

# INTEGRATED RECOGNITION SYSTEM FOR MUSIC SCORES

Artur Capela, Jaime S. Cardoso  
FEUP and INESC Porto  
Portugal

gcapela@inescporto.pt  
jaime.cardoso@inescporto.pt

Ana Rebelo  
FCUP and INESC Porto  
Portugal

arebelo@inescporto.pt

Carlos Guedes  
ESMAE and INESC Porto  
Portugal

carlosguedes@mac.com

## ABSTRACT

Many music works produced in the last century still exist only as original manuscripts or as photocopies. Preserving them entails their digitalization and consequent accessibility in a digital format easy-to-manage which encourages browsing, retrieval, search and analysis while providing a generalized access to the digital material. The manual process to carry out this task is very time consuming and error prone. Automatic optical music recognition (OMR) has emerged as a partial solution to this problem. However, the full potential of this process only reveals itself when integrated in a system that provides seamless access to browsing, retrieval, search and analysis. We address this demand by proposing a modular, flexible and scalable framework that fully integrates the abovementioned functionalities. A web based system to carry out the automatic recognition process, allowing the creation and management of a music *corpus*, while providing generalized access to it, is a unique and innovative approach to the problem. A prototype has been implemented and is being used as a test platform for OMR algorithms.

## 1. INTRODUCTION

The impact of music in our lives can hardly be overestimated. Music is a pivotal part of our cultural heritage and its preservation, in all of its forms, must be pursued.

Portugal has a notorious lack in music publishing from virtually all eras of its musical history. However, whereas most of the known original music manuscripts before the twentieth century are kept at the National Library Archive in Lisbon, there is virtually no national repository for the Portuguese music from the twentieth century. Although there are recent efforts in order to catalogue and preserve in digital form the Portuguese music from the late twentieth century—notably the Music Information Center [10] and the section on musical heritage from the Institute of the Arts website [8]—most of the music pre-dating computer notation software was never published and still exists as manuscripts or photocopies spread out all over the country in inconspicuous places.

For example, all the music composed by Jorge Peixinho (1940-1995), an internationally-renowned composer who epitomized the Portuguese *avant-garde* in the 1960s and 1970s, was never published in Portugal (few of his

scores were published abroad), and almost his entire *oeuvre* consists of manuscript paper [6]. Almost fourteen years past his death, his music is already catalogued, although not published. Unfortunately, this case is not unique, and this situation is common with other great Portuguese composers from the twentieth century. The risk of irreversibly losing this rich cultural heritage is thus a reality.

The project “Optical recognition system for handwritten music scores” initiated in 2007 by INESC Porto and ESMAE is the point of departure for creating a web-based system of music manuscripts of Portuguese composers from the twentieth century. This database will provide generalized access of a wide *corpus* of handwritten unpublished music encoded in MusicXML that can be accessed remotely via the Internet. The database will not only centralize as much information as possible but will also serve to preserve this *corpus* in a way that is easily accessible for browsing, analysis, and ultimately, for performing this *repertoire*, therefore help keeping the Portuguese music alive.

The ambitious goal of providing generalized access to handwritten scores that have never been published has been severely hampered by the current state-of-the-art of handwritten music recognition. There are currently various commercial OMR software solutions [4, 16, 14] and a few open source solutions [1, 15, 2], but they are all offline standalone applications. The existing online archives of music scores [9, 5, 13] usually provide them in inadequate formats—usually only as the scanned score image—for retrieval or automatic analysis. These online archives are mere standard websites, without facilities for optical recognition, editing and searching through the scores’ musical content. The creation of an OMR system, integrating optical recognition, storage, search, browsing and downloading capabilities, while keeping the scores in their original format along with their digital counterpart, would therefore be extremely beneficial.

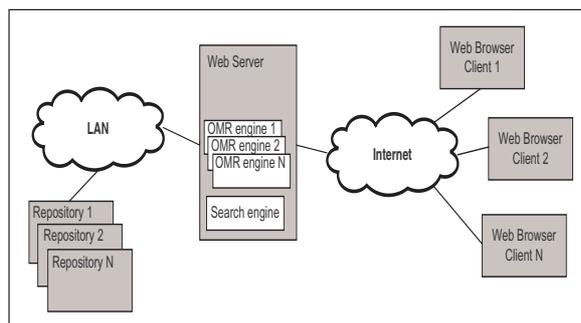
Using this background, we present the specification and implementation of a system integrating all the required features. It uniquely combines OMR technology in a system, easing the conversion of scores to the Music Extended Markup Language format [12]—MusicXML—as it is being widely adopted and meets our needs. In Section 2 the proposed system is described. We continue in Section 3 by presenting a usage scenario. Finally, conclusions are drawn in Section 4.

