



Project Acronym: Europeana Sounds
Grant Agreement no: 620591
Project Title: Europeana Sounds

MS10 Evaluation of Linking Music to Scores Pilot

Revision: Final

Date: 02/03/2015

Authors: Alexander Schindler (AIT)

David Haskiya (EF)

Abstract: This document summarises the findings of the evaluation of the linking music to scores pilot. The evaluation comprises user, content and technical perspectives based on which it provides a conclusion as to the potential of integrating score following in the Europeana infrastructure and Music Channel.

Dissemination level	
Public	X
Confidential, only for the members of the Consortium and Commission Services	



Coordinated by the British Library, the Europeana Sounds project is co-funded by the European Union, through the ICT Policy Support Programme as part of the Competitiveness and Innovation Framework Programme (CIP) http://ec.europa.eu/information_society/activities/ict_psp/



Revision history

Version	Status	Name, organisation	Date	Changes
0.1	ToC, 1st draft	David Haskiya, EF	24/02/2015	
0.2	2nd draft	Alexander Schindler, AIT David Haskiya, EF	25/02/2015	Technical & content requirements, collections investigated, prototype write-up
1.0	Final draft	Alexander Schindler, AIT David Haskiya, EF	28/02/2015	Input from PMB
1.1	Final	Richard Ranft, BL	02/03/2015	Layout, minor changes

Review and approval

Action	Name, organisation	Date
Reviewed by	PMB	27/02/2015
Approved by	Coordinator and PMB	02/03/2015

Distribution

No.	Date	Comment	Partner / WP
1	02/03/2015	Submitted to the European Commission	BL/WP7
2	02/03/2015	Posted on Europeana Pro website	BL/WP7
3	02/03/2015	Distributed to project consortium	BL/WP7

Application area

This document is a formal output for the European Commission, applicable to all members of the Europeana Sounds project and beneficiaries. This document reflects only the author's views and the European Union is not liable for any use that might be made of information contained therein.

Statement of originality

This document contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Project summary

Europeana Sounds is Europeana's 'missing' fifth domain aggregator, joining APEX (Archives), EUscreen (television), the Europeana film Gateway (film) and TEL (libraries). It will increase the opportunities for access to and creative re-use of Europeana's audio and audio-related content and will build a sustainable best practice network of stakeholders in the content value chain to aggregate, enrich and share a critical mass of audio that meets the needs of public audiences, the creative industries (notably publishers) and researchers. The consortium of 24 partners will:

- Double the number of audio items accessible through Europeana to over 1 million and improve geographical and thematic coverage by aggregating items with widespread popular appeal such as contemporary and classical music, traditional and folk music, the natural world, oral memory and languages and dialects.
- Add meaningful contextual knowledge and medium-specific metadata to 2 million items in Europeana's audio and audio-related collections, developing techniques for cross-media and cross-collection linking.
- Develop and validate audience specific sound channels and a distributed crowd-sourcing infrastructure for end-users that will improve Europeana's search facility, navigation and user experience. These can then be used for other communities and other media.
- Engage music publishers and rights holders in efforts to make more material accessible online through Europeana by resolving domain constraints and lack of access to commercially unviable (i.e. out-of-commerce) content.

These outcomes will be achieved through a network of leading sound archives working with specialists in audiovisual technology, rights issues, and software development. The network will expand to include other data-providers and mainstream distribution platforms (Historypin, Spotify, SoundCloud) to ensure the widest possible availability of their content.

For more information, visit <http://pro.europeana.eu/web/europeana-sounds> and <http://www.europeanasounds.eu>

Copyright notice

Copyright © Members of the Europeana Sounds Consortium, 2014-2017. This work is licensed under the Creative Commons CC-BY License: <https://creativecommons.org/licenses/by/4.0/>.

Contents

Executive summary: MS10 Evaluation of Linking Music to Scores Pilot	5
1 Background	5
2 Evaluation	5
2.1 User experience analysis	6
2.1.1 User scenarios	6
2.1.2 Matching user scenarios with the personas	6
2.1.3 Wireframe	6
2.1.4 Prototype score following player	7
2.2 Content requirements and analysis	9
2.2.1 Technical challenges	9
2.3 Content availability in Europeana	14
2.3.1 Existing collections in Europeana	15
2.4 Integration potential in Europeana and the Music Channel	16
3 Cost Benefit Analysis	16
4 Conclusions	17
4.1 Recommendations for future work	17
5 References	18
Appendix A: Terminology	18

Executive summary: MS10 Evaluation of Linking Music to Scores Pilot

This document summarises the findings of the evaluation of the linking music to scores pilot. The evaluation comprises user, content and technical perspectives. The conclusion of the pilot evaluation is that while score following would be an attractive feature in the Music Channel the current collections available in Europeana which meet the requirements for score following are too few to motivate developing the feature in the portal or Music Channel. It is currently unknown whether the efforts of Europeana Sounds will radically change that and it suggested that the evaluation is followed up once the project has made substantial ingestions into Europeana.

