TROMPA

TROMPA: Towards Richer Online Music Public-domain Archives

Deliverable 6.7

Working Prototype for Music Enthusiasts v2

Grant Agreement nr	770376
Project runtime	May 2018 - April 2021
Document Reference	TR-D6.7-Working Prototype for Music Enthusiasts v2
Work Package	WP6 - End User Pilots
Deliverable Type	Demonstrator
Dissemination Level	PU- Public
Document due date	28 February 2021
Date of submission	February
Leader	UPF
Contact Person	Nicolas Gutierrez (UPF)
Authors	Nicolas Gutierrez, Juan Gomez Lorenzo Porcaro, Aggelos Gkiokas (UPF)
Reviewers	Vladimir Viro (PN), Cynthia Liem (TUD)

Executive Summary

This deliverable is the second version of the demonstrator deliverable for the music enthusiasts pilot submitted on M24 of the project, that went through intermediate usability and functionality evaluation with relevant audiences (**Deliverable 6.8 - Mid-term evaluation**[1]). This deliverable is submitted concurrently with the other WP6 pilot deliverables that are the main deliverables related to **Milestone 3 - Working integrated prototypes ready v1.0**. Although these deliverables are demonstrators rather than detailed reports, the main purpose of this document is to present the functionalities of the pilot and a link to its actual implementation.

The deliverable contains three main sections. Section 2 presents the main functionalities of the pilot by providing screenshots, a URL where we can access the pilot software, a demo video of the pilot along with instructions on how to use the pilot. We present the user interface, its different options and the available functionalities. The final version of the prototype includes three main sections which implement different functionalities: annotate music, where users annotate music excerpts of 30 seconds with the perceived emotions, following the circumplex model (arousal and valence); the musical profile, where users can explore all the discovered music using emotion filters (current version allows users to filter the music based on their own annotations for the annotated music, and based on algorithmically-extracted annotations for the recommended music); and the user settings, where users can modify personal and contextual information important for further recommendation and research purposes, as well as the privacy settings for their accounts.

Section 3 describes the integration with WP3 technologies, the link with WP4 human-generated data gathering strategies, and the use of the Contributor Environment and WP5 components. We observe that these integrations are inline with TROMPA **Deliverable 2.2 - Complete Requirements** submitted on M18. The prototype integrates the TROMPA Processing Library (see **Deliverable 5.3 - TROMPA processing library**) to the pilot in order to trigger music recommendation and emotion recognition algorithms. Likewise, the prototype communicates through the TROMPA Contributor Environment specified by **Deliverable 5.1 - Data Infrastructure**, and incorporates TROMPA's components such as the **Deliverable 5.5 - Annotation Tools**.

Section 4 presents the different recruitment strategies to evaluate the prototype and to engage participation, the activities done so far for user evaluation of the pilot and the plan for the next months regarding actions to evaluate the prototype, engage participants and increase the user base. The COVID-19 crisis had an impact on all the face-to-face recruitment strategies. Thus, virtual strategies have been implemented to evaluate the prototype, while online strategies have been delayed. Section 4.4 explains the procedure and results obtained during the online evaluation workshop performed from 6th to 13th of April 2020 with 15 participants, to evaluate the functionality and usability of the english version of the prototype. Preliminary results of the workshop gave insights of minor functional errors that have been fixed for the next iteration of the workshop (testing the spanish version). Regarding the usability, we could validate that participants were able to use the platform without major problems. We also measured the System Usability Scale (SUS) through a survey with 13 of the participants, and we obtained an average score of 74.6/100. Results also highlighted the importance of the recommendation system and the need for transparency (i.e. clear explanations about how every section of the platform works) within the platform. Section 5 presents the conclusions of the deliverable and future work in the context of the

[1]

final user evaluation, including the evaluation of the usability and the engagement strategies implemented in the pilot.

Version Log		
#	Date	Description
v0.1	22 February 2021	First version for internal review
v0.2	24 February 2021	Final version including corrections from internal review
v1.0	28 February 2021	Final version submitted

Table of Contents

1. Introduction

This deliverable presents the final version of the demonstrator deliverable of the pilot for Music Enthusiasts use case, submitted on M34 of the project. The deliverable follows the same structure as the other WP6 pilot deliverables. These deliverables are demonstrators rather than detailed reports, and the main purpose of the document is to present the functionalities of the pilot and the met requirements. Additionally, we present a detailed description of the relation of the pilot with other Work Packages and the integration of the TROMPA tools, as well as the insights taken based on the user evaluations.

For sake of completeness to the reader, this final version is considered as an extension of the 1st version of the documents; thus some sections of the 1st version can be repeated or slightly modified. We keep a similar structure to the 1st version, which is shared amongst all deliverables D6.3 - D6.7 and contains three main sections. Section 2 presents the main functionalities of the pilot by providing screenshots, a URL where we can access the pilot software, a demo video of the pilot along with instructions on how to use the pilot. Section 3 is related to the integration of WP3 technologies to the pilot as well as the integration of the pilot to WP5 components. Section 4 is dedicated to the user evaluation of the pilot, while section 5 presents the conclusions.

2. Main functionalities of the prototype

The prototype allows users to annotate the perceived emotions in different music pieces (grouped in "campaigns"), following Russell's circumplex model of emotion (arousal and valence) [1]. Users can compare their annotations with the community results and receive musical recommendations based on the provided annotations. The prototype implements gamification mechanisms (scoring systems, rankings and musical recommendations as rewards) to incentivize users to contribute. Likewise, admin users are able to create new campaigns.

2.1. Access information

- The **Pilot Prototype** is accessible through a TROMPA subdomain[2]. The pilot prototype is currently translated in English, Spanish, Italian and Dutch.
- A **Demo video** showing the main functionalities of the pilot is reachable online[3]
- A **Demo video** showing examples of automatic emotion recognition algorithms is also available online[4].
- Infographics describing the workflow in English and Spanish are available in the home page of the prototype[5] (see Figure 2.1)

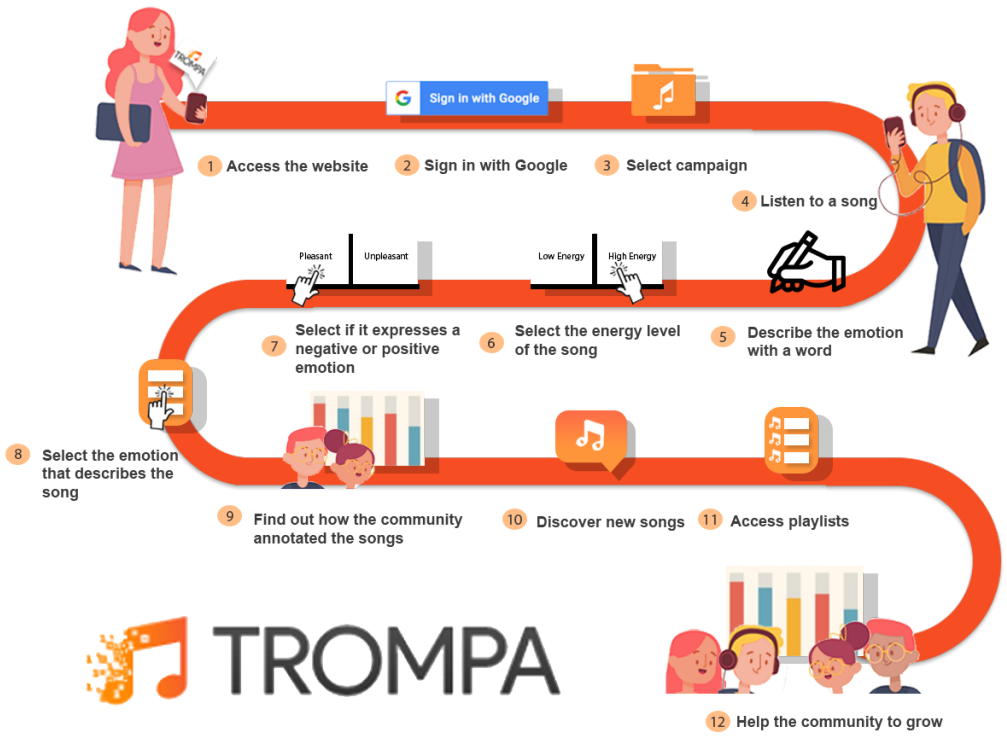


Figure 2.1. TROMPA music enthusiast pilot workflow.

