D6.2 – First version of the TELMI Platform

Project acronym	TELMI
Grant Agreement nr	688269
Project title	Technology Enhanced Learning of Musical Instrument Performance
Call identifier	H2020-ICT-2015
Topic	ICT-20-2015
Type of action	Research and Innovation Action
Start date of project (dur.)	February 1st, 2016 (36 months)
Report availability	Public
Document due Date	July 31th, 2017
Actual date of delivery	July 31th, 2017
Leader	UNIGE
Reply to	Gualtiero Volpe [UNIGE] (gualtiero.volpe@unige.it)
Additional main contributors (author's name / partner acr.)	Corrado Canepa, Giorgio Gnecco, and Erica Volta [UNIGE], Oscar Mayor [UPF], Ayman Moghnieh [SAICO]
Document reference	TELMI-D-WP6–UNIGE-20160504-D6.2 FirstVersionOfTheTELMIPlatform_v0.2
Document status	Final

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

Executive Summary

Deliverable D6.2 presents the first version of the TELMI platform, released at the end of the first half of the project (Month 18). The document consists of two major parts. The first one introduces the overall logical architecture of the TELMI platform, as defined in Deliverable D6.1 (Month 6), and briefly discusses each of its components. The second one illustrates the first prototype of the platform, showing which modules were implemented and to which extent. Finally, the future work planned for the second half of the project is discussed.

The TELMI platform is available on the TELMI website (http://telmi.upf.edu), in the section devoted to software outputs of the project.

The platform will be demonstrated at the second public event of the TELMI project (Deliverable D7.4, Month 20), which will be held in Reykjavik (Iceland) during the International Symposium on Performance Science (ISPS2017, 30 August – 2 September 2017).

To understand this document the following deliverables have to be read.

Number	Title	Description
--	Description of Work	Annex 1 to the Grant Agreement
D2.2	Definition of exercises and record of expert models of success	Description of the recordings carried out to create the TELMI reference archive of violin performances, to be used with the platform.
D3.2	First working prototypes of the measuring systems	Description of the first prototype of measuring systems, providing details on the recording and analysis components of the TELMI platform.
D4.2	Metrics for evaluation of skill performance and progress assessment	Description of the features computed and analysis techniques implemented in the first version of the TELMI platform.
D5.2	Initial design of the gamified TELMI platform with MVP	Description of the front-end of the TELMI platform, consisting of the TELMI app.
D6.1	Definition of Use Cases, Requirements, and Specifications of the TELMI Platform	Definition of the overall logical architecture of the TELMI platform, and of the requirements for each of its components.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

Table of contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

Table of figures

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

1. Introduction

Deliverable D6.2 presents the first version of the TELMI platform, released at the end of the first half of the project (Month 18). The final version of the platform will be released towards the end of the project at Month 30 (Deliverable D6.3).

The document consists of two major sections. Section 2 introduces the overall logical architecture of the TELMI platform and briefly discusses each of its components. The TELMI platform was defined in Deliverable D6.1 (Month 6) and is here refined on the basis of the work carried out in the first half of the project. Section 3 illustrates the first prototype of the platform, showing which modules were implemented and to which extent. Finally, conclusions in Section 4 discuss future work planned for the second half of the project.

The TELMI platform collects and integrates software modules developed within the research and development work-packages of the TELMI project, namely WP3, WP4, and WP5. Hence, this deliverable presents an overview of the platform and refers the reader to the relevant deliverable of WP3, WP4, and WP5 for more details about the developed techniques.

From an implementation point of view, the first prototype of the TELMI platform is made by four major interconnected software components:

- The *EyesWeb XMI Platform* (UNIGE), especially customized for the purposes of TELMI project. This includes specific EyesWeb modules (blocks) and applications (patches) developed within TELMI.
- The *ViolinRT app* by UPF, a real-time play-along app able to record a user performance and carry out audio analysis giving visual feedback to the user. This app is integrated as the PlayTools in the HIGHSKILLZ TELMI App.
- The *TELMI app* by HIGHSKILLZ, representing the front-end for users and implementing the student interface and session management.
- The *repoVizz repository* (UPF), where multi-modal performance data recordings from experts are archived (https://repovizz.upf.edu).

The TELMI platform is available on the TELMI website (http://telmi.upf.edu), in the section devoted to software outputs of the project.

The current version of the TELMI platform along with some demos of how it can be used will be publicly demonstrated in the framework of the second public event of the project (Deliverable D7.4, Month 20), which will be held in Reykjavik (Iceland) during the International Symposium on Performance Science (ISPS2017, 30 August – 2 September 2017).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

2. Architecture of the TELMI platform

2.1 Different solutions for different targets

Following the requirements in Deliverable D6.1, the TELMI platform is conceived as a remote repository located in remote servers, possibly in the cloud and a collection of clients, connected via network with the remote repository, and adapted to the single conditions (e.g., the physical environments, available devices, and computational resources) envisaged in the use cases. The prototypical fully-featured TELMI client is thus deployed in three major different solutions. These mainly differs on the input devices that are supported:

- *High-end solution*: this includes top-level and highly expensive motion capture systems, such as Qualysis, which can be available at research centres and very high-end cultural and educational institutions. The high-end solution of the fully-featured TELMI client enables and supports a very accurate analysis of performance data and is mainly intended for research purposes and professional users.
- *Intermediate solution*: this includes mid-level motion capture systems, which are less expensive than the previous ones, such as e.g., Polhemus. The intermediate solution of the fully-featured TELMI client enables and supports accurate analysis of performance data at a cheaper cost with respect to the high-end solution. This is mainly intended for cultural and educational educations and also for research purposes.
- *Low-end solution*: motion capture systems are replaced with affordable RGB-D sensors (e.g., Microsoft Kinect or Intel Real-Sense) or, possibly, video cameras (e.g., webcams). The low-end solution of the TELMI client enables and supports analysis of performance data, which is less accurate than in the other two cases, but still good enough to get valuable feedback and affordable for most users, including violin students.

