documentación de Música y Danza* estatal al territori català consten 116 centres amb fons musicals⁹ amb un alt grau de casuístiques en les dimensions i capacitat logística per a emprendre projectes digitals i obrir els seus inventaris en línia.

En analitzar en detall la tipologia documental que els centres han publicat en accés lliure al centre (Figura 1), s'observa com la característica més habitual és posar en disposició de l'usuari, un inventari del registres disponibles per a consulta, l'acompanyament d'una imatge digital, i en menor mesura, l'ús d'àudio, vídeo o documents nadius digitals (ex. midi, musicXML).

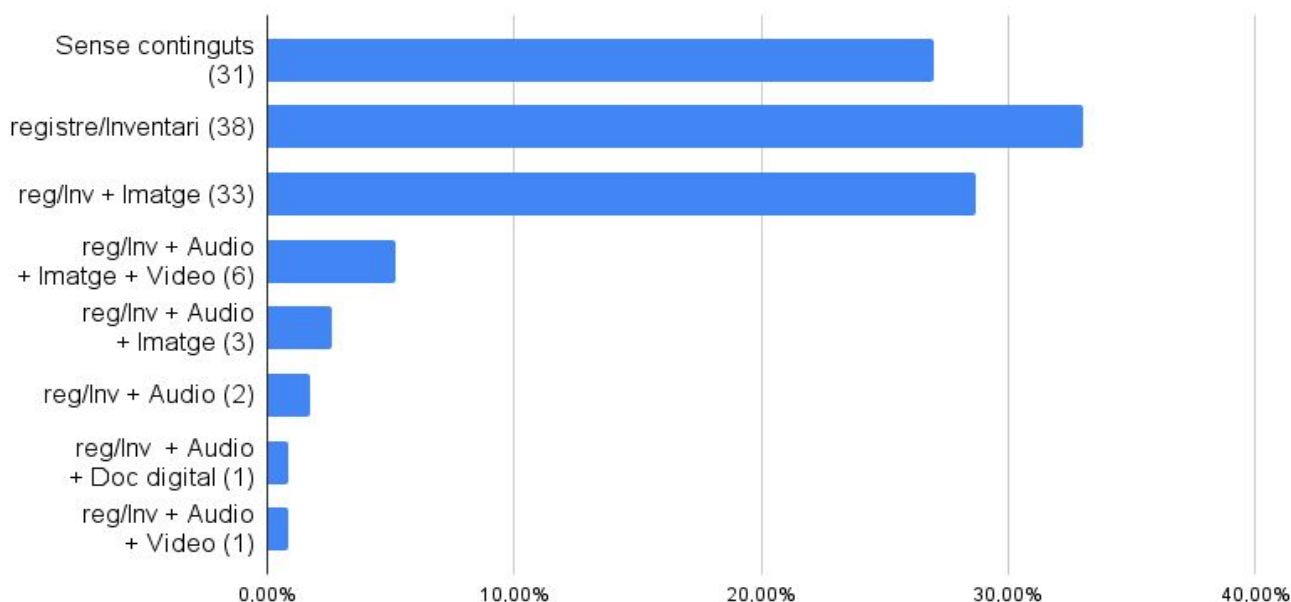


Figura 1: Tipologies documental a plataformes digitals llistades per Mapa de patrimoni digital español

Comparant el llistat de centres que custodien patrimoni musical del *Mapa* amb el que proporciona l'*Inventari de Fons Musicals de Catalunya (IFMUC)*, la plataforma *Patrimoni Musical Català* i el catàleg d'abast internacional *RISM* s'observa també la dispersió que es produeix en intentar accedir a la documentació dels centres des d'un únic repositori temàtic (Figura 2), incidint en els mateixos problemes que en el casos exposats anteriorment.

⁹ Centro de Documentación de Música y Danza. Mapa del patrimonio musical en España · Centro de Documentación de Música y Danza [Internet]. 2020 [citado 3 setembre 2020]. Disponible a: <http://cdmyd.mcu.es/mapatrimoniomusical/patrimoniomusical.php>

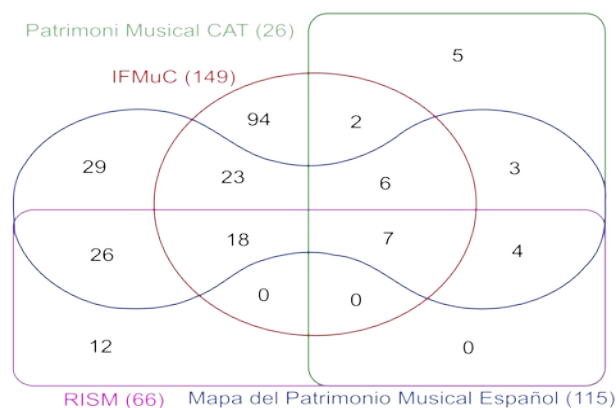


Figura 2: Coincidència de Centres llistats a Mapa del patrimoni musical, IFMUC, Patrimoni musical català i RISM

Sistematitzant les causes, molt similars a

Una de les raons principals per aquesta dispersió es basa en els circuits de difusió que hi ha en funcionament al territori. Per una banda, el circuit acadèmic (on podem englobar centres universitaris i la Biblioteca de Catalunya) foren els primers en introduir plataformes de difusió dels continguts com la *Memòria Digital de Catalunya* (basada en ContentDM) i el desenvolupament en Dspace, principalment (amb menció especial per l'*Inventari de Fons Musicals de Catalunya*). El circuit d'arxius públics, per la seva banda, va prioritzar l'ús d'un programari creat *ad-hoc* en Java per desenvolupar la plataforma *Arxius en línia* i la resta, un conjunt força heterogeni d'institucions locals, privades i eclesiàstiques, ha optat per integrar els continguts de l'arxiu en un apartat de la web institucional creada en CMS d'ús generalista (Wordpress, Drupal), tot i que cal destacar en aquest grup els desenvolupaments punters de l'INSPAI, *Centre de la Imatge de la Diputació de Girona*¹⁰ i la gran labor realitzada en el Cançoner virtual i interactiu Prodiemus¹¹.

La tria de CMS d'ús generalista en aquest últim grup no és baladí. Seguint els preus públics de 2022 publicats a la web del Consorci de Serveis Universitaris de Catalunya (CSUC), el cost mínim per al desenvolupament d'un repositori d'altres prestacions és 18.076,80€ d'alta i 1.300,83€ mensual per al seu manteniment que pot variar segons el volum de dades i la personalització del servei.

Moltes de les institucions no tenen establert en el pressupost una partida estable amb aquest import per a poder dedicar a la creació i manteniment d'una plataforma virtual adequada a la presentació de patrimoni documental, no tenen personal tècnic amb dedicació completa i externalitzen el desenvolupament sense tenir

10 INSPAI, Centre de la Imatge, Diputació de Girona, 2022. [Internet]. Disponible a: <https://www.inspai.cat/Inspai/ca/centre-imatge-diputacio-de-girona>

11 Orts Alís M. Cançoner virtual i interactiu Prodiemus. Prodiemus, 2005 [Internet]. Disponible a: <http://www.prodiemus.com/canconer/>

en compte l'accés digital a les col·leccions. En aquests casos, la plataforma queda ancorada fins al proper cicle de renovació de la infraestructura, situació que pot comportar anys d'estancament i pot seguir sense tenir en compte la difusió del patrimoni als nivells dels dos altres grups.

2.4 Solucions modulars per a la interoperabilitat de plataformes digitals de recerca i difusió del patrimoni musical

2.4.1 Plataformes JAMStack

Seguint la descripció que planteja Lavoie (2020) per a la planificació de projectes digitals en el marc GLAM¹² es pot observar com aquesta necessitat d'innovació en els sistemes d'accés a les col·leccions i la sostenibilitat econòmica de cada projecte es veuen cada cop més condicionats per la infraestructura tècnica que suporta la instal·lació, i com una mala planificació en aquest sentit pot obstruir nous desenvolupament. Si aquestes estructures no estan preparades per fer accessibles les col·leccions en un flux de treball modular, escalable i amb perspectives de suportar noves actualitzacions en un futur, els costos per plantejar la logística necessària per a la millora/migració a un nou desenvolupament poden restringir (en excés) la inversió econòmica per a fer-ho possible.

Aquest plantejament actualitza els plantejaments més generalistes per a la posada en marxa de nous serveis documentals (Abadal, 2004) i condicionen tant la anàlisi interna com els aspectes amb què aquestes variables interactuen amb el disseny i els objectius inicials que es plantegen renovar. El seu èxit, a més, s'atribueix a la col·laboració entre aquells agents amb objectius comuns que en la mera interoperabilitat entre els seus sistemes (Lavoie, 2020).

Per evitar aquests tipus de situacions molts desenvolupaments recents trenquen amb el model clàssic que incorpora el *frontend* i el *backend* en el mateix conjunt per explorar models més modulars, fluxos de treball basats en capes dedicades a funcions concretes (preservació, serveis de descoberta, etc.) que es comuniquen entre sí a partir del *backend* i que són gestionats de manera autònoma per minimitzar el risc de comprometre la resta de serveis amb que conviu (Figura 3). Aquests models parteixen d'un sistema heretat, sovint ja parametrizat cap a estructures OAIS¹³ que dirigeixen esforços a la millora dels sistemes de descoberta i la integració i normalització de l'accés als materials dipositats al DAM¹⁴ provinents d'esquemes de descripció sense una base comuna, amb la perspectiva de ser modelats semànticament i vocació integradora (Figura 4) (Gall, 2021).

12 Acrònim anglòfon de *Galleries Libraries, Archives, Museums*

13 Acrònim de *Open Archival Information System*

14 Acrònim de *Digital Asset Manager*

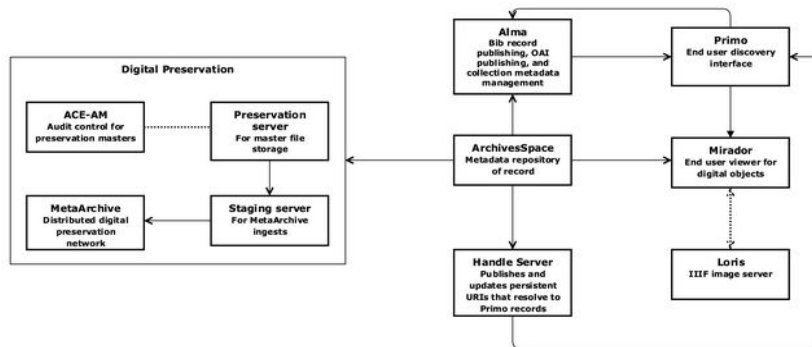


Figura 3: Diagrama de flux de treball a Mayo et al. (2019)

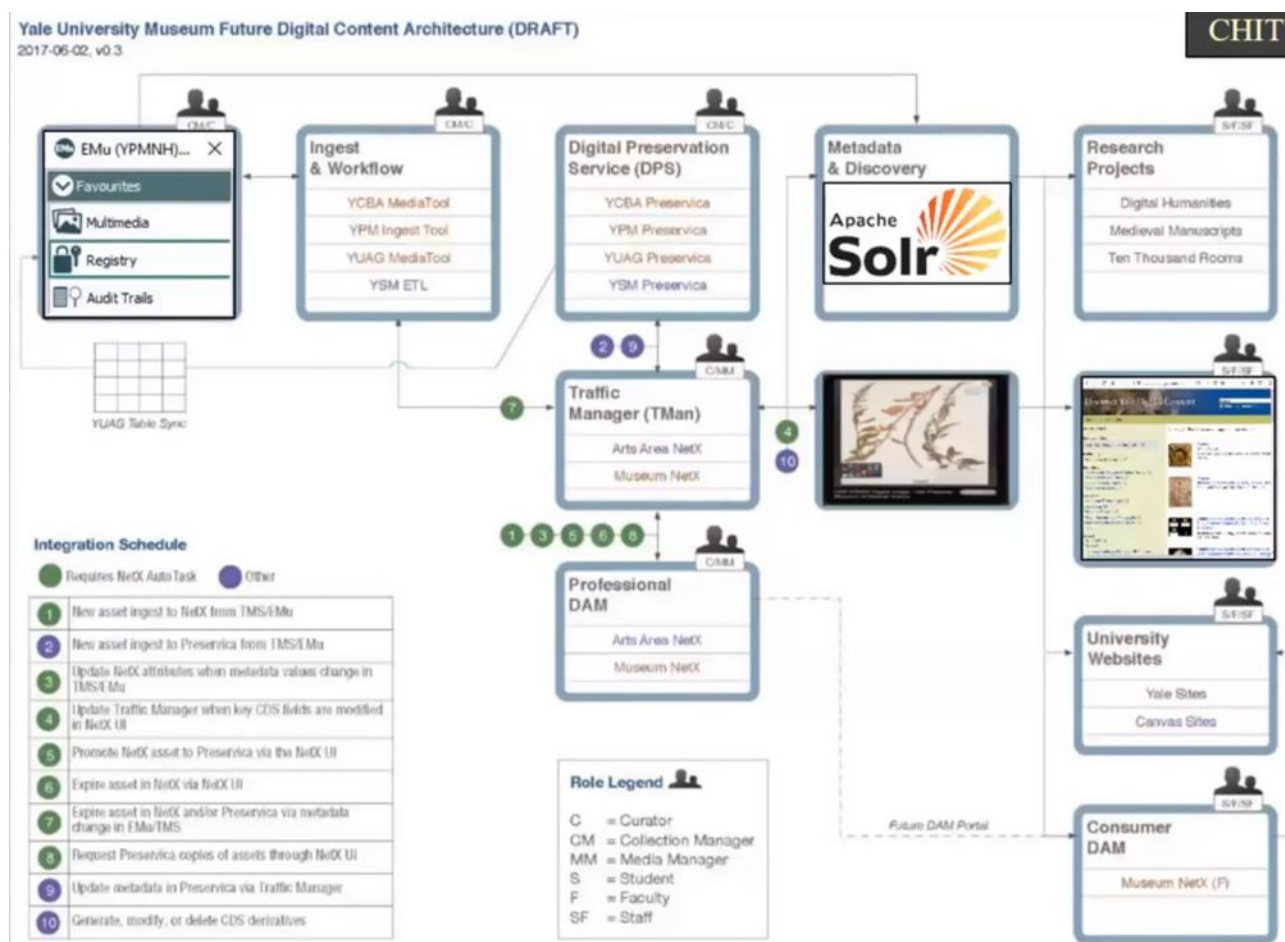


Figura 4: Diagrama de flux per a integració de recursos digitals a Yale University Museum (Gall, 2021)

Una segona aproximació planteja el retorn al desenvolupament de webs estàtiques, a partir de desenvolupaments basats en JAMstack¹⁵, entorns de programació àgil que combinen un seguit de codi javascript, APIs i microserveis per a la generació de plataformes digitals sense la necessitat d'utilitzar un servidor (Tramullas, 2020).

15 Acrònim de Javascript, Apis i Markup language

L'argumentació que es fa en aquesta proposta incideix en que el seu ús simplifica, i per tant accelera, la posada en funcionament de la interfície que sense la necessitat de desplegar plataformes més complexes (Mayo et al, 2019; Babcock et al. 2020). A partir de les altes capacitats per modelar funcionalitats del codi HTML i CSS es permet arribar als estàndards de presentació en sistemes CMS dissenyats per aquest tipologia documental (Omeka, Dspace, Drupal) sense la necessitat de configurar sistemes complexos d'administració i bases de dades associades al servidor, amb el corresponent descens de riscos en qüestions de seguretat d'accés i a la detecció/actualització de vulnerabilitats en els seus components o *plugins* (Newson, 2017); la càrrega de contingut s'executa des del client web, alliberant i accelerant la càrrega de treball al servidor (Figura 5); i, finalment, també fa més senzilla la preservació digital de tot el conjunt.

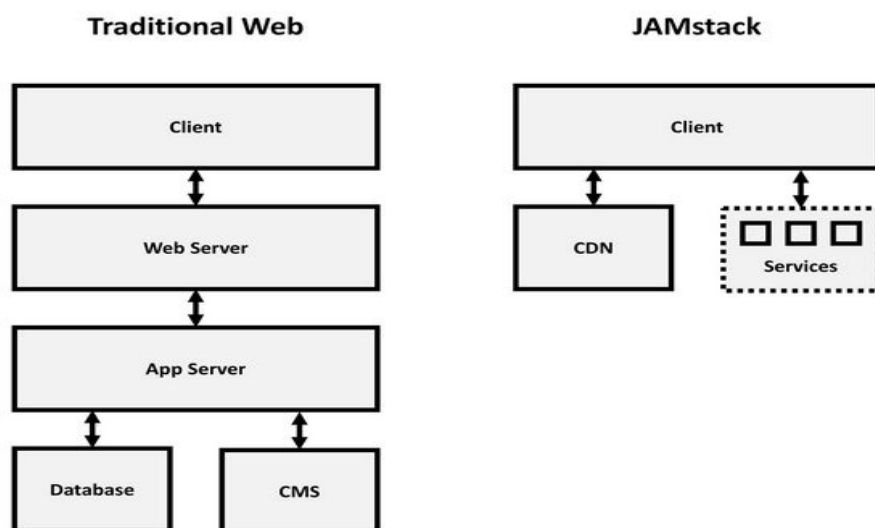


Figura 5: Arquitectures basades en CMS i JAMstack (Tramullas, 2020)

Els mèrits que argumenta Lavoie (2020) per a la acció coordinada en el desenvolupament d'infraestructures, ja sigui des d'aquest sistema de publicació com en propostes més globals és, també, corretja de transmissió per adreçar les experiències d'interacció i descripció dels materials d'estudi destinats a la recerca, sempre des de la integració i la perspectiva modular.

