

100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la Biblioteca de Catalunya

Irene Valle de Lope - Paula Molina González - Gerard Vidal Santos

Paraules clau:

Patrimoni musical, Humanitats digitals, cuplet, Biblioteca de Catalunya, *Crowdsourcing*, *Optical music recognition (OMR)*,
Computer-aided musicology



Col·legi Oficial
de Bibliotecaris-
Documentalistes
de Catalunya



SEMINARI
D'EXPERIÈNCIES '20
EL FUTUR ÉS NOSTRE

100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

Resum

Els progressos i la democratització en l'accés a les eines electròniques per a músics han experimentat un augment considerable en els últims anys, tant en la diversitat com en la versatilitat d'escenaris. En aquest context digital, les biblioteques poden actuar amb el propòsit d'oferir serveis força improbables en anys anteriors. Amb motiu de l'efemèride del centenari del cuplet en català es va proposar una prova pilot amb aquest patrimoni musical a la Biblioteca de Catalunya.

L'objectiu principal d'aquest projecte va ser la conversió d'un conjunt de partitures d'aquest gènere al seu equivalent digital. D'aquesta forma, es van desenvolupar les següents tasques: l'exploració de la maduresa de tecnologies de reconeixement òptic musical (OMR), l'avaluació de programes d'edició musical de codi obert i la planificació tant d'escenaris com de productes digitals, els quals puguin ser d'ajuda en la promoció de la recerca i la difusió del patrimoni musical català.

Índex

1. Objectius i condicionants
2. Digitalització i OMR
3. Transcripció i edició de partitures
4. Transformació de formats
5. Exploració de sortides
6. Conclusions
7. Bibliografia
8. Programari

100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

1. Objectius i condicionants

L'**objectiu** del projecte va ser la realització d'una prova pilot per integrar eines digitals en el tractament de les col·leccions del centre i avaluar-ne, en la mesura que es pogués, el seu impacte sota els següents paràmetres:

- Avaluar la maduresa d'eines de reconeixement de notació musical (OMR) i eines d'edició de notació musical de codi obert
- Avaluar els processos de transformació a formats digitals (MusicXML, MIDI, MP3, PDF) per generar materials alternatius (com MusicXML o MIDI) a la seva versió física
- Explorar vies per donar a conèixer aquests tipus de formats alternatius a l'usuari

Aquests objectius es van establir a partir dels següents **condicionants**:

- La càrrega de treball no havia d'anar en detriment de les funcions que ja teníem assignades als nostres respectius departaments
- Les eines utilitzades en la prova no havien de generar cap cost addicional al centre
- El projecte havia de ser desenvolupat durant la nostra estada a la BC (març - desembre 2019)

100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

2. Digitalització i OMR



Original

A digital scan of the same musical score as the original, but with significant image processing. The text and musical notation are much clearer, with reduced noise and better contrast. The title 'LA BARBIANA' is centered at the top. Below it, 'Letra de Juan Planas (Couplet) Música de José Voltas Viñas' is printed. The music is on a grand staff with 'ALLEGRETTO' and 'VOZ' markings. The lyrics 'Cuan, de de pa. se. o por las ca. lles voy' are visible at the bottom.

B/N 300 dpi + *Despeckling* + Gaussian i Reenfocat

- Digitalització d'imatges i tractament de la imatge per millorar la detecció
- Taula d'avaluació indica una millora puntal en el reconeixement d'aspectes concrets de les partitures, cap d'elles sobresurt significativament de la resta

100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

2. Digitalització i OMR

LA SUERTE

Letra de Magl'n Tobella

COUplet

Música de J.

The image shows a musical score for piano accompaniment in 2/4 time. The tempo is marked 'Maestoso'. The score is divided into two sections. The first section, labeled 'Compàs mal formatat', shows a bar with a whole note and a half note, which is not a valid 2/4 bar. The second section, labeled 'Compàs formatat correctament (amb omissions)', shows a bar with a whole note, which is a valid 2/4 bar. The score includes various musical notations such as treble and bass clefs, time signature, dynamics (piano), and articulation marks.

