Umanistica Digitale – ISSN: 2532-8816 – n. 20, 2025: Special Issue
N. Orio, D. Zilio, A. Micheletti – *Listening2Painting: Teaching Multimedia Skills Through an Audio Augmented Reality Approach for Arts*
DOI: http://doi.org/10.6092/issn.2532-8816/21249

Listening2Painting: Teaching Multimedia Skills Through an Audio Augmented Reality Approach for Arts

Nicola Orio

Department of Cultural Heritage, University of Padua, Padua, Italy
nicola.orio@unipd.it

Daniel Zilio

Department of Cultural Heritage, University of Padua, Padua, Italy
daniel.zilio@unipd.it

Andrea Micheletti

Department of Cultural Heritage, University of Padua, Padua, Italy
andrea.micheletti@unipd.it

Abstract

This paper presents the *Listening2Painting* project, an interdisciplinary initiative that explores Audio Augmented Reality (AAR) to enhance artwork appreciation in museum settings. The project's main goal is to bridge the experience of observers and paintings through sonification, while this manuscript particularly emphasizes its application as an educational resource for students of the humanities focusing on multimedia skills. The research details the initial implementation of the project and its integration into two subsequent laboratory sessions, with special emphasis on one session that engaged 195 university students from humanities disciplines. By involving participants in the development of a real-world application, *Listening2Painting* exemplifies a practical case study aimed at improving multimedia competencies. This method effectively connects theoretical learning with practical digital skills, providing students with valuable experience applying technology to cultural heritage. One of the approach's strengths was that students were involved in a project to increase their engagement towards artworks, so they perceived that their activities had a purpose related to their studies. The paper further analyzes educational outcomes and assesses the effectiveness of AAR in enhancing digital literacy within the humanities field. Through this innovative approach, the

project illustrates how technology can be utilized to enrich museum experiences and, at the same time, to foster essential digital skills among students from diverse academic backgrounds.

Keywords: Audio Augmented Reality -- Teaching -- Digital Skills -- Arts -- App -- #AIUCD2024

Questo articolo presenta il progetto Listening2Painting, un progetto interdisciplinare che esplora l'uso l'Audio Augmented Reality (AAR) per migliorare l'apprezzamento delle opere d'arte nei musei. L'obiettivo principale del progetto è quello di creare un ponte tra l'esperienza degli osservatori e i dipinti attraverso la sonificazione, mentre questo manoscritto sottolinea in particolare la sua applicazione come risorsa educativa per gli studenti di materie umanistiche che si concentrano sulle competenze multimediali. La ricerca illustra l'implementazione iniziale del progetto e la sua integrazione in due successive sessioni di laboratorio, con particolare attenzione a una sessione che ha coinvolto 195 studenti universitari di discipline umanistiche. Coinvolgendo i partecipanti nello sviluppo di un'applicazione reale, Listening2Painting esemplifica un caso di studio pratico volto a migliorare le competenze multimediali. Questo metodo collega efficacemente l'apprendimento teorico con le competenze digitali pratiche, fornendo agli studenti una preziosa esperienza di applicazione della tecnologia al patrimonio culturale. Uno dei punti di forza dell'approccio è che gli studenti sono stati coinvolti in un progetto per aumentare l'impegno nei confronti delle opere d'arte, percependo così che le loro attività avevano uno scopo legato ai loro studi. Il documento analizza inoltre i risultati educativi e valuta l'efficacia di AAR nel migliorare l'alfabetizzazione digitale nel campo delle scienze umane. Attraverso questo approccio innovativo, il progetto illustra come la tecnologia possa essere utilizzata per arricchire le esperienze museali e, allo stesso tempo, per promuovere le competenze digitali essenziali tra gli studenti provenienti da diversi contesti accademici.

Parole chiave: Audio Augmented Reality -- Teaching -- Digital Skills -- Arte -- App -- #AIUCD2024

Introduction

Painting has been one of the most effective forms of communication and expression since the origins of the human species. Visiting art museums and observing the marvelous artworks of the past and present is a hallmark of cultural tourism. However, due to a lack of either knowledge or time, one might fail to fully comprehend the elements artists intended to communicate through artworks displayed in museums and expositions. A relevant contribution could be offered by using technologies such as Augmented Reality, which has been exploited in several applications. We focus on Audio-Augmented Reality (AAR) as a tool to reduce this distance between the observers and the paintings without inserting further information in the visual channel that is already devoted to the direct interaction. Utilizing the definition provided in [1], AAR presents an additional layer of contextual information in addition to the user's experience with the real world, and the additional layer is presented in audio form. In this paper, we present the initial stage of the project Listening2Painting (L2P), whose main goal is to explore how AAR can be utilized in the art world to ensure that a wider audience can comprehend the meaning of a painting. The study encompasses several areas of interest, including enhancing art appreciation and understanding, improving the museum experience for visitors, and introducing an innovative approach to teaching computer skills to students of humanities. This project is centered around the intention to integrate a participatory approach with the development of a model for artistic engagement that appeals to a diverse audience. The methodology employed is participatory in nature, particularly engaging university students, who are invited to take an active role in the development process.

A cornerstone of the project is the mobile application called L2P, designed and developed by the authors, that allows users to interact with a collection of selected artworks in various ways. This app will provide a specific experience by enabling users to explore several paintings using, at the same time, their hearing and sight senses. Basically, while watching a painting, visitors can hear the soundscape of the depicted environment and interact with a mobile device that represents the same image and listen to sounds associated with the individual elements.

The project aims to involve users in each step of the designing phase, from creating a participatory audio-labeled collection of artworks, to being testers and peer-reviewers of the community work.