2.4.2 International Image Interoperability Framework (IIIF)

L'*International Image Interoperability Framework* (IIIF) és un conjunt d'estàndards oberts per a la creació de interfícies de programació d'aplicacions (API) enfocades a la millora de de la interacció i la interoperabilitat

a repositoris digitals amb contingut multimèdia (DAM), en especial atenció a aquelles dedicades a la gestió d'imatges i la interoperabilitat entre sistemes d'informació.

La iniciativa, nascuda l'any 2011, sorgeix de la necessitat de poder establir unes directrius tècniques que poguessin resoldre la falta de consistència en el tractament documental dels objectes digitals (nomenclatura dels fitxers, disposició d'objectes compostos), les particularitats de la gestió de drets, i la necessitat d'implementar avenços tècnics cada vegada més costosos per a la millora de la qualitat en la seva visualització. Les directrius fixen els pilars bàsics per formar una estructura modular capaç de definir un estàndard de metadades per a la presentació d'imatges d'alta qualitat, proveir d'un nucli d'APIS per accedir a les imatges i la seva informació bibliogràfica associada i establir possibles jerarquies en objectes compostos, sense plantejar el desenvolupament tècnic per a fer-ho possible des de zero i minimitzar els costos per a poder implantar-ho. (Sanderson, 2015)

L'entrada en aquest ecosistema de treball, a més, planteja el benefici de poder compartir les col·leccions digitalitzades de manera més horitzontal amb altres institucions, evitant la dependència de grans agregadors o els efectes *silo* mencionats en apartats anteriors.

Funcionalitats destacades:

- Provisió d'imatges d'alta resolució de manera ràpida i eficient a les col·leccions digitals
- Alta capacitat d'interacció i manipulació d'imatges
- Compatibilitat nativa amb *Web annotation Data Model* de W3C
- Interoperabilitat amb recursos disponibles en altres sistemes, opcions per comparar o interactuar amb altres col·leccions des del mateix visor
- Suport a la integració de *JSON-LD*, dades enllaçades i protocols W3c com *web annotation* per facilitar la creació de col·leccions digitals
- Especial motivació per a l'ús de URIs persistents; agilitzant la redistribució de col·leccions existents, migracions i desenvolupaments al *front-end* sense la necessitat de reformular les estructures internes del DAM

La documentació tècnica especifica amb més detall les funcionalitats de les quatre APIS desenvolupades que interactuen amb el servidor que gestiona les comandes a partir del *manifest*, un fitxer amb estructura *JSON* on es configura cada paràmetre que es desitja aplicar (Figura 6):

- *Image API* desenvolupa un servei web que retorna una imatge en resposta a les peticions d'un client. La URI pot especificar la regió, mida, rotació, qualitat i altres dades tècniques de la imatge

- *Presentation API* permet estipular la estructura i les característiques descriptives d'una col·lecció d'imatges digitals (extensible a vídeo, so i altres formats digitals) a través d'un manifest en format JSON. D'aquesta manera l'usuari pot navegar de manera seqüencial o jeràrquicament a través dels objectes, visualitzar informació relacionada amb el recurs o la col·lecció i poder incorporar nova informació (anotacions) en el visor a partir d'aquest protocol
- *Content Search API* permet la cerca a aquell contingut present a les imatges a la API de presentació. El protocol permet interactuar amb text generat per OCR o anotacions inserides i facilita la navegació per l'estructura jeràrquica dels objectes depositats o la col·lecció
- *Authentication API* permet la integració d'objectes dipositats en dominis amb diferents controls d'accés i permet que no interfereixin en la entrega. El servei disposa en el manifest la URI correcta per poder accedir al recurs sense interferir en la *cache* del navegador on s'executa el visor IIIF, independentment del sistema d'accés al repositori on s'allotja

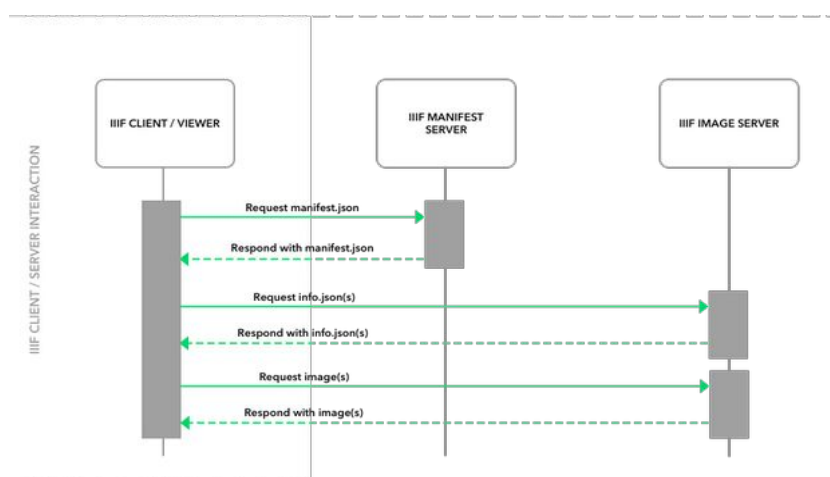


Figura 6: Interacció d'APIs a IIIF (Ronallo, 2018)

2.4.3 Music Encoding Initiative

L'aparició de la iniciativa *Music Encoding Initiative* també presenta una interessant aproximació a la gestió de col·leccions musicals des d'aquest punt de vista i suposa una iniciativa més completa a la transcripció digital de música notada en aquests moments en permetre fer accessibles la descripció bibliogràfica del fons, codificar paràmetres analítics necessaris per a la seva interpretació i fusionar-ho tot en un mateix paquet que permet reproduir la peça musical a partir del mateix fitxer.

El format està especialment dissenyat per actuar en un ventall d'escenaris molt ampli a partir d'una estructura interna modulable basada en XML que dona cobertura a la codificació dels aspectes lògics,

gestuals, visuals i analítics de la notació musical, en casuístiques que es van reforçant en cada actualització del model conceptual (actualment 4.0). El seu ús permet i predisposa a les institucions que en fan ús a ser més efectius en el rang i l'abast de la recerca que es desenvolupa a partir dels fons musicals que custodien i és, actualment, un dels formats prioritaris per a l'adquisició i preservació d'aquest tipus de fons en grans institucions (*Library of Congress, 2019*).

Principals característiques:

- Suport per a notació musical occidental, mensural i neuma
- Estructura interna XML modulable , disposició gràfica en SVG
- Suport per a la codificació d'anotacions crítiques/editorial
- Coordinació de la notació musical amb àudio i imatge digital
- Adaptació a formats i convencions normatives internacionals (FRBR, METS, TEI)
- Capacitació per a l'enriquiment a partir de fons *Linked Open Data* (LOD)

La seva orientació pràctica deriva de les iniciatives TEI i de la necessitat de poder establir un estàndard *de facto* per a l'intercanvi efectiu d'informació en una comunitat de recerca que ha vist desenvolupaments en aquest sentit massa orientats a un projecte en concret, la preservació de la informació i de la interfície de consulta i afavorir la creació de projectes col·laboratius en la disciplina que minimitzin els problemes relatius a la obsolescència per mitjà de dades codificades independents del seu substrat tecnològic (*platform-agnostic*) (*TEI consortium, 2017; Merchant Sánchez-Jara, 2019*)

En totes les facetes de representació del model (taula 1), el format adopta una estructura modular que li permet adaptar la exhaustivitat i la especificitat de cada element descrit segons les necessitats de cada centre.

Taula 1: Facetes conceptuais de l'esquema Music Encoding Initiative (*Merchant Sánchez-Jara, 2019*)

FACETES DEL MODEL MEI					
Sistemes de notació	Aspectes bibliogràfics	Aspectes editorials	Elements estructurals del sistema	Reproduccions digitals (material gràfic i multimèdia)	Lògica musical
Mòduls implementats a l'esquema					
Notació musical comuna	FRBR	Aspectes editorials	Elements compartits, models i atributs	Facsimils	Harmonia
Notació mensura	Noms i dates	Marcatge editorial	Text vocal	Figures i taules	Informació analítica
Notació neuma	<i>Corpora</i> musical	Textos	Enllaços i referències	Interpretacions	
Ornamentació en notació musical comuna		Símbols definits per l'usuari		<i>Musical Instruments Digital Interface</i> (MIDI)	
Notació en					

tabulatura					
------------	--	--	--	--	--

D'aquesta manera, l'arquitectura de l'esquema permet situar a l'encapçalament del fitxer els elements descriptius de l'obra (autor, data, edició..) com també informació analítica que pot ser d'interès en funció de la seva posterior recuperació com l'incipit, la tonalitat i el tempo.

El segon element destacat per a la canònica formació del fitxer és l'element <music> on es detallen les dimensions estructurals de l'obra i les seves jerarquies (si s'escau), la representació de la notació musical i les particularitats de la seva transcripció digital per a la seva visualització i reproducció sonora: sigui en la seva dimensió visual (variacions textuais, anotacions, marques); dimensió lògica (explicitació de distincions harmòniques, rítmiques o melòdiques); o a nivell analític o gestual.

2.4.4 Exemples d'aplicació

Burns antiphone¹⁶

Burns antiphone és un recurs interactiu en accés obert que permet la consulta del llibre litúrgic franciscà del mateix nom a partir de la disposició de capes dinàmiques per cercar, explorar i visualitzar el contingut a partir de les metadades descriptives, notació musical i textual i els incipits prèviament inventariats.

La formulació tècnica del projecte utilitza el visor diva.js (compatible amb IIIF), el servidor d'imatges IIF Loris, la plataforma JAMStack Jekyll, Verovio per a la visualització dels components musicals basats en MEI i el motor de cerca lunr. La plataforma disposa una presentació funcional en un menú de navegació superior amb pocs elements jeràrquics que prioritza la interfície d'accés al contingut envers la resta, orientada a explicar amb llenguatge planer i directe els objectius del recurs i a oferir enllaços a les tecnologies de base o al contacte amb l'equip investigador.

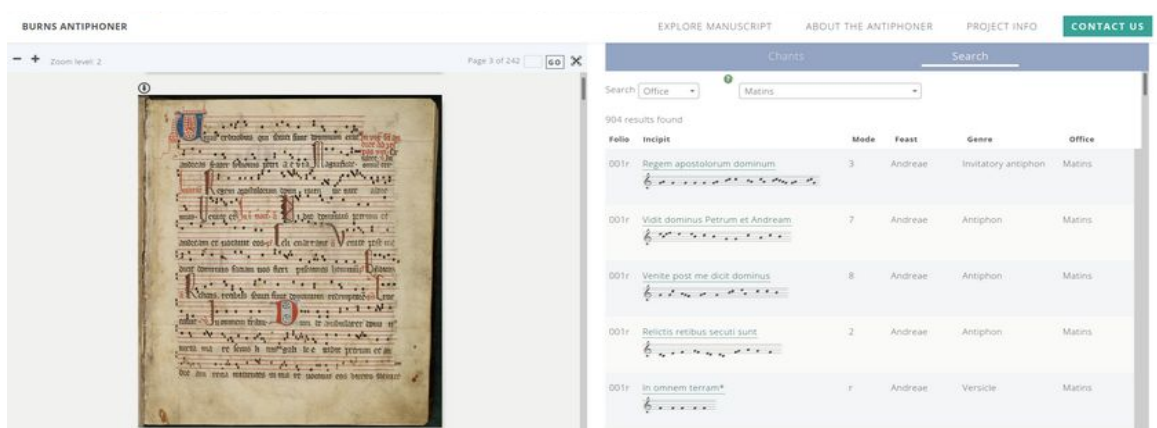


Figura 7: Detall de interfície de consulta a Burns antiphone

16 Noone M, Skinner G, Boston College Libraries. The Burns Antiphoner : Methodology. 2016 [citat 14 gener 2021]; Disponible a: <https://burnsantiphoner.bc.edu/>

Per la seva part, la interfície de consulta divideix la pàgina web en dos panells. un dedicat a la visualització de les imatges digitalitzades de la obra , l'altre dedicada a la presentació de les transcripcions del cos textual a la cerca i exploració de la obra a partir de les metadades codificades segons íncipit, gènere musical, celebració, ofici i tonalitat. (Figura 7).

Tasso in Music Project. Digital Edition of the Settings of Torquato Tasso's Poetry, c. 1570–1640¹⁷

El projecte Tasso és un projecte per a consulta en línia de la obra de Torquato Tasso de 2016 fruit de la col·laboració d'universitats i centres de recerca especialitzats en musicologia d'Estats Units i Europa coordinada per la University of Massachusetts Amherst i el Center for Computer Assisted Research in the Humanities de la universitat de Stanford . La formulació tècnica del projecte utilitza també la plataforma de generació de webs estàtiques Jekyll per gestionar el recurs, en combinació amb l'ús de plantilles de transformació de fitxers JSON basats en el programari ATON i Handlebar i l'ús de Verovio-Hundrum per la presentació de textos i partitures musicals.

Prima parte.

Canto
Alto
Tenore
Quinto
Basso

Ca - ro, a - mo - ro - so Ne - o, Ca - ro, a - mo - ro - so Ne - o, <Ca - ro, a - mo - ro - so

Ca - ro, a - mo - ro - so Ne - o, Ca - ro, a - mo - ro - so Ne - o, <Ca - ro, a - mo -

Ca - ro, a - mo - ro - so Ne - o,

Ca - ro, a - mo - ro - so Ne - o,

- PDF
- MIDI
- MP3
- Humdrum
- Humdrum (on Github)
- Humdrum (in VHV)
- MEI
- Musedata
- MusicXML (for Finale/MusicScore/Sibelius)
- Piano Roll

COMPOSER

Muzio Effrem (1549–aft. 1626)

Active areas (alphabetical) Bari, Naples

Figura 8: Detall de la plataforma de consulta a Tasso in Music Project

De la mateixa manera que el projecte anterior, el desenvolupament ofereix un estil funcional de navegació i la disposició concisa de la presentació del projecte, els elements de context i les polítiques editorials de transcripció i edició del projecte per prioritzar els esforços en la visualització de continguts i les opcions de cerca. Cal destacar l'esforç per concentrar en cada registre tota la informació relacionada amb la peça musical, la gran varietat de tipologies de fitxer en les opcions de descàrrega i les opcions de configuració del visor , que permeten delimitar la visualització de secció de la partitura o la cerca de text i elements concrets en la notació musical (Figura 8).

Minicomp Ed i Wax¹⁸

17 Ricciardi E, Sapp CS. Tasso in Music Project. Digital Edition of the Settings of Torquato Tasso's Poetry, c. 1570–1640. 2016 [citad 15 gener 2021]. Disponible a: <https://www.tassomusic.org/>

18 Gil A. Ed. A jekyll theme for minimal editions [Internet]. 2016 [citad 15 gener 2021]. Disponible a: <http://minicomp.github.io/ed/>

Finalment, cal fer una menció especial a les plantilles per Jekyll Ed. i WAX per a la creació de col·leccions digitals i textos literaris que es basen en els principis de computació mínima promoguts pel grup de treball del mateix nom associat a la Global Outlook::Digital Humanities de la Alliance of Digital Humanities Organisations (ADHO) ¹⁹.

Aquests desenvolupaments, especialment dissenyats per a la docència universitària durant el transcurs d'un semestre, permeten desplegar de manera eficaç i relativament senzilla projectes de recerca i difusió de continguts a partir de poques passes, fent un ús intensiu de plantilles específicament desenvolupades per la composició de textos i la presentació d'imatges sota els protocols IIIF, respectivament, i una serie de tutorials ben estructurats per a posar en marxa l'eina sense gaires complicacions.