1 Background

In Europeana Sounds Work Package 2, Task T2.4 *Innovative exploration pilots* specifies that the project will develop innovative discovery pilots for audio related content based on Music Information Retrieval (MIR). In the first sub-task, T2.4.1. *Linking music to scores*, it is specifically defined that a pilot linking sheet music to corresponding music recordings will be developed.

As a pilot it is not a set goal to develop score-following to a state of production ready software. Instead it is to develop a pilot that can be used as a basis of evaluation to decide whether further resources should be invested in developing a production-ready version. The state-of-the-art of the score-to-audio alignment research task is evaluated for its applicability in the Europeana infrastructure for ingestion flow, data infrastructure and Portal/Channels publication. This document summarises the evaluation which has been carried out from three perspectives: the user experience perspective, the content availability perspective and the perspective of technical integration.

Note that in this document the score following pilot is evaluated from the perspective of integrating it in Europeana's technical infrastructure and making the feature available in the planned Music Channel. The score following solution could also conceivably be implemented by Europeana Sounds partners on their own collection databases and websites. Each such integration scenario would be unique to each partner and collections management system and is out of scope/budget for this evaluation.

2 Evaluation

The evaluation comprises the user perspective, the content perspective and the technical integration perspective.

2.1 User experience analysis

2.1.1 User scenarios

The user scenarios are based on discussions with practicing beginner and advanced musicians within the network.

Scenario 1: Beginner musician who wants to learn how to play a piece of music

For a beginner musician who wants to practice or rehearse playing a piece of music, score following can be a very valuable tool.

Scenario 2: Advanced musician who wants to analyse/review a long musical work

For an advanced musician the value of score following would be as a support tool in reviewing or analysing a longer piece of music. The value would lie in the ability to quickly “jump” both in the sheet music and in the sound recording, while both remain in sync.

The advanced musician, or composer, could also be interested in the MusicXML¹ file produced by feature extraction process necessary for the score following. The MusicXML file can be imported into score-writing programs and sequencers for editing and processing.

Unlike the beginner musician, for the advanced musician score following would not be useful for practice or rehearsal purposes, as they would not be interested in closely copying the performance of another musician.

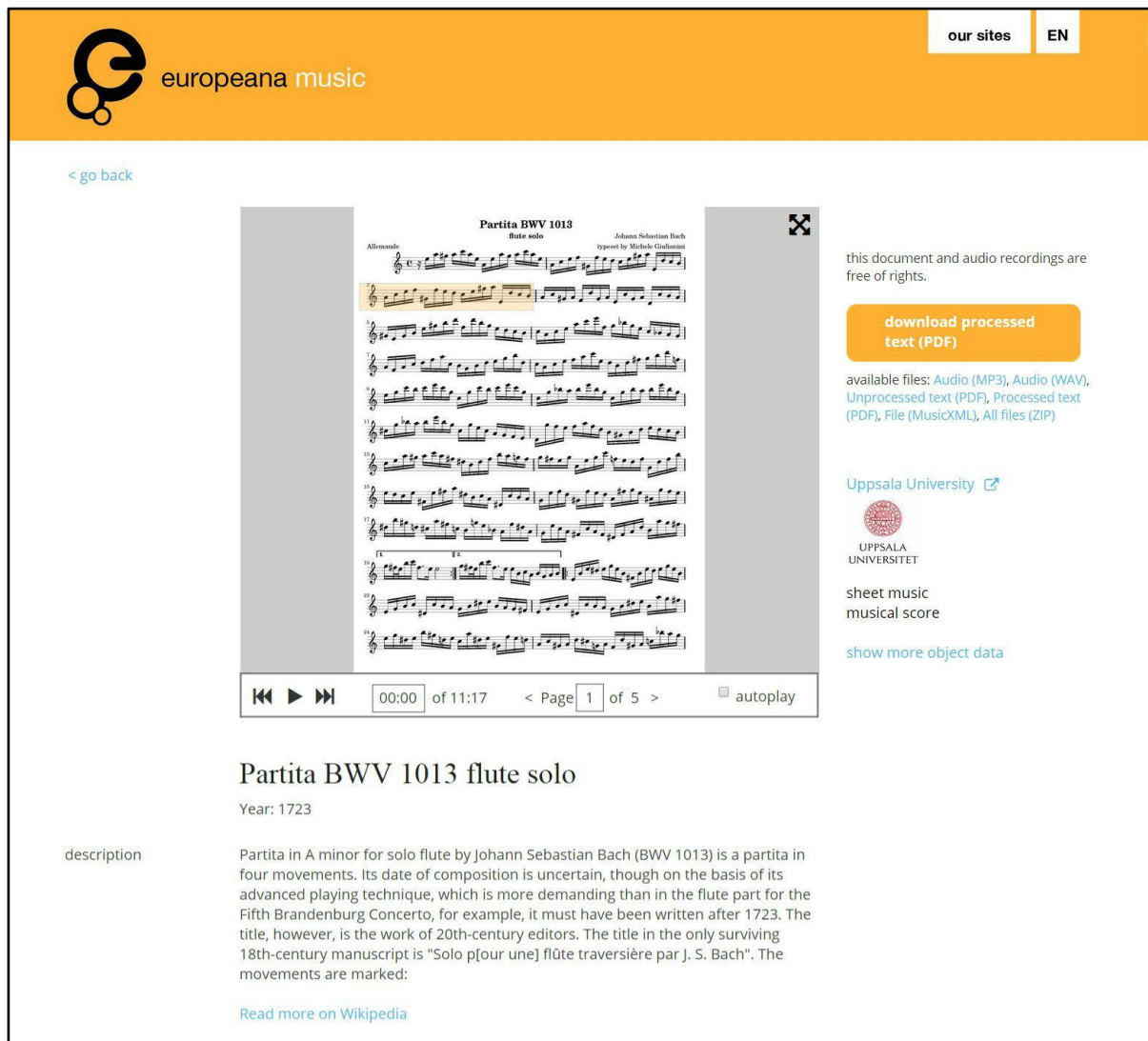
2.1.2 Matching user scenarios with the personas