## 2. SYSTEM ARCHITECTURE AND IMPLEMENTATION

The system that we propose on this paper comprises the creation of a database of music scores and a web application mainly featuring:

- Addition of music scores to the system, performing their recognition and conversion to MusicXML in an integrated manner, allowing the user to confirm and correct the conversion results at the last stage of this process.
- Complete maintenance of a fully navigable music scores archive, including both the original version and the digital version obtained from the optical recognition.
- Browsing and searching the database, as well as the MusicXML contents. Visualization, downloading and edition of the selected music scores.
- Complete system management.

The architecture for the proposed system is based on a client-server model. The system is intended to be accessible through the Internet. There are three different entities present in this system, as it can be seen in Figure 1.



**Figure 1.** Generic system architecture

The *Repository* module stores the original scanned score, the digital counterpart in MusicXML and all the descriptive metadata inserted by the user, as detailed latter. All the remaining system contents, such as the user information, are also stored in this entity.

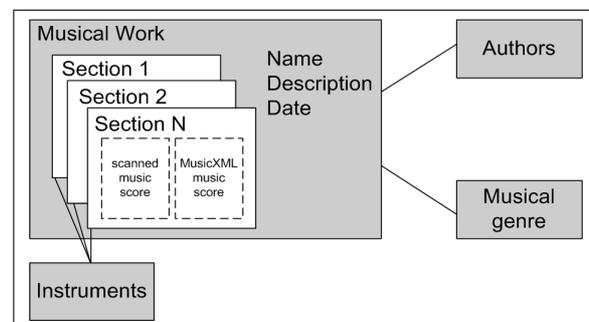
The *Web Server* is the user access point to the system as well as to all of its processing modules run on the server, encompassing the search engine and the optical recognition engine for the music scores. There is support for the inclusion of several OMR Engines, aiming to provide the ability to meet different needs (e.g. different music notation systems). The most adequate OMR Engine can be chosen manually by the user or automatically by the system, by detecting the scores notation and type (i.e. handwritten or printed). Our Search Engine allows not only generic searches throughout all of the system contents, but it also provides the capability of searching throughout the music scores MusicXML information in an innovative manner. The Web Server interacts with the Repository and

with the Web Browser, which establishes the interface between the user and the system.

The user interface on a *Web Browser* allows the complete management of the music scores and associated metadata, as well as carrying out the system administration. Generally speaking, the user interface provides the user the ability to execute all the necessary tasks to fully use the proposed system. On the administration side it is possible to manage the users, as well as the whole system contents and validate new ones.

There are four user types which can access the system: General User, Registered User, Privileged User and the Administrator. The General User represents a visitor and may only consult and download contents. The remaining types are registered users and according to their level they may add/edit/remove certain contents with or without restrictions. The Privileged User is similar to the Administrator and has full access to all functionalities, though it cannot manage the users from its own level. The contents added by Registered Users have to be validated by Privileged Users or the Administrator, although the later two are able to add any contents without the need to be validated, they are considered to be trustful users.

For content management, functionalities include the addition of music scores to the Repository, their automatic recognition, visualization and edition, searching, and browsing. It is also possible to insert and browse information related to the music scores—name, authors, instruments, musical genres—providing the user with a Repository containing all the necessary information to keep a complete music *corpus*. This metadata can then be used by the user on search queries or for a more flexible browsing experience. Finally, a music work is organized into sections, where each section is a music score which represents a part from the whole music work. This flexible structure allows accommodating either simple or complex works on the system. Each score can be visualized and its representation in MusicXML edited, side-by-side with the original score directly on the Web Browser. Both the visualization and the edition of the music scores are done in a graphical easy-to-use editor available through the Web Browser. Figure 2 illustrates the information recorded in the system associated with a music work.



**Figure 2.** Content stored in the Repository associated with a music work

## 2.1. Prototype Implementation – OMRSYS

We developed a prototype—OMRSYS—taking the system architecture shown in Figure 1 as a basis. Currently, the prototype supports only a single Repository collocated with the Web Server and a single OMR Engine was used to prove this concept.

The Repository is implemented as a PostgreSQL<sup>1</sup> database, an open source Database Management System (DBMS). The main reasons for choosing PostgreSQL were its native XML support, needed for creating a search engine that allows searching on the scores' MusicXML counterparts stored on the database. This is an important aspect as it is a major feature for the proposed system. It is also a very mature and well documented popular DBMS. Another important feature is the great ease of integration on the framework we have selected to develop the prototype, which we will describe next. MySQL<sup>2</sup> is the default DBMS used on the framework chosen for developing the Web Application, but it lacks on XML support. Other DBMSs were also considered but PostgreSQL was the one that best fitted our needs.