The three different solutions include audio capturing and processing. Moreover, clients developed for specific purposes may implement only a subset of the functionalities of the TELMI platform.

All the clients share an underlying general logical architecture and are conceived as instances of such a general architecture. Such overall platform architecture is described in the next section.

2.2 Platform architecture

Figure 1 shows the overall logical architecture of the TELMI platform, as defined in Deliverable D6.1 (Month 6). On the top, a remote repository in the cloud is represented, including an archive of exercises, a reference archive of performances, and the archive of the performance sessions carried out by students with the platform. Clients are represented around the remote repository, the lower part of the figure representing the fully-featured TELMI client, which can be declined in its high-level, intermediate, and low-level solutions. Blue arrows in the figure represent streams of data the components of the platform exchange. Dashed green arrows represent parameters controlling the processing of single components of the platform. Wide orange arrows represent data from/to the users. Some connections are omitted for sake of clarity.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

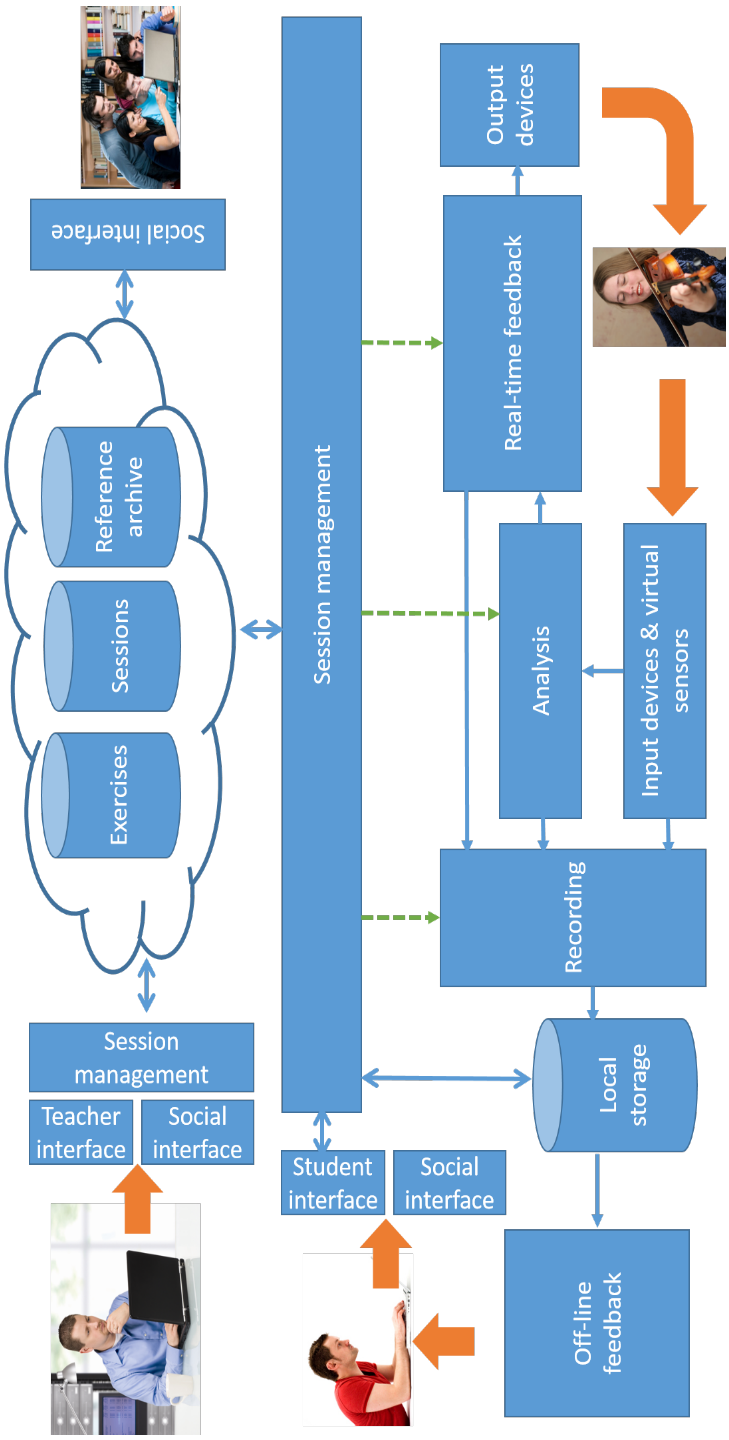


Figure 1. Overall logical architecture of the TELMI Platform.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

The role each single component of the platform has in the logical architecture and the set of functionalities it has to provide are briefly summarized in the following. A more detailed description is available in Deliverable D6.1.