Based on the use cases above, Europeana’s UX Designer reviewed the use cases against the personas developed in Work Package 4 *Channels Development* (Ref 1). Scenario 1 matches well with the Marion persona, who is described as an amateur musician. While none of the personas are described as advanced musicians there is nothing in the underlying user research that would rule out features aimed towards advanced musicians. Both personas could conceivably be advanced musicians.

2.1.3 Wireframe

A simple wireframe has been created that shows how a “playable score” could be displayed. The wireframe is in the same medium-fidelity style as in *MS20 Second Audio Channels Prototype* (Ref 1).

¹ [MusicXML](#) is an industry standard exchange format for scores.



our sites EN

Europeana music

< go back

Partita BWV 1013
flute solo
Johann Sebastian Bach
Arranged by Michele Giulianini

download processed text (PDF)

available files: Audio (MP3), Audio (WAV), Unprocessed text (PDF), Processed text (PDF), File (MusicXML), All files (ZIP)

Uppsala University

UPPSALA UNIVERSITET

sheet music
musical score

show more object data

00:00 of 11:17 < Page 1 of 5 > autoplay

Partita BWV 1013 flute solo

Year: 1723

description

Partita in A minor for solo flute by Johann Sebastian Bach (BWV 1013) is a partita in four movements. Its date of composition is uncertain, though on the basis of its advanced playing technique, which is more demanding than in the flute part for the Fifth Brandenburg Concerto, for example, it must have been written after 1723. The title, however, is the work of 20th-century editors. The title in the only surviving 18th-century manuscript is "Solo p[our une] flûte traversière par J. S. Bach". The movements are marked:

[Read more on Wikipedia](#)

Figure 1. Screenshot of a wireframe displaying the score-following player incl. download options.

As part of the score-following, the current movement through the musical work played is highlighted in the sheet music. The player is provided with controls both for the audio recording (play, pause, fast forward, fast backward, time code, autoplay) and the sheet music (pagination, full screen).

The various files produced during the feature extraction needed to support score-following are made available for download. While simple, the wireframe hence supports both use cases listed above.

2.1.4 Prototype score following player

An [overview of playable prototype examples](#) has been created in which you can play the sound recording with score following.