The development of the Web Application was supported on Ruby on Rails<sup>3</sup>. Rails is an open source, full-stack framework for developing dynamic database-backed web applications according to the Model-View-Control (MVC) pattern. It is an almost complete platform which requires only a DBMS and a server. These features present us with a suitable choice for supporting the development of our prototype. Ruby is the programming language at its core, a flexible and powerful Object Oriented language, with a vast array of powerful characteristics. Another strong advantage is the database manipulation, as it is greatly simplified, which is ideal to develop a system of this kind. Other frameworks [17, 7] were considered but fell short compared to the features of Ruby on Rails.

The Web Server selected for the prototype was the Apache HTTP Server<sup>4</sup> with Mongrel<sup>5</sup> to execute the Web Application. This solution is suggested by Ruby on Rails and was found suitable, being both efficient and open source.

The Search Engine developed for this prototype allows the usual queries on the database contents, although the groundbreaking search through the MusicXML contents is still not possible at this stage.

The OMR Engine on the Web Server is the module responsible for the automatic recognition of the submitted music scores. We initially adapted it from the open source OpenOMR project [15]. The OMR Engine performs an automatic conversion of a submitted music score to a digital easy-to-use representation, the MusicXML format. This digital format allows representing sheet music by its musically-relevant parts, sections, phrases and motives, thus easing the access to the relevant portions of the score while browsing that score in a computer monitor.

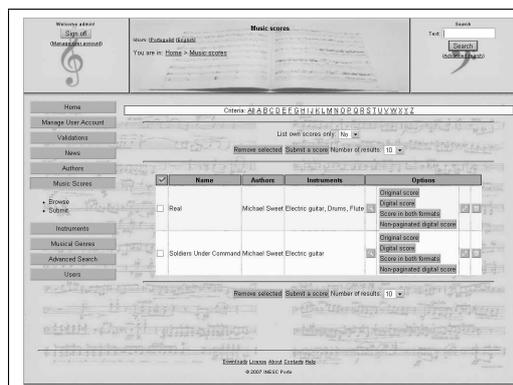


Figure 3. The OMRSYS user interface

Simultaneously, MusicXML enables the retrieval of relevant musical information for analysis, thus facilitating certain types of computational analysis to be performed on a *corpus* of scores. Nevertheless, it also provides an adequate way to restore old sheet music, preventing them from oblivion. The other OMR applications we have analysed were left aside because they were either less complete than OpenOMR at the time or they were commercial solutions. Our main goal at this stage was to prove the concept.

The User Interface has a great impact in the user experience and is divided into several sections. There is the authentication, title and quick search sections on the upper portion of the screen, followed by the middle and largest portion which includes the main menu and the contents area. The main menu works as a two-level expansible menu and allows access to all the system's functionalities by grouping them in a logical manner. Similar functionalities follow similar designs to keep the interface consistent, intuitive and easy to learn. Some of the functionalities are the common Create/Read/Update/Delete (CRUD) and the listing of the database contents based on a chosen criteria, all following a familiar behaviour. The main differences and most unique aspects rely on the submission and the update of music scores, which is discussed in Section 3. As an interface example, Figure 3 shows the Graphical User Interface (GUI) being used for browsing the music scores available in the Repository. Both the Privileged Users and the Administrator have additional options in the main menu for validation purposes.

The MusicXML Editor for this prototype is still implemented as a plain text editor and viewer, but already showing side-by-side the original score with its MusicXML counterpart. However, the development of a fully integrated graphical MusicXML editor is being pursued. Such editor would allow a higher level and intuitive edition and could be developed for example in Flash, in the likes of MusicRain [11], an online interactive sheet music viewer. The main purpose for the editor in this initial prototype was to give the end-user the possibility to at least view and edit the music scores in MusicXML. The existing editors and visualizers [11] usually have a high maturity level but they are offline applications.

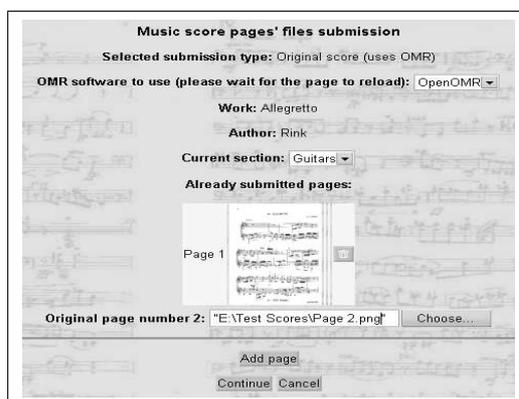
<sup>1</sup> <http://www.postgresql.org>

<sup>2</sup> <http://www.mysql.com>

<sup>3</sup> <http://www.rubyonrails.org>

<sup>4</sup> <http://httpd.apache.org>

<sup>5</sup> <http://mongrel.rubyforge.org>



**Figure 4.** Score submission scenario

The Digital Rights Management control is done in two ways: the acceptance of a license agreement at the registration process and the validation of submitted music scores by a Privileged User before they become available on the system.