The project's objective can be summarized in three points. Firstly, to enhance the visitor's experience at the museum. Secondly, to provide a tool that enables the reading of a painting by uncovering details and making it easier for visitors to understand the choices made by the painter. Finally, to serve as a way to teach computer skills to students of humanities.

There are several ways to enrich the experience of exploring an artwork through audio. In [2] an interactive system is described that uses machine learning to recognize objects inside Claude Monet's painting automatically. In this research, the authors manually created a training dataset and enriched the user experience with soft music and natural sounds played in response to mouse positioning. The automatic sonification of a group of four artworks is the topic discussed in [3] [4]. The music is produced using a specific algorithm [5] and further developed by a musician. Finally, a laboratory test is carried out in which participants view the reproduction of the painting while listening to the proposed audio in a setting that mimics a museum gallery. A sentiment analysis was conducted to explore participants' reactions. The approach proved useful in enhancing their experience. The effects of background music on the aesthetic experience of visual art are presented in [6]. In this paper, the authors explore the emotional impact of background music while observing an abstract painting by Wassily Kandinsky, reporting the effects on the experience reported by a group of visitors and how it affected their judgments about the artwork. The use of sound in pictorial observation has also been exploited for persons who are blind or visually impaired. For instance, in [7] the relationship between sound and color is analyzed. The goal is to codify a set of colors with different melodies that enhance the experience for people with visual impairments. A touchscreen exploration and verbal feedback are investigated in [8], where two different approaches are presented for segmenting elements within a painting for presentation. Finally, proxemic audio is the approach involved in [9], using a Microsoft Kinect to detect the observer's position and so provide specific audio in response.

Teaching of digital skills

In the contemporary education landscape, particularly within universities and higher education institutions, digital skills development has emerged as an essential requirement that warrants careful attention. While many students are often characterized as *digital natives* [10] due to their extensive exposure to technology from a young age, significant gaps in their understanding remain that inhibit their attainment of true digital proficiency. This proficiency encompasses more than the essential operation of standard technologies such as smartphones and tablets; it requires a nuanced and comprehensive skill set. The terms *digital literacy* [11] and *digital competence* [12] are increasingly prevalent, reflecting the growing necessity for educational initiatives that systematically cultivate these skills among students. The research [13] [14] underscores the critical role that digital skills play in achieving academic excellence and professional success. Notably, integrated information processing skills are recognized as distinct from general academic abilities and are pivotal for thriving in the digital era. The COVID-19 pandemic has further underscored the imperative of leveraging technology within educational settings, revealing the transformative

potential of digital learning. This approach involves integrating new media into the educational framework while thoughtfully considering power dynamics, cultural narratives, and effective pedagogical practices [15]. The incorporation of technology is aimed at creating an interactive and engaging learning environment [16], which may include the utilization of mobile applications and the establishment of task-based courses [17, 4] to enhance student engagement and participation.

An illustrative example of this innovative integration is the collaborative effort to develop a digital library tailored to specific academic projects. This initiative aligns with experiential learning principles, particularly emphasizing project-based learning [18], where students actively participate in initiatives that enrich their understanding and practical application of digital skills in real-world contexts. Such methodologies not only enhance the educational experience but also equip students to navigate the complexities of an increasingly digitized world effectively.

The manuscript is organized as follows. Section The Pilot Study presents the first approach, prior to the development of the app and the web portal, to the enjoyment of certain works of art through sound. Section The Web Platform and the App introduce the tools developed, while the section Workshop presents the learning approach inside a university course. The results and the observations about this workshop are described in the Peer Review Analysis and Questionnaire sections. Finally, an extract on the use of the presented approach with individuals with cognitive disabilities and autism spectrum disorders is presented in the A Twin Project: L2P for "More Than Words" section.

The Pilot Study

A prototype version of the L2P has been created to enable visitors of the Civic Museum of Palazzo Chiericati in Vicenza to interact with four selected paintings, selected by the curator of the museum. The aim was to make visitors more engaged during their visit to Palazzo Chiericati and measure the effectiveness of the approach in a real scenario.

From each painting, the most significant elements have been highlighted and associated with a sound. For instance, a bright metallic sound represented an unsheathed sword, the calm babble of a river represented the house of Nereids at the court of Diana, and the sound of wind strokes on a flag represented the banner of an army. An attempt was made to reproduce the sounds in such a way that they were as representative as possible of the figurative representation of the work, thus taking into account the impact of the individual elements within the work itself. Moreover, some of the paintings were also associated with a sound that represented the overall landscape: a forest, a windy open space, a square. The relevance of the pictorial element guided the choice of whether or not to associate an acoustic description. However, it was not always possible to associate all the relevant elements with a sound. Sounds can be listened individually, by touching the corresponding element on a screen, or in a predefined sequence that follows the order in which the painting is expected to be *read*, according to an art historian.

After completing the acoustic description, the sounds have been added to the interactive app. A tablet was positioned in front of each painting, and we asked 12 visitors of the museum to participate to the experiment.

Participants were divided into three groups to balance the stimuli and their order. One painting was used to allow participants to familiarize themselves with the interface: participants were asked to watch the painting without any sound, listen to the predefined sequence, and then

interact. After this training, each group had a different order of stimuli (no-sound, sequence, and interactive) and paintings. Participants, aged between 18 and 30 and half female, freely agreed to participate in the test. They were provided with headphones to ensure that they could perceive the sounds without interference from other noises.