[2] https://enthusiasts.trompamusic.eu
[3] https://trompamusic.eu/demos/enthusiasts
[4] https://ilde.upf.edu/trompa/rc/about
[5] https://ilde.upf.edu/trompa/

2.2. Requirements

The pilot needs the following requirements to be used:

1. A computer, tablet or smartphone.
2. Suggested browser: Chrome with a recent version.
3. Headphones, earphones or speakers (for listening to audio fragments).
4. A Google account (for login)

2.3. Pilot Functionalities

The main functionalities of the application are:

- ❖ Participate in music emotion annotation campaigns:
  - ➢ Select a campaign
  - ➢ Provide information about your current mood
  - ➢ Listen to a song
  - ➢ Select the energy level (arousal)
  - ➢ Select negative/positive emotions (valence)
  - ➢ Select the emotion that describes the song
  - ➢ Find out how the community annotated the song
- ❖ Explore the music discovered through the annotation campaigns
  - ➢ Discover new songs
  - ➢ Access your playlists
- ❖ Configure platform settings
  - ➢ Edit your profile

2.3.1. Annotate music

The pilot core functionality is the pipeline contained in the annotation campaigns. From the *Home* page clicking on *Annotate Music!* in the top menu, the user is redirected to the campaign selection page, where using a search bar or manually, it is possible to start an annotation campaign, or to continue a campaign already started. The details for participating in an annotation campaign are described in Section 2.4.

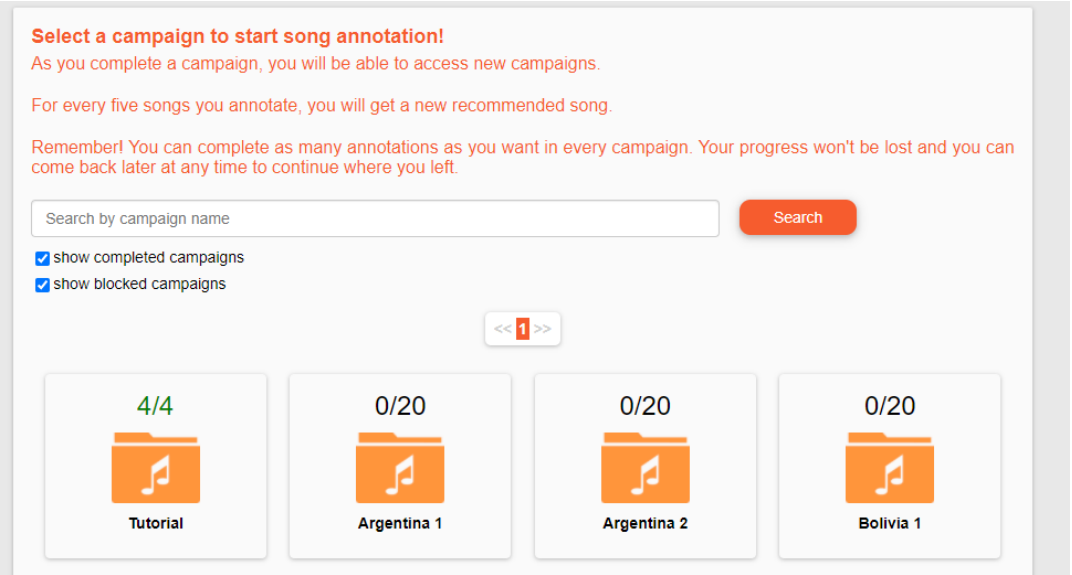


Figure 2.2. *Annotate Music!* page. Annotation campaigns can be selected by clicking on the folder icon, and searched using the search bar. Users can limit the search results by filtering completed or blocked campaigns.

2.3.2. The Musical Profile

Users may navigate within the music they annotated and/or discovered in the platform accessing their musical profile. From the *Home* page clicking on *My musical profile* in the top menu, the user can explore three lists of songs:

- *Music I have discovered:* it contains all the songs recommended to the users after completing the annotation tasks.
- *Music I have annotated and I like*: it contains all the songs annotated and marked as preferred during the annotation task.
- *All the music I have annotated:* it contains all the songs annotated during the annotation task.

Furthermore, the user can filter the list of songs by emotion. Emotions can be selected individually, or the filter can be applied selecting a specific quadrant (see Figure 2.3). The users can filter using emotion annotated by themselves (for Music I have annotated and Music I have annotated and I like). In the case of music discovery, they are presented with an algorithmically-extracted annotation.

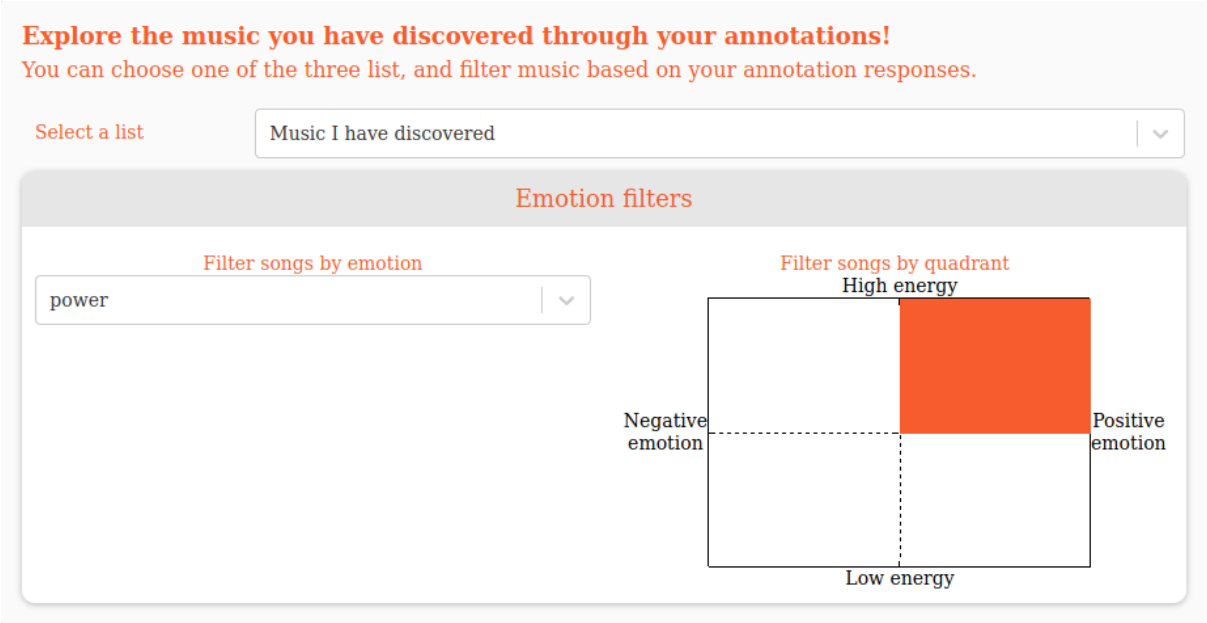


Figure 2.3. Musical profile page. In the top scroll menu users can select the list in which to navigate songs. Emotion filters can be applied, on the left, selecting an emotion from the scroll menu, and on the right by clicking on the quadrant.

2.3.3. User Settings

After the login, users are able to modify in their account the following data:

- Username
- Birth Place
- Mother Language
- Other Spoken Languages

In addition, they are able to opt-in or opt-out to the following settings:

- Receive emails with news about the platform
- Authorize the use of personal data for research purposes

For doing that, from the *Home* page users have to click on *User settings* in the top menu.

2.4 Using the Music Enthusiasts Pilot

The main objective of the application is to learn about musical properties that relate to emotion while annotating, and to receive music recommendations in return.