The screenshot shows a web interface for 'El corrido mexicano'. At the top, the title 'El corrido mexicano' is displayed in a serif font, with the subtitle 'Encoded text and music for analysis, display, and playback' below it. There are two buttons: 'Back' and 'Click to play'. Below the buttons, the title 'Del peligro' is centered. The main content is a musical score for a corrido. The score is written on a single staff with a treble clef and a 3/8 time signature. The lyrics are: '¡Ma-dre mía de Gua-da-lu-pe é - cha-me tu ben-di - ción, yo - ya ma voy a la gue-rra, ya - viene la In-ter-ven-ción'. Below the score, the title 'Del peligro de la Intervención americana' is displayed, followed by the lyrics: '¡Madre mía Guadalupe, échame tu bendición,'.

Figura 9: Detall de la interfície de consulta a El Corrido Mexicano

Entre les característiques llistades destaquen el disseny *responsive*, la generació automàtica de taules de contingut (Ed) o registres a partir de fulls de càlcul (wax) i la incorporació d'eines de cerca simple i citacions bibliogràfiques al contingut (Ed i Wax) però en la exploració a repositoris d'accés obert es possible inspeccionar desenvolupaments que estenen les seves capacitats originals, orientant-les a materials no previstos inicialment.

En aquest sentit destaquem la edició digital anotada de la obra el *Lazarillo de Tormes*²⁰ generada pels alumnes del departament de cultures llatinoamericanes i ibèriques de la Universitat de Columbia, la

Nyröp M, Gil A. Wax: An extensible workflow for creating minimal scholarly exhibitions with Jekyll and Rake [Internet]. 2019 [citad 15 gener 2021]. Disponible a: <https://minicomp.github.io/wax/>

19 Els textos que originen aquesta metodologia de treball es poden consultar a la web del grup de treball del mateix nom. GO:DH Minimal Computing group. Thought Pieces [Internet]. Minimal computing: a working group of GO::DH. 2021 [citad 14 gener 2021]. Disponible a: <https://go-dh.github.io/minicomp/thoughts/>

20 Equipo Mini Lazarillo. *Mini Lazarillo: Edición mínima del Lazarillo de Tormes*. Columbia University, 2016. Web. <https://minilazarillo.github.io/>

col·lecció digital de gravats originals pertanyents a la revista de moda *Journal des dames et des modes*²¹, de la biblioteca Morgan, que pertany també a la Universitat de Columbia o, d'acord amb les intencions d'aquest projecte, la variació de Mark Saccomano per a la publicació d'una col·lecció de *corridos* mexicans²², que incorpora la visualització i reproducció de les partitures musicals de cada composició a partir de la conversió al format MEI de les partitures originals (Figura 9).

3. Objectius del projecte

Objectiu de projecte:

Creació de prototip de plataforma de difusió multimèdia amb la integració de visors d'imatge IIIF i notació musical en format MEI

Subjectius:

- *Disseny i implementació de la seqüència de digitalització de la obra*
- *Disseny i implementació de la transformació de formats digitals per a la lectura/reproducció de la obra*
- *Disseny funcional de la plataforma web estàtica en arquitectura JAMstack*
- *Integració de visor d'imatges segons IIIF*
- *Integració de visor de partitures musicals en format MEI*
- *Avaluació de paràmetres correspostos al model conceptual*

El projecte de digitalització del *Cancionero musical popular español* pretén integrar en la planificació del recurs aquelles tecnologies de reproducció multimèdia que puguin facilitar al màxim la difusió dels textos i la música notada de l'obra i, a la vegada, donar peu a facilitar noves vies en la recerca musicològica en els cançoners populars de la península en un disseny capaç de poder integrar-se amb serveis o plataformes digitals amb objectius similars.

La intenció de l'estudi és formular la possibilitat d'establir un sistema de visualització web amb les mínimes capacitats funcionals per a la creació del projecte que pugui ser prototipat i avaluar-ne els resultats a partir de les següents funcionalitats bàsiques:

- Infraestructura tècnica mínima per al desenvolupament d'una plataforma web d'accés al contingut
- Allotjament, transformació i visualització dels materials originals en múltiples formats

21 The Morgan Library. Style Revolution. Columbia University Libraries, Disponible a: <https://stylerevolution.github.io/>

22 Saccomano M. El corrido mexicano. [citad 15 gener 2021]. Disponible a: <https://mss2221.github.io/corridosEd>

Per aquest motiu, el projecte intenta resoldre aquesta concepció de disseny a partir dels estàndards *IIIF* i *MEI* i la creació d'una plataforma web estàtica per a la seva consulta basada en la plataforma *Jekyll*. L'estudi es centrarà, primer, en l'anàlisi de les particularitats de la obra i les opcions de disseny necessàries per a implementar un símil digital; la seqüenciació de les etapes necessàries per a poder convertir la edició en paper a els formats idonis per a la plataforma, estructurant els mecanismes en un ordre sistematitzat que argumenta els criteris utilitzats per la transformació i control de qualitat per, finalment, dur a terme un prototipatge de la plataforma de consulta i considerar la seva validesa/viabilitat per al projecte final.

4. Consideracions metodològiques, tècniques i logístiques per a la creació del projecte

4.1 Anàlisi de la estructura formal de l'obra

Felip Pedrell (1841-1922) fou un dels actors més destacats en la obertura del període modernista musical a l'estranger i a la seva renovació, apuntant les tendències de futur que molts dels seus il·lustres deixebles portaran després a la pràctica. Gràcies a la seva extensa obra i al convenciment en la necessitat de divulgar i formar culturalment aquella Catalunya de principi de segle XX --marcadament influïda per les teories romàntiques i nacionalistes de l'època-- ara som capaços de reconstruir amb tot luxe de detalls una etapa prolífica i plena d'avenços en la recerca musicològica al territori a partir de la seva investigació el camp de la etnomusicologia des d'on seguir traçant nous camins a partir dels sòlids esforços en la recopilació de la tradició musical.

Aprofitant l'avinentesa del proper centenari de la seva mort l'any 2022 es dissenya un projecte per reivindicar la vàlua de la seva figura en el camp de la musicologia a la península que, a la vegada faciliti la consulta digital i la recerca aprofitant els desenvolupaments més actuals en la difusió de col·leccions digitals. Per aquest motiu, el projecte pretén fer accessible digitalment el contingut íntegre del *Cancionero musical popular español*.

L'obra, compilada en quatre volums publicats entre 1918 i 1922, i ja de domini públic, és un dels exponents més importants de la recerca en musicologia del principis de segle XX a l'estat i explora amb detall la variants de la tradició musical a la península, aportant extensa documentació i bibliografia de reforç en una edició on les dissertacions de l'estudi són aprofundides per extensos annexos que recullen les tonades exposades en l'estudi que aproximen des d'un rigor metodològic la melodia melismàtica al cànon de transcripció musical occidental.

La digitalització d'aquesta obra presenta un repte a múltiples nivells que la fan ser un gran candidat per als objectius del treball. La clara especialització de l'abast temàtic, eminentment musical, la combinació de

materials textuais en un mateix contenidor i l'ús de formats de reproducció musical i visualització d'imatges es combinen per a intentar formar un recurs que pugui millorar la seva difusió i impulsar la creació de nova recerca per aprofundir en les bases del cant popular a la península ibèrica i les seves connexions amb altres cultures del mediterrani.

La orientació del projecte intentarà recrear de manera fidel tots els elements descrits a continuació, però no contempla incloure addendes a la edició original:

- L'obra combina el text escrit amb el recull de cançons i tonades que originen la publicació en tres parts: *El canto popular en la vida domèstica (Vol. I) - El canto popular en la vida pública (Vol II) - El canto popular i la tècnica musical espanyola (Vol III i IV)*, que a la vegada es subdivideix en diverses seccions segons el tipus de tonada o el tipus de registre formal
- El text desenvolupa una anàlisi de cada branca temàtica de la obra i incorpora detalls sobre les composicions musicals que s'enllacen a partir de la numeració del cançoner
- El text també incorpora citacions bibliogràfiques a les notes a peu amb una seqüència *autor + títol + lloc de publicació + editor + pàgina* (o sense seqüència estructurada, alternativament)
- S'insereixen exemples de versos per il·lustrar la composició d'una tonada o la motivació temàtica d'un gènere en el desenvolupament de la part textual
- Cada volum incorpora sumaris i índexs onomàstics a inici o final de cada volum, indistintament
- El cançoner es divideix en les mateixes seccions del text escrit i són titulats amb la seqüència *títol de la composició + Menció de responsabilitat* precedits d'una seqüència numeral acumulativa en les 451 composicions que formen el repertori.
- S'observa que la secció lírica de part del cançoner es distribueix en la part superior o inferior del pentagrama, alternativament

El disseny, a la vegada, introdueix la possibilitat d'enllaçar materials derivats o de nova creació que ampliïn o explorin de manera crítica els continguts de la obra original en la plataforma de consulta i reconstruir la edició per incorporar aquests elements a una nova edició en futurs desenvolupaments, si es creu oportú.

4.2 Establiment de fases de treball i parametrització de processos

El projecte delimita en 4 fases les accions a dur a terme i les agrupa a partir de les similituds en el mètode de treball (Taula 2) .

Aquesta disposició, delimitada a efectes pràctics per a l'avaluació de cada fase segons el subobjectius marcats en seccions anteriors, s'organitza de manera seqüencial malgrat que les fases 2 i 3 (dedicades a la transformació de formats i la creació facsímil) poden ser iniciades de manera independent un cop finalitzat el procés de digitalització de la obra.

El diagrama de flux de treball (Figura 8) permet copsar de manera més intuïtiva aquesta disposició. En ell, a la vegada, s'indiquen aquells estadis de la seqüència designades a avaluar els resultats de cada transformació²³.

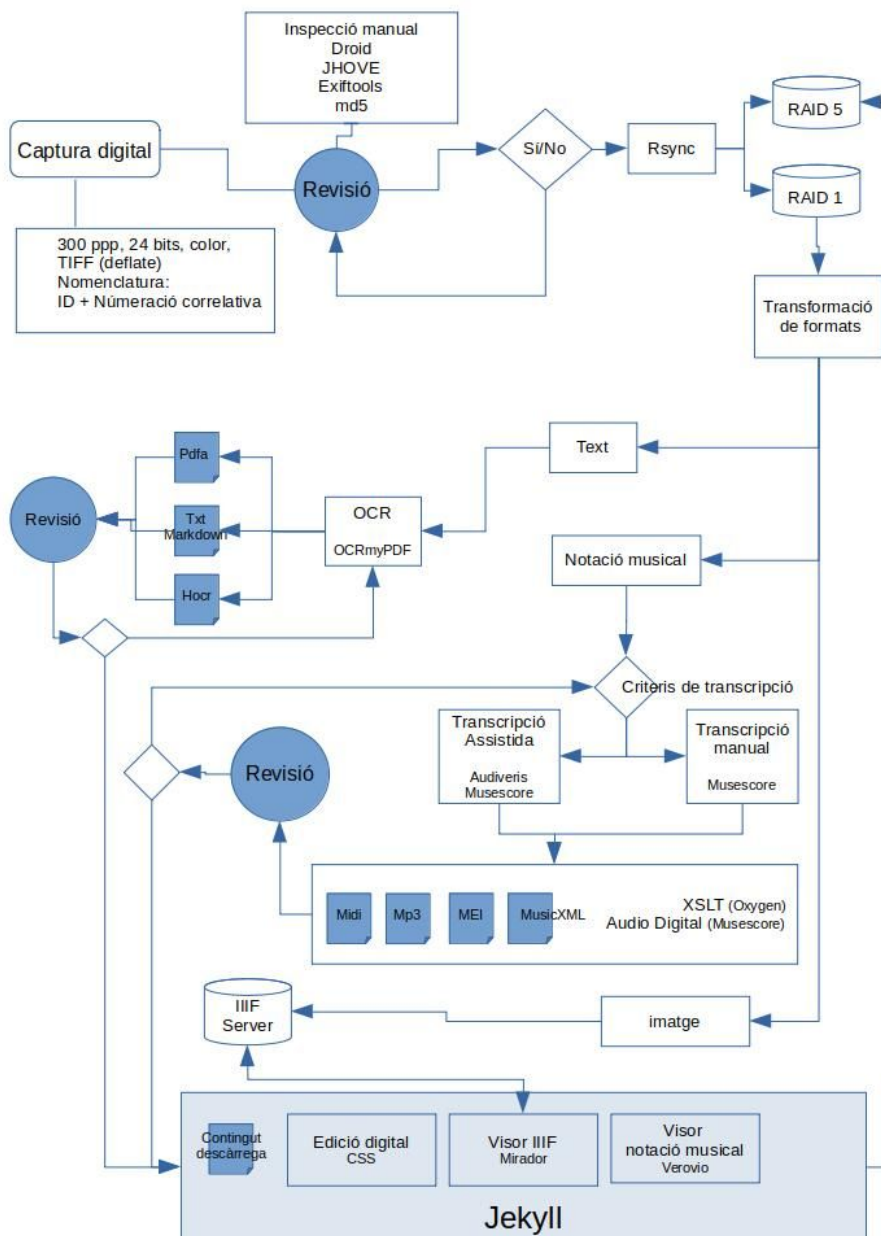
Taula 2: Establiment de tasques i eines seleccionades per al projecte

Fase 1 - Digitalització	
Tasques	<ul style="list-style-type: none"> • Revisió dels paràmetres de digitalització, calibratge i comprovació de l'estat de les obres • Digitalització dels documents • Revisió dels lots de digitalització segons mecanismes de verificació i redundància de dades • Incorporació de metadades descriptives a les imatges digitalitzades • Emmagatzematge en sistemes RAID 5 i RAID 1 durant el transcurs d'un any a partir de la finalització del projecte, extensible a tota la documentació generada per al projecte
Eines	<ul style="list-style-type: none"> • Escàner (especificacions per determinar) • ScanTailor Advanced https://github.com/4lex4/scantailor-advanced • DROID https://digital-preservation.github.io/droid/ • Jhove https://jhove.openpreservation.org/ • Exiftools https://exiftool.org/ • Rsync https://rsync.samba.org/
Fase 2 - Transformació de formats digitals	
Tasques	<ul style="list-style-type: none"> • Establiment de criteris de política editorial i transcripció de música notada • Transcripció automàtica de cos textual de la obra a formats digitals (TXT, hOCR, Markdown) • Transcripció automatitzada per sistema ORM de partitures impreses (MusicXML, MEI) • Transcripció manual de partitures impreses (melodia melismàtica) (MusicXML, MEI) • Revisió de textos i música notada • Transformació de formats a difusió (PDF, MIDI, MP3, WAV)
Eines	<ul style="list-style-type: none"> • Audiveris https://github.com/Audiveris/audiveris • OCRmyPDF https://ocrmypdf.readthedocs.io/en/latest/ • MuseScore https://musescore.org/en • OxygenXML Editor https://www.oxygenxml.com/ • Atom https://atom.io/ • Atom MEI Tools https://github.com/nCoda/mei-tools-atom • Exiftools https://exiftool.org/ • Rsync https://rsync.samba.org/
Fase 3 - Creació de la edició facsímil	
Tasques	<ul style="list-style-type: none"> • Creació de servidor d'imatges IIIF • Creació de manifest IIIF per a la consulta en línia de la obra • Revisió automàtica i manual de la estructura JSON i la correcta visualització de la obra

²³ El diagrama de flux a la Figura 10 és una versió simplificada del projecte. A l'apartat Implementació s'ofereix una versió més detallada de cada procés.

Eines	<ul style="list-style-type: none"> • Cantaloupe https://cantaloupe-project.github.io/ • Bodlean IIF manifest editor http://dmt.bodleian.ox.ac.uk/manifest-editor/ • OxygenXML Editor https://www.oxygenxml.com/ • Atom https://atom.io/ • Rsync https://rsync.samba.org/
Fase 4 - Disseny i prototipatge de la plataforma d'accés digital	
Tasques	<ul style="list-style-type: none"> • Disseny de la interfície de consulta • Implementació en plataforma • Incorporació de visors MEI i IIF • Revisió automàtica i manual de la estructura JSON
Eines	<ul style="list-style-type: none"> • Ed. http://minicomp.github.io/ed/ • Jekyll https://jekyllrb.com/ • Mirador https://projectmirador.org/ • Verovio https://www.verovio.org/ • Rsync https://rsync.samba.org/

Figura 10: Diagrama de flux del projecte



4.3 Criteris de selecció i altres consideracions logístiques

Els criteris de selecció d'eines llistades a la taula anterior es basen en experiències prèvies per a la realització de tasques similars, la facilitat d'instal·lació i/o utilització i en la gran majoria de casos, en la seva gratuïtat.

Per a la eines seleccionades en la fase 1 s'utilitzen aquell programari d'ús habitual per a la generació de perfils als fitxers de sortida a efectes de ser incorporats en rutines de preservació digital que no seran descrites en aquest treball, malgrat disposar en la planificació d'un servei d'emmagatzematge per al projecte. A la vegada, també s'ometen les característiques del dispositiu de captura malgrat haver aconseguit una còpia digital per a les proves a partir dels equips de digitalització de la Biblioteca de Catalunya i utilitzar les versions digitals de la obra disponibles a Internet Archive dipositades per la Boston Public Library²⁴.