Results have been encouraging. In a short interview carried out after the experience, participants showed appreciation for the experience, saying that it was involving and informative at the same time. When asked which was the preferred modality, 17.5% preferred to watch the painting in silence, 41% preferred the interactive application, and 41.5% preferred the prerecorded sequence. This slight prevalence of the prerecorded sequence may be due to the fact that it was a passive experience, not requiring searching for the sounds in the painting, which might have been considered complex, with the risk of missing some sonic elements. As regards the perception of the painting, in general, the use of AAR (both prerecorded sequence and interactive) raised the perception of some emotions. When asked to define, in a Likert scale, the level of 'touching', 'lively', and 'joyful', participants gave higher values with AAR than without sounds. Adjectives 'sad' and 'dramatic' did not change.

No further textual information was provided in this initial prototype, so the user experience was limited to listening to the sounds. However, some of the users expressed an interest in knowing more about the reason why some elements have been sonified.

The Web Platform and the App

The Web Platform

The web platform serves as the starting point for the project, enabling users to populate a digital collection. A dedicated database has been designed to represent this collection, and the web platform ensures its connectivity. Currently, the platform is hosted on a university server.

After completing a mandatory registration, participants gain access to the platform where they can contribute to and evaluate the activities of their peers and other project participants. Additionally, a mock-up example is provided as a guideline.[1]

To upload images of the artworks, the format allowed are either PNG or JPG/JPEG format. They must also provide a specific set of information, including the title, artist name, year of creation, and a brief description of the paintings. Furthermore, it is possible to include an audio file as a landscape element. The process of sonification requires users to supply several parameters, including an audio file, rectangle coordinates, a volume level, and a brief description. Users also need to indicate whether the sounds will play in the background without user intervention. This feature allows the integration of individual sounds, such as the rustling of leaves in a forest or the chatter of a crowd. For each sonified element, we request a brief description. Although this feature is not yet visible on the app, it will be an additional element available to users in the future. Finally, after completing the sonification process, participants are required to upload a copy of the selected painting with a set of icons positioned at the locations of the audio elements.

[1] To see an excerpt and example of the portal: https://www.game4culture.com/l2p/.

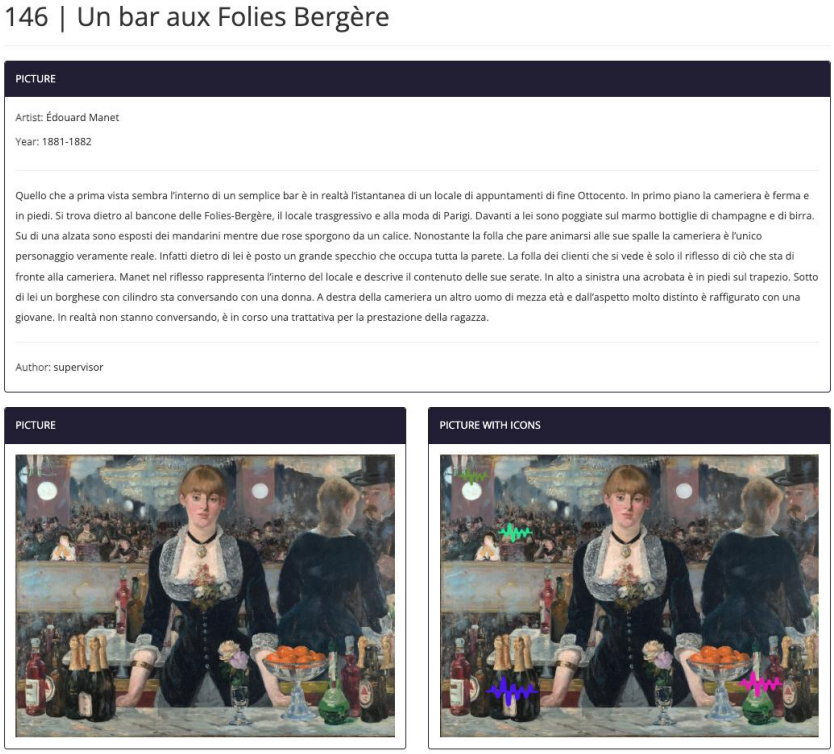

Figure 1. Example screen of the web portal

For the workshop we did not provide suggestions for audio source searches but recommended using Audacity for sound manipulation. Typical sound processing activities include: down-mixing to a mono channel, changing the sampling frequency, extracting the most relevant parts, and applying fade-in and fade-out effects. Loudness can be adjusted directly on the web platform to normalize the effects of different sources. To detect the pixel coordinates associated with the picture elements to be sonified, we recommend using GIMP. This software is also suggested for normalizing image size and resolution.