- *Input devices and virtual sensors component*: this component is responsible for capturing and pre-processing of multimodal data. Data is captured by an array of virtual sensors, associated to a broad range of physical sensors, including motion capture, video cameras, microphones, and possible other sensors (e.g., IMUs and physiological sensors). Data is processed to get representations suitable for the recording component and for the data analysis component.
- *Analysis component*: this component is responsible for extracting features from the music performance data captured by the input devices and virtual sensors component. Analysis concerns audio, bow controls, body postures, and body movements and is organised on multiple layers. The output is a collection of values of features characterising the music performance under several aspects and at different time scales.
- *Real-time feedback component*: this component takes as input the values of the features computed by the analysis modules and generates a real-time feedback to be presented to the performer. Given the specific task of violin playing, feedback mainly exploits the visual channel in order not to interfere with the played sound.
- *Output devices component*: this receives the feedback representations produced by the real-time feedback component and generates the corresponding outputs on the selected output devices (ranging, e.g., from the screen of a tablet to a video projection on a wall).
- *Recording component*: this is responsible of the synchronised recording of all the data collected, computed, or generated during a violin performance with the TELMI platform. The recording component stores data in a local storage device and produces appropriate metadata to identify the performance (e.g., date, player, session, music piece, and so on). Consolidated standard formats are used to store the data.
- *Offline feedback component*: this component is responsible of providing the user (either student or teacher) with views on pre-recorded performance data. Views can consist of the audio, video, or audiovisual playback of a recorded performance and of possible complex representations of performance data (e.g., the drawing of the trajectory of the tip of the bow overimposed to the performance video). Data is retrieved from a local storage device, possibly after downloading it from the remote (cloud) repository.
- *Session management component*: This component is responsible of keeping track of the sessions users perform with the platform. It manages the creation of a new session, opening an existing session, closing an existing session, and association of metadata to sessions. The component communicates with the remote TELMI repository where material for exercises, data of existing sessions, and a reference archive of recorded performances are collected and stored. As such, the session management component is a kind of gateway between each local client of the TELMI platform and a remote TELMI server.
- *Teacher interface*: this component is used by teachers to manage sessions with their students. It enables teachers to create a new session for a student, to configure the session by selecting exercises / pieces to be performed, feedback to be provided by the system, data to be recorded, material to be displayed to the student.
- *Student interface*: this component is used by students to manage their sessions with the TELMI platform. Students can open a session created by their teacher, check which exercises / pieces should be performed, fine-tune and calibrate the TELMI platform so that it is customised to their specific needs, start and stop recording, upload performance data in the TELMI remote repository, review data of their current or past performances by

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

enabling views in the off-line feedback component. Students can also self-monitor their progresses and receive ratings and comments by their teacher.

- *Social interface*: This component is used by both teachers, students, and their peers to share data, materials, and experiences. Students can, for example, share selected violin performance data with their peers to evaluate each other performances. Teachers can share materials and sample performances, can provide advice to students, and can notify them communications. A teacher can also share data with other teachers to possibly get their comments and suggestions.

As an outcome of the refinement of the requirements for the TELMI platform, it emerged that a *Learning analysis component* could be useful to implement the metrics for assessing students' performances. This component can help also teachers to follow the advancements of students by providing parameters that will be displayed in the teacher interface. Even if not initially included in the original overall logical architecture, this component will be taken into account for the final version of the TELMI platform.

Next section briefly describes how and to what extend these components have been implemented in the first version of the TELMI platform.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

3. First prototype of the TELMI platform

Given the overall logical architecture presented in Section 2, this section describes in some more detail the first prototype of the TELMI platform, illustrating the components that were implemented. The components collect and integrate software developed in the research and development work-packages of the TELMI project, namely WP3, WP4, and WP5. Here, we provide a brief description of each component and we refer the reader to the relevant deliverables of WP3, WP4, and WP5 for a more detailed discussion of the algorithms and of their implementation.

3.1 Input devices and virtual sensors component

The input devices and virtual sensors component has been fully implemented. This component integrates software modules mainly developed in WP3, and supports at the moment the following input channels:

- High-quality motion capture system (Qualysis)
- Intermediate-quality motion capture system (Polhemus)
- RGB-D devices: Microsoft Kinect
- Multichannel audio (either microphones or violin pick-ups)
- Video (both professional video cameras and webcams)
- IMU sensors (e.g., X-OSC)
- EMG sensors (Myo).

The clients of the three different solution of the TELMI platform are endowed with an input devices and virtual sensors component, which could support only a subset of the above mentioned input channels. In particular:

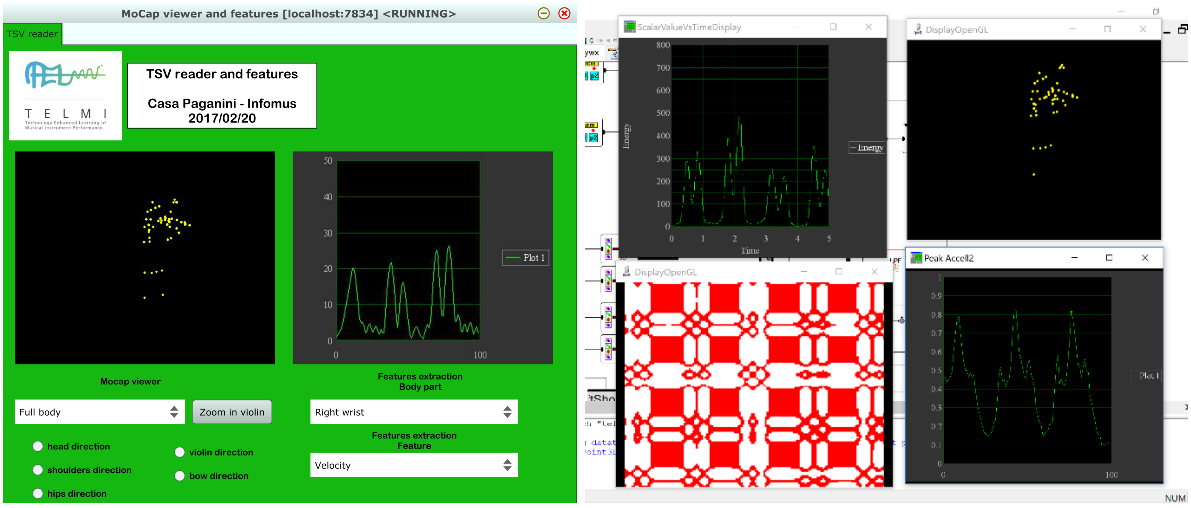
- The *high-end solution* includes a high-quality motion capture system (e.g., Qualysis) and all the other input channels (excluding the intermediate-quality motion capture system).
- The *intermediate solution* includes an intermediate-quality motion capture system (e.g., Polhemus) and all the other input channels (excluding the high-quality motion capture system).
- The *low-end solution* includes an RGB-D device (Microsoft Kinect) and stereophonic audio (with possible extension to multichannel audio).