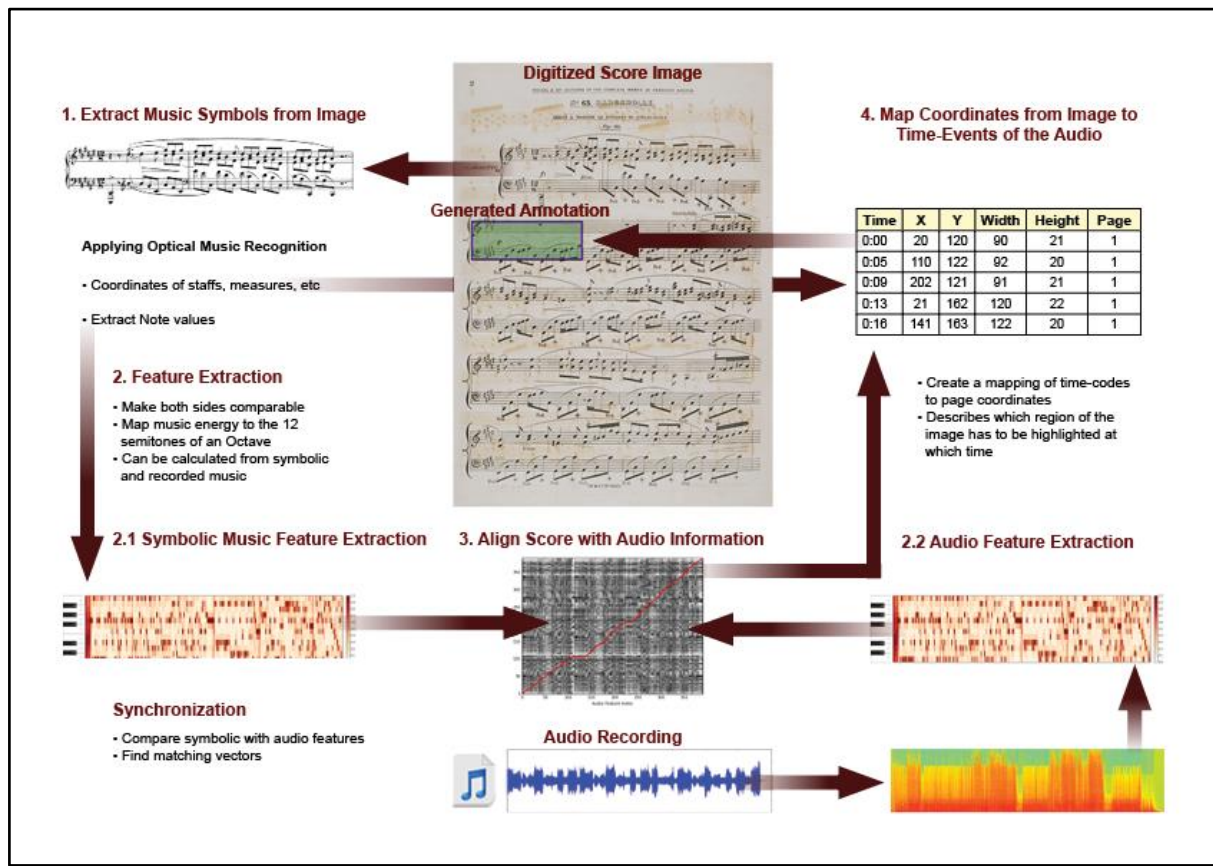


Figure 2. An overview of the process behind score following.

An extensive description of the technological background of audio-to-score alignment will be provided in the forthcoming D2.3 *Linking Music to Scores delivery report*. The required workflow as depicted in Figure 2 is subdivided into the following tasks:

- 1. Symbolic music information extraction:** In this step the symbolic description of the music is extracted from the music sheet images. This includes the semantic description (e.g. notes, measures, clefs, etc.) and the page layout (e.g. coordinates of the staves, measures, systems, etc.).
- 2. Feature extraction:** In terms of audio processing, symbolic music differs greatly from recorded audio. The two music representations have to be converted into a comparable form which is accomplished through mapping the music energy to the twelve semitones of the musical octave. This procedure is applied to both forms in small subsequent time intervals, providing two lists of audio features.
- 3. Audio to score alignment:** The previously calculated lists of features can be used to align the score information to the recorded audio. The feature vectors are compared and most similar consecutive sequences are identified.
- 4. Audio to image mapping:** The indexes of the aligned features are back-referenced to their page numbers and coordinates in the score image as well as to the corresponding time-events of the audio file.

5. **Visual representation of the results:** To visualize the computational results an HTML5 Webpage was created. The required functionality to show the music sheet images, interact on audio events, highlight the currently played measure and handle page turns appropriately, was implemented in JavaScript.

2.2 Content requirements and analysis

A reliable implementation of a score-following feature can only be realized if **both** of the following two content related criteria are met:

- Access² to high-quality scans of printed³ sheet music for a particular musical work, that are processable by optical music recognition systems.
- Access to a music recording of the same musical work as represented in the sheet music

The score following feature would then have what it requires to match sheet music print to the corresponding music recording. Please note that the two objects do not need to originate in the same collection - one could come from institution X and the other from institution Y.

2.2.1 Technical challenges

Optical Music Recognition (OMR) is a relatively young research topic compared to Optical Character Recognition (OCR). While recent OCR systems are mature and capable of detecting and extracting text from digitized images of low quality, OMR systems still require higher quality images as input. The following paragraphs summarize the most severe problems concerning the application of OMR to scans of old and low quality music scores:

General image content restrictions

Suitable quality (for OMR) can be considered as images of music sheets using modern staff notation with contemporary fonts and symbols. Staff lines are straight and horizontally aligned. The sheet is cropped and centred to show only the content of the score, having all overlapping space of the scan removed. Music scores curated and provided by national libraries are usually older. Piano or orchestral music printed before 1900 in particular creates many OMR problems. The following content type is currently unsupported or only supported by highly specialized software - most of them were developed in other research projects and require considerable adaptation effort.

- **Handwritten scores:** Contrary to OCR where some handwritten content can already be recognized, no OMR system is currently capable of processing non-printed sheet music.
- **Early music notation:** Specialized versions of OMR are available to process early typographic music. The most advanced system is ARUSPIX⁴ which is also evaluated in the Europeana Cloud project. An

² In the context of Europeana this means a direct link to to the high resolution scan in edm:isShownBy (as a PDF) or edm:hasView (as a sequential series of JPGs or PNGs).

³ Hand-written or older pre-18th century print notations are not supported due to the much higher difficulty in extracting the notation.

⁴ <http://www.aruspix.net/>

integrated system providing the capability to process multiple types of music scores is currently not available.