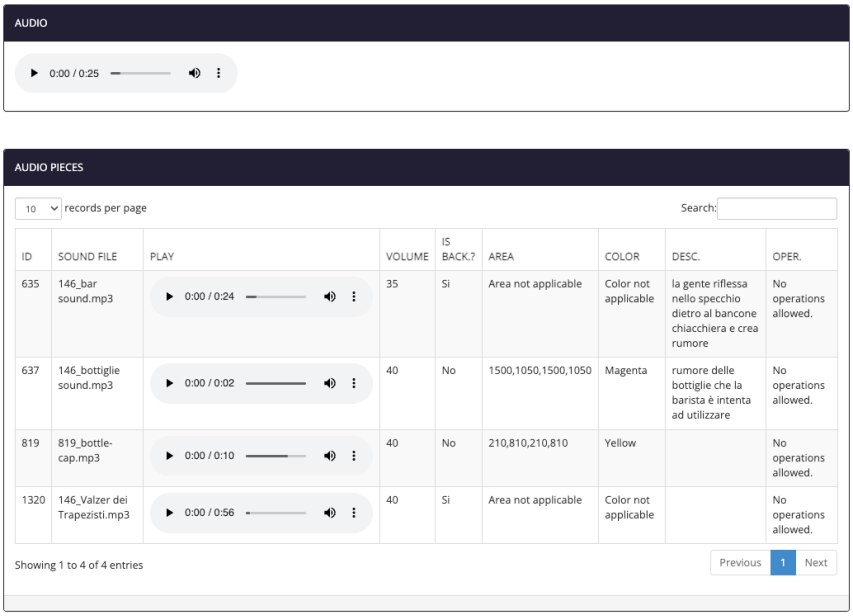

Figure 2. Example screen of the web portal

L2P App

Compared to the previous lab work, this project involved a larger number of participants, so it was necessary to release the app not as a simple internal test but as an alpha test, creating a dedicated group, since the number of concurrent accesses exceeded one hundred, within the Google Android Dashboard. Currently, the app has been designed only for Android, with plans for an iOS version in future updates. This choice is primarily motivated by the possibility of reaching a higher audience among the students.

Figure 3 displays two screens from the app that feature a dynamic list showcasing all the artworks added to the database. Users can retrieve this list by clicking the red button or when the app is launched, allowing students to quickly find and interact with their artwork.

Once an artwork is selected, a dedicated screen appears that enables participants to engage with the painting. The soundscape of the painting starts immediately, while all sound related to elements starts when the user touches the corresponding element on the screen. There are two basic modes. A testing mode, where the responsive parts of the image are highlighted in a different color, and all the changes made on the web portal (sounds, volume, positions on the screen) are automatically transferred to the app. This allows users to test the effectiveness of the AAR and fine-tune the parameters. A normal mode, where the only element shown on the screen is the original painting. In both cases, the image is resized in order to fill the screen and maintain the aspect ratio of the original image.

The mock-up of the upcoming version, intended for a wider audience, is also included.

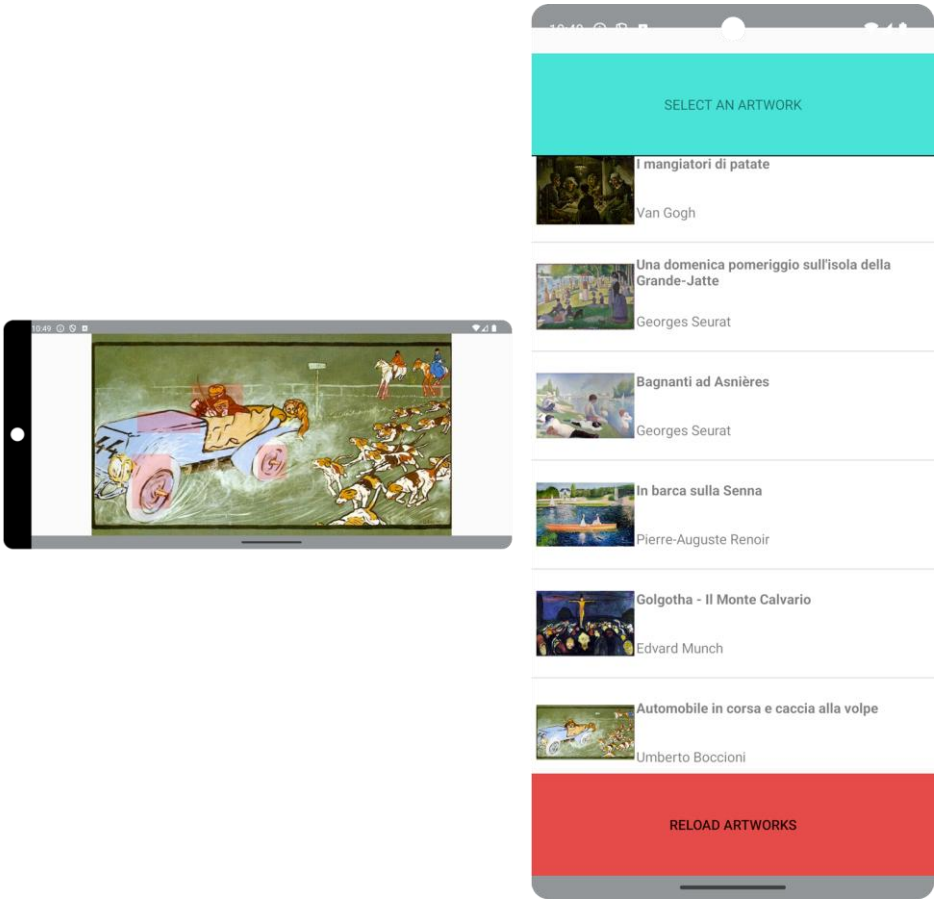

Figure 3. L2P alpha version screenshots.