2.4.1 Annotation explanation

Since the subjectivity in emotion annotation is critical, our approach is to explain which musical features relate to certain emotions, in order to collect annotations with higher quality. The *Home* page contains the *How do I do my annotations?* section that guides the users through basic musical concepts they can use to associate certain emotions in the arousal-valence space. Each quadrant contains information about the aforementioned musical properties.

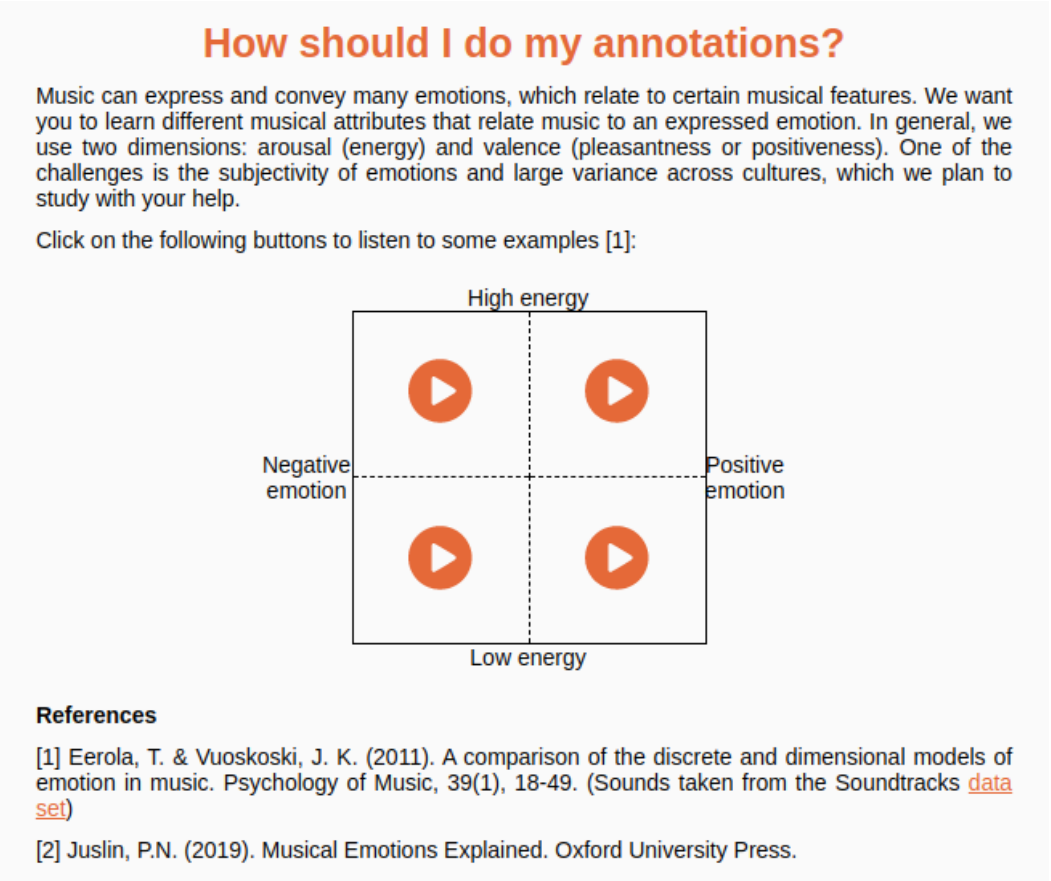


Figure 2.4. Annotation explanation.

2.4.2. Gamification strategy

In order to incentivize the use of the platform, we offer the users a game that will eventually lead to rewards. Since we expect users to utilize objective musical features to produce the annotations, those annotations that belong to the most rated categories will result in more points. The user is presented with a dashboard containing the current score, general statistics about the amount of campaigns and songs annotated, and the top 10 ranking of users. In general, the scores depend on the following factors:

- The similarity of the user's answer with the most frequent answers from the ME community.
- Each annotation can give the user 0 to 6 points.
- With respect to quadrants (4 quadrants of arousal - valence space), the user may receive 2 points when annotating the most voted quadrant for a particular excerpt and 1 point for the second most voted quadrant.
- With respect to emotion categories (11 emotion words), the user may receive 4 points when annotating the first, 3 points for the second, 2 for the third, and 1 for third most voted emotions.

Figure 2.5. User stats and top ranking users from the platform.

A new ranking section has been added, where users can explore the whole ranking filtering by date ranges:

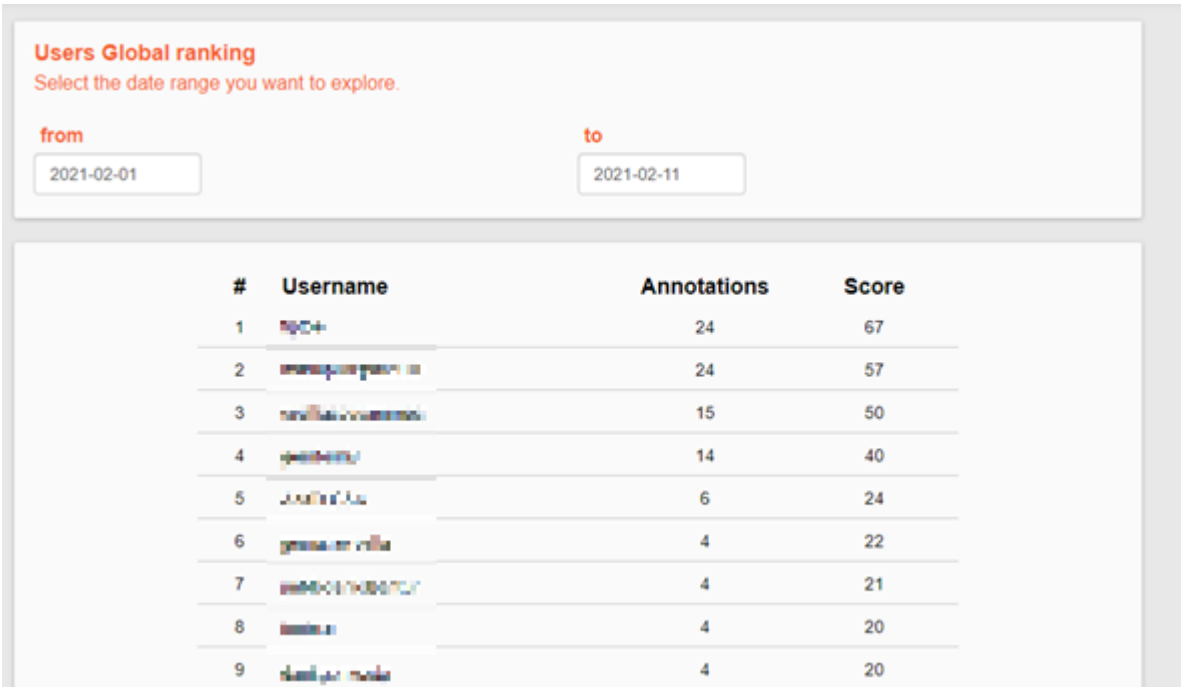


Figure 2.6. Ranking section with date range filter.

2.4.3. Annotation procedure

Currently, the ME use case is centered on annotating music in different languages. The user selects the “Annotate music!” button from the menu and will be presented with the available annotation

campaigns. Each 6 hours of inactivity (or first time annotating) users will be asked to express their current mood (figure 2.7) though a free text box and using the Pick-A-Mood scale [2].