La transcripció de textos per a la edició laboral ha prioritzat el programari OCRmyPDF per davant d'altres eines en oferir les sortides a fitxers txt, hOCR i Pdf-A, el motor OCR de codi obert Tesseract, la seva gratuïtat i la possibilitat de processament de lots a partir d'una interfície de línia d'ordres (CLI). Durant el procés d'avaluació es contempla també *Irfanview* per a la generació de PDFs en malgrat ser un desenvolupament únicament disponible per al sistema operatiu Windows²⁵

La selecció d'eines per a la transcripció de notació musical responen a la familiaritat adquirida en l'ús previ per al projecte *100 anys de cuplet en català: prova pilot de dinamització digital dels fons musicals de la Biblioteca de Catalunya*²⁶, d'on gran part dels objectius del projecte tenen origen.

Per a la creació de la edició facsimil segons el model IIF s'avaluen les característiques tècniques dels servidors i visors llistats a la web oficial del consorci²⁷. En la selecció inicial es valora la possibilitat d'utilitzar el programari *iiif-tiler*²⁸ per a la generació de tesselles i mantenir el servei de la Api d'imatge sense implementar un servidor dedicat. A partir de l'anàlisi de la documentació disponible s'escull el servidor basat en java Cantaloupe per la fàcil instal·lació i configuració de la eina i l'editor gràfic d'edició de manifest desenvolupat per la *Boodlean libraries* per assistir en la creació dels manifest IIF

24 Els 4 volums es poden trobar en obert al perfil de la institució a Internet Archive:
<https://archive.org/details/bostonpubliclibrary?tab=about>

25 Skiljan I. IrfanView [Internet]. [citad 16 gener 2021]. Disponible a: <https://www.irfanview.com/>

26 Valle-de-Lope I, Molina-González P, Vidal-Santos G. 100 anys de cuplet en català: prova pilot de dinamització digital dels fons musicals de la Biblioteca de Catalunya. En Col·legi Oficial de Bibliotecaris-Dokumentalistes de Catalunya; 2020 [citad 10 gener 2021]. Disponible a: <http://eprints.rclis.org/40784/>

27 International Image Interoperability Framework Consortium. Awesome International Image Interoperability Framework (IIF) [Internet]. International Image Interoperability Framework; 2020 [citad 16 gener 2021]. Disponible a: <https://github.com/IIF/awesome-iiif>

28 Robson G. *iiif-tiler* [Internet]. 2020 [citad 16 gener 2021]. Disponible a: <https://github.com/glenrobson/iiif-tiler>

Finalment, per a la creació de la plataforma de difusió es fa un de la implementació de Mark Saccomano de la plantilla Ed. per la seva similitud i adequació als objectius del projecte en la seva aplicació del visor Verovio i les llibreries MIDI per a la reproducció via MIDI de partitures MEI.

5. Implementació del projecte

5.1 Digitalització, transformació de formats i edició de continguts

5.1.1 Procès de digitalització

El procés de captura manté la línia de recomanacions del departament de cultura i les institucions professionals catalanes (Térmens, 2017) per aconseguir una imatge d'alta qualitat en 24 bits de color anàloga al estàndards de preservació de textos. Per aconseguir la versió final de la imatge s'utilitza l'aplicació ScanTailor per corregir detalls com la orientació i la disposició dels marges en la versió destinada a ser el màster digital que inicia el procés de transformació (Figura 11).

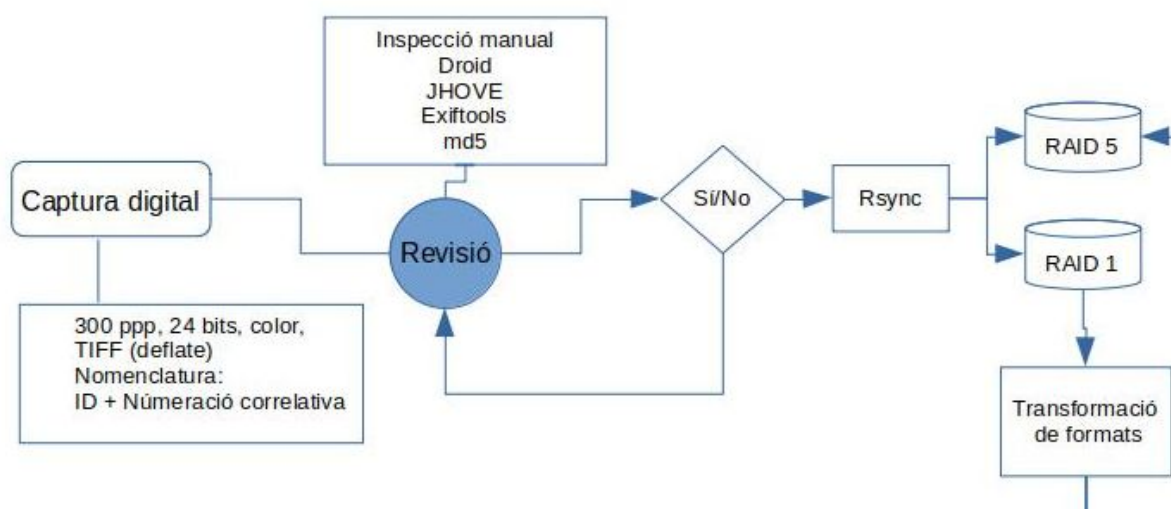


Figura 11: Diagrama de flux per a la digitalització de la obra

Aquesta imatge s'analitza de manera manual i per tres inspectors automàtics (*DROID*, *Jhove* i *Exiftools*) en previsió d'utilitzar els paràmetres detectats als informes com a perfil digital de la imatge en processos de preservació digital. Si no apareixen irregularitats en la imatge, s'envia a dos dispositius digitals (RAID-5 i RAID-1, el primer actuant com a còpia de preservació i el segon com a espai de treball) per mitjà del protocol *rsync* per aconseguir estabilitat en la migració de les dades.

La inserció de la bibliografia citada al text original es situa al final de cada capítol per no destorbar la lectura en *scroll* a partir d'enllaços interns HTML i la crida a la referència bibliogràfica albergada al fitxer BIBTEXT per mitjà del plugin Jekyll-scholar²⁹ de manera molt similar a la construccions emprades en la creació de llibres electrònics (Figura 14) .

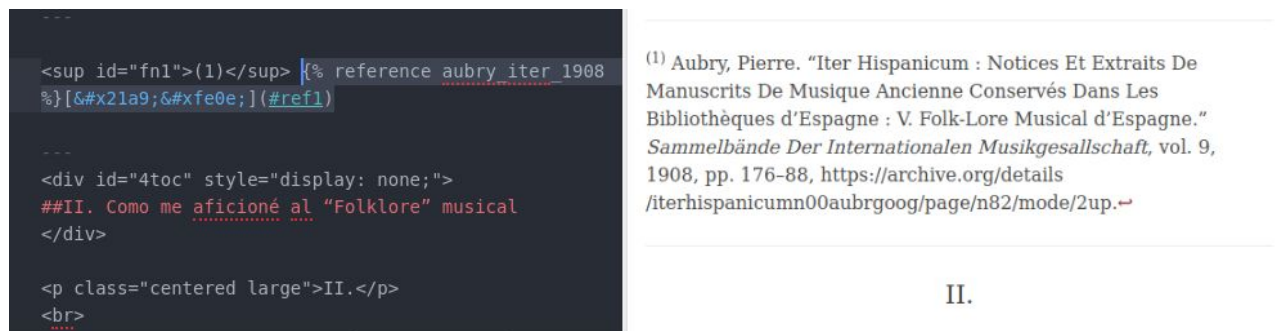


Figura 14: Inserció de referències bibliogràfiques i marcatge a sumari

5.1.3 Transcripció de música notada³⁰

El reconeixement òptic de música notada, en canvi, s'estableix primer en la anàlisi del corpus per establir quines partitures compleixen amb els cànons de notació occidental que poden ser utilitzades amb el motor OMR del programa Audiveris, quines són aproximacions melismàtiques que caldrà transcriure digitalment per satisfer les intencions del cant registrat i altres qüestions formals com la transcripció al cànon modern d'aquelles convencions de representació atribuïbles al context d'edició³¹, com l'exportació a un fitxer sonor per a ser comparat amb altres variacions territorials o temàtiques del cant popular. La selecció prèvia permet la construcció de lots i un flux de treball adient per a cada cas (Figura 15).

En el primer cas es crea un primer lot d'entrenament per el motor de reconeixement a les particularitats de la edició, prioritzant aquells accidents i elements que es consideren més rellevants en la detecció per poder ser incorporats al conjunt d'imatges digitalitzades i reconèixer de manera més eficaç les partitures.

29 Keil, Sylvester. Jekyll-Scholar [Internet]. Citat 24 maig 2021]. Disponible a <https://github.com/inukshuk/jekyll-scholar>

30 Exemples del procés de transcripció disponibles a Annex A, B I C

31 Aquesta fase replica els treballs realitzats al projecte *100 anys de cuplet a Catalunya*. (2020). Valle-de-Lope I, Molina-González P, Vidal-Santos G. 100 anys de cuplet en català: prova pilot de dinamització digital dels fons musicals de la Biblioteca de Catalunya. En Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya; 2020 [citat 10 gener 2021]. Disponible a: <http://eprints.rclis.org/40784/>

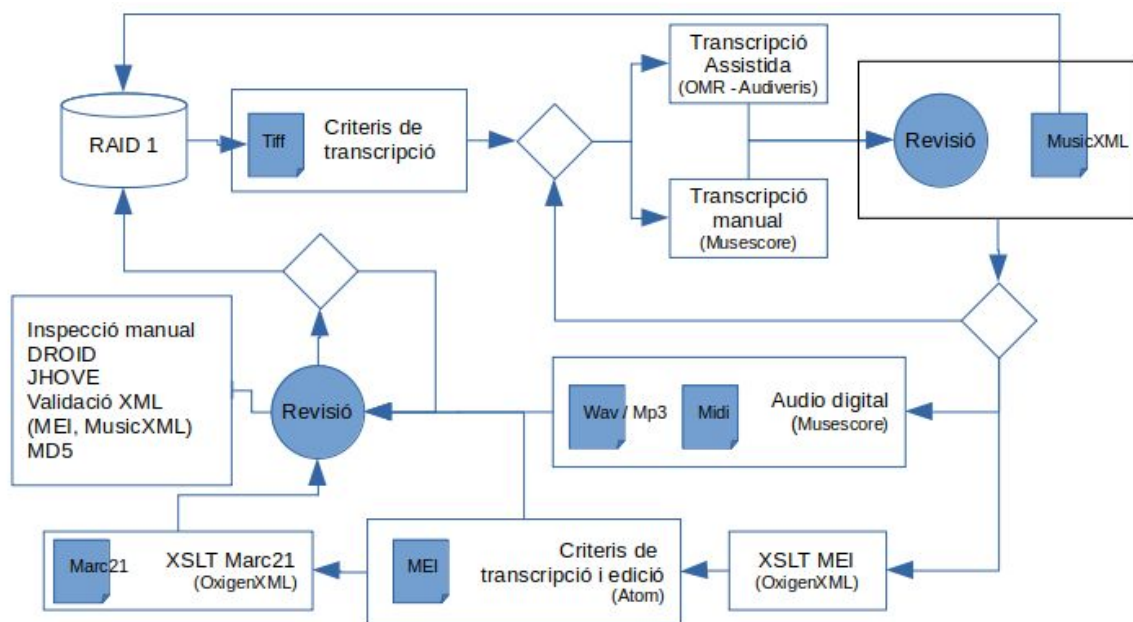


Figura 15: Diagrama de flux per a la transcripció de música notada

A diferència dels sistemes de reconeixement de caràcters (OCR), el fitxer de sortida MusicXML necessita una coherència interna que no se li exigeix necessàriament als fitxers de text generats a partir de la detecció de textos. Per aquest motiu el procés de revisió és imprescindible per establir la validesa i el rigor en la transcripció segons els criteris d'edició, qualsevol error de codificació o incoherència en el format XML d'una partitura digital pot comportar errors d'interpretació dels paràmetres musicals, lectures incorrectes als programes d'edició musicals, o fins i tot no permetre la seva càrrega al sistema d'edició.

El procés de reconeixement a Audiveris és capaç d'identificar jerarquies entre els elements de la partitura i establir-ne les relacions, però el sistema prioritza la detecció d'elements simples als compostos i aquelles relacions complexes que s'estableixen visualment en el pentagrama són les més propenses a contenir errors, (Figura 16). L'anotació manual per a la correcció d'aquests errors, realitat per mantenir la coherència dels fitxers MusicXML per exportar posteriorment a MEI, serveix també de base per a l'ampliació de corpus de glifs que s'utilitzen en la classificació i adaptar-les a l'estil tipogràfic de la obra.

Un cop corregides la selecció de partitures, es carreguen els glifs identificats a cada plana i es seleccionen aquells on la figura detectada per el motor de reconeixement té un índex més alt en el sistema de classificació. El procés permet una millora relativa en la detecció automàtica del procés però és especialment significativa en la identificació de formes i expressions associats a signes no presents anteriorment en el corpus original.

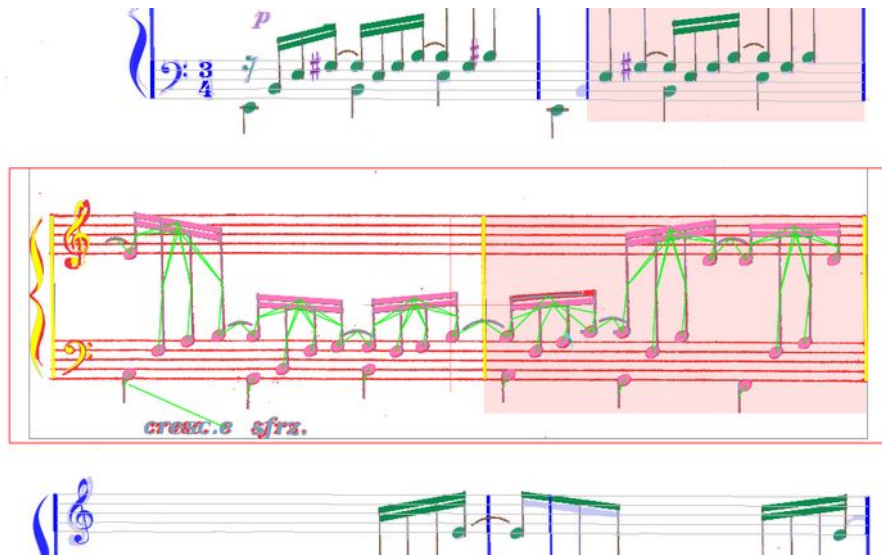


Figura 16: Detall del procés de reconeixement OMR a Audiveris

Per provar aquesta funcionalitat s'introdueixen les formes que indiquen la dinàmica *Sforzando* (*sfz*) en la interpretació presents a les partitures 310, 317 i 318 i es realitza el test en la núm. 385. Els signes en aquestes partitures diferent de la forma prevista en el corpus (*sfz.*) i en introduir els nous glifs s'aconsegueix que el sistema reconegui el nou glif en les noves fases de reconeixement (Figura 17).