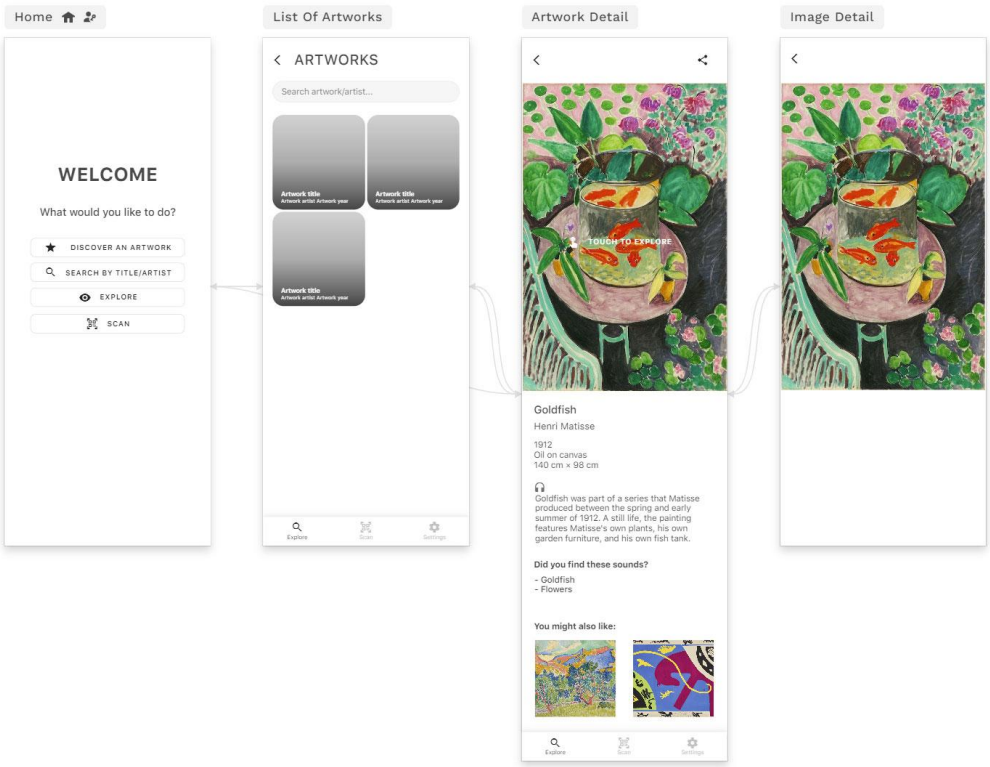

Figure 4. L2P Mockup of the final version.

Workshop

As described in the introduction of this paper, one of the goals of this research was to explore how contributing to the creation an experience that promotes users' engagement can provide the motivation to learn new skills. In particular, we focused on the computer skills required for basic image and sound processing to prepare the AAR experience, and on the basic interaction with a web portal to upload the multimedia content.

To this end, a workshop was organized in 2024, starting on November the 3rd and closing on December the 15th. From the initial number of 208 participants, only 13 decide to quit before completing the task. Participants were undergraduate students of humanities at the Department of Cultural Heritage of the University of Padova, most belonging to degrees in tourism and in history of art.

The workshop was organized in a number of tasks.

1. Select and process a digital image of a painting, ensuring its dimensions and resolution are normalized, and export it in a compressed format.
2. Identify the relevant elements of the painting and create a second digital image where small colored icons was placed on these elements.

3. Gather relevant audio files related to the subjects of the painting, perform some audio processing cut the irrelevant parts, add fades, normalize the volume, downmix the channels to mono, and finally export them in a compressed format.
4. Create textual metadata to describe the painting and the individual elements with their function withing the painting.
5. Upload the processed multimedia content to the web platform already described, along with created metadata.
6. Interact with the platform to define the parameters for the interaction: volume of audio files, associated areas in the screen, color to be shown to highlight the areas during testing.
7. Install and test an alpha version of the L2P mobile app in order to fine-tune the parameters of the interaction and eventually share their work with other participants.
8. Evaluate works made by five other participants using a double-blind peer review system. The platform used was Moodle, we did not provide any information about the authors of the evaluated artworks and the distribution was random. During the evaluation participants were asked to give a score to all the activities: audio processing, textual relevance, quality of the interaction.

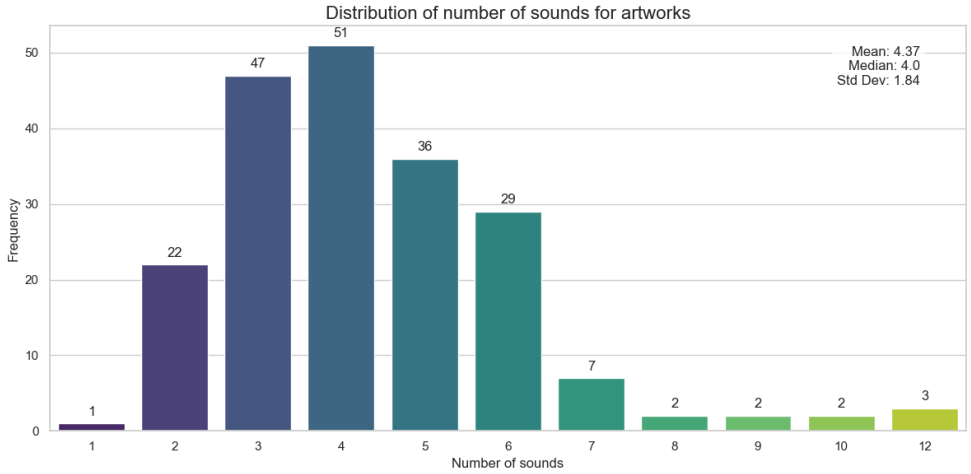

Figure 5. Distribution of number of sounds for painting.

Peer Review Analysis

Figure 6 illustrates the trend of peer evaluations among students. The scores are generally positive, ranging from a high of 97.1 to a low of 41.4, scale from 0 to 100. Given the random assignment of the works to be assessed and the anonymity of those works, we believe the evaluations were as honest as possible. While it is possible that some students received works from peers with whom they had a conflict of interest, we consider this situation negligible due to the large number of participants involved.