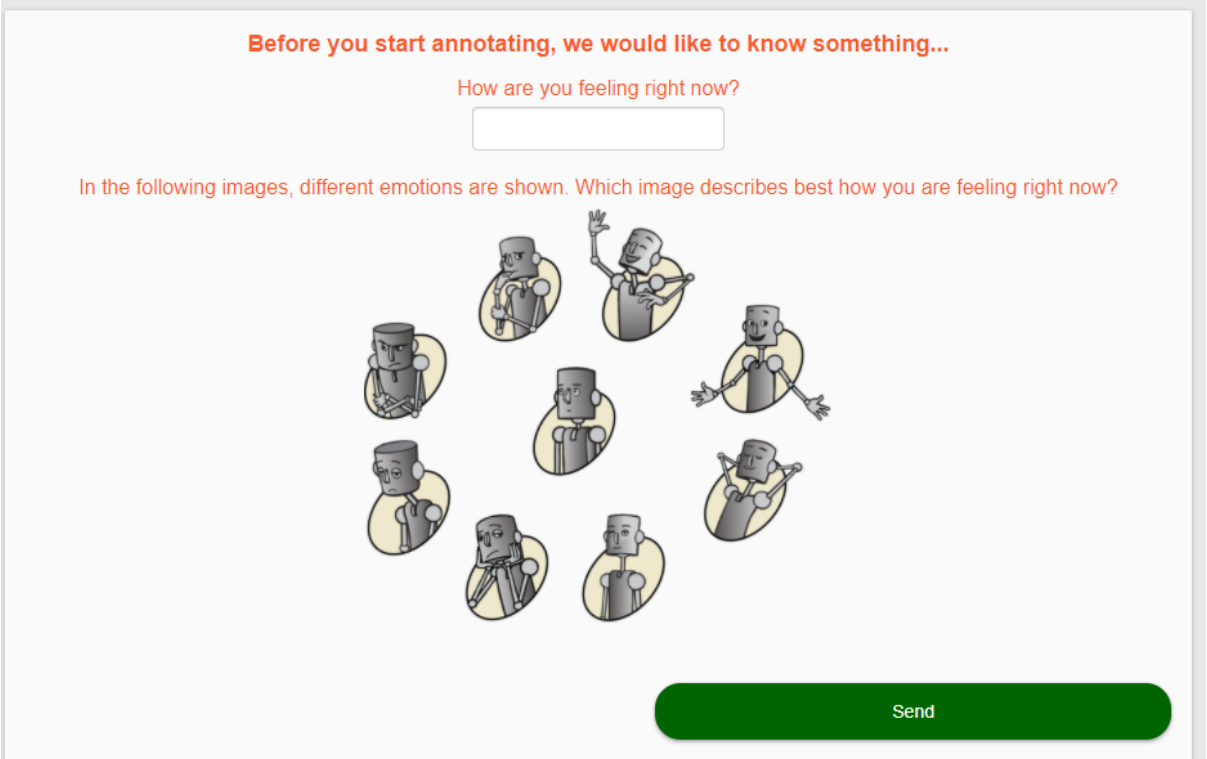


Figure 2.7. User mood interface sample.

After providing the current mood information, the user is redirected to the song annotation screen. The interface contains: the number of songs pending to complete the campaign, a playback control, selectors for preference and familiarity with the excerpt, a free text box to enter the perceived emotion, and the different categories presented per fragment. As mentioned previously, we use **arousal** (high or low energy), **valence** (positive or negative emotion), and the corresponding categories for each quadrant. The resulting emotions per quadrant results as follows:

- Q1 (positive arousal and valence) - joy, surprise, power;
- Q2 (positive arousal and negative valence) - anger, fear, tension;
- Q3 (negative arousal and valence) - sadness, bitterness;
- Q4 (negative arousal and positive valence) - peace, tenderness, transcendence.

Additionally, the user has a text field in order to justify their annotations independently (i.e., once for energy, valence, and emotion category).

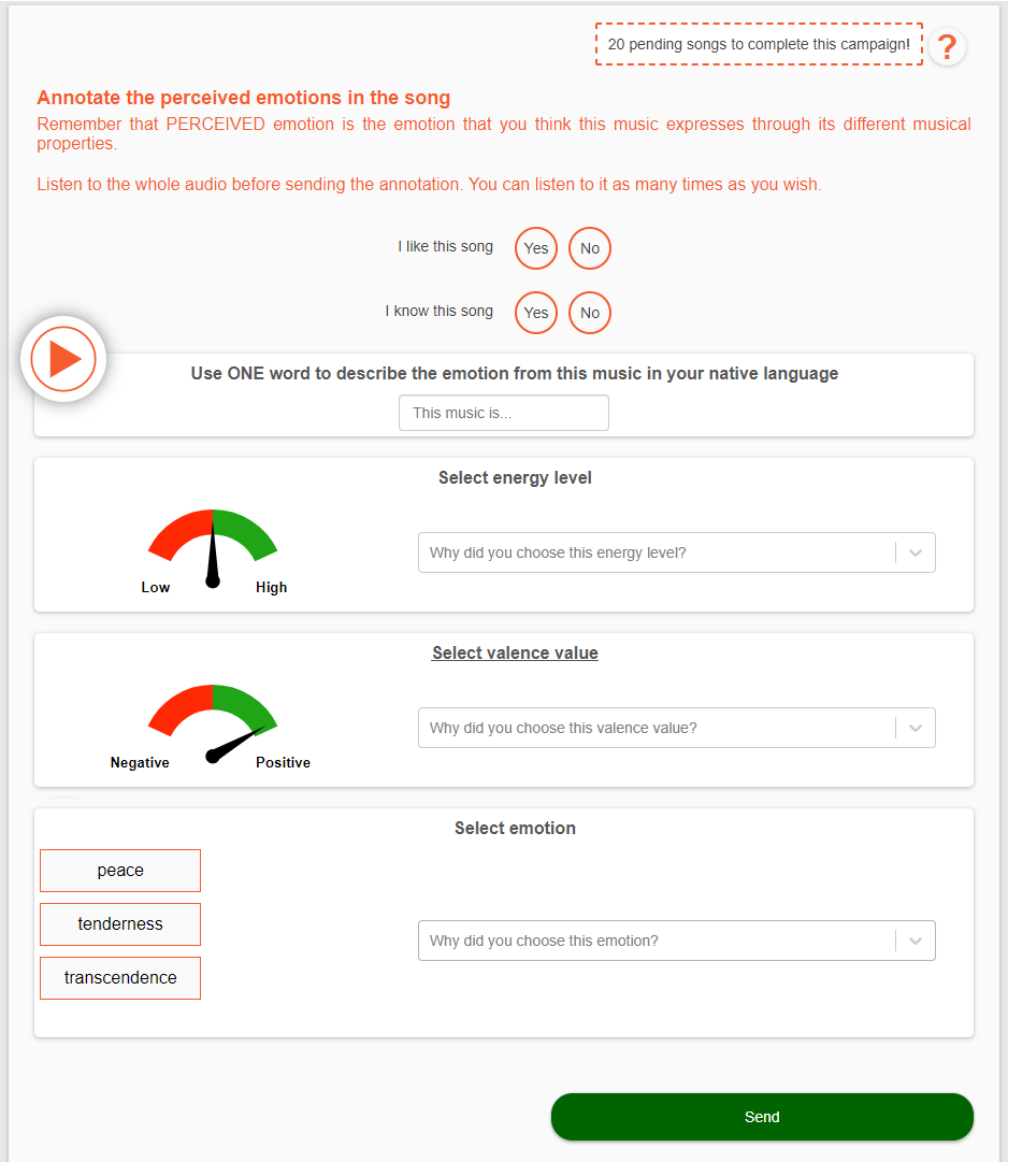


Figure 2.8. Annotation interface sample.

2.4.4. Accessing the Help tutorial

A tutorial section with simple explanations on how to perform a correct annotation. Users will access this tutorial automatically during the first annotation. Then, they will be able to access it again at any time, clicking on the help icon ('?') in the annotation interface.