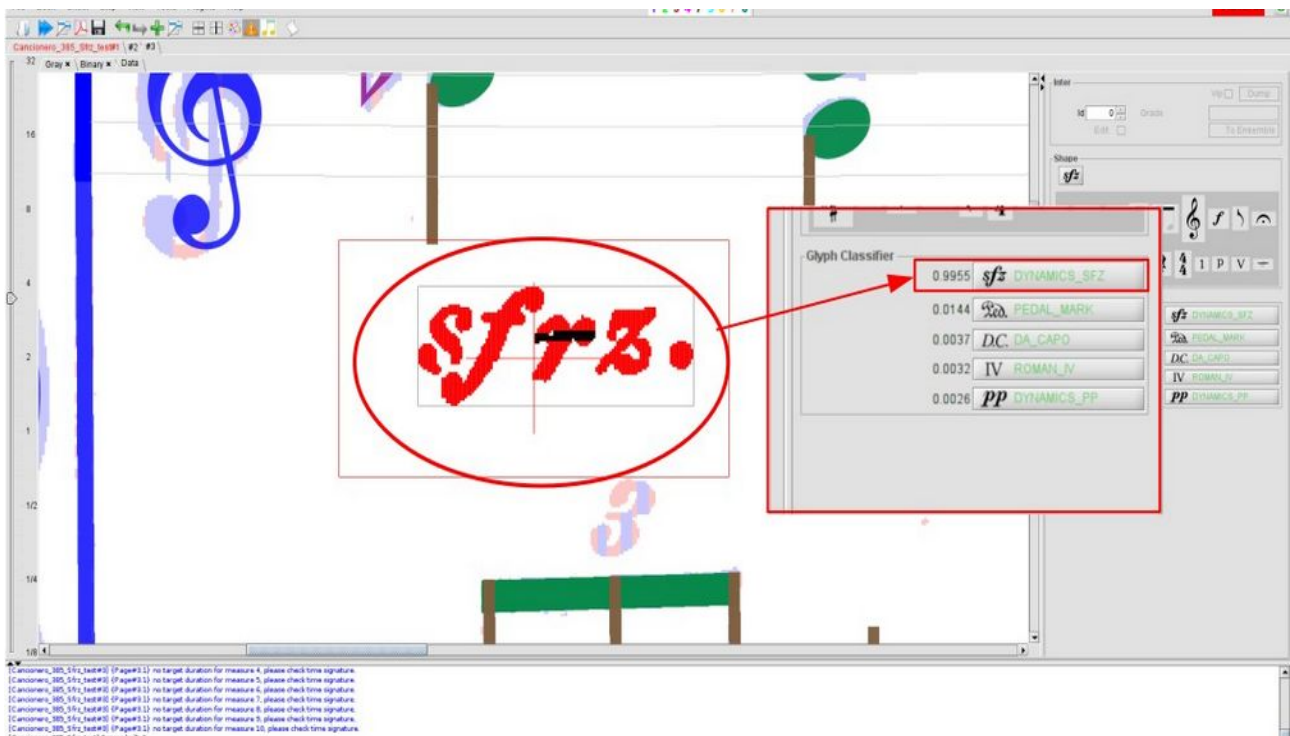


Figura 17: Exemple de millora en la classificació automàtica en ampliar el corpus de glifs

Aquest procés permet accelerar el treball amb la partitura digitat a l'editor MuseScore, on s'introdueixen les correccions que el sistema a Audiveris no permet modificar fins establir el fitxer MusicXML definitiu que serà la base de les transformacions en format sonor i la transformació via XLST a MEI.

El procés per a la transcripció manual d'aquelles partitures que no poden ser correctament transcrits a partir de la partitura original, (les aproximacions a la melodia melismàtica, per exemple) es realitza directament a l'editor MuseScore amb la dedicació de personal expert en musicologia, que prèviament, analitza les característiques de cada peça musical de la obra i s'encarrega de la revisió dels fitxers MusicXML i del control de versions amb els canvis necessaris a la versió original per a la correcta reproducció en la fase de transformació a altres formats (MEI, audio, MIDI).

Després de procés de revisió s'extreuen les versions d'àudio i MIDI i s'importen els fitxers MusicXML resultants a l'editor de text Atom on es realitza el procés de transformació a MEI a partir dels complements MEITools, que permeten definir i corregir la notació visual a partir del visor que previsualitza el codi XML (Figura 18).

Mentre el primer d'aquests processos es pot generar a partir de la eina d'edició musical MuseScore i l'ús de llibreries de so que incorpora, la transformació al format MEI és el moment on s'incorpora al fitxer la informació bibliogràfica que permetrà identificar la obra (Figura 19).

Figura 18: Exemple de la interfície de validació de MEI a Atom

The image shows a screenshot of the Atom text editor interface. The top half displays XML code for MEI (Music Encoding Initiative) format, including elements like <beam>, <note>, <slur>, <measure>, <staff>, and <layer>. The bottom half shows a musical score preview for 'Cançó de batre' (Drum Song). The score is written for Piano (Pno.) and includes lyrics in Catalan: 'D'ací es tant veig u - na estre... lla, - d'ací es - tant veig u na es - tre lla... d'ací es - tant la veig llu - ir... jo no se si es l'es... tre lla... jo no se si es l'es tre lla...'. The interface includes a sidebar on the left with 'Your project is currently empty' and a bottom toolbar with 'MEI Tools', 'Scale', 'Highlight color', and 'Page 6 of 44'.

Malgrat la potencia de la sintaxis per a la formació d'entitats i agregació d'informació analítica, en aquest test s'introdueixen les metadades bàsiques però no s'incorporen aquelles consideracions que cal establir en els

critèris d'edició per la transcripció de la obra o els procediments FRBR per generar els registres analítics de cada composició musical que permetrien codificar en format Marc21 i ser incorporats directament al catàleg bibliogràfic i, encara que el fitxer resultant pot ser parcialment incomplet, és un gran complement a efectes d'establir punts d'accés al material en línia i fer més permeable la connexió entre els materials primaris i les versions digitals que prenen com a punt de partida l'obra original.

```

<meiHead>
  <fileDesc>
    <titleStmt>
      <title>Cançó de batre</title>
    </titleStmt>
    <pubStmt>
      <publisher>Eduardo Castells</publisher>
      <pubPlace>Valls</pubPlace>
      <date isodate="1918">1918</date>
      <respStmt>
        <persName enddate="1922-08-19" startdate="1841-0-19" role="editor">Felip Pedrell</persName>
        <persName role="encoder">Irene Valle de Lope</persName>
      </respStmt>
    </pubStmt>
  </fileDesc>
  <encodingDesc xml:id="encodingdesc-0000000323472510">
    <applInfo xml:id="appinfo-0000000881334946">
      <application xml:id="application-0000000095964350" isodate="2020-07-31T20:36:37" version="3.0.0-dev-298ab95">
        <name xml:id="name-0000000396718436">Verovio</name>
        <p xml:id="p-0000000768184996">Transcoded from MusicXML</p>
      </application>
    </applInfo>
  </encodingDesc>
</meiHead>

```

Figura 19: Encapçalament de prova a MEI

5.2 Creació de la edició facsímil

La reproducció de les imatges digitals per a la consulta de l'obra en el seu format original s'origina en la migració de les imatges Tiff en alta resolució al servidor *Cantaloupe* per mitjà del protocol rsync (Figura 20).

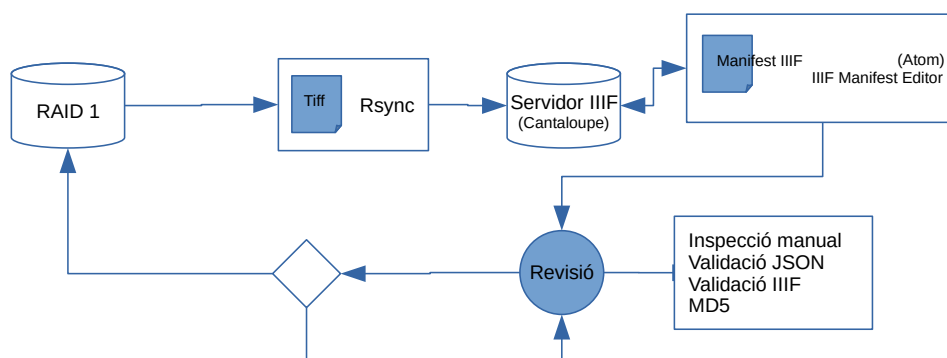


Figura 20: Diagrama de flux per a la creació de la edició facsímil

La configuració dels paràmetres de sortida faciliten l'escalat de la imatge original a 6 nivells basats en tesselles de 512 punts a 150 dpi que seran formatades pel servidor a partir de les ordres que es donen en la parametrització de la url d'accés, que per agilitzar la càrrega de treball en les demandes s'especificarà en

format jpg amb compressió de 20% en aquesta fase de prototipatge. (Figures 21 i 22). Cal tenir en compte que la versió escollida limita l'ús dels paràmetres a la versió 2 de les especificacions de la API.

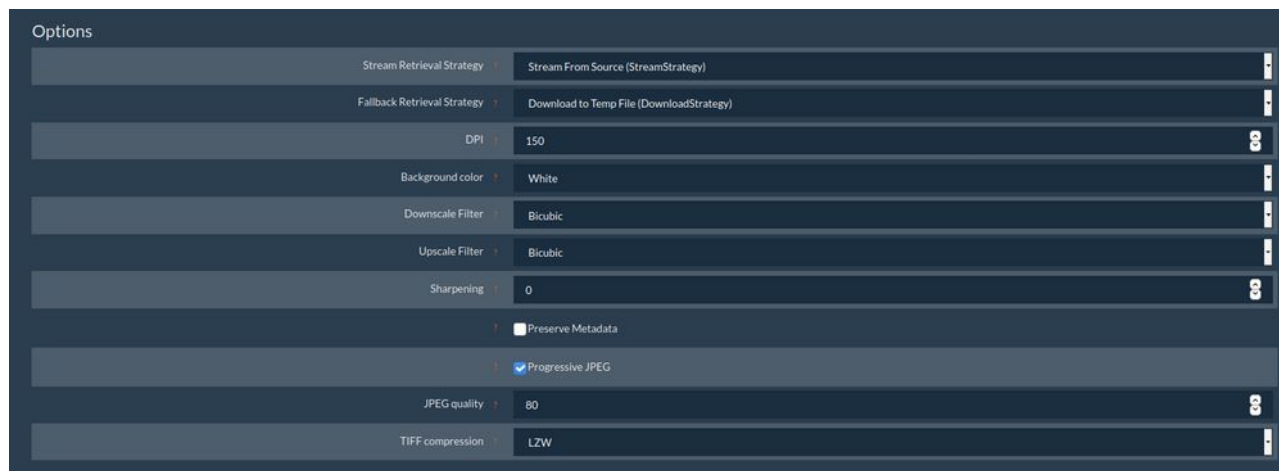


Figura 21: Paràmetres de configuració d'imatges de sortida a servidor IIF

La construcció del manifest de presentació IIF es realitza en tres fases. En primer lloc s'introdueix la seqüència de canvas i el conjunt de metadades descriptives de cada volum a partir de l'eina *IIF manifest editor*³² per la seva comoditat en la construcció del manifest a partir de la seva interfície gràfica. De la mateixa manera que en la formació dels fitxers de text i notació musical, s'incorporen les metadades descriptives bàsiques per testejar la viabilitat de la eina en el procés.

	IIF parameter	
http	{scheme}	http
127.0.0.1:8182	{server}	127.0.0.1:8182
iiif/2	{prefix}	iiif/2
DSC5042_1 L.tif	{identifier}	DSC5042_1 L.tif
full	{region}	800,780,850,1400
full	{size}	full
0	{rotation}	0
default	{quality}	default
jpg	{format}	jpg

Figura 22: Exemple de parametrització a Image API de

32 Bodleian Libraries. IIF manifest editor [citat 14 novembre 2020]. Disponible a: <https://github.com/bodleian/iiif-manifest-editor>

```

"@type": "sc:Manifest",
  "label": "Cancionero musical popular español",
  "metadata": [
    { "label": "Creator", "value": "Felip Pedrell" },
    { "label": "Title", "value": "Cancionero musical popular español" },
    { "label": "source", "value": "b18641799" },
    { "label": "bibliographicCitation", "value": "Pedrell F. Cancionero musical popular español. Valls: Eduardo
Castells; 1918. 4 p." }
  ],
  "description": [
    { "@value": "Test - Afegir descripció de la obra", "@language": "ca" },
    { "@value": "Edit description", "@language": "en" },
    { "value": "Test - Edició facsímil digital de la obra", "@language": "es" }
  ],
  "license": "https://creativecommons.org/licenses/by/3.0/",
  "attribution": "TEST",
  {...},
  "structures": [
    {...},
    {
      "@id": "http://127.0.0.1:4000/corridosEd/Manifests/range/r1",
      "@type": "sc:Range",
      "label": "Proemio I",
      "ranges": [ "http://127.0.0.1:4000/corridosEd/Manifests/range/r1-1" ],
      "canvases": [
        "http://4f1e2a77-c61b-412a-a305-701a4c99581c",
        "http://6851e326-0f27-42cb-878d-22fc0968045f",
        "http://843f99d8-8554-48ca-bae0-ca771b2110d7"
      ]
    },
    {...},
  ]
}

```

Figura 23: Detall de les seqüències de descripció i rangs a manifest IIIF

Finalment, s'introdueix a partir d'un editor de text Atom les seccions *ranges* i el text de cada imatge que permetran establir els apartats de la obra en una taula de continguts i la possibilitat de seleccionar el contingut de cada pàgina, sense èxit, la possibilitat de poder de-referenciar les urls que apunten directament a les transcripcions musicals per permetre l'accés directe al visor a partir de l'enllaç. En el primer cas la seqüència es codifica en el mateix manifest (Figura 23) i la incorporació del text original, es fa a partir del reconeixement de text de les imatges originals en la fase anterior, la conversió a ALTO-XML i una última transformació XSLT en un nou document que converteix el text i la situació espacial al protocol de context API de IIIF i és enllaçat al manifest original (Figura 24 i 25)

Figura 24: Detall de la incorporació de text al manifest IIIF

```

(Fitxer JSON extern)
{...}
{
  "@type": "oa:Annotation",
  "motivation": "sc:painting",
  "resource": {
    "@type": "cnt:ContentAsText",
    "format": "text/plain",
    "chars": "Las dos Canciones de llaurar y de batre, que se insertan"
  },
  "on": "https://83f46288-b699-4193-9ad1-bb87e8768d39#xywh=569,1699,1747,67"
  "within":
  {"@id": "https://vidalsantos.github.io/iiif_test.github.io/iiif/manifest/manifest_static_ranges.json",
  "@type": "sc:Manifest"
  }, {...},
  ...
}
(Manifest)
{
  "@id": "https://83f46288-b699-4193-9ad1-bb87e8768d39",
  "@type": "sc:Canvas",
  "label": "7 - Portada",
  {...},
  {
    "@id": "https://vidalsantos.github.io/iiif_test.github.io/iiif/Altojson/anno_list07.json",
    "@type": "sc:AnnotationList",
    "label": "Transcript"
  }
}

```

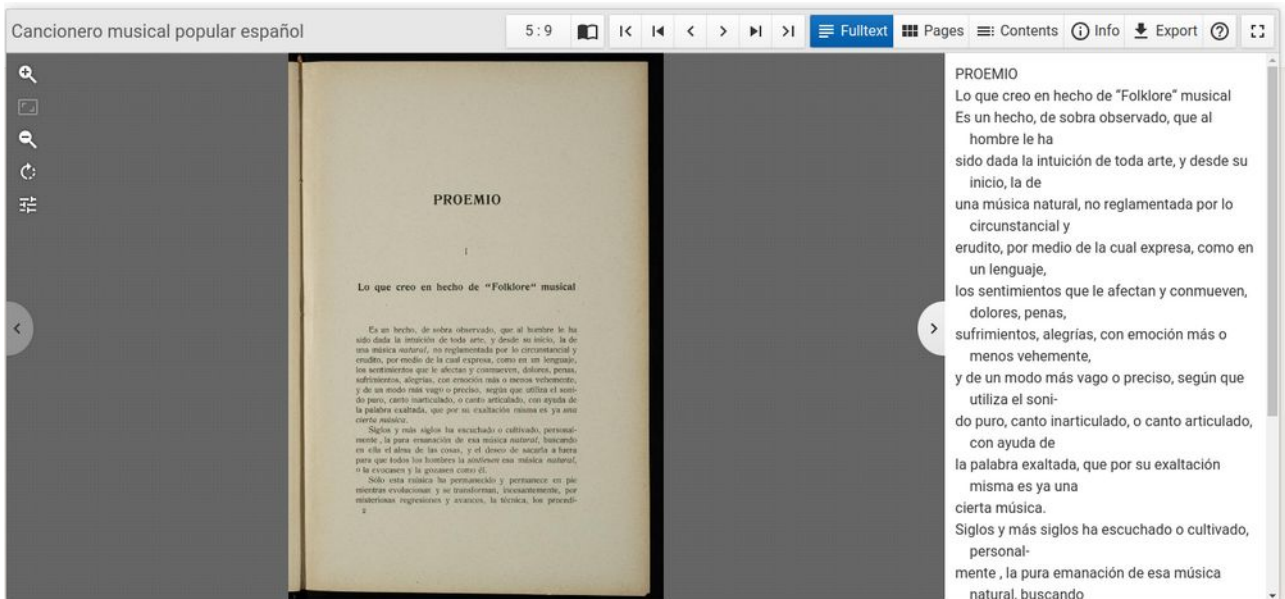



Figura 25: Detall de la capa de text sobreimpres al manifest IIIF

5.3 Disseny i prototipatge de la plataforma d'accés digital

Per a la creació del prototip per a la plataforma de difusió es segueixen les instruccions per a la instal·lació de Jekyll desenvolupats per al projecte «El corrido mexicano» que incorpora una versió del visor Verovio per a la al que s'afegeix una el visor Mirador, que no era present en el desenvolupament original.