Compàs mal formatat

Compàs formatat correctament
(amb omissions)

OMR no és OCR

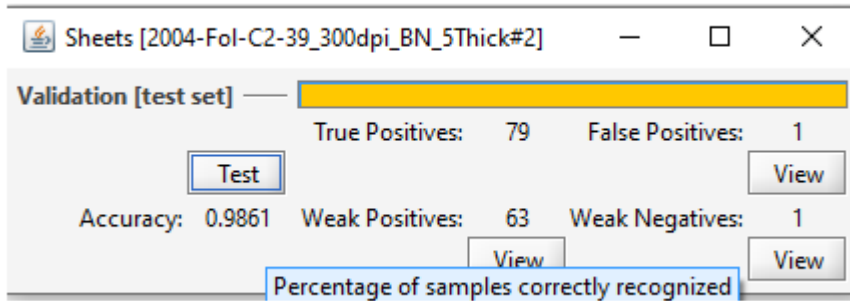
El fitxer de sortida (MusicXML) necessita una coherència interna que no se li exigeix necessàriament als fitxers de text generats per OCR

(* 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'n la reproducció en àudio

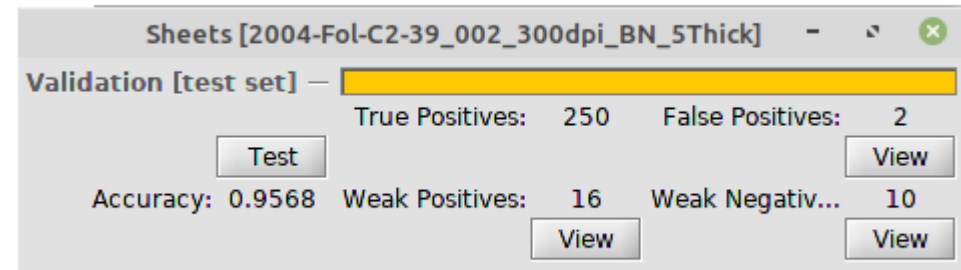
100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

2. Digitalització i DMR



Model 1 : Corpus per defecte



Model 2 : Corpus *ad hoc*

Procediment:

1. Correcció d'errors i omissions en la detecció
2. Incorporació de la correcció/ anotació assistida al model original
3. DMR segons model *ad hoc* permet al sistema detectar amb més encert cada partitura

100 anys de cuplet en català

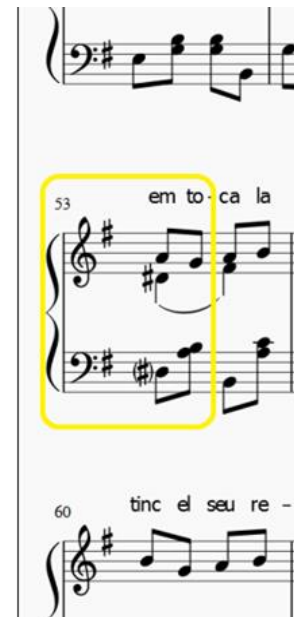
Prova pilot de dinamització digital dels fons musicals de la BC

3. Partitures impreses - Establiment de criteris d'edició

Els fitxers generats per OMR van necessitar correccions al programa d'edició musical (MuseScore 3.2.3), per generar la versió final en MusicXML

Criteris d'edició

1. Normalització del text, correcció d'errors ortogràfics, incorporació d'accentos quan calia així com la introducció de guions entre síl·labes quan no hi eren
2. Correcció d'errors harmònics, indicant la correcció entre parèntesis ()
3. Omissió dels símbols que no són a MuseScore, 3.2.3 (no s'ha intentat modificar aquells que posteriorment l'arxiu MusicXML no podia interpretar)
4. Manteniment de la separació dels ganxos de les figures rítmiques inferiors a la negra



Exemple de correcció d'errors harmònics

100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

3. Partitures Manuscrites - Establiment de criteris de transcripció i edició

Criteris de transcripció (*)