3.2 Analysis component

Analysis concerns audio, Analysis concerns audio, bowing, body postures, and body movements. The software developments integrated in this component mainly come from WP3 and WP4. The list of features, which are currently computed, is available in Deliverable D4.2 (Metrics for evaluation of skill performance and progress assessment). The same Deliverable D4.2 describes such features in more details. Just to provide a brief summary, regarding audio analysis, this is performed in real-time and some meaningful descriptors are extracted like pitch and energy. From these descriptors, we can extract high level features like tuning, timing, articulation and dynamics. With respect of analysis of body movements and posture, the set of features available in the different solutions of the TELMI platform are partially different due to the different motion capture devices involved (namely a motion capture system vs. an RGBD device). Some features are at the moment available only in the high-end solution of the TELMI client. In the other solutions, their

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

accuracy is appropriately scaled down, depending on the specific motion capture device that is available in a given set-up. For example, a precise measure of kinetic energy may be replaced by an overall measure of movement activity (e.g., the Motion Index), if only the blob and the depth-map provided by Kinect is reliable enough (whereas e.g., the Kinect skeleton may be too noisy). As an example of the functionalities the analysis component makes available, Figure 2 shows a couple of EyesWeb XMI applications (patches) for real-time extraction of low-level and mid-level movement features. The left panel shows the EyesWeb patch for extraction and analysis of kinematic features and its interface developed in the EyesWeb Mobile tool. The right panel shows the EyesWeb patch for extraction of movement coordination (intra-personal synchrony), using Recurrence Quantification Analysis.



(a) (b)

Figure 2. An example of analysis of body movement in violin performance, as provided by the analysis component of the TELMI platform. (a) real-time extraction of kinematic features from motion capture data. (b) real-time extraction of mid-level features from motion capture data: coordination (intra-personal synchrony) is computed using Recurrence Quantification Analysis; in the top, the kinetic energy signal of the right wrist, and, in the bottom, the recurrence plot of such kinetic energy and its probabilities of recurrence.

Future work in this component for the second half of TELMI, will focus on relaxing constraints that were currently assumed to simplify the analysis (see also Deliverable D4.2 for more details). For example, the current settings expect some reflexive tape to be put on the bow in order to reliably localize and track it. In the second half of TELMI, computer vision techniques will be investigated to remove this constraint.

3.3 Real-time feedback component

Software in this component comes mainly from WP4 and WP6. From the real-time audio analysis performed within the ViolinRT prototype, real-time feedback is given to the user regarding pitch, timing, and timbre characteristics of the user performance. This is explained with more detail in Deliverable D4.2, but we enumerate here the main features.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

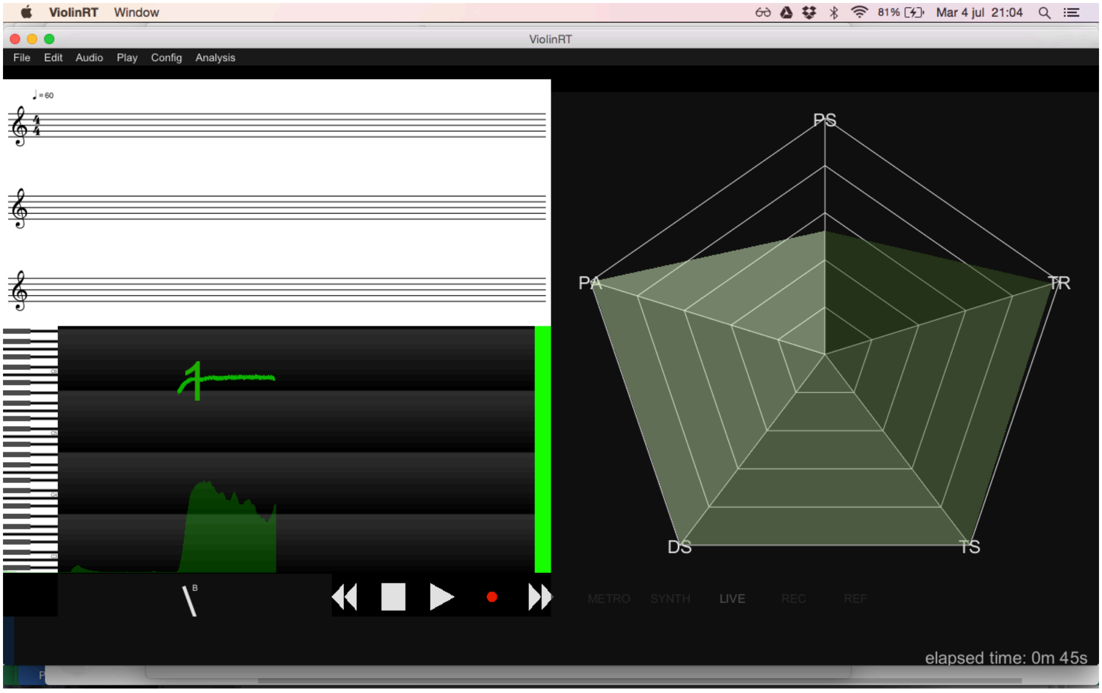


Figure 3: Widget for real-time feedback at single note level.