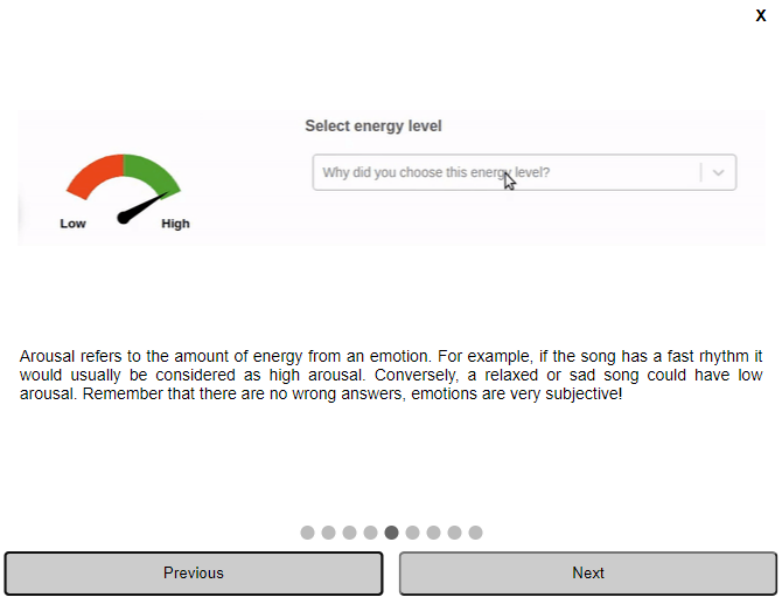
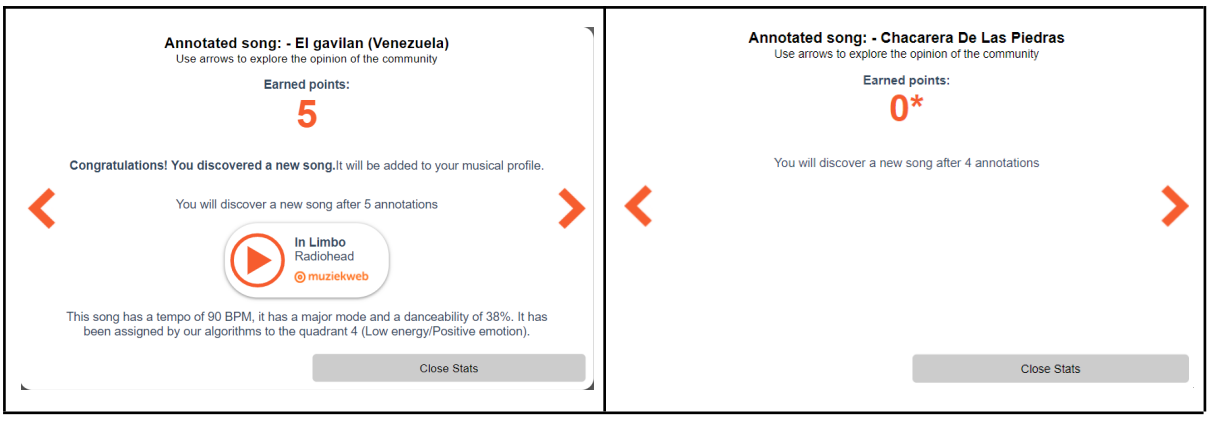


Figure 2.9. Help guide interface example. The guide presents a simple explanation about the annotation process, step by step.

2.4.5. Annotation results

After completing the annotation for a particular excerpt, the user is presented with a pop-up containing three slides: 1) information regarding earned points, and the recommended song (if the user completed 5 annotations) or the pending annotations to get a recommendation. If a recommended song is presented, information about its tempo, mode and danceability is shown; 2) a heat map with the answers from the community regarding quadrants, and 3) a ranking of emotions (see figure 2.10).



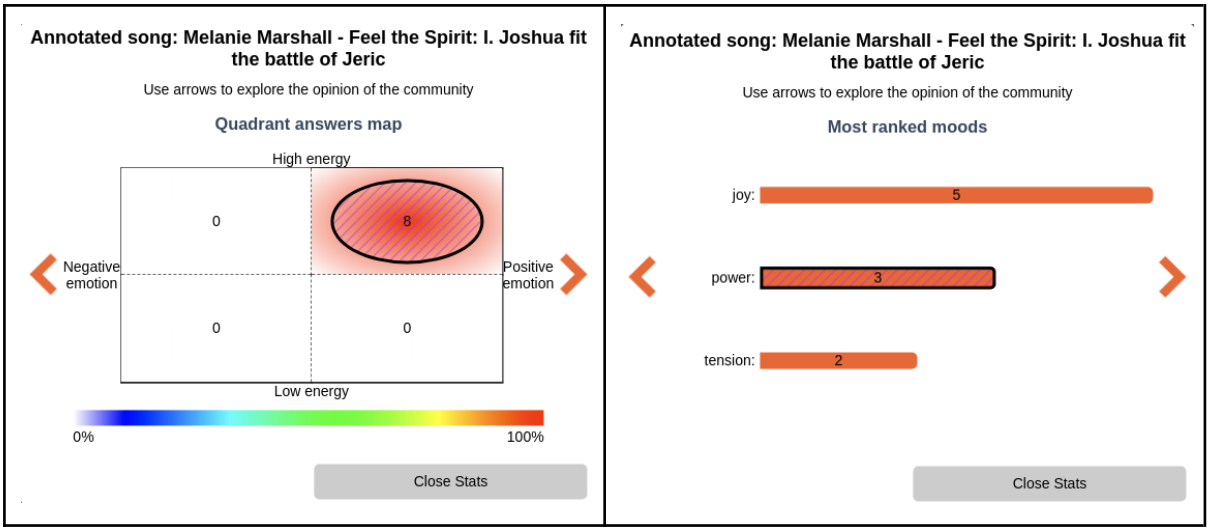


Figure 2.10. Annotation statistics sample.

2.4.6. Musical profile

The musical profile contains information about discovered and annotated music as mentioned on Section 2.3.2. In general, we present the user with three options: discovered music, music that has been annotated and preferred, and all annotations. Additionally, the user may select emotion filters in order to select songs belonging to a particular quadrant or emotion category.

2.4.7. User settings

The users can also access a form in which they can change their username, fill additional demographic information, and give express authorization for contact and use of data for research purposes (see Section 2.3.3).

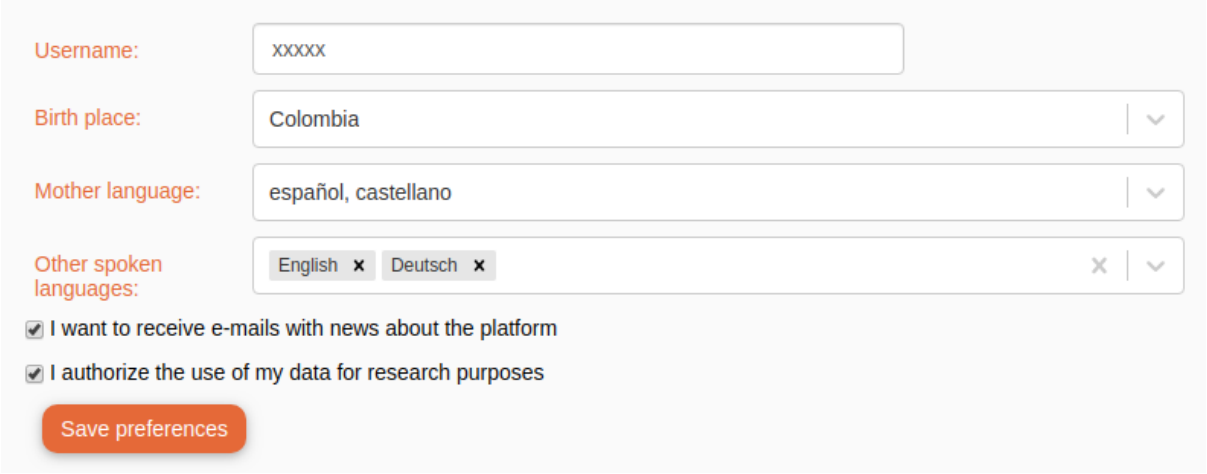


Figure 2.11. User settings sample.

2.5 Connection to requirements

The functional requirements for the pilot were described in **Deliverable 2.2 - Complete Requirements[6].** The current prototype matches most of the requirements from the technical and user perspective. We have focused on registered users and prioritized the basic functionalities related to annotation gathering and basic music recommendation, since these are the sections where all types of users will be involved. Further functionalities such as customized campaign creation, detailed community statistics or more complex recommendation systems will be included in further versions after a full validation of basic functionalities.