### 3. USAGE SCENARIO: SCORE SUBMISSION

When the insertion of a music score in the system is requested, the user inserts the metadata associated with the music score, of which some is optional—name, year, description, etc—and associates it with one or more authors and a musical genre. Each section of a music work has to be associated with the instruments present on the music score. In the last submission step, after the insertion of the requested metadata, the user submits the various pages for each section, as illustrated in Figure 4.

After validating the inserted data, the user then triggers the automatic recognition process by calling a suitable OMR engine. Afterward, an overview of the submitted score is shown by listing its contents allowing the user to view the result of the automatic process on the built-in editor side-by-side with the original scanned image, offering the user the possibility to manually correct the automatic results. After confirming the results and making the necessary corrections, the user then finalizes the music score submission by accepting it. If the score was submitted by the Administrator or a Privileged User it then becomes immediately available on the system; if it was submitted by a standard Registered User it is kept on queue for validation and will only become available once a user with administration privileges validates it.

### 4. CONCLUSION

The proposed system offers a complete solution for the preservation of our musical heritage. It includes an optical recognition engine integrated with an archiving system and a user-friendly interface for searching, browsing and edition. The digitized scores are stored in MusicXML, a recent and expanding music interchange format designed for notation, analysis, retrieval, and performance applica-

tions. An additional benefit of the automatic conversion of the music score to MusicXML is the possibility of encoding the manuscript score in MX format, an XML-base, multi-layered format for music representation [3]. MX synchronizes several layers belonging to the description of a piece of music, e.g. an audio recording and score of the same piece.

A system of this kind promotes the creation of a full *corpus* of music documents, promoting its preservation and study. This project will culminate in the creation of a repository of the handwritten scores, accessible online. The database will be available for enjoyment, educational and musicological purposes, thus preserving this *corpus* of music in an unprecedented way.

### Acknowledgments

This work was partially funded by Fundação para a Ciência e a Tecnologia (FCT) - Portugal through project PTDC/EIA/71225/2006.

### 5. REFERENCES

- [1] AOMR2. [Online]. Available: [http://www.bzzt.net/~arnouten/wiki/index.php/Gamera#AOMR2:\\_omr\\_toolkit](http://www.bzzt.net/~arnouten/wiki/index.php/Gamera#AOMR2:_omr_toolkit)
- [2] Audiveris. [Online]. Available: <http://audiveris.dev.java.net>
- [3] A. Baraté, G. Haus, and L. Ludovico, "An XML-based format for advanced music fruition," in *Proceedings of the Third Sound and Music Computing Conference*, 2006, pp. 141–147.
- [4] Capella-scan. [Online]. Available: <http://www.capella-software.com/capscan.htm>
- [5] Classical Sheet Music and MIDI Files. [Online]. Available: <http://www.music-scores.com>
- [6] Delgado, Cristina, Machado, Jorge, and J. Machado, *Catálogos da obra de Jorge Peixinho*, ser. José Machado (ed.) Jorge Peixinho: In Memoriam. Lisbon: Caminho, 2002.
- [7] Google Web Toolkit. [Online]. Available: <http://code.google.com/webtoolkit>
- [8] Institute of the Arts. [Online]. Available: <http://patrimonio.dgartes.pt/?lang=pt>
- [9] The Lester S. Levi Collection of Sheet Music. [Online]. Available: <http://levysheetmusic.mse.jhu.edu/>
- [10] Music Information Center. [Online]. Available: <http://www.mic.pt>
- [11] MusicRain. [Online]. Available: <http://musicrain.us>
- [12] The MusicXML Format. [Online]. Available: <http://www.musicxml.org>
- [13] The Mutopia Project. [Online]. Available: <http://sca.uwaterloo.ca/Mutopia>
- [14] OMeR. [Online]. Available: <http://www.myriad-online.com/en/products/omer.htm>
- [15] OpenOMR. [Online]. Available: <http://sourceforge.net/projects/openomr>
- [16] SharpEye Music Reader. [Online]. Available: <http://www.visiv.co.uk>
- [17] Tacos. [Online]. Available: <http://tacos.sourceforge.net>