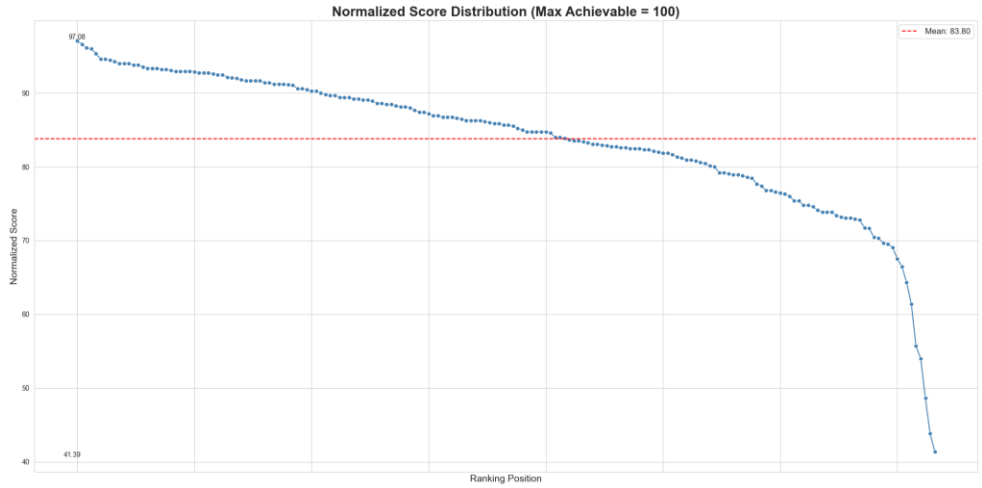

Figure 6. Distribution of peer review evaluation marks.

After observing the marks assigned, we analyzed the reviews' text. The analysis employs comprehensive techniques to extract meaningful insights from the peer review corpus. At its core, the study utilizes a pre-trained BERT-based model calibrated for Italian language sentiment analysis, enabling a nuanced interpretation of the emotional tone in each review. This deep learning approach is complemented by a thorough lexical analysis, which identifies and quantifies the most frequent words and phrases, offering a window into the key themes and descriptors prevalent in the reviews. The algorithm leverages spaCy's Italian language model for part-of-speech tagging, with a particular focus on extracting and analyzing adjectives, providing a granular view of the descriptive language used by reviewers. Statistical analysis of sentiment scores, including measures such as mean, median, and standard deviation, offers a quantitative characterization of the overall sentiment distribution.

This study uses a dedicated script integrating several Python libraries for comprehensive text analysis. We utilize NLTK and spaCy for linguistic processing, with NLTK addressing Italian stopwords and spaCy conducting both morphological and syntactic analysis. The Hugging Face Transformers library is utilized for sentiment analysis, leveraging access to pre-trained BERT-based models.

The process starts with pre-processing, which cleans the text, removes punctuation, converts it to lowercase, and eliminates Italian stopwords including contracted articles. Additionally, any HTML and CSS code present is removed before these steps are applied. The lemmatization process uses spaCy's Italian language model to reduce words to their base or dictionary form, converting inflected forms (such as plural nouns or conjugated verbs) to their lemma.

Results

Sentiment analysis indicates a predominantly positive perspective. The average sentiment score is 0.861269, reflecting a strong favorable trend. The distribution of sentiment categories supports this observation: out of 791 reviews, 675 (85.3%) are positive, 99 (12.5%) are negative, and only 17 (2.2%) are neutral. Further descriptive statistics reveal that the median sentiment score is 0.997935. These figures highlight the concentration of reviews toward the positive end of the scale.

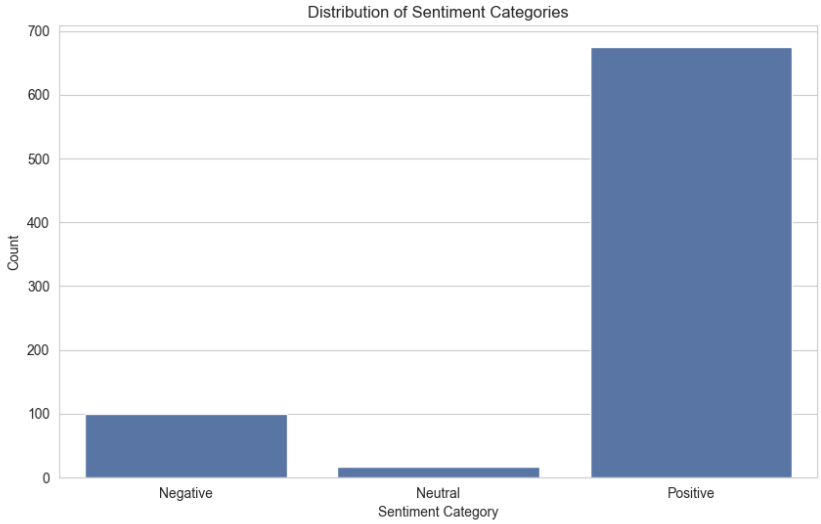

Figure 7. Distribution of the peer review sentiment.

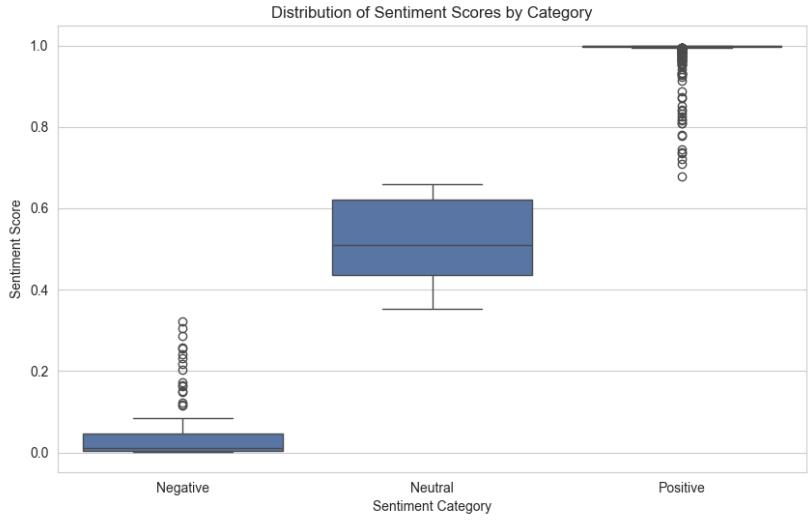

Figure 8. Distribution of peer review sentiment score.