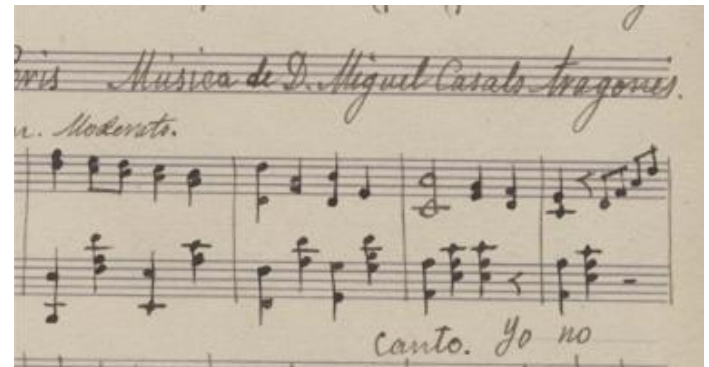
1. Manteniment de la distribució dels compassos i pentagrames
2. No normalització del text
3. No correcció d'errors harmònics.
4. Manteniment de la direcció de pliques i separació dels ganxos

Criteris d'edició

1. Modificació de la distribució dels compassos i pentagrames
2. Normalització del text, correcció d'errors ortogràfics, introducció de sinalefes i guions
3. Correcció d'errors harmònics, indicant la correcció entre parèntesis ()
4. Manteniment de la separació dels ganxos de les notes musicals en la part del cant, però no en la veu de l'acompanyament
5. Modificació de la direcció de les pliques d'acord amb els paràmetres actuals

OMR desestimat – transcripció manual

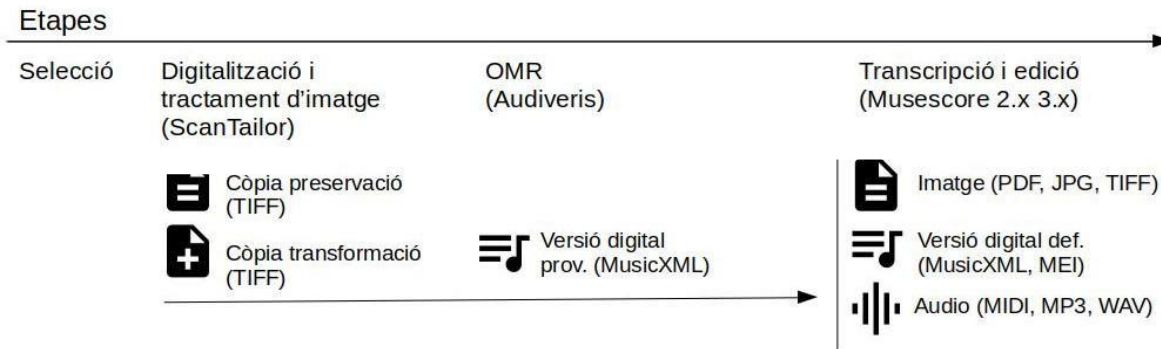
Com que els símbols de repetició dels originals no s'han pogut afegir al MusicXML, perquè la reproducció de la música sigui com s'indica vam canviar-los per aquells que ofería l'editor de música utilitzat



100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

4. Transformació de formats



Un cop generades les versions finals de les partitures impreses i manuscrites es poden generar perfils d'àudio, documents PDF amb criteris moderns d'edició i formats digitals per a la recerca musicològica

El ventall de formats de sortida permet pensar en noves formes per fomentar la difusió i la recerca del patrimoni musical

100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

5. Exploració de sortides per a la difusió i la recerca

Les possibilitats de recerca i difusió que se'n deriven de la transformació digital del patrimoni musical són múltiples. Per aquesta raó, aquest tipus de material digital obre la porta a:

- **La investigació:** pel que fa als formats digitals, a partir del MusicXML es poden exportar altres també necessaris en la investigació musical, com ara PDF, MIDI o arxius de so. Per exemple, formats d'àudio permeten la reproducció sonora dels continguts que en format físic eren menys accessibles al gran públic. També és possible, a partir del MusicXML, manipular i transformar la informació d'aquestes partitures amb programes com algun dels més de 250 softwares que s'ofereixen a aquesta pàgina: <https://www.musicxml.com/software>, i que poden ser útils a la comunitat científica, docent i divulgadora.
- **Crowdsourcing:** projectes de transcripció col·laborativa. Aquestes iniciatives suposen un alleujament de la feina, donat que la generació dels formats digitals esmentats pot convertir-se en tot un repte logístic i humà per als centres interessats en aquests, a causa de la complexitat tècnica i la càrrega de treball durant el procés. Segons la complexitat de la música o la notació, aquests projectes es podrien adreçar a un públic general (amb un editor senzill i intuïtiu) o a un públic especialitzat.

100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

6. Conclusions

- A la fi d'aquest projecte hem generat sis partitures (tres impreses i tres manuscrites) de cuplet en català del fons de la Biblioteca de Catalunya en els formats MusicXML, PDF i l'arxiu d'àudio WAV mitjançant eines de reconeixement de notació musical (Audiveris) i eines de notació musical de codi obert (MuseScore).
- Els dos principals problemes que ens hem trobat han estat el nul reconeixement dels caràcters de les partitures manuscrites per part dels programes de reconeixement òptic, i el poc temps per entendre i resoldre els problemes dels arxius generats (codificació del MusicXML i la seva transformació a MEI).
- Com a persones que es troben dins del món de les humanitats digitals, podem afirmar com a **conclusió** que l'oferta de fonts en nous formats digitals per part de les institucions suposen un avenç necessari i una ajuda considerable en els **serveis d'ajuda a la recerca, a la difusió del patrimoni, i al suport a la docència**. Ara bé, un projecte d'aquestes característiques demana un equip tecnològic i humà interdisciplinari, amb especialistes en digitalització, sonologia i musicologia computacional que sàpiguen fer front als problemes que la generació de documents musicals digitals pugui causar.

100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

7. Bibliografia

- Burghardt, Manuel; Spanner, Sebastian. "Allegro: User-centered Design of a Tool for the Crowdsourced Transcription of Handwritten Music Scores". Proceedings of the 2nd International Conference on Digital Access to Textual Cultural Heritage. 2017
- Bodleian Library. "What's the Score at the Bodleian?". [citat 29 abril 2020]. Disponible a: <https://whatsthescoreatthebodleian.wordpress.com/>
- Meier F.; Bazo A.; Burghardt M.; Wolff C. "A Crowdsourced Encoding Approach for Handwritten Sheet Music". Proceedings of the Music Encoding Conference (MEC) 2013 and 2014. 2015
- Merchán-Sánchez Jara, Javier. "Music Encoding Initiative (MEI), un estándar para la edición, codificación y la descripción de documentos musicales en el entorno digital". *Education in the knowledge society*. Núm. 20 (2019)
- Merchán-Sánchez Jara, Javier. *Partituras electrónicas: lectoescritura musical en el ámbito digital*. Barcelona: Editorial UOC, 2019

100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

8. Programari

- Bacchushlg. Audiveris [Internet]. GitHub. [citat 29 abril 2020]. Disponible a: <https://github.com/Audiveris/audiveris>
- Music Encoding Initiative. Tools [Internet]. [citat 29 abril 2020]. Disponible a: <https://music-encoding.org/resources/tools.html>
- MuseScore: Free music composition and notation software [Internet]. [citat 29 abril 2020]. Disponible a: <https://musescore.org/en4lex4>
- ScanTailor-Advanced [Internet]. GitHub. [citat 29 abril 2020]. Disponible a: <https://github.com/4lex4/scantailor-advanced>

100 anys de cuplet en català

Prova pilot de dinamització digital dels fons musicals de la BC

Agraïm a la Biblioteca de Catalunya i l'Oficina de Suport a la Iniciativa Cultural (OSIC) per haver propiciat aquesta prova pilot, especialment a Eugènia Serra, Rosa Montalt i Paquita Navarro pel suport donat durant el seu desenvolupament

Irene, Paula i Gerard

Desembre de 2020