- The pilot currently integrates third party authentication systems for user login. It provides a safe environment to manage user personal data and annotations.
- Registered users can participate in annotation campaigns designed specifically for training purposes. Music repertoire for these campaigns include choral music (annotations) as well as popular music (recommendations).
- Users can access personal statistics to his/her progress, with metrics such as current score, amount of annotated songs, completed campaigns.
- Users can access community statistics. Every annotation provided by the user can be fully compared with the responses of the community. Likewise, users can track the top 10 contributors based on the scoring system.
- Users have a "Music profile" section, where they can access all the music they have discovered through their annotations. Emotional content filtering is possible.

2.6 Final outlook

The final prototype includes a Solid (WebID) authentication system, improves the recommendation algorithm, including additional contextual information and open text emotion categories, and includes a new filter for the discovered music according to: user annotations, community perception, algorithms. Furthermore, the Contributor Environment (CE) is fully integrated within the prototype to create campaigns (through the Multimodal Component, users will be able to search for specific recordings to add to a campaign) and store annotations. It also provides community statistics for every music piece.

[6] This deliverable is confidential to the consortium only

3. Integration with other TROMPA WPs

This section summarizes the integration of the prototype with the components and technologies developed in other WPs, including WP3 (automatic description), WP4 (crowd-sourcing and human computation technologies) and WP5 (contributor environment and core components).

3.1. Relation with TROMPA WP3

WP3 tasks	Integrated in prototype v1	Integrated in prototype v2
Music description	X	X
Audio processing	n/a	n/a
Visual analysis of scores	n/a	n/a
Alignment of musical resources	n/a	n/a
Multimodal cross-linking	n/a	n/a

Table 3.1. Summary of the relation of the prototype with TROMPA WP3.

3.1.1 Music Description

The prototype uses the TROMPA Processing Library (TPL) to use the algorithms for emotion-based recommendations and to access training models for emotion-based classification of music. The prototype makes use of the music descriptors defined in **Deliverable 3.2 - Music Description**[7] to collect Emotion tags, and includes a free-text description for the user to tag the song with their own words. Additionally, it includes emotion-based music recommendations. Recommendations are computed analyzing users' emotion annotations. For each recommendation, an explanation of part of the features used for the emotion-based classification is provided.

The music emotion recognition system relies on handcrafted musically-relevant features, in order to perform classwise predictions of emotions. In the context of the prototype, we use the annotations of arousal/energy (positive and negative) and valence/pleasantness (positive and negative) to assemble 4 classes with respect to the corresponding quadrants in cartesian coordinates: Q1 (positive arousal and valence), Q2 (positive arousal and negative valence), Q3 (negative arousal and valence), and Q4 (negative arousal and positive valence).

3.2. Relation with TROMPA WP4

WP4 tasks	Integrated in prototype v1	Integrated in prototype v2

[7] https://trompamusic.eu/deliverables/TR-D3.2-Music_Description_v1.pdf

Crowd-powered improvement	X	X
Annotators	n/a	X
Incentivisation of TROMPA crowds	X	X
Campaign design	X	X

Table 3.2. Summary of the relation of the prototype with TROMPA WP4.

3.2.1 Crowd-powered improvement

Currently, we use users annotations to produce datasets and agreement analysis about the emotional content of music. With the annotations provided by users, recommendation and classification algorithms will be periodically fine-tuned in order to obtain more accurate results for the community. In next user evaluations of the prototype, we will use crowds to assess the automatic recommendations.

We perform active learning to allow our classification models to improve with new annotations from particular users - with the aim of personalizing these models. The idea is that song instances that are difficult to assess should be prioritized in order to be annotated first. *Uncertainty sampling using entropy* is used over the prediction probabilities from a collection of classifiers, in order to measure the uncertainty produced by particular predictions: instances with low entropy are assumed to be the *most informative*, while low entropy highlights the *least informative* instances that should be annotated by our users.

3.2.2 Incentivisation of TROMPA crowds

We have included both intrinsic and extrinsic incentives for users to participate in the community, following the Incentivisation model proposed in **Deliverable 4.3 - Crowd incentivisation mechanisms**[8]. We have adjusted the prototype to be a platform to get recommendations based on emotional content. Thus, recommendations became an incentive to provide the requested information, proposing new musical references for facilitating the task of emotion annotation. We also provide participants knowledge about how certain musical features relate with the emotional content in music and engage participants to make use of wisdom of crowds to understand how other community members perceive music. We included feedback for each annotation provided by users so they are able to compare themself with the community. We also included gamification mechanisms such as a scoring system, user progress statistics (number of annotations provided, number of campaigns completed) and a top 10 ranking of users based on the scoring system. We also included other ranking systems and user progress statistics to engage participation.

3.2.3 Campaign design

The prototype currently allows administrators to design their own annotation campaigns to engage participants in contributing. The prototype workflows provide an scalable platform that can contain multiple campaigns. This version of the prototype includes research results of the currently existing campaigns and includes new campaigns with different repertoires. As explained in section 3.3.1 of

[8] This deliverable is confidential to the consortium only.

this deliverable, the Multimodal Component can be used as a tool to create new annotation campaigns of the desired pieces contained within the CE. The campaigns will be customized according to the desired target users, adjusting the repertoire to be annotated as well as the one for recommendations.

3.3. Relation with TROMPA WP5

WP5 components	Currently integrated	To be integrated in next version
Score edition component	n/a	n/a
Processing library	n/a	X
Multimodal integration	n/a	X
Performance assessment	n/a	n/a
Annotation tools	X	X

Table 3.3. Summary of the relation of the prototype with TROMPA WP5.

3.3.1 Multimodal integration

The current version of the prototype does not include the use of the multimodal component (see **Deliverable 5.1 - Data Infrastructure**[9]) to explore resources stored in the CE. Nevertheless, this feature provides a way to create new customized campaigns and use this feature as an incentive mechanism for expert users. A detailed workflow for the integration is provided.

3.3.2 Annotation tools

Following the descriptions of the **Deliverable 5.5 - Annotator Tools**[10], the annotations made through the prototype follow a compatible structure to be stored within the CE. The final version of the prototype is fully integrated to the CE for getting the musical resources and to store the annotations produced by users.

3.3.3 TROMPA Processing Library

We integrated the TROMPA Processing Library (see **Deliverable 5.3 - TROMPA processing library**[11]) to the pilot in order to trigger the music recommendation and the emotion recognition algorithms (see Section 3.1.1)

4. User evaluations

4.1. Initial target audience & recruitment strategies

The initial planning of target audience and recruitment strategies had been as following:

[9] https://trompamusic.eu/deliverables/TR-D5.1-Data_Infrastructure_v2.pdf
[10] https://trompamusic.eu/deliverables/TR-D5.5-Annotation_Tools_v2.pdf
[11] https://trompamusic.eu/deliverables/TR-D5.3-TROMPA_Processing_Library_v2.pdf