En ella s'insereixen la versió editada en *Markdown* del text de la obra (on es repliquen les convencions que permeten la disposició de la taula de continguts i les referències a peu de pàgina a partir del full d'estils CSS), el manifest IIIF (per a la presentació de la edició facsímil, les partitures generades en MEI de les tonades de la obra, el plugin jekyll-scholar i els fitxers de referències bibliogràfiques (Figura 26)

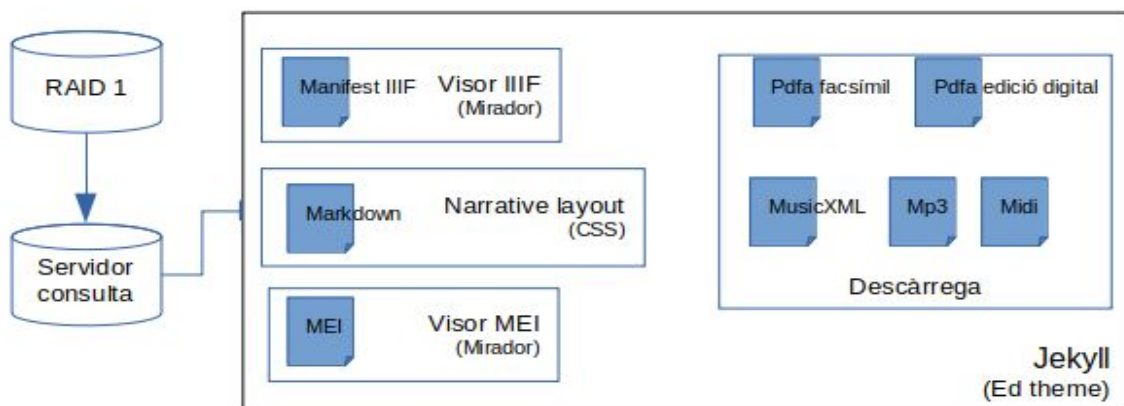


Figura 26: Diagrama de flux per a construcció de la plataforma d'accés digital

La configuració dels paràmetres per a la generació de la web estàtica situats a l'arxiu YAML de configuració permeten definir la codificació de text, la estructura de la *url* de sortida i *permalinks*, les opcions de la llibreria *kramdown* per a la conversió del marcatge de text i la incorporació de metadades descriptives o sistemes de d'anàlítica web a la pàgina més comuns per a la traçabilitat del tràfic del projecte en línia. ANNEX?

La generació de les planes web estàtiques s'estableix en un sistema de plantilles que permet popular el contingut de la web de manera modular a partir de documents que descriuen cada un dels apartats que la componen (text, visor, plantilla). Aquests fitxers incorporen els paràmetres YAML a l'inici del document que serveixen per identificar cada element en la codificació de variables del sistema de plantilles *líquid* per unir-les en la versió final (Figura 27) i permeten incorporar les metadades descriptives de la obra.

```
/_layouts/narrative_mirador.html
---
layout: default_Mirador
---
{% include mirador-viewer.html manifest=page.manifest %}

/_texts/IIIFmirador.md
---
title: Cancionero musical popular español (edició facsímil)
author: Felip Pedrell
editor: Vidal Santos
rights: Public Domain
source: Biblioteca de Catalunya
publication-date: 2021

layout: narrative_mirador
manifest: 'http://127.0.0.1:4000/cancionero/Manifests/Pedrell_CancioneroFacsímil_manifest.json'
---
```

Figura 27: Exemple de construcció modular a Jekyll

El disseny de la interfície parteix del original sense plantejar canvis en la disposició dels elements de la plantilla. Ja en les fases de transformació de textos s'utilitzen les convencions per adaptar els versos, la taula de continguts i la inserció de citacions bibliogràfiques segons els esquemes previstos en la plantilla (Figura 28). També la interfície IIIF per a la presentació de la edició facsímil respon a les expectatives del disseny i la generació de les partitures en SVG i el so generat en MIDI es reproduïx sense problemes amb el programari Verovio integrat a Jekyll.

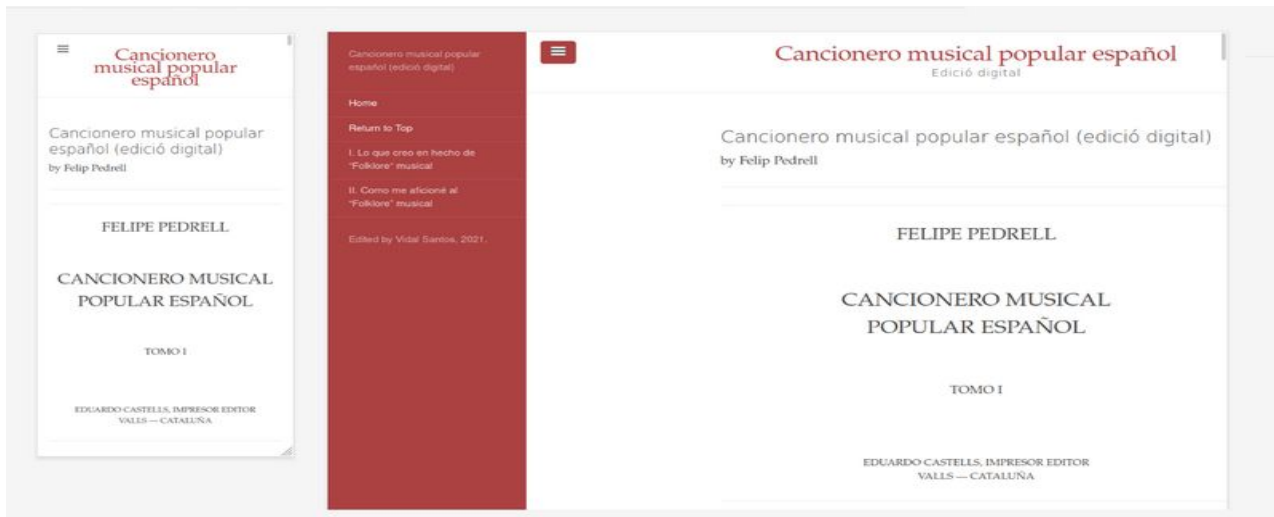


Figura 28: Exemple de disposició responsive de la edició digital del text

6. Discussió del resultat

6.1 Limitacions del disseny i la parametrització de processos plantejats

A expenses d'una avaluació amb més profunditat de les eines utilitzades i la incorporació d'idees omeses per motius de calendari, aquest treball es declara prou definit per actuar com a prova de concepte a plantejar l'ús de les eines exposades en un desenvolupament complet del projecte. D'aquesta manera s'aconsegueix generar un espai virtual accessori i de caràcter transitori que permet expandir les capacitats ja incorporades en el procés de digitalització dels *surrogates* digitals sense comprometre desenvolupaments de caràcter més general i amb cicles de producció més llargs i complexos.

L'exercici de mimètica en les possibles aproximacions al text des d'un punt de vista més proper a la recerca ha permès que la notació musical i les relacions bibliogràfiques i hipertextuals codificades al paper puguin ser explorades de manera més convenient al entorn virtual. Permetent, a la vegada, que els aspectes documentals de la obra viatgin còmodament empaquetats en el mateix objecte digital, aliens a les particularitats del expositor.

En aquest sentit el projecte és, conceptualment, un símil digital a la *ephemera*. El caràcter modular en el disseny de la plataforma estàtica, recolzada en el sistema traçabilitat GIT i el desenvolupament en Ruby, és totalment independent dels materials a que dona accés i, a la vegada, és justament aquest aspecte el que ha de ser considerat per tractar com a punt d'ancoratge en les cadenes de preservació digital del centre i

Els detalls que s'exposen a continuació permeten establir les prioritats que la propera fase de desenvolupament necessita per seguir experimentant en les funcionalitats dels esquemes MEI i IIIF, la incorporació de nous formats de difusió per al projecte i el desenvolupament de la plataforma de difusió.

- El projecte defineix específicament la necessitat d'establir processos d'avaluació i validació de resultats que durant la fase d'implementació han estat utilitzades en les mostres de text, imatge i notació musical de mostra, però es possible que algunes particularitats de la obra completa puguin necessitar adaptar algun procés de codificació o la validació per un esquema *ad-hoc* en la edició final del projecte, molt possiblement a efectes d'una correcta visualització i reproducció dels fitxers MEI.
- La reutilització de la plantilla Ed. de *El corrido mexicano* representa un bon punt de partida per a la integració de text i música notada del projecte però cal ser modificat per un disseny d'arquitectura de la informació i usabilitat de la plataforma que ha estat desestimat per motius de calendari
- Tot i no ser haver estat utilitzat en el plantejament del projecte, és necessari contemplar, la possibilitat d'agilitzar la introducció dels materials de consulta a partir de la eina wax-tasks³³ o una solució similar.

Aquesta eina permet la creació de l'encapçalament de les pàgines en Markdown a la plataforma Jekyll a partir de taules on es pot definir les metadades descriptives del recurs, incorporar la ubicació del fitxer a carregar i assignar la plantilla de visualització (*layout*) per a generar-ne la versió HTML, especialment útil per a la incorporació de les composicions musicals de manera automatitzada.

- La possibilitat d'integrar en els rangs del manifest de la edició facsímil totes les composicions musicals de la obra i aprofitar les opcions de dereferencialització per poder ser enllaçada en altres seccions de la plataforma de consulta necessita un treball més acurat en l'establiment de url permanents per a complir amb les especificacions del sistema.

Malgrat la completa funcionalitat del manifest, aquesta opció, juntament amb la possibilitat d'inserir el text extret per OCR en una capa de visualització que permetés la selecció de passatges per mitja de la Content Search API haurien de ser prioritàries en tasques posteriors.

- Un dels principals problemes per a la correcta disposició de la secció musical de la obra eren els paràmetres de transcripció per aquelles composicions que incorporen melismes. Durant el procés de transcripció de les mostres s'obvien per a poder establir la validesa del procés però és necessari

33 Minicomp. wax_tasks [Internet]. minicomp; 2021 [citad 18 gener 2021]. Disponible a: https://github.com/minicomp/wax_tasks

recalcar la necessitat de personal expert en musicologia per a la correcta transcripció i edició de les partitures del cançoner es fa indispensable per aconseguir desplegar totes les funcionalitats que la versió 4.0 del esquema pot oferir i els requeriments del visor Verovio per a visualitzar correctament els elements en pantalla.

També cal explorar altres vies de codificació dels elements bibliogràfics de la obra musical que permetin agilitzar el procés de catalogació manual i la correcció de les partitures. En aquest sentit s'observa l'assistència del programari MerMEId³⁴, en l'ús de la seva interfície gràfica per assistir en la incorporació dels elements de descripció, i en termes de disseny conceptual, valorar la possibilitat de concebre l'obra de manera unitària en la codificació (FRBR) i establir punts d'accés per a cada part component a la plataforma digital des del mateix fitxer font.

- La utilització de Markdown (kramdow) en la disposició de la secció de text de la edició digital de la obra seguint els paràmetres de la plantilla Ed. compleix amb la seva funció i, fins i tot, des del disseny original pot implementar el complement jekyll-scholar³⁵ per a la correcta presentació de la bibliografia de la obra, pendent d'implementar.

La decisió en utilitzar aquest llenguatge de marques ve condicionat en la selecció de la plantilla Ed. i comporta l'ús de convencions que poden no afavorir la reutilització del material per altres usos o fer ferragosa la edició dels 4 volums de la obra. En aquest sentit és necessari plantejar l'ús de un editor específic per a la codificació Markdown³⁶ o, idealment, establir el format TEI per beneficiar la incorporació d'informació formatada a altres formats de difusió que automatitzi de manera més eficient el procés de transformació a altres formats de lectura o la creació d'una futura edició crítica.

6.2 Integració del prototip en altres sistemes de gestió de continguts

Malgrat les correccions necessàries exposades en l'apartat anterior l'aplicació d'aquests nous estàndards es demostra eficaç en els plantejaments establerts en els objectius de l'estudi, tot i que no justifiquen per si sols la orientació que es pretén assolir en aquest projecte.

34 Dansk Center for Musikudgivelse. MerMEId [Internet]. [citad 18 gener 2021]. Disponible a: <https://github.com/kb-dk/MerMEId>

35 Keil S. jekyll-scholar [Internet]. 2021 [citad 19 gener 2021]. Disponible a: <https://github.com/inukshuk/jekyll-scholar>

36 Tramullas J. Editores de escritorio para Markdown [Internet]. [citad 18 gener 2021]. Disponible a: <https://tramullas.com/editores-de-escritorio-para-markdown/>

Un cop resolta la implementació és vital assegurar que els materials creats per a la millora de la consulta/estudi de la obra puguin ser localitzats fora de la plataforma creada *ad-hoc*. Aquest tipus de desenvolupament permet la creació de plataformes més flexibles a l'hora d'adequar formats de presentació i usabilitat a continguts molt específics però té el risc de fragmentar conjunts d'arrel temàtica i aprofundir la erràtica situació de la gestió de biblioteques digitals a institucions patrimonials i de recerca en la creació d'un *silo* més que engronsi la partida de recursos dispersos al territori.

Els esforços en la transformació XSLT a format Marc21 i Dublin Core (no generats expressament però assumible a partir de les indicacions en la documentació tècnica de MEI) poden incrementar la vies d'accés al recurs per mitjà d'altres sistemes presents al territori, ja sigui com a conjunt o a cada una de les peces musical recollides a la obra. Per aquest motiu s'ofereix, a tall d'exemple, una possible via per a la integració dels derivats digitals a l'actual ecosistema de difusió de patrimoni musical al territori.

Com hem vist a l'apartat 2.3, el Mapa de patrimoni musical desenvolupat pel *Centro de Documentación de Música y Danza* estatal al territori català consten 116 centres amb fons musicals³⁷ amb una gran disparitat de casuístiques en les dimensions i capacitat logística per a emprendre projectes digitals i obrir els seus inventaris en línia. De la mateixa manera que l'adopció de sistemes CMS ha estat gradual i subjecta a barreres d'entrada tant econòmiques com tècniques, l'ús de les tècniques descrites en aquest treball poden ser un impediment per a la generació d'un ecosistema de treball que vehiculi els avenços en el camps OMR i edició digital a tots els nodes de la xarxa de patrimoni musical i és vital apuntalar un node que pugui fer partícips altres centres, ja sigui a partir de la generació de guies i la reutilització de materials i tècniques com en la projecció temàtica de la disciplina.

Prenent com a punt de partida la participació del Departament de Cultura de la Generalitat de Cultura en el projecte i les plataformes d'accés que ha creat o participa activament en l'actualitat (COFRE, Calaix, Memòria Digital de Catalunya, Patrimoni musical), pot ser relativament fàcil i convenient disposar d'aquests espais per emmagatzemar els continguts generats i utilitzar la web estàtica i els visors IIF i MEI com agents vehiculars per a la integració del conjunt.

D'aquesta manera es garanteix que l'esforç en la descripció dels materials i la reformulació digital de les partitures puguin ser descobertes i contextualitzades en espais més heterogenis ja integrants en eines de descoberta i circuits de difusió a escala per sistemes d'agregació com Europeana o MDC (Figura 29).

En aquest sentit, es replica el model generat per al projecte substituint COFRE com a repositori de preservació (1); allotjar les imatges digitalitzades a MDC, els derivats en el repositori Calaix i utilitzar la catalogació analítica de cada peça musical per oferir un punt d'accés via catàleg (2) i, finalment, utilitzar la

³⁷ Centro de Documentación de Música y Danza. Mapa del patrimonio musical en España · Centro de Documentación de Música y Danza [Internet]. 2020 [citad 3 setembre 2020]. Disponible a: <http://cdmyd.mcu.es/mapatrimoniomusical/patrimonio-musical.php>

plataforma Patrimoni musical Català³⁸, creada recentment pel centre d'estudis musicals Robert Gerhard i que assumeix el rol d'unificar temàticament el contingut dispers en altres institucions en una interfície preparada per a fer conuiuere imatge i contingut audiovisual, com a destí alternatiu a la plataforma dedicada al Cancionero per potenciar-ne l'accés a llarg termini en el moment en que es decideix actualitzar els continguts o tancar-ne l'accés i ser desplaçat al sistema COFRE per a preservar-lo digitalment (3).

La idea de desballestar els continguts derivats i les imatges digitalitzades pren com a punt de partida facilitar l'accés en tot moment a l'obra a partir de les pròpies URI permanents d'aquests sistemes i no comprometre la construcció i l'enllaç de nous textos d'anàlisi que puguin referenciar la obra o les seves parts components.