Table 1 reports the three most positive and negative reviews. We translated them from the original language (Italian).

Positive	**Negative**

Through the chosen sounds, I was able to enjoy even better what Monet wanted to convey. The sounds are right, and the description is also good.	Unfortunately, I found errors in the image size and in the mixing of sounds, which turns out to be an audio shorter than the background sound and doesn't seem like a set of mixed sounds. It seems like an incomplete work with great potential.
The images are of the right size and have good quality, the waves are correctly positioned (I would only change the colors and the size of the wave in the sky). The sounds are consistent with the image, they are well described, and the mixing becomes a good description of the overall work. There is a slight discomfort in the sound of the party on the shore, but otherwise, all the sounds are clear and balanced.	Overall, the work is sufficient but presents several gaps and imperfections, such as the lack of the second image with the small waves or the somewhat superficial description of the artwork and its elements. Finally, in my opinion, the experience on the application is not the best as the areas concerning the sounds are not easy to find, and most of the sounds related to the elements don't start.
Well-done work, I particularly liked the choice to include a background sound. I also think the individual sounds representing the elements are very accurate. The image quality is perfect, and the initial description is very careful and detailed.	Many errors, not very detailed, and superficial. You could have done an excellent job; it's a shame about the second photo with the photos that is not present. The work is incomplete in too many aspects.

Table 1. Three most positive and negative peer reviews.

Word frequency analysis provides additional insights into the content of the reviews. The most frequently mentioned words include "essere" (to be, 1746 occurrences), "il" (the, 1477 occurrences), "di" (of, 742 occurrences), and "suono" (sound, 735 occurrences). This distribution suggests that the reviews primarily focus on existential aspects and sonic elements of the work in question.

Figure 9. Peer Review adjectives Wordcloud.

An examination of the most common adjectives adds depth to the analysis. Leading the list are "buono" (good, 127 occurrences), "singolo" (single, 96 occurrences), "completo" (complete, 95 occurrences), and "ottimo" (excellent, 86 occurrences). These adjectives indicate a general appreciation for the quality and completeness of the work. The presence of "chiaro" (clear, 72 occurrences) and "coerente" (coherent, 58 occurrences) suggests that clarity and coherence are valued aspects of the work. The frequent use of positive adjectives such as "ottimo" and "buono" confirms the generally positive tone of the reviews. Furthermore, the presence of "unico" (unique, 68 occurrences) among the most frequent adjectives suggests that the originality or singularity of the work is often mentioned and appreciated.

Questionnaire

This questionnaire aims to thoroughly evaluate participants' experiences in a workshop centered on sound and image processing, as well as the associated web portal and application. The questionnaire, which was non-compulsory and anonymous, was attended by 135 students out of 195 participants to the workshop, i.e. about 70% of the total. Utilizing a 7-point Likert scale for most of the questions, the survey addresses various facets of the workshop, including the ease of selecting and modifying works, integrating audio and images, and navigating the web portal and application. Additionally, the questionnaire evaluates participants' understanding of audio-visual processing techniques. It also investigates the effectiveness of peer review and self-assessment, providing participants the opportunity to reflect on their work with their peers. The final open-ended question solicits constructive suggestions for enhancing the workshop, portal, and application. This comprehensive approach seeks to capture a complete picture of the participants' experiences and offers valuable insights for future improvements.

Results

The workshop evaluation questionnaire revealed generally positive feedback from students. Most aspects received mean scores above 5, indicating a good level of satisfaction and ease of use across various components of the workshop. The highest-rated aspects shed light on its most successful elements. Students found the ability to listen to their sounds on the app particularly

helpful, with this aspect scoring the highest mean of 6.11 (median 7, mode 7). This suggests that practical application and auditory feedback were highly valued by participants. Additionally, students felt that the workshop significantly enhanced their understanding of sound and image processing programs, reflected in a mean score of 6.07 (median 7, mode 7). The process of selecting an appropriate artwork for the project was also considered relatively easy, achieving a mean score of 5.97 (median 6, mode 7). This implies that the initial stage of the project was well-structured and accessible to most students.

Several aspects of the workshop received mid-range scores, which still indicate positive reception but highlight areas for potential improvement. Identifying elements of the artwork to associate with sound and viewing colored rectangles on the image both scored a mean of 5.93, suggesting these visual aids were helpful, yet not universally easy to use. The tasks of inserting images and audio files were also generally well-received, with mean scores of 5.86 and 5.68, respectively, indicating that most students found these core tasks manageable. Interestingly, students found evaluating their peers' work engaging, with a mean score of 5.68. On the other hand, some aspects of the workshop received comparatively lower scores, indicating specific areas for improvement.

The use of the web portal and understanding archiving interfaces both received mean scores of 5.54, suggesting that, while generally positive, there is room to enhance the user experience of these tools. The effectiveness of the workshop in helping students understand peer review processes scored a mean of 5.43, indicating that this aspect could benefit from further development. To this end, it has to be noted this group of students is not used to participate to peer-review processes, which are seldom exploited in their degrees. Technical tasks such as modifying images, editing audio files, and inserting areas associated with audio received mean scores ranging from 5.32 to 5.11, indicating that these were among the more challenging aspects for students. The acquisition of computer skills in multimedia processing for students in humanities was one of the main goals of the project, so it is not surprising that students find the task difficult. The quality of the final results show that students developed the required competences. Notably, the task of finding audio files to insert received the lowest rating, with a mean score of 4.61, suggesting it was the most difficult aspect for many students.