- **Workshop in Maker Faire Barcelona (October 5th and 6th 2019)**. The goal of this workshop with 7 participants was to evaluate user perception of basic functionalities. Participants were adults with interest in state-of-art technologies and community involvement in science. A first prototype of the ME pilot was presented as a tool to support research in emotion recognition in music. The workshop aimed to highlight the importance of massive involvement of persons in tasks such as collecting training data and evaluating the results of designed algorithms.
- **Workshop in Barcelona Science week (November 12th 2019)**. The goal of this workshop was to complement and validate the results obtained in the previous workshop and to disseminate the Music Enthusiasts use case research objectives. The workshop was done with 20 school children under 14 years old. For this reason, we did not collect any data for research or analysis purposes. During the session, some features identified by science to objectively classify and identify emotions in music were presented, highlighting the importance of music enthusiasts involvement to improve current knowledge and technologies.
- **YoMo participation (February 22th to 25th 2020. POSTPONED).** YoMo is part of the Mobile World Congress in Barcelona. The goal of this workshop was to test the first working version of the ME pilot with young people (from 14 to 17 years old) interested in science and new technologies, and validate the impact of the implemented incentives within the platform, namely the recommendation of new music, the gamification aspects (scoring system and user rankings) and the annotation feedback. We planned to run an annotation campaign offering external tangible rewards to engage participation. During 4 days it was expected that approximately 1 thousand students (14-17 y.o.) per day would attend the event.
- **UPF Open days (March 2020. POSTPONED).** We planned to launch new campaigns such as the one implemented during YoMo within the UPF Open Days to collect multiple annotations and test the effect of the engagement mechanisms such as music recommendations. The goal was to validate the functional requirements of the pilot.
- **Mahler Festival (May 8th to 17th. CANCELED).** The goal of this event was to promote the platform within the Mahler enthusiasts community. Experts and non-experts would have the opportunity to explore Mahler's music through the emotional content while they contribute by annotating different repertoires.
- **Online campaigns through UPF community channels.** After testing the impact and effectiveness of the prototype in the online evaluation testing, we plan to create larger campaigns to be promoted through different online communication channels within the UPF community, such as internal events, university postmaster, etc.
- **Online campaigns through Muziekweb**. After validating the functional requirements of the pilot, we plan to launch a massive annotation campaign within the Muziekweb community to evaluate the participation and to define future engagement actions.

4.2. Impact of COVID-19 crisis

Due to COVID-19 crisis, the Mobile World Congress 2020 and all the parallel activities were cancelled. Likewise, universities were closed and all the activities were translated to virtual format. Thus, part of the planned campaigns couldn't be launched and the evaluation of the full functionalities couldn't be performed following the face to face strategies. For that reason, first we have designed an online evaluation study protocol to test the ME pilot. Once we analyse the feedback collected from the

online evaluation, we will launch the planned online campaigns through Muziekweb and UPF communities.

4.3. Adjusted target audience & recruitment strategies

Our target audience has not changed, but the recruitment strategies have been adjusted, as follows (as long as face to face activities are restricted):

- **Online evaluation workshops (April - May 2020)**. The goal of these workshops is to validate usability and functionality of the first working version of the ME pilot. The target audience are master students from UPF which are interested in music and technology. We plan to do two iterations: first iteration evaluates prototype functionalities for the english version (15 participants); second iteration will evaluate prototype functionalities for the spanish version (this iteration is still in progress, with 17 registered participants).
- **Online campaigns through UPF community channels (October 2020).** These campaigns keep the same objectives and structure as planned, but they will be delayed until the online evaluation workshops end.
- **Online campaigns through Muziekweb (ONGOING - February 2021)**. These campaigns keep the same objectives and structure as planned, but it will be delayed until the first online campaigns with UPF community are launched and validated.

4.4. Insights taken along for prototype iterations

4.4.1 Online evaluation study protocol

The objective of this study is to evaluate the functionality and usability of the TROMPA Music Enthusiasts Pilot, determine the effectiveness of the incentives created within the platform and evaluate compliance with the design requirements. We created three campaigns of choral music within the ME pilot, to be completed by the participants. All the responses gathered are handled confidentially, and there are no right or wrong answers. Participants work individually from home to complete different assigned tasks. The duration of the study is approximately two hours distributed over several days, and the content of the material was made available in Spanish and English. The activity takes place virtually and is designed as a relaxed and informal activity. Participation is rewarded with 15 euros. The study is divided in two phases:

Phase 1. Exploring the platform: For this phase participants spend approximately 1 hour without interruptions. A form with the instructions and detailed description of the tasks to be performed was designed to be shared (english version available online[12]). Half of the participants perform the requested tasks using a computer, while the other half perform them using a mobile device (smartphone or tablet). Users also answer questions related to the usability and the difficulty associated with the tasks and the concepts. The requested tasks are associated to the following features:

- Pilot goals and description
  - ➢ Explore the **home** section.

[12] https://docs.google.com/forms/d/e/1FAIpQLScUgWEdIDZdOmysOdacTcRsTcsC767FodMO4wKgz4WVtYmyDw

- ➢ Explore **About us** section.

❖ User Registration and login and Terms and conditions formulary
- ➢ Register using Google account and explore the Terms and Conditions.

❖ User Settings
- ➢ Complete user settings.

❖ Annotate emotions in music (Music player, Annotation procedure, Feedback)
- ➢ Annotate all the songs in the campaign *"Tutorial"*, paying special attention to the scoring system
- ➢ Annotate all the songs in the campaign *"Campaign 1"*, paying special attention to the recommendations generated through the annotations.

❖ Explore music based on the annotated emotions
- ➢ Annotate songs in the campaign *"Campaign 1"* until you get a new music discovery.
- ➢ Explore the music discovered through the platform using the provided filters.

Phase 2. Using the platform: After the familiarization with the ME pilot platform, participants are requested to complete a long annotation campaign (60 songs) during several days (spending in total around 1 hour). After finishing the campaign, users are requested to complete a usability survey[13]. We used the System Usability Scale [3] to quantify and estimate the level of usability of the ME pilot.

This online evaluation study protocol is divided in two iterations. First iteration took place from 6th to 13th of april and 15 participants tested the english version of the prototype. In the following days, the second iteration will be held, and we expect to have 17 participants. They will test the spanish version of the prototype plus some adjustments made based on the results of the first iteration.

4.2.2 Results (first iteration: English version testing)

We present here some results of the evaluation exercise done during the **Online evaluation workshops** (see section 4.3). We gathered a total of 15 participants, 5 (33.3%) women and 10 (66.6%) men, with an average age of 26 years old (rank=12). As it was explained in the protocol, 7 of the participants used a mobile device to test the pilot, while the other 8 used a computer. The great majority of the participants (14) stated that they have had musical training (e.g. music theory lessons, guitar lessons, piano lessons, etc.). Likewise, 9 (52.9%) participants have received formal training in singing or have belonged to a choir.

[13] https://docs.google.com/forms/d/e/1FAIpQLSfcI8kkxqnnTxhiWreDQrCVgDb3Ri9Id5HjD6xKanrJDJ30kg/

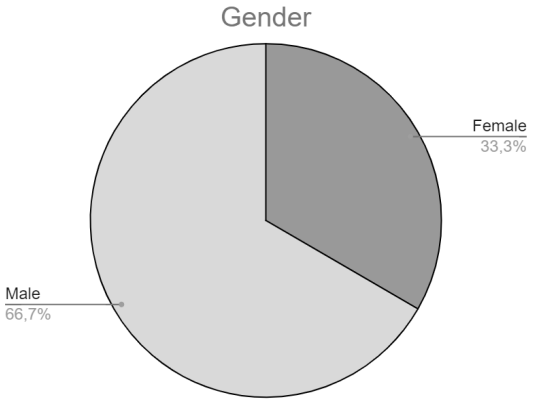


Figure 4.1. Participants' gender.

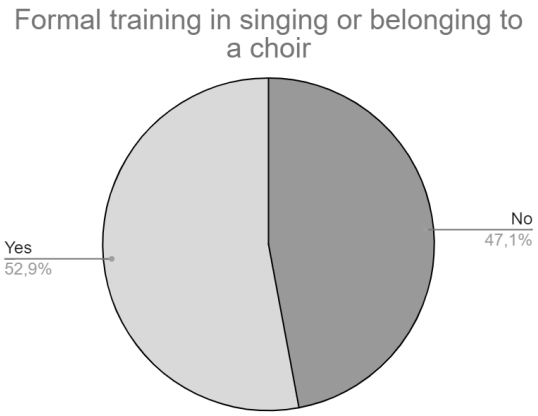


Figure 4.2. Participants' formal training in singing or belonging to a choir.

4.2.2.1 Performed Tasks

In general, participants indicated that they were able to complete all the proposed tasks and that language and terms used in the platform are easy to understand. Regarding the issues found by participants to complete the tasks, one user identified functional problems of the pilot using Firefox v.74 for Linux/MAC. This issue is specifically related to the CDR integrated player. Other issues are associated with the mobile version and the way the menu is displayed.