Quality of the paper

A set of definitions and two models concerning document image quality and degradation are provided by Baird [Ref 3]. Balk and Contech [Ref 4] summarise the finding of the IMProving ACcess to Text (IMPACT) project, which focused on the development of new approaches to the extraction of text content from historical documents. Thirty-seven characteristics which can affect OCR performance were identified, including bleed-through, stains, page, curl, broken characters, low contrast, skew, and presence of watermarks. All of them have the same influence on OMR performance.

- **Bleed-through:** caused by degradations of ink and paper, the content of the adjacent page becomes visible. Binarization is a typical pre-processing step to remove such artefacts. Based on the dimension of the effect, specialized approaches might be required. The provided example depicts the problems of strong bleed-through artefacts. While the binarized image still contains fragments of the undesired content, necessary information such as staff lines, which are fundamental to the OMR extraction process, is destroyed.



- **Stains, scratches and watermarks:** Missing or additional content on music sheets cannot be interpreted correctly or will provide erroneous results.
- **Low contrast:** Low contrast becomes a problem in combination with other artefacts such as bleed through. Pre-processing steps to remove these artefacts also degrade the quality of low contrast scores. The example image shows eight notes with bleached ink. Binarization can remove the inner regions of the notes which makes them look similar to semi notes.



Quality of the scans

- **Layout detection:** Experiences from evaluating OMR engines in task 2.4.1 showed that most systems have problems with additional content in scans such as black borders which typically arise at non-page regions of scans. Consequently, digitized scores showing these kinds of artefacts require either manual pre-processing or the implementation of an automatic cropping algorithm.



- **Image rectification:** Images are often distorted non-linearly, especially in bow areas of books. Built-in algorithms to locate lines of text such as implemented in OCR systems are still ongoing research.

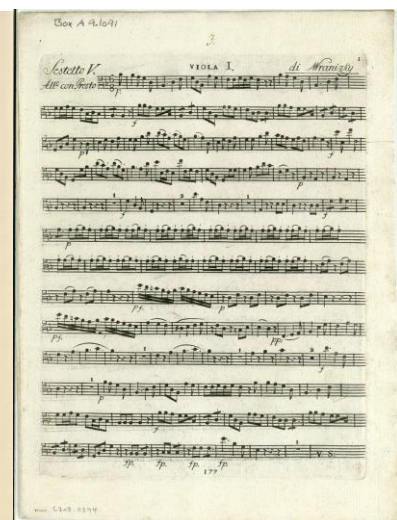


Quality of the scores

- **Atypical fonts:** OMR systems are optimized to printed music sheets, preferably printed with a standard computer font which is available since the late 1960s. Older scores for example use different symbols for clefs and also the layout of the notes differs.
- **Manual annotations:** Manual annotations of historically important persons are worth to preserve, but in terms of OMR they fall into the same category as stains and scratches.
- **Complexity of the score:** Piano music is often quite complex and dense, which makes it hard for OMR systems. Orchestral scores are often small, resulting in not enough pixels for distinguishing different symbols.



- **Examples of digitized scores of appropriate quality:**



Quality of audio recordings

The problems of sub-optimal audio quality are less severe than those of erroneous extracted music information from images. Urbano et al. [Ref 5] analysed the effects of audio quality on Chroma features

and concluded that they appear to be robust against variations from different sampling rates, codecs and bitrates. Mauch and Ewert [Ref 6] evaluated the effect of audio degradation on different music information retrieval tasks including audio-to-score alignment. They developed a Matlab toolbox to artificially apply degradation effects to audio files, thus facilitating controlled experiments. Simulated degradation effects included:

- **Live recordings:** Echo of the room and pink noise of the audience
- **Radio broadcast:** Dynamic Range compression, Speedup (commonly applied to music in mainstream radio)
- **Smartphone playback:** pink noise
- **Smartphone recording:** Dynamic Range Compression (auto-gain effect), Clipping and pink noise
- **Strong MP3 compression:** 64kbps
- **Vinyl:** Crackling sounds, wow-and-flutter, light pink noise

The authors based the evaluation on a similar approach as evaluated in Task 2.4.1. They used a simplified setting without optical music recognition. Symbolic music was available in high quality from MIDI files. For the evaluation the Saarland Music Data set [Ref 7] was used, consisting of 50 piano tracks played on a Yamaha Disklavier MIDI piano.

Based on the results of the experiments the authors concluded that the applied audio-to-score alignment method is generally robust against audio degradation effects, except for the categories *Live* and *Smartphone Playback*.