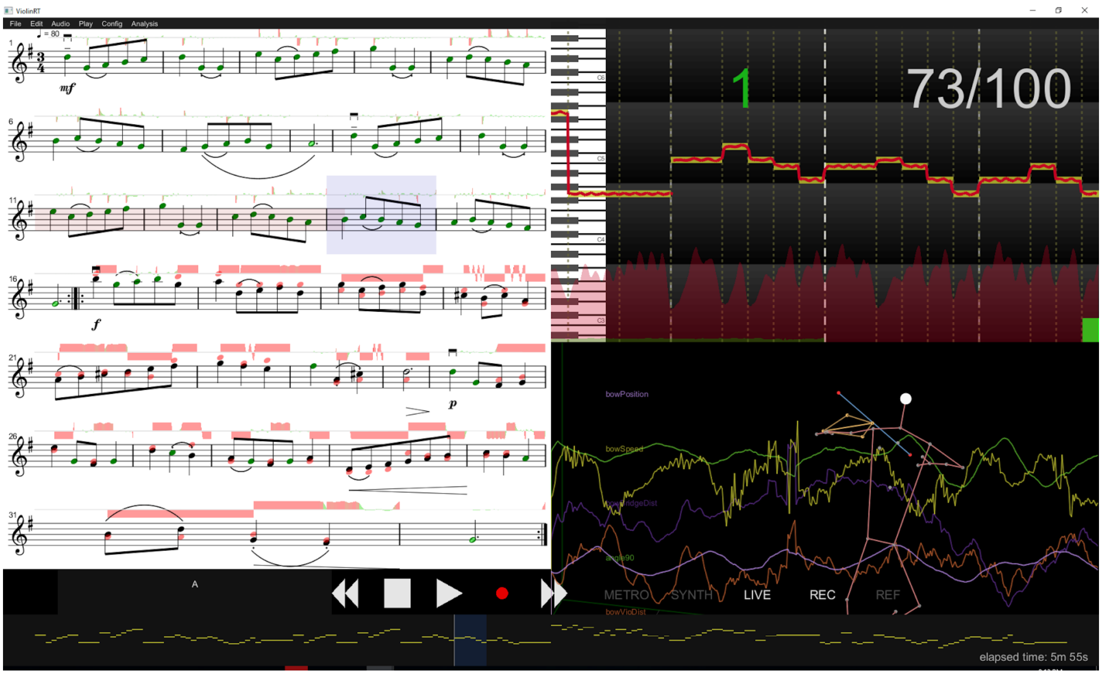


Figure 4: Widget for real-time feedback at piece level.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

At single note level (see also figure 3), feedback concerns:

- Intonation (pitch accuracy and pitch stability)
- Dynamic stability rating
- Tone (timbre stability and timbre richness).

At phrase level (see also figure 4), feedback concerns:

- Intonation (pitch linear regression and pitch deviation for each note)
- Dynamics curve to be compared with expert's curves
- Timing (note deviations over time).

3.4 Output devices component

The first version of the TELMI Platform supports a broad range of visualization devices, from small hand-held devices (e.g., tablets), to large projections on walls. For example, the TELMI app can run on smartphones, tablets, and desktop PCs, enabling a wide variety of set-ups in different environments and conditions (e.g., at home, at the music school, at a research centre, and so on), as envisaged in the TELMI use cases (see Deliverable D6.1).

3.5 Recording component

The recording component was fully implemented within the EyesWeb XMI platform and the EyesWeb Mobile, with patches and interfaces developed for the purposes of the TELMI project. The component supports synchronized recordings of multichannel audio signals, multiple video sources, motion capture data, and physiological signals. EyesWeb Mobile interfaces are provided to simplify management of recordings (e.g., creation of recording sessions, start and stop) and play back of the recorded data. Figure 5 displays an EyesWeb Mobile interface for play back of recorded data, and shows the live frontal video (including ambient audio), the tracked positions of the motion capture markers, and the EMG signals from two MYO sensors on the performer's arms.

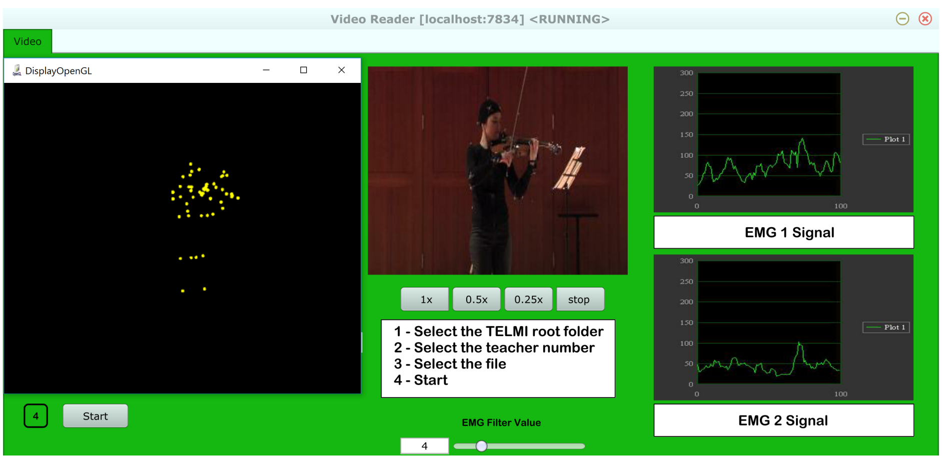


Figure 5. An EyesWeb Mobile Interface for the recording component, enabling synchronized play back of the live frontal video (including ambient audio), of the tracked positions of the motion capture markers, and of the EMG signals from two MYO sensors on the performer's arms.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