4.2.2.2 Participants first impressions of the pilot

Regarding participants' motivations for using a platform such as the ME pilot (Figure 4.3), discover music based on its emotional content and learn about musical properties and the emotions associated with them are the most important aspects for users. On the other hand, some participants are not entirely comfortable with the scoring system and the community feedback.

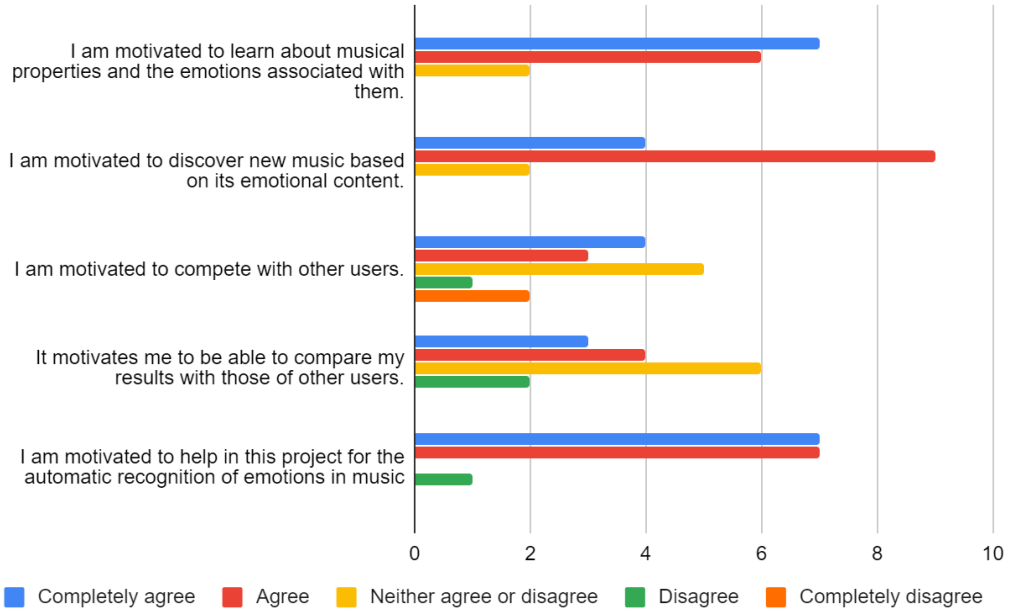


Figure 4.3. Participants' motivations to use an application such as the ME pilot.

When we asked participants if they would use a platform such as the music enthusiasts pilot, around 80% answered they would use or probably use it (Figure 4.4). According to users opinions, instead of being a stand alone application, the pilot should be presented as a complement to existing streaming platforms such as Spotify, and annotations should be as simple as possible in order to engage participants to use it to collect recommendations.

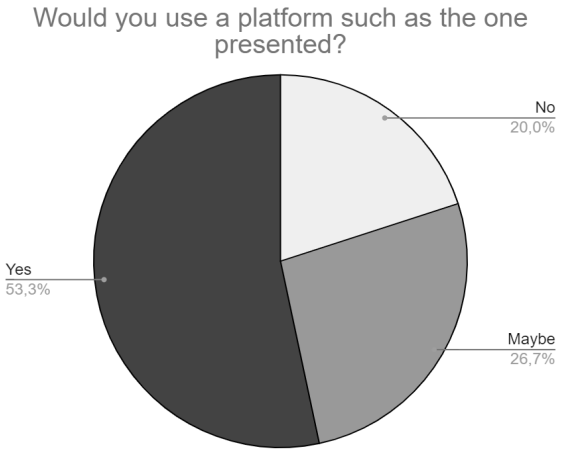


Figure 4.4. Perception of participants about using an application such as the ME pilot.

Participants were also asked to assess the scoring system and the amount of annotations they need to provide to obtain a new recommendation. 10 of the participants did not consider the scoring system as fair, since they considered emotional content in music is subjective. All these responses

can be explained because users still do not understand the differences between perceived emotions and evoked emotions.

4.2.2.2 Usability survey

From the 15 participants that participated in the workshop, 13 completed the System Usability Scale (SUS) survey. The results of the SUS Scale (Table 4.1) show that participants' perception about the usability of the pilot is good (SUS score between 68 and 80.3).

Metrics	Value
Average SUS Score	74.6 / 100
Standard deviation	9.3
Rank	30
Usability	68.8
Learnability	80.4

Table 4.1. SUS Score table

Participants also provided suggestions and comments to improve the pilot. Likewise, based on the performance of the users and the usability indicators, we summarize the adjustments and requirements for the next versions of the pilot application:

- The ME pilot should provide a list of supported web browsers to avoid problems for users.
- The ME pilot should be presented as a complement to streaming platforms such as Spotify.
- It is important to provide a tutorial where users understand the differences between perceived and evoked emotions when listening to musical pieces.
- Scoring system rules need to be better explained and users' feedback should include a detailed explanation for users in order to understand their scores.
- Variables such music genre and user age should be added in the recommendation system.
- The section for exploring the discovered music should include additional filters for the music so the user can select between "algorithm based classification", "community classification" and "user perception classification".
- Annotation of arousal/valence should be done in the same step of the annotation.
- Provide more emotion tags to select.
- Provide descriptions and/or examples in each part of the annotation (arousal, valence, emotion tag).

4.5. Points of attention for final evaluation

Final user evaluation stage will focus on the evaluation of the usefulness of the different incentivisation mechanisms that have been implemented, namely scoring and ranking systems, recommendation systems and learning systems. Likewise, we will focus the evaluation on the different user behavior metrics that can be obtained from the pilot in order to predict and estimate

how engaged participants are, and which future engagement actions should be necessary to improve it. Within the final evaluation, a better performance in the SUS is also expected.

5. Conclusion

In this deliverable we have presented the final version of the Pilot Prototype for Music Enthusiasts. The Pilot is already online, and the current version already satisfies the basic features described in the WP2, and also includes refinements based on the outcomes of **Deliverable D6.8 - Mid-term Evaluation**[14]. In addition, we have provided some usability evaluation with 17 users, who provided usability feedback and comments on the user interface and technologies of the pilot.

The final version of the music enthusiasts pilot includes all the functionalities defined in the user requirements, and several user evaluation stages have been defined from the beginning of the development phase. Users' feedback provided useful information to prioritize and refine requirements. Future user evaluation stages will focus on the final evaluation of the usability of the pilot, as well as the usefulness of the different incentivisation mechanisms that have been implemented, namely, scoring and ranking systems, recommendation systems and learning systems.

Before the COVID-19 crisis, we had planned different campaigns at different levels to reach a bigger audience in every campaign. First, we planned to run campaigns within the student community (youngsters and young adults) through different local events such as the YoMo festival and the UPF community. Later, we would launch online campaigns for specific events such as the Mahler Festival, where a specific Mahler campaign would be designed, to expand our audience. Finally, after some refinements of the platform, our final idea was to launch online campaigns and to promote them through different only communities, such as Muziekweb users. Taking into account the health situation, we focused our plans to scale the TROMPA user base in online campaigns that are being promoted through different groups. Online campaigns are the main strategy to reach a large number of participants, but other local events have been used to carry out usability tests with real users, like a workshop with 20 students of a vocational education course in web applications design.

References

[1] Russell, J.A. (1980). A circumplex model of affect. Personality and Social Psychology, vol. 39, no. 6, pp. 1161–1178.

[2] Vastenburg, M.H., Romero, N., van Bel, D., & Desmet, P.M.A. (2011). PMRI: development of a pictorial mood reporting instrument. In proceedings of CHI 2011, May 7-12, 2011, Vancouver, BC, Canada.

[3] Bangor, A., Miller, J. & Kortum, P. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of Usability Studies, 4(3), 114–123.

[14] https://trompamusic.eu/deliverables/TR-D6.8-Mid_Term_Evaluation.pdf