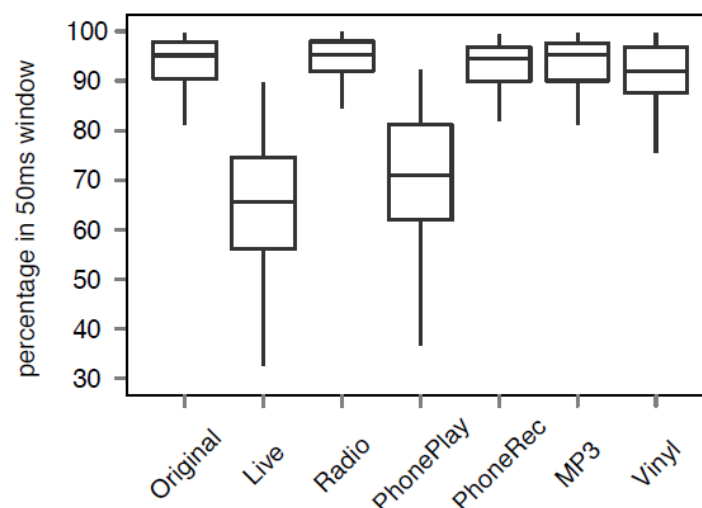


Figure 3. Score-to-audio alignment accuracy under the applied audio degradation effects. The boxes indicate the 1st, 2nd (median) and 3rd quartiles, the whiskers extend to ‘the most extreme data point which is no more than 1.5 times the interquartile range’ (Source: Ref 6)

While the latter category might not be of concern for Europeana Sounds content, audio properties described by the *Live* category are more likely to be present. Based on these results a general reliable

applicability of the automated mapping approach evaluated in Task 2.4.1 cannot be stated and further adaptations to normalize audio recordings towards their quality should be considered.

Error categories

The following table lists some issues that are expected to arise frequently using Europeana content for score-following.

Table 1: expected OMR and audio error categories

Type	Description	Consequence	Severity
OMR	Measure boundary not detected	The sum of note values exceeds the expected time signature. This might lead to erroneous alignment and incorrect visualization.	Middle
	Systems not detected correctly	Systems interprets that the staff lines of a system are independent and should be played in consecutive sequence instead of being played in parallel. This leads to misalignments due to the additional measures.	High
	Repeat sign not detected	Leads to misalignments due to missing measures. If complete score has to be repeated and the sign is missed, the complete audio has to be mapped to half of the measures.	Very High
	Note lengths are misinterpreted	The sum of note values varies from the expected time signature. This might lead to erroneous alignment and incorrect visualization.	Middle
	Note values are misinterpreted	This leads to deviations in the music features and might affect correct alignment. Because the features are normalized, a small amount of misinterpretations are neglectable.	Minor
	Missing staff lines or measures	Leads to misalignments and wrong timing in the visualization.	Middle
	Additional staff lines or measures	Attempting to map non-existing content to recording leads to misalignments.	High
	Wrong time signature	Sum of note lengths within measures does not correspond to the detected time signature. Might result in erroneous results.	High
	Wrong clef or wrong key / accidentals	Leads to wrong calculation of music features. The feature values are shifted. If the clef is globally misinterpreted, the alignment will still work.	Medium
Audio	Constant noise from old records	Affects audio feature calculation and mapping. If noise remains the same for the whole recording the mapping will still work.	Low
	Varying noise	Affects audio feature calculation and mapping. Might lead to wrong mapping.	Middle

	Audio watermarks	Artificially altered content used to protect property rights. Will lead to wrong mapping.	High
	Fragmented audio, no overlaps	Such as two sides of a record. Additional effort has to be provided to implement proper handling of audio concatenation and partitioning of the score	Low
	Fragmented audio with overlaps	Such as magnetic tape recordings with redundant overlaps. Additional effort has to be provided to correctly align the truncated audio pieces to combine them into a complete recording.	High
	Fragmented audio, missing content	Will lead to misalignments. Additional effort has to be planned to implement methods to identify missing content and correctly annotate these regions in the score images.	High

2.3 Content availability in Europeana

At the time of writing this evaluation it is unknown how much content the Europeana Sounds consortium will provide to Europeana that meets the content requirements for successful score following. The existing inventories made in WP1 [Ref 2] are not sufficiently detailed to make an estimate at the time of writing this evaluation. For existing collections in Europeana, the numbers are very low to non-existent (see below). Future work on evaluating the content (to be) aggregated by Europeana Sound against the requirements for score following will need to happen in collaboration with WP1.

From the perspective of Europeana and integration in the music channel making the investment in developing, integrating and maintaining the score following feature would only be meaningful if we know that we would have matching objects (where there are sheet music objects with matching music recordings) in the high hundreds or low thousands. If suitable content is available only in smaller numbers it would be economically more sensible to feature score following only in manually created exhibitions or similar.