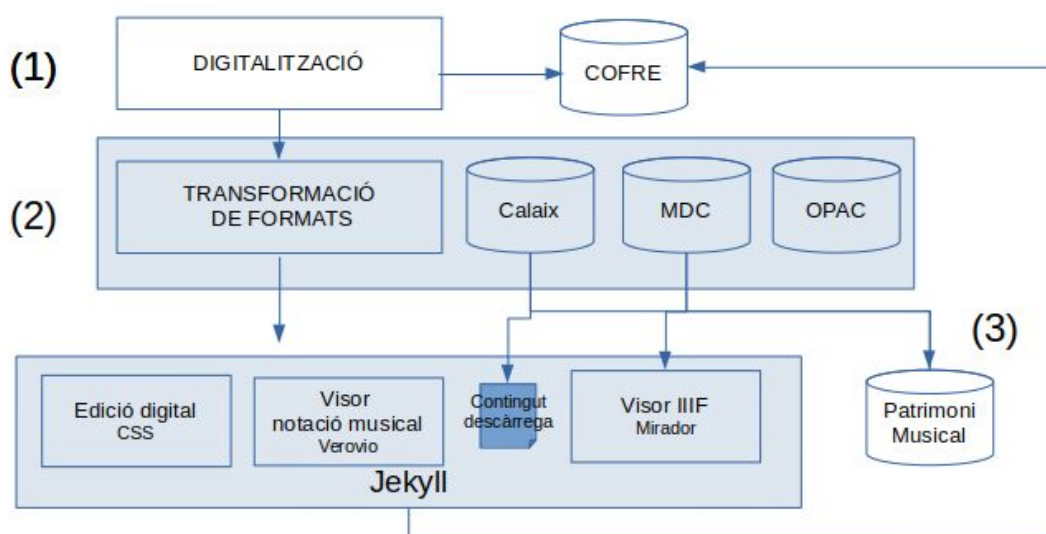


Figura 29: Proposta d'integració del projecte a infraestructures gencat.cat

6.4 Pressupost estimat

Els resultats d'aquesta experiència de prototipatge permeten calcular un cost estimat per a la producció definitiva de la edició digital del *Cancionero musical popular español*.

En un pressupost ajustat i un càlcul de temps previsiblement assumible, es preveu que un equip reduït de 3 persones (2 musicòlogues, 1 documentalista) poden realitzar aquest treball en un termini de 6 mesos (Taula 3 i 4).

Les fases del projecte varien lleugerament a les del prototip per donar més pes a la contextualització de la obra i per separar la transformació de formats de la descripció bibliogràfica i analítica de la obra dirigida a les possibles tasques de recerca i a la integració en repositoris temàtics dedicats a la musicologia. A la vegada

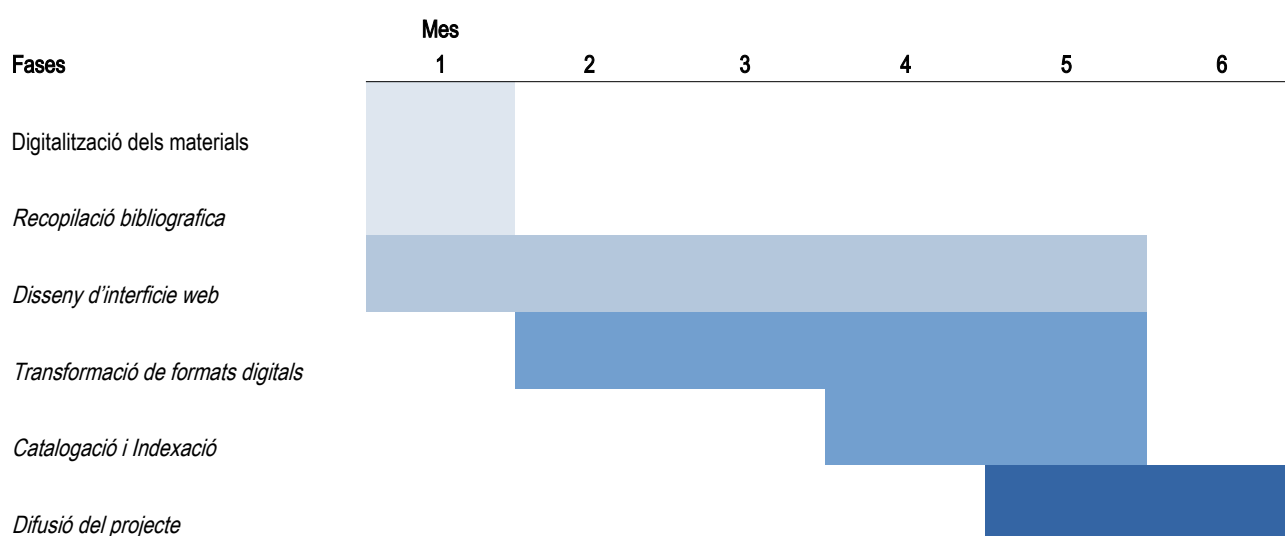
38 Centre Robert Gerhard. Patrimoni musical català [Internet]. 2019 [citat 8 juny 2021]. Disponible a: <http://www.patrimonimusical.cat/>

s'introdueix una partida sota aquest concepte a la estimació pressupostaria per fer valer els esforços en aquesta direcció, destinada a dietes de transport i despeses associades.

L'apartat d'equipament disposa la partida més gran per al procés de digitalització, i pressuposa un sistema de d'emmagatzematge en Raid 1 de baix cost i un servei d'allotjament amb domini per a l'intercanvi de fitxers treball intern i la gestió de correus i sistemes de gestió ofimàtica en línia per a la coordinació dels integrants del equip.

Per la concepció inicial del projecte a partir d'eines de programari lliure o de codi obert no es preveuen partides per a la compra de programari especialitzat.

Taula 3: Calendari estimat per al projecte



Taula 4: Pressupost estimat per al projecte

Concepte		Import (€)
Equip humà	1 Musicòloga	18000
	1 Musicòloga	18000
	1 Documentalista	18000
Gestió del projecte amb agents externs		3000
Equipament	Digitalització de materials	1200
	Sistema d'emmagatzematge de còpies de seguretat i manteniment	120
	Compra de material fungible	400
Allotjament web i gestió del domini web (1 any)		550
		59270

6.3 Ús dels principis ODD en la consolidació d'iteracions en el tractament documental

Els resultats obtinguts posen en valor les possibilitats que ofereix el sistema ODD en la construcció de biblioteques digitals i sistemes de recuperació més efectius en l'àmbit del patrimoni documental, malgrat no ser explorats a fons en el desenvolupament del projecte.

La integració de contingut de la obra, les dades que componen el registre bibliogràfic i aquella informació analítica en la creació d'objectes digitals beneficia la modelació de fluxos de treball per a la curació de continguts digitals de manera similar als dedicats a les vessants de preservació digital.

La iteració en la millora de la informació descriptiva o analítica --poc explorades a pesar de la insistència dels models DCC i altres iniciatives similars com el projecte KONDA i el model LIDO³⁹, les tendències en la consolidació dels models semàntics per a la millora de l'accés a les col·leccions en el marc de les RDA i les LRM exposats en la revisió bibliogràfica⁴⁰ o els principis FAIR TRUST (*Lin et al., 2020*) promoguts i altres variants de sindicació de models OAIS-- tindrà una importància vital en aquest àmbit i, consegüentment en el desenvolupament de col·leccions dels centres de documentació a llarg termini.

En aquest sentit cal posar en valor la predisposició d'aquests models en fer conviure de manera harmònica en un catàleg aquells sense una orientació específicament bibliogràfica. Les transformacions en base XLST proposades per a la generació de registres Marc21 a partir dels apartats de descripció MEI i la possibilitat d'establir un flux similar amb IIIF pot ser una via d'entrada del model semàntic en la construcció de sistemes de recuperació més robustos i preparats per a la convivència de múltiples fonts i orígens.

Com es pregunten en un article recent al blog *Hanging together* que explora les necessitats dels centres en la construcció del projecte *Next Generation Metadata* (Dortmund, 2021), pot aquesta metodologia de treball agilitzar els processos de descripció i accés a la documentació? És possible que les destacades aportacions de l'ecosistema de software lliure, codi obert i moviments a l'entorn del coneixement obert reorientin els fluxos de treball i la percepció de la catalogació i el set d'habilitats específiques atribuïdes al personal bibliotecari (Frederick, 2016). No seria extremadament temerari pronosticar una tendència en aquest sentit si s'observa amb detall l'evolució dels sistemes de catalogació consorciada, la integració de grans paquets per subscripció als sistemes de descoberta i la necessitat d'emmagatzemar grans conjunts de dades a grans plataformes per a la reutilització de dades de recerca.

39 Acronim de DCC Curation Lifecycle Model, Continuous Quality Management of Dynamic Research Data on Cultural Heritage Objects using the LIDO standard i Lightweight Information Describing Objects, respectivament

40 El cas més representatiu dels exposats és la presentació de Gall a CU Boulder Libraries, on estableixen els requeriments per al model semàntic que ha de guiar el sistema de descoberta LUX (Gall, 2021)

Segons la documentació recollida en la revisió bibliogràfica, és molt possible que aquesta tendència sigui prioritària en la planificació estratègica digital dels molts centres de documentació a llarg termini, permetent que les construccions que actuen com a plataformes d'agregació enriqueixin les seves vies d'interacció amb els continguts des de una perspectiva confederada.

La opció que presenta el projecte, mentrestant, és una via fàcilment adaptable a les necessitats d'un objecte digital concret o a necessitats que requereixen de funcionalitats específiques i permet no comprometre ecosistemes interns o consorciats amb un alt grau de dependència econòmica per afrontar una reconversió total de la seva estratègia digital o la falta de perfils tècnics per dur-la a terme. La metodologia emprada per a la creació dels objectes digitals permet que puguin ser tractats de manera immediata per aquells agents interessants en fer-hi recerca sense perdre el pols a les pràctiques de descripció i indexació de la unitat d'informació, sempre codificada en el mateix contenidor i amb la garanties de poder ser reutilitzada en el futur des d'una òptica associada als processos de transformació necessaris per establir un circuit més complet com els mencionats anteriorment, capaç de trencar el temut efecte *silo* a partir de fitxers ODD i la resolució de IDs persistents que permetin accedir als recursos amb independència de la plataforma que els allotji.

En futurs desenvolupaments caldria estar atents en aquest sentit davant els avanços en el desenvolupament de BIBFRAME i els requeriments de les RDA i la seva predisposició a l'ús de models de descripció externs com FOAF o OCDM⁴¹ o Music ontology per establir relacions més sòlides en el sistema de relacions de les citacions, les peces del cançoner popular (en aquest cas) i, a llarg termini, fent via cap a models consolidats OAIS i considerar la creació i consolidació de documentació i objectes digitals segons metodologies properes a la traçabilitat dels sistemes de control de versions GIT ([Lousky, 2021](#)).

41 Acrònims de *Friend of a Friend* i *Open Citations Data Model*, respectivament

7. Conclusions i futures actuacions

L'aplicació d'aquests nous estàndards es demostra eficaç en els plantejaments establerts en els objectius de l'estudi, tot i que no justifiquen per si sols la orientació que aquest projecte pretén assolir.

Per un cantó, aquest tipus de desenvolupament permet la creació de plataformes més flexibles a l'hora d'adequar formats de presentació i usabilitat a continguts molt específics a partir dels plantejaments *One Document Does it all* (ODD). Segons la documentació recollida en la revisió bibliogràfica, aquesta tendència serà prioritària en la planificació estratègica digital dels molts centres de documentació a llarg termini, permetent que les construccions que actuen com a plataformes d'agregació enriqueixin les seves vies d'interacció amb els continguts des de una perspectiva federada.

D'altra banda, són les ramificacions que se'n poden derivar, en clau investigadora o en la seva difusió al gran públic, el que dirigeix l'esforç de fer més accessible el *Cancionero popular musical español* a partir d'aquestes tecnologies i, per aquest motiu, és vital acompanyar l'usuari en l'ús.

Les intencions del projecte van encaminades a estrènyer els llaços existents entre la comunitat investigadora i la xarxa de centres de documentació amb fons musicals. La sintonia entre un centre de documentació i l'usuari depèn de manera directa en la seva capacitat per mantenir un diàleg fluid en les metodologies de recerca i les tècniques de tractament documental actuals. Si aquestes no coincideixen, es difícil que els fruits generats per una investigació siguin reaprofitats pel centre i que aquest, a la vegada, pugui assistir millor en consultes similars.

És aquest l'esperit del treball que aquí s'exposa. Proveir de transcripcions digitals a usuaris acostumats a tractar digitalment i facilitar, d'aquesta manera, l'entrada de noves perspectives de recerca i difusió de patrimoni als centres de documentació especialitzats en música i el material que se'n deriva per facilitar l'adopció d'aquests materials i les tècniques de gestió de la recerca que se'n deriven a aquells investigadors que vulguin integrar-les en el seu objecte d'estudi o que ja treballin plenament amb elles.

El seguit de processos tècnics de transformació de formats, especialment els dedicats a la generació de música notada en aquests projecte no pretenen suplantar la tasques pròpies dels col·lectius investigadors que planifiquen, cimenten i apliquen els estàndards aquí exposats i és per això que la participació d'experts en la planificació d'aquests tipus de projectes és essencial per a fer créixer les possibilitats de millora en els serveis que pot oferir un centre de documentació en el marc de les humanitats digitals i cal concentrar esforços en que els seus objectius conflueixin amb els objectius del centre.

8. Bibliografia

- Abadal E. Gestión de proyectos en información y documentación. Ediciones Trea, S. L.; 2004 [citad 7 juliol 2020]. Disponible a: <http://diposit.ub.edu/dspace/handle/2445/61044>
- Alcaraz Martínez R, Vázquez Puig E. TEI: un estàndard per codificar textos en l'àmbit de les humanitats digitals. BiD: textos universitaris de biblioteconomia i documentació. 2016;(37). Disponible a: <http://bid.ub.edu/37/alcaraz.htm>
- Allés-Torrent S, Riande G del R. The Switchover: Teaching and Learning the Text Encoding Initiative in Spanish. Journal of the Text Encoding Initiative. 2020 ;(12) [citad 4 setembre 2020]. Disponible a: <https://doi.org/10/gg9q9ss>
- Apollon D, Belisle C, Régnier P, editors. Digital critical editions. Urbana: University of Illinois Press; 2014. 357 p. (Topics in the digital humanities).
- Babcock K, Lee S, Rajakumar J, Wagner A. Where Do We Go From Here: A Review of Technology Solutions for Providing Access to Digital Collections. The Code4Lib Journal. 17 febrer 2020;(47).
- Barbuti N. Thinking digital libraries for preservation as digital cultural heritage: by R to R4 facet of FAIR principles. Int J Digit Libr [Internet]. 27 agost 2020 [citad 8 setembre 2020]; Disponible a: [10/gg92zd](https://doi.org/10/gg92zd)
- Boté Vericad J-J, Térmens i Graells M. Reusing Data: Technical and Ethical Challenges. 16 desembre 2019 [citad 10 agost 2020]; Disponible a: <http://diposit.ub.edu/dspace/handle/2445/151341>
- Fuhr J. Developing Data Services Skills in Academic Libraries. College & Research Libraries. 2 maig 2022 [citad 12 maig 2022]; Disponible a: <https://crl.acrl.org/index.php/crl/article/view/24863>
- Castro P. Roles and Competencies of CRIS Managers. Medical Library Forum, 2021 [citad 12 maig 2022]; Disponible a: <https://doi.org/10.34738/mlf.0056>
- Centro de Documentación de Música y Danza. Mapa del patrimonio musical en España · Centro de Documentación de Música y Danza [Internet]. [citad 3 setembre 2020]. Disponible a: <http://cdmyd.mcu.es/mapatrimoniomusical/patrimonio-musical.php>
- Diaz C. Jekyll and Institutional Repositories. 2018.
- Diaz C. Using Static Site Generators for Scholarly Publications and Open Educational Resources. The Code4Lib Journal [Internet]. 8 novembre 2018 [citad 7 agost 2020];(42). Disponible a: <https://journal.code4lib.org/articles/13861>
- Eco U. Apocalípticos e integrados. Barcelona: Editorial Lumen; 1977.
- Gil A. Ed. A jekyll theme for minimal editions [Internet]. 2016 [citad 15 gener 2021]. Disponible a: <http://minicomp.github.io/ed/>
- Gil-Leiva I, Fujita M, Díaz Ortuño P, Dos Reis D. The Abandonment of the Assignment of Subject Headings and Classification Codes in University Libraries due to the Massive Emergence of Electronic Books // El abandono de la asignación de encabezamientos de materia y códigos de clasificación en las bibliotecas universitarias debido a la incorporación masiva del libro electrónico. Vol. 47, Knowledge Organization. ISKO; 2020 [citad 18 març 2021]. p. 646-67. Disponible a: <http://eprints.rclis.org/40981/>
- Gomez J, Clarke KS, Vuong A. IIF by the Numbers. The Code4Lib Journal [Internet]. 11 maig 2020 [citad 4 setembre 2020];(48). Disponible a: <https://journal.code4lib.org/articles/15217>
- Grevtsova I. Informe sobre el estado de las tecnologías en los museos catalanes durante la crisis de la covid-19 • Digital Heritage & Culture [Internet]. Digital Heritage & Culture. 2020 [citad 31 agost 2020]. Disponible a: <https://irinagrevtsova.com/informe-tecnologias-museos-catalanes-covid-19/>
- Grusin R. The Dark Side of Digital Humanities: Dispatches from Two Recent MLA Conventions. differences. 1 maig 2014;25(1):79-92.