The SMPTE protocol is used to synchronize data sources supporting it (e.g., the Qualysis motion capture systems and professional video cameras). Further data sources (e.g., Microsoft Kinect) are synchronized by EyesWeb using a custom timestamp mechanism. The recording component integrates modules mainly developed in WP3. More details are available in Deliverable D3.2 (First working prototypes of the measuring systems).

3.6 Session management component and student interface

Figure 6 shows a snapshot of the current student interface as it appears in the TELMI app. The TELMI app also integrates the session management component. Students can plan their sessions with the system, they can manage their sessions (for example, by selecting which exercises to rehearse and how long to rehearse them), they can view reports of how sessions are developing.

The TELMI app has been developed in the framework of WP5 and its current version is described in detail in Deliverable D5.2 (Initial design of the gamified TELMI platform with MVP). Here, we just provide a brief summary of its major functionalities.

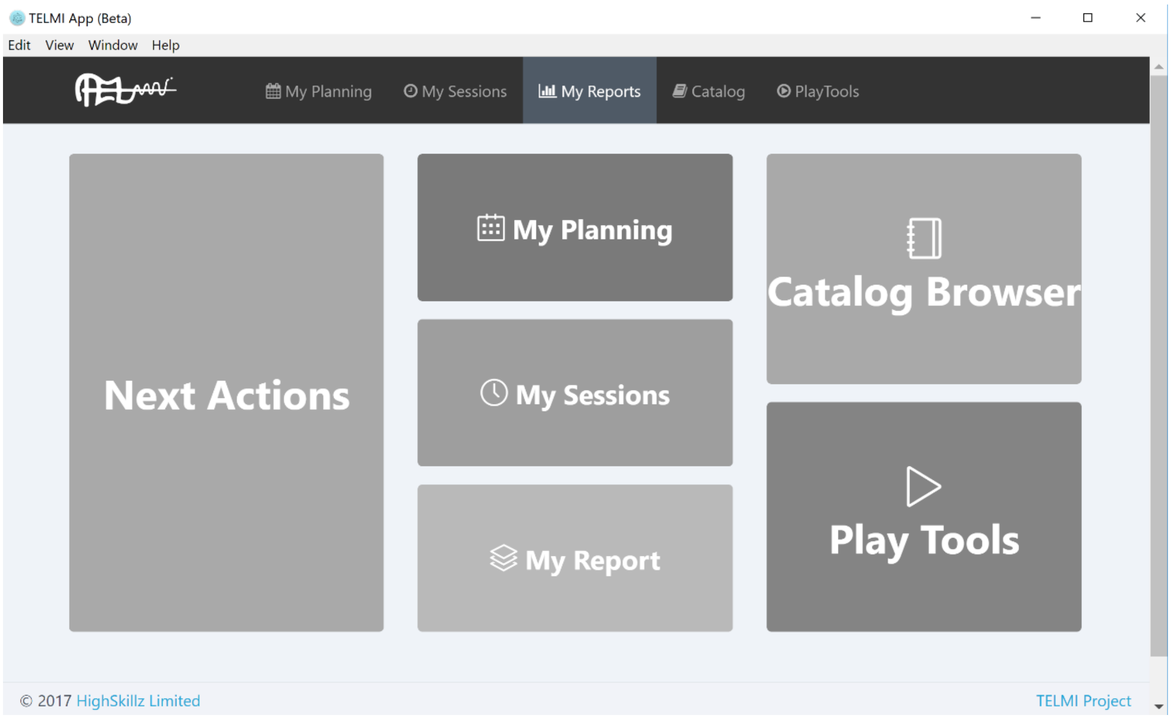


Figure 6. A snapshot of the student interface in the TELMI app as the student sees it.

Besides storing/retrieving data in a local device, the TELMI app can also connect with repoVizz to provide students with the data available in the TELMI reference archive. Students can listen to the recorded audio, watch the recorded videos, look at the music sheet and get the visualizations available in the real-time feedback component (see Section 3.3). Figure 7 shows an example of how students can select and access content through the student interface in the TELMI app.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