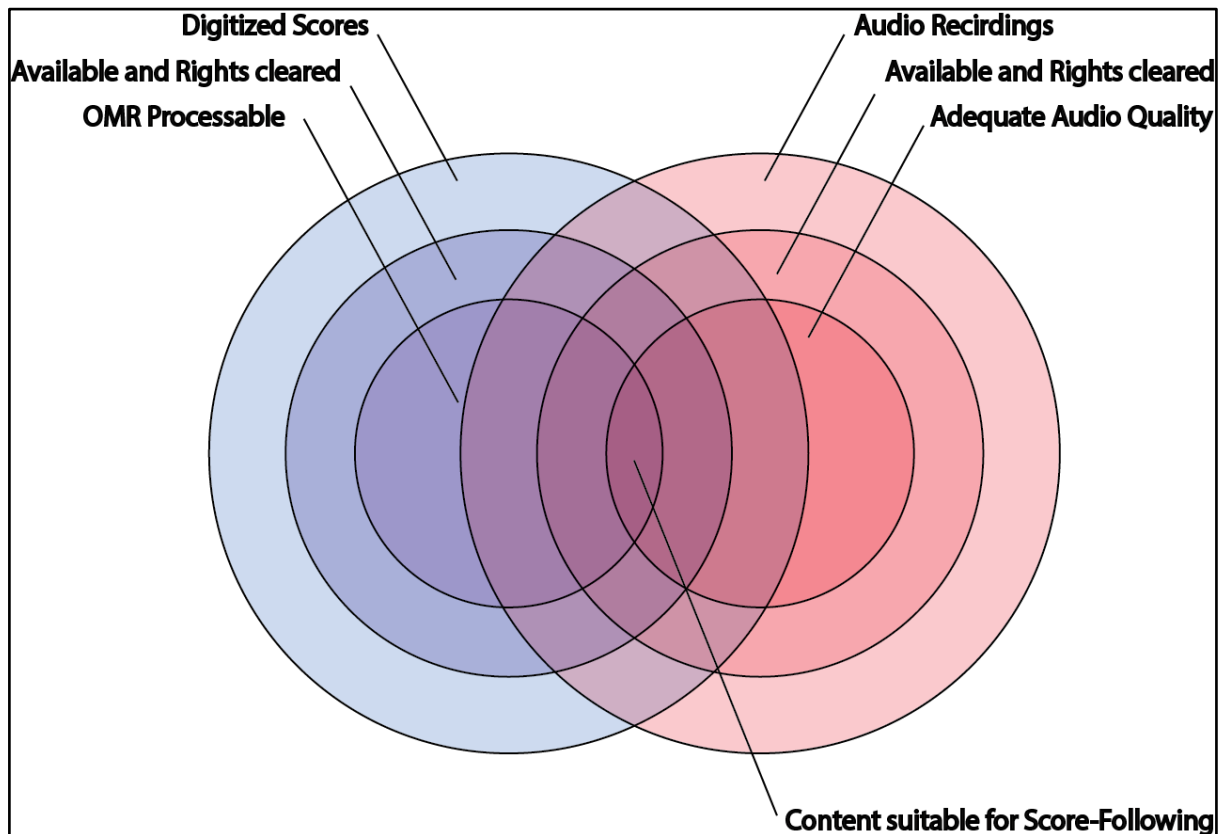


Figure 4. Venn diagram of content requirements that all need to be met for successful score-following.

2.3.1 Existing collections in Europeana

During the evaluation of the score following pilot we attempted to find existing collections of sheet music in Europeana⁵ that could be suitable for OMR. Note that this research has not been exhaustive but is indicative as to how difficult it is to find sheet music in Europeana that meets all requirements (see Figure 4).

List of sheet music collections identified and investigated:

- [Berlin Library collection of First World War era sheet music](#). In terms of the quality of the scans and direct access to the scans this is the collection we identified that is most suitable for OMR. However, we have found no corresponding music recordings to match the sheet music.
- [Central Institute for the Union Catalogue of Italian Libraries collection of First World War era sheet music](#). The quality of the scans is good but the pixel dimensions of the images is sometimes too small. As with the Berlin State Library collection we have found no corresponding music recordings to match the sheet music.
- [Danish National Library collection "Gieddes Samling"](#). The quality of the scans in this collection are borderline for OMR and is likely to depend on the individual sheet music. The musical pieces are

⁵ Exploratory search for locating [sheet music](#). Note the use of the word for sheet music in multiple languages

quite well-known so there is at least a small chance that it has corresponding music recordings in Europeana.

These three collections are the only collections of sheet music identified so far where the institution has provided to Europeana direct links to scanned images or PDF. While many other institutions have provided sheet music collections they have not provided direct links to the scanned images or PDF files, making it impossible to OMR process the sheet music.

2.4 Integration potential in Europeana and the Music Channel

While a deployment architecture of the Score Following Pilot has not been formalised, we will make the assumption that the back-end modules of the Score Following Pilot would need to be deployed centrally in the Europeana infrastructure. While it is theoretically possible to install the Score Following Pilot in the collection management systems of all relevant partners in Europeana Sounds, this would require significant additional development effort for each such partner.

A detailed overview of all the technical component/libraries used in the Score Following Pilot will be provided in D2.3. In this document we will not evaluate the technical choices or quality of the Score Following Pilot, but simply evaluate it from the perspective of how easy or difficult it would be to deploy in the Europeana infrastructure.

Note that this is a “desktop evaluation”. Europeana has not spent any effort on hands-on prototyping of integrating the Score Following functionality into the core infrastructure.

3 Cost Benefit Analysis

On the **cost side** if the equation it is estimated that the development investment needed to integrate the score following feature in the Europeana infrastructure is significant. Not so much in the front-end of the Music Channel but in its integration with the data ingestion and data enrichment pipeline.