- Higgins S. The DCC Curation Lifecycle Model. *International Journal of Digital Curation*. 6 agost 2008;3(1):134-40. Disponible a: <http://www.ijdc.net/article/view/69>
- ICOM International Committee for Documentation. LIDO [Internet]. ICOM CIDOC. 2021 [citat 16 juny 2021]. Disponible a: <http://cidoc.mini.icom.museum/working-groups/lido/>
- Lavoie B. Social interoperability builds a stronger three-legged stool [Internet]. *Hanging Together*. 2020 [citat 31 agost 2020]. Disponible a: <https://hangingtogether.org/?p=8144>
- Library of Congress. Sustainability of Digital Formats: Planning for Library of Congress Collections. Music Encoding Initiative (MEI) Format Family [Internet]. 2019 [citat 9 setembre 2020]. Disponible a: <https://www.loc.gov/preservation/digital/formats/fdd/fdd000502.shtml>
- Lin D, Crabtree J, Dillo I, Downs RR, Edmunds R, Giaretta D, et al. The TRUST Principles for digital repositories. *Scientific Data*; 7 (1):144. [citat 15 maig 2020]. Disponible a: <https://doi.org/10.1038/s41597-020-0486-7>
- Lousky S. Datasets should behave like git repositories [Internet]. *DAGsHub Blog*. 2021 [citat 26 maig 2021]. Disponible a: <https://dagshub.com/blog/datasets-should-behave-like-git-repositories/>
- Mayo C, Jazairi A, Walker P, Gaudreau L. BC Digitized Collections: Towards a Microservices-based Solution to an Intractable Repository Problem. *The Code4Lib Journal*. 6 maig 2019 [citat 28 juliol 2020];(44). Disponible a: <https://journal.code4lib.org/articles/14445>
- Morillas JLH. Fons antic i repositoris universitaris a Espanya. BiD: textos universitaris de biblioteconomia i documentació. 2015;
- Newson K. Tools and Workflows for Collaborating on Static Website Projects. *The Code4Lib Journal* [Internet]. 18 octubre 2017 [citat 28 juliol 2020];(38). Disponible a: <https://journal.code4lib.org/articles/12779>
- Nualart Vilaplana J. Text Analysis and Visualisation: creating deep interfaces to read textual document collections [Internet] [Doctoral thesis]. [Canberra]: University of Canberra; 2016 [citat 26 juliol 2020]. Disponible a: <http://research.nualart.cat/wp-content/uploads/2016/JaumeNualart-UC-thesis-2016.pdf>
- Nyröp M, Gil A. Wax: An extensible workflow for creating minimal scholarly exhibitions with Jekyll and Rake [Internet]. 2019 [citat 15 gener 2021]. Disponible a: <https://minicomp.github.io/wax/>
- Ordine N, Flexner A. *La utilitat de l'inútil: manifest*. Barcelona: Quaderns Crema; 2013.
- Raju J. Information Professional or IT Professional?: The Knowledge and Skills Required by Academic Librarians in the Digital Library Environment. *Portal*. 2017;17(4):739-57.
- Ronallo J. IIF Technical Workshop: IIF Technical Workshop. 2018 [citat 7 juny 2021]. Disponible a: <http://ronallo.com/iif-workshop-new/>
- Saccomano M. El corrido mexicano. [citat 15 gener 2021]. Disponible a: <https://mss2221.github.io/corridosEd>
- Sälgö M. Carl Larsson who is that - sadly Europeana doesnt know --> #Metadatadebt [Internet]. *LinkedData in action a la Sälgö*. 2022 [citat 12 maig 2022]. Disponible a: <https://minancestry.blogspot.com/2020/03/carl-larsson-who-is-that-sadly.html>
- Sánchez-Jara JM. Music Encoding Initiative (MEI), un estándar para la edición, codificación y la descripción de documentos musicales en el entorno digital: características, prestaciones y ámbitos de aplicación. *Education in the Knowledge Society (EKS)*. 24 maig 2019;20(0):13.
- Schmidt D. Towards an Interoperable Digital Scholarly Edition. *Journal of the Text Encoding Initiative* [Internet]. 12 novembre 2014 [citat 6 agost 2020];(Issue 7). Disponible a: <http://journals.openedition.org/jtei/979>
- Tarver H, Gentry S. What is Metadata Assessment? [Internet]. *Digital Library Federation*. 2021 [citat 17 març 2021]. Disponible a: <https://www.diglib.org/what-is-metadata-assessment/>

Térmens i Graells M. Els estàndards de digitalització: una revisió de Catalunya al món. Item: revista de biblioteconomia i documentació. juny 2017;(62):140-52.[citat 15 gener 2021]. Disponible a: <http://diposit.ub.edu/dspace/handle/2445/122653>

Tramullas J. Elaboración de productos de información con JAMstack: del sistema de gestión de contenidos al web estático | Anuario ThinkEPI. [citat 19 agost 2020]; Disponible a: <https://recyt.fecyt.es/index.php/ThinkEPI/article/view/thinkepi.2020.e14f05>

TSOU J, Davison S. Review of Sheet Music Consortium, Davison Stephen. Journal of the American Musicological Society. 2016;69(1):255-63.

Valle-de-Lope I, Molina-González P, Vidal-Santos G. 100 anys de cuplet en català: prova pilot de dinamització digital dels fons musicals de la Biblioteca de Catalunya. En Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya; 2020 [citat 10 gener 2021]. Disponible a: <http://eprints.rclis.org/40784/>

Zhang Y, Liu S, Mathews E. Convergence of digital humanities and digital libraries. Library Management. 8 juny 2015;36(4/5):362-77.

Table of Figures

Figura 1: Tipologies documental a plataformes digitals llistades per Mapa de patrimoni digital español.....	12
Figura 2: Coincidència de Centres llistats a Mapa del patrimoni musical, IFMUC, Patrimoni musical català i RISM. .	13
Figura 3: Diagrama de flux de treball a Mayo et al. (2019).....	15
Figura 4: Diagrama de flux per a integració de recursos digitals a Yale University Museum (Gall, 2021).....	15
Figura 5: Arquitectures basades en CMS i JAMstack (Tramullas, 2020).....	16
Figura 6: Interacció d'APIs a IIF (Ronallo, 2018).....	18
Figura 7: Detall de interfície de consulta a Burns antiphone.....	20
Figura 8: Detall de la plataforma de consulta a Tasso in Music Project.....	21
Figura 9: Detall de la interfície de consulta a El Corrido Mexicano.....	22
Figura 10: Diagrama de flux del projecte.....	27
Figura 11: Diagrama de flux per a la digitalització de la obra.....	29
Figura 12: Diagrama de flux per a la transcripció de textos.....	30
Figura 13: Exemple d'indentació en la construcció del text a Markdown.....	30
Figura 14: Inserció de referències bibliogràfiques i marcatge a sumari.....	31
Figura 15: Diagrama de flux per a la transcripció de música notada.....	32
Figura 16: Detall del procés de reconeixement OMR a Audiveris.....	33
Figura 17: Exemple de millora en la classificació automàtica en ampliar el corpus de glifs.....	33
Figura 18: Exemple de la interfície de validació de MEI a Atom.....	34
Figura 19: Encapçalament de prova a MEI.....	35
Figura 20: Diagrama de flux per a la creació de la edició facsímil.....	35
Figura 21: Paràmetres de configuració d'imatges de sortida a servidor IIF.....	36
Figura 22: Exemple de parametrització a Image API de.....	36
Figura 23: Detall de les seqüències de descripció i rangs a manifest IIF.....	37
Figura 24: Detall de la incorporació de text al manifest IIF.....	37
Figura 25: Detall de la capa de text sobreimpres al manifest IIF.....	38
Figura 26: Diagrama de flux per a construcció de la plataforma d'accés digital.....	38
Figura 27: Exemple de construcció modular a Jekyll.....	39
Figura 28: Exemple de disposició responsive de la edició digital del text.....	40
Figura 29: Proposta d'integració del projecte a infraestructures gencat.cat.....	44

9. Annexos

Annex A. Exemple de transcripció manual sobre font original

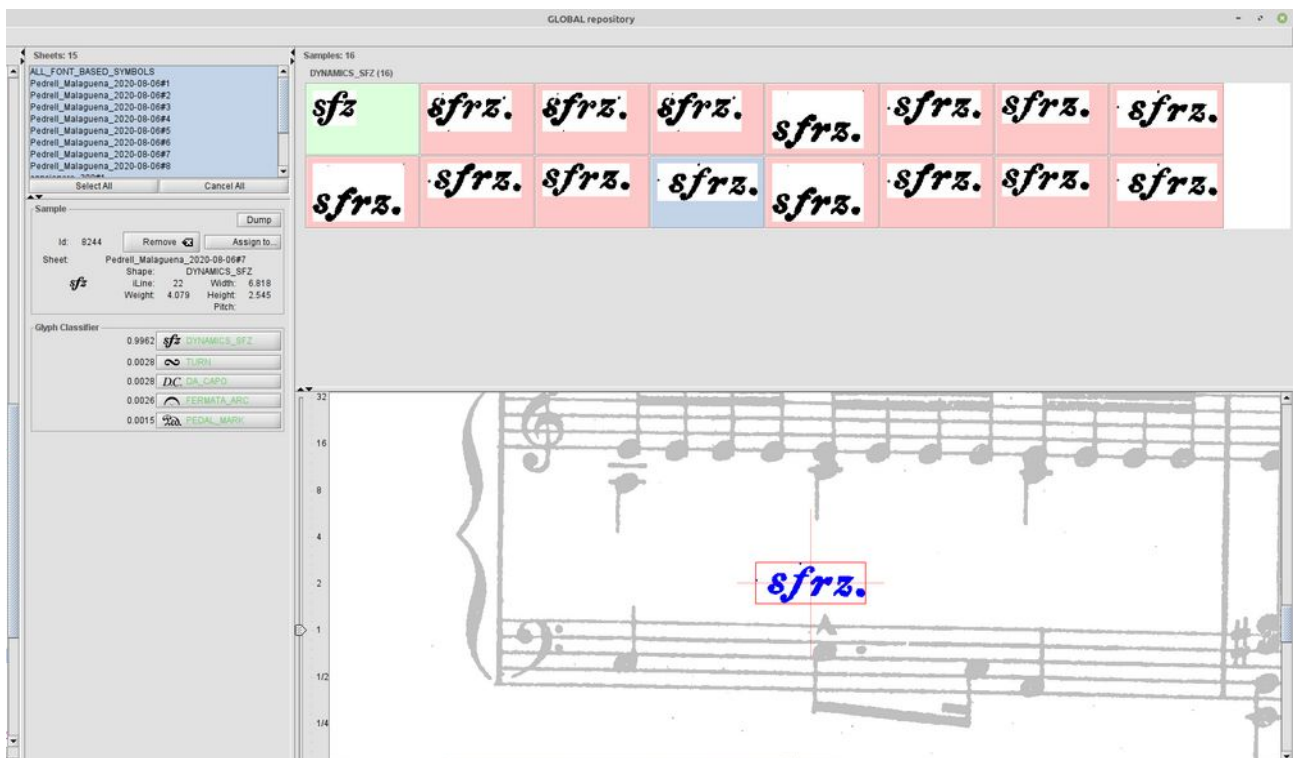
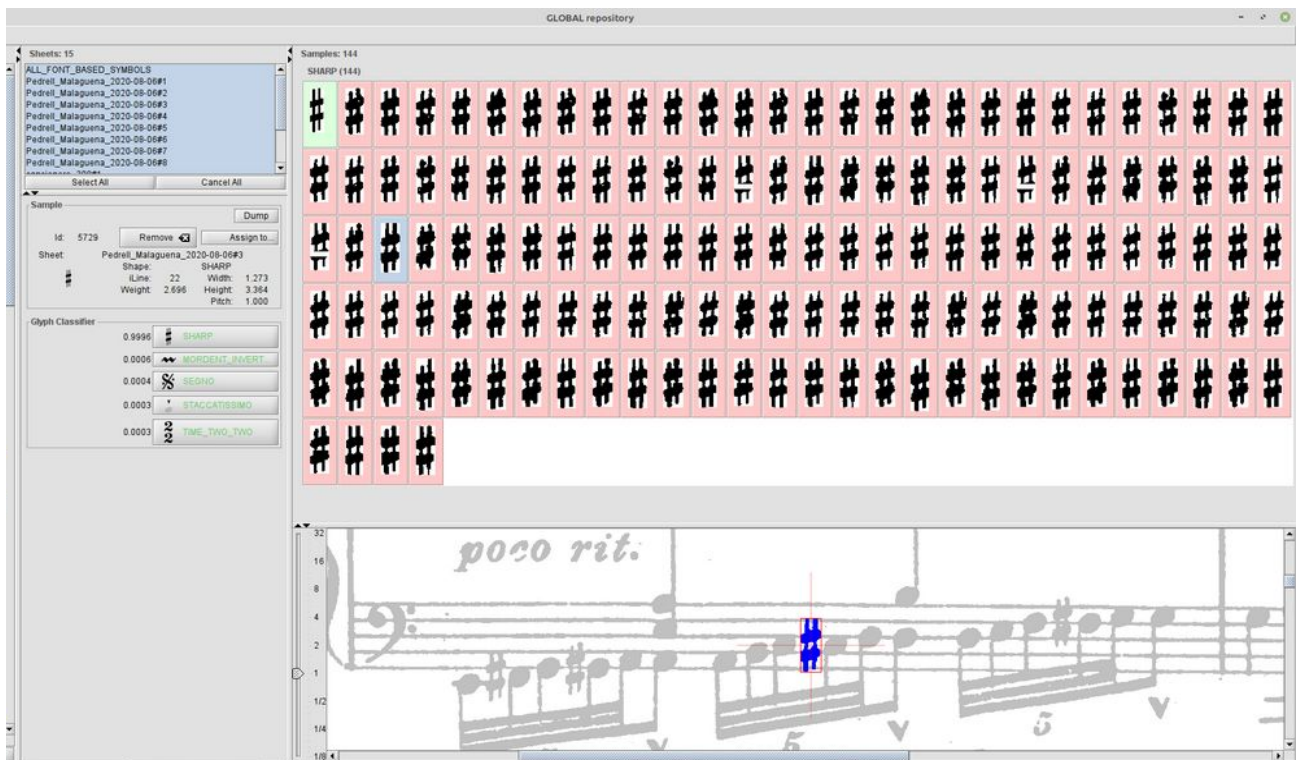
178

326)
Canço de batre

Ulldemolins (Tarragona) *Recogida por*
D. Higinio Anglés.

D'acies - tant veig u - naestre -
lla. — d'aci es - tant veig u - naes - tre - lla —
d'aci es - tant la veig
llu - ir — jo no
se si es l'es - tre lla — jo no
se si es l'es - tre lla —
o l'a - mor que'ns fa
pa - tir — ai — jo no se si
es — l'es - tre - lla. —

Annex B. Audiveris. Exemples de càrrega de glifs a repositori global



130 305)
Malagueña
(Andalucia)

Allegretto

p

cresc. e sfrz.

p

poco cresc.

p

Annex D. Prototip funcional en proves

Per millorar la resposta a un entorn real de proves el prototip no disposa d'un servidor IIIF associat per a proveir les imatges digitalitzades de la versió facsímil. Per aquest motiu es genera un manifest estàtic, produït per mitja de l'aplicació IIIF-Tiler allotjat en un servidor web públic⁴².

Instruccions per visualitzar la plataforma web:

1. Descomprimir el contingut de l'arxiu .zip a una carpeta
2. Posar en marxa un servidor local amb CORS per accedir a la plataforma⁴³
3. Obrir a un navegador l'adreça `http://http://127.0.0.1:[port seleccionat]/CancioneroDigital` (ex. `http://127.0.0.1:8000/`)

Alternativament s'ofereix una versió en línia provisional a: <https://ficus.maadix.org/CancioneroDigital/>

⁴² El manifest es troba allotjat a https://github.com/vidalsantos/iiif_test.github.io/blob/main/iiif/manifest/manifest_static_ranges.json

⁴³ Instruccions a:
IIIF Technical Workshop: IIIF Technical Workshop. 2018. Ronallo J. Setting up a simple local web server. [citat 7 juny 2021].
Disponible a: <http://ronallo.com/iiif-workshop-new/preparation/web-server.html>