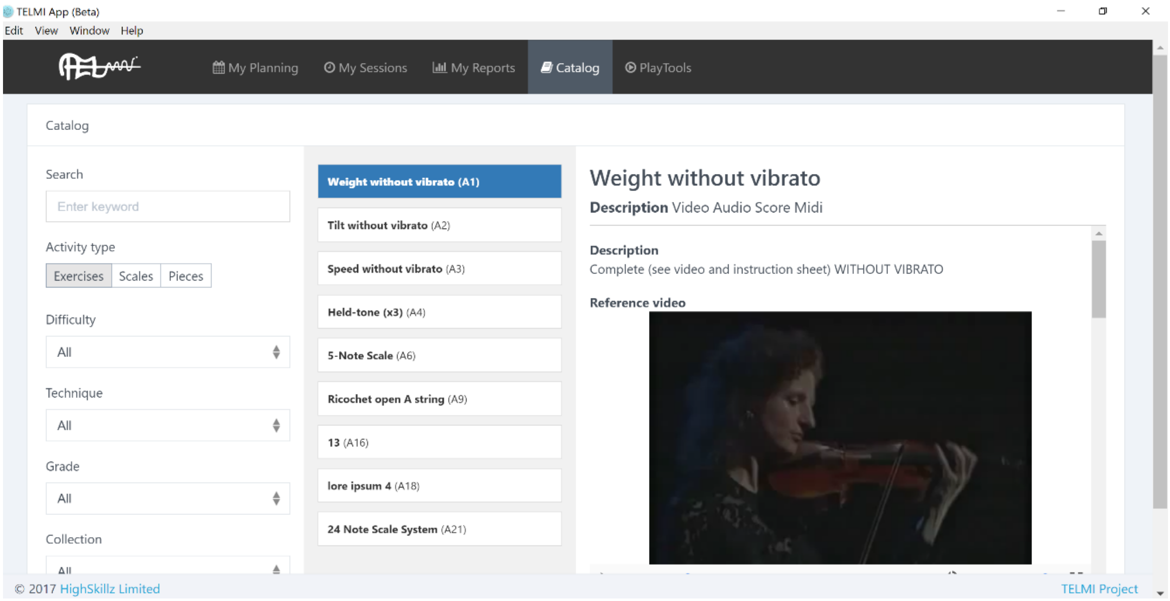


Figure 7. A snapshot of how students can access content through the student interface in the TELMI app.

3.7 Teacher interface and social interface

These components are not yet included in the first version of the TELMI platform. They will be subject of work in the second half of the project. Moreover, the consortium started to implement the features that will be used to assess students' performances. An initial overview of such features is available in Deliverable D4.2 (Metrics for evaluation of skill performance and progress assessment). These features are candidate to be included in the learning analysis component, which will be added in the second and final version of the TELMI platform.

3.8 Implementation details

As mentioned above, from the point of view of implementation the first prototype of the TELMI platform is made of four major software components:

1. The *EyesWeb XMI platform* (UNIGE), providing software modules for (i) capturing data from different input channels (e.g., motion capture systems, audio, video, RGBD devices, and physiological sensors), (ii) performing analysis of body posture and movement both off-line and in real-time, (iii) synchronized recording of raw data and extracted features. Specific EyesWeb modules (blocks) and applications (patches) were developed for the purposes of the TELMI project.
2. The *ViolinRT app* (UPF), providing functionalities for real-time and off-line processing of audio and bow control data and feature analysis and feedback visualization. ViolinRT also provides visualization and synchronized playback of musical scores in musicxml format and recording of student exercises to be reviewed later.
3. The *TELMI app* (HIGHSKILLZ), providing functionalities related to session management, and for connection with the ViolinRT app and the TELMI repository.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

4. *RepoVizz* (UPF), used for storing and indexing the recordings of professional violin players representing the reference archive for the project. The recordings actually consist of 41 exercises performed by 4 internationally renowned professional violin players and include motion capture, audio, video, and physiological sensors (EMG). See Deliverable D2.2 for more details.

EyesWeb XMI and ViolinRT are connected through Open Sound Control (OSC). EyesWeb blocks were specifically developed to enable communication of motion capture and Kinect skeleton data through OSC. ViolinRT and the TELMI app are also connected. The TELMI app can invoke ViolinRT when a user selects the PlayTools option, and data communication between both apps is guaranteed by storing data in a json format file with information regarding session, exercise, duration, and so on, that both apps are able to read and write.

Figure 9 in the next page shows how the three software components of the TELMI platform are related to and implement the components of the logical architecture of the TELMI platform.

The Social Interface and the Teacher Interface are not yet included in the current version of the platform and will be subject of future work in the second half of the project.

3.9 Additional development: web-app wrapper for repoVizz

In parallel to the work invested on integrating the components of the TELMI platform, a web-app wrapper was developed for repoVizz to allow different devices to access it under different permission configurations. Although this solution is versatile, it is geared to:

- Host and organize access to the open databases that will house the open-data media collections produced in the context of TELMI.
- Provide reliable and stable interfacing with repoVizz for a variety of cases.
- Identify and implement functionalities that expand the current API with more user-oriented functions (e.g., share datapacks, download datapacks, and multi-dimensional search)
- Develop an exploitation-oriented demonstration of repoVizz for a powerful solution that manages multi-dimensional media assets.

Figure 8 here below shows the components incorporated on top of repoVizz as part of the web-app wrapper. The Extension component adds user-oriented and platform-oriented functionalities to sustain a user-friendly and platform-independent usage. The Web server component provides and streamlines public access to repoVizz via the API interface of the Extension component.

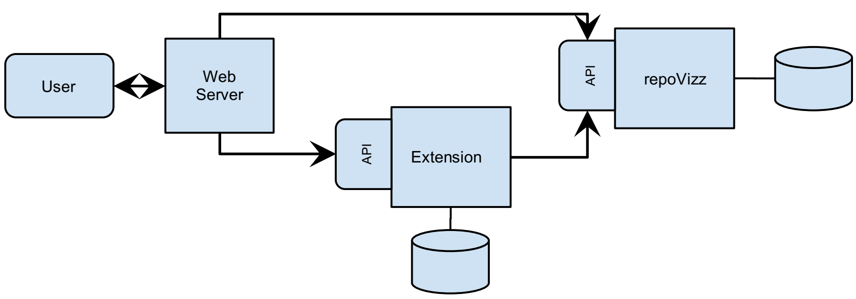


Figure 8. Components of the repoVizz web-app wrapper.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