Continual operational costs would be noticeable as Europeana does not have the domain specific skills needed to continually keep the solution at levels of best practices within the field of Music Information Retrieval. While Europeana is not accustomed to maintaining Python based software this would be of comparatively smaller concern than the lack of domain expertise.

On the **benefit side** our analysis of user needs places score following as non-essential to the needs of our users. Essential features are all discovery related: search, facet, browse and retrieve⁶ are the fundamentals. While not essential to the functioning of the Music Channel as a discovery tool it can be viewed as an “attractive quality”⁷ that sets the Music Channel apart from other music discovery services.

The big unknown on the benefit side is that we do not know how much content would be available to fuel it. For the feature to be meaningful to users, scores that can be played with integrated score

⁶ See the personas Marcel and Marion.

⁷ See the [Kano product development model](#)

following need to be easy to find and numerous (i.e. covering a substantial number of musical works). Current analysis indicates Europeana currently has too little suitable content.

4 Conclusions

The conclusion of the pilot evaluation is that current collections in Europeana that meet the requirements for score following are too few to motivate developing the feature for the Music Channel. It is currently unknown whether the efforts of Europeana Sounds will radically change that.

Score-following fits the defined users and use cases of the Music Channel well and is an attractive feature. However, the cost benefit analysis shows that the cost of development, integration and operation pushes it down the overall backlog of features to develop for the Music Channel. The uncertainty in regards to whether the Sounds consortium have and are able to make available meaningful amounts of content that meet the criteria further lowers the priority of the feature.

4.1 Recommendations for future work

The conclusion of this evaluation should be revisited when Europeana Sounds have ingested substantial amounts of data in Europeana.

Concrete future tasks:

- Once the Sounds project has published large amounts of audio related content to Europeana **make an estimate of the amounts of sheet music/music recordings that potentially meet the content requirements**
 - Investigate with the data providing partners in WP1 if they can modify their plans for data provision to include more sheet music and with direct links to digital files
- **Investigate alternative deployment architectures**
 - for example, setting up the service as an online service ([SaaS](#)) which can be used by both Europeana, all Europeana Sounds partners, other institutions/users with holdings of sheet music or even individual users.
 - Individual institutions may wish to evaluate the Score Following Pilot for potential integration in their own collection managements and publications systems
- Investigate whether Europeana can acquire substantial amounts of sheet music with corresponding music recordings by **harvesting the crowdsourced [IMSLP/Petrucci Library](#)** and other crowdsourced repositories of sheet music and music recordings.
- Apart from this milestone and deliverable D2.3 the Score Following Pilot will be **documented for publication** on [Europeana Labs](#). Europeana can also host the source code on its GitHub account.

Main lessons learned:

- The usage of the score following pilot on a larger scale and its usefulness **depends on the availability of appropriate content in terms of format, size and quality**. The highest limitation encountered during the evaluation is the lack of printed music scores using the classical music notation, which have the visual quality required by the OMR tools.
- Having **direct links to files representing scanned sheet music/scores and music recordings is essential** to the *Music Information Retrieval* Task 2.4 and this should be more strongly reflected in the aggregation strategy and mapping guidelines of the project.

5 References

Ref 1	MS20 Second Audio Channels Prototype http://pro.europeana.eu/web/europeana-sounds/documents/-/document_library_display/0Hv5/view/2034829
Ref 2	D1.5 Aggregation Report 1 http://pro.europeana.eu/web/europeana-sounds/documents/
Ref 3	H. S. Baird, "The state of the art of document image degradation modelling," in <i>Digital Document Processing</i> , pp. 261-279, Springer, 2007
Ref 4	H. Balk and A. Conteh, "Impact: centre of competence in text digitisation", in <i>Proc. Workshop on Historical Document Imaging and Processing, HIP '11</i> , (New York, NY, USA), pp. 155-160, ACM, 2011.
Ref 5	Urbano, J., Bogdanov, D., Herrera, P., Gómez, E., & Serra, X. (2014). What is the effect of audio quality on the robustness of MFCCs and chroma features. In <i>International Society for Music Information Retrieval Conference</i> .
Ref 6	Mauch, Matthias, and Sebastian Ewert. "The Audio Degradation Toolbox and Its Application to Robustness Evaluation." <i>ISMIR</i> . 2013.
Ref 7	M. Mueller, V. Konz, W. Bogler, and V. Arifi-Mueller. Saarland music data (SMD). In <i>Late-breaking session, Intl. Society for Music Information Retrieval Conf. (ISMIR)</i> , 2011.

Appendix A: Terminology

A project glossary is provided at: <http://pro.europeana.eu/web/guest/glossary>.

Additional terms are defined below:

Term	Definition
AB	Advisory Board
AIT	Austrian Institute of Technology
APEX	Archives Portal Europe network of excellence
EF	Europeana Foundation

EC-GA	Grant Agreement (including Annex I, the Description of Work) signed with the European Commission
PMB	Project Management Board
TEL	The European Library
UAP	User Advisory Panel
WP	Work Package