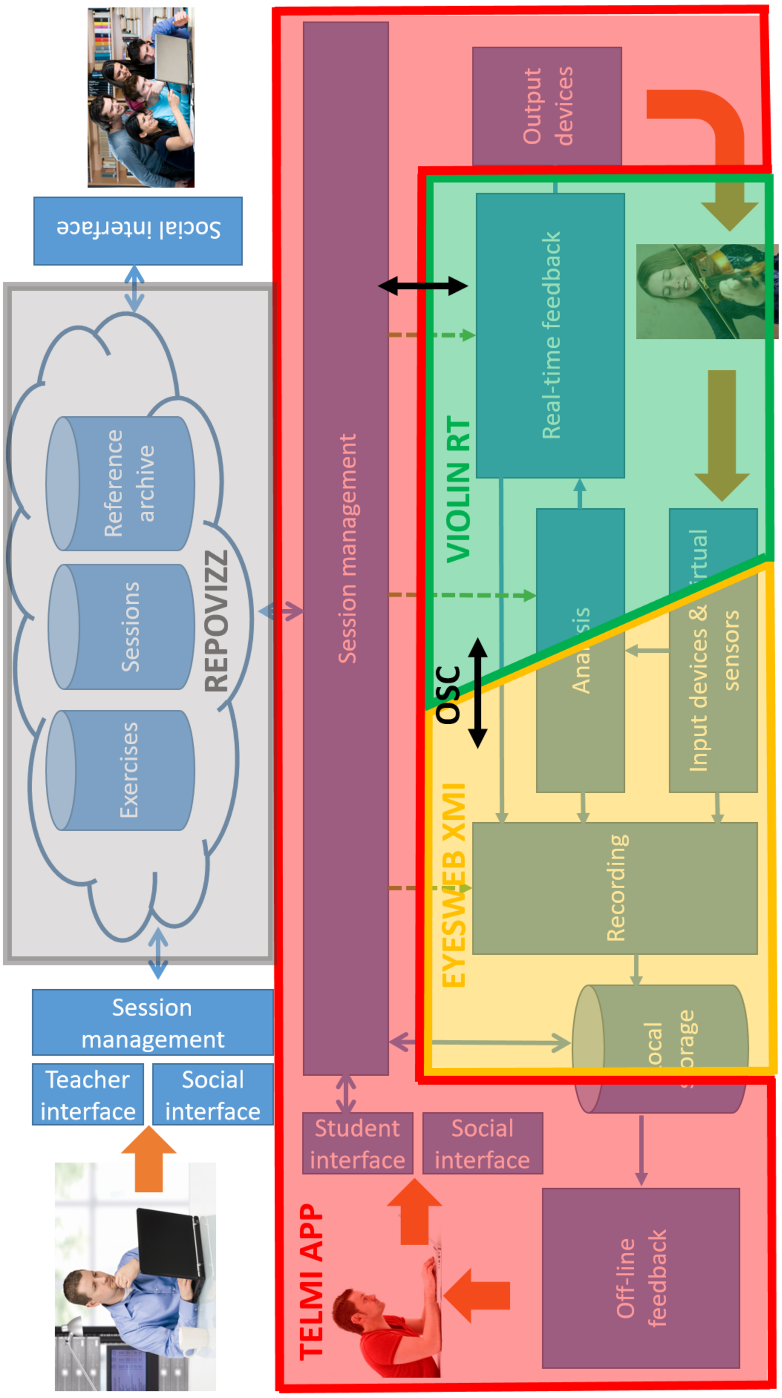


Figure 9. How the developed software components cover the logical architecture of the TELMI Platform.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269

4. Conclusions

This Deliverable presented the first version of the TELMI platform. Whilst already including many of the expected functionalities, e.g., tools and facilities for recording and storing performance data, techniques for both offline and real-time analysis of such data, techniques and tools for real-time feedback to students, interfaces for the students, session management, and so on, the platform still needs further work to be carried out in the second half of the project. This will mainly concern the following major directions:

- Integration in the TELMI platform of computational models and techniques for automatic learning assessment. These will build upon the data captured in WP3 and the features computed in WP4 and will mainly be developed in WP4. As for the logical architecture of the TELMI platform, such techniques may be collected in a dedicated *Learning analysis component*. Assessment data will be made available to students and teachers through the student and the teacher interfaces, respectively.
- Improvement of the capturing and analysis techniques developed in WP3 and in WP4, aiming at removing some constraints that are actually needed, such as for example putting reflective markers on the bow and possibly on the violin. More details about the planned improvements in this direction are discussed in Deliverable D4.2.
- Implementation of the teacher and of the social interfaces. The social interface of the TELMI platform, in particular, is mainly addressed in WP5, which is developing the social TELMI platform expected at Month 30.

Future work will develop across the second half of the TELMI project, leading to the final release of the TELMI platform, which is expected at Month 30 (Deliverable D6.3, Final version of TELMI System Platform). During this time, the current version of the TELMI platform will be extensively tested in different frameworks, such as for example at the Royal College of Music in London, at conservatories and music institutions the TELMI partners are in contact with, and at public events where the results of the TELMI project will be presented (e.g., at the already mentioned second public event of the TELMI project in Reykjavik during the 2017 International Symposium on Performance Science, at the International Festival of Science in Genova, at the International Workshop on Machine Learning and Music 2017 in Barcelona, and so on). Following the outcomes of such a testing phase, the need for further improvements and refinements may emerge. These will be addressed by the Consortium and included in the final version of the TELMI platform.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688269