The impact of peer evaluation on self-assessment received a mean score of 5.07, indicating that while some students found this process valuable for reassessing their work, others may not have experienced as significant an impact. This suggests there are opportunities to enhance the reflective aspects of the peer review process. Overall, the data presents a picture of a well-received workshop with strengths in practical application and understanding of multimedia tools. While most aspects received positive ratings, the varying scores highlight areas where targeted improvements could further enhance the student experience, particularly concerning technical tasks related to audio manipulation and maximizing the benefits of the peer review process.

The final question in the survey was open-ended, inviting participants to provide suggestions for improving the workshop, portal, app, and related components. This question offered valuable qualitative feedback. Below in Table 2 a summary of the key suggestions and comments:

App Accessibility and Functionality	A number of participants emphasized the importance of making the app available across all operating systems, not solely on Android. This concern was prevalent among many iOS users who felt disadvantaged. Additionally, some respondents reported issues related to the app's stability and functionality, indicating that enhancements in these areas would significantly improve the overall user experience.

Technical Improvements	Several requests were made for improved audio playback features, including the capability to hear overlapping sounds simultaneously. Additionally, some participants experienced challenges with saving and loading data in the portal, particularly while editing individual sounds or descriptions. These technical difficulties were a source of frustration for some users.
Clearer Instructions and Examples	A contingent of respondents suggested that clearer instructions be provided, especially concerning the peer review process. Recommendations included offering additional examples or demonstrations earlier in the workshop, akin to the provided Zoom lesson, to better clarify expectations and processes from the outset.
Peer Review System	Views on the peer review system were varied. While some participants found it beneficial, others expressed the need for modifications or potential removal. Concerns regarding the subjectivity of peer evaluations and the possibility of unfair scoring were raised. To address these issues, some participants proposed making written feedback mandatory to provide greater context for the scores awarded.
Workshop Structure and Content	There were suggestions to expand the range of artworks eligible for selection in the project, including sculptures and other forms of art beyond paintings. Additionally, some participants expressed an interest in receiving further instruction on audio and image editing tools as part of the workshop curriculum.
User Interface Improvements	Participants recommended enhancements to the user interface of both the app and the portal. Suggestions included incorporating a search function within the app to facilitate easier navigation and improving the visibility of sound-associated areas within images.
Real-world Application	A few participants proposed a closer linkage between the workshop content and real-world applications, such as developing audio guides for actual museum exhibitions. This approach could enhance motivation and provide a clearer context for the skills being cultivated.
Time Management	Some respondents noted that the workshop required considerable time investment, which posed challenges in balancing other commitments. This feedback indicates that establishing clearer time expectations or a more structured timeline might be beneficial.

Table 2. Analysis of the questionnaire free text field.

In conclusion, while many participants found the workshop to be valuable and engaging, these suggestions underscore areas where focused improvements could enhance the learning

experience, particularly regarding technical functionality, clarity of instructions, and the structure of the peer review process.

A Twin Project: L2P for "More Than Words"

The digital infrastructure of the L2P project (web portal and app) has been exploited within another initiative tailored to a diverse group of individuals with cognitive disabilities and autism spectrum disorders. It was part of a broader project titled "More Than Words", led by the University of Padua, specifically under the auspices of the *University Museum Centre*. The main aim of this initiative is to improve accessibility to university museums, ensuring they are welcoming environments for everyone. Participants, who were referred to as *talents* because they all have significant abilities in using digital media to communicate either with images or with text, contribute to the co-creation of tools to introduce the collections and help during the visit.

L2P has been used during several sessions as a tool for representing the internal space of the museum to aid persons with autism in managing the stimuli of the environment, transforming it in an interactive *sensory map* of the museum, and for interacting with the exhibits. The goal was to foster opportunities for creative exploration and expression to describe in a novel way the collections of the museums. For instance, in order to promote the Museum of Machines "Enrico Bernardi", a number of historical photos of designers and their prototypes have been sonified. Participants recorded and edited sounds that described the sense of these images, supported by the guidance of educators involved in the project. This experience was showcased at the university's closing event and included in the guided visits to the museum.

The project "More Than Words" was presented at the annual CIMUSET Conference 2024 in Vienna, themed "Within the Digital: Opportunities and Challenges in Science Museums".

Conclusions and Future Work

In conclusion, the Listening2Painting project demonstrates the potential of Audio Augmented Reality (AAR) not only as a tool for the valorization of cultural heritage, but also as an effective educational method for teaching multimedia skills, particularly aimed at students of the humanities. The project's dual nature – combining artistic enhancement with the teaching of digital skills – seems to be particularly effective. The creation and implementation process provided students with the opportunity to develop practical digital skills that are essential in today's academic and professional landscapes. Feedback from the workshop participants indicated a general appreciation for this practice-based teaching methodology. Additionally, it provided an opportunity to enhance awareness of peer review processes. This approach not only enriched their technical skills but also stimulated critical reflection on the interaction between technology and cultural heritage.

The L2P application demonstrated its potential to enhance the experience of museum visitors, illustrating how artworks can be made more accessible and engaging through sonification. This aspect will be a major focus of future research and development. On the one hand, the basic interaction through which users can listen to the sounds will be inserted in a complete experience, where users are able to access also to textual information related to the artwork and its elements. The basic idea is that, once the user focuses her/his attention on a particular element, she/he is more willing to receive information on that element. On the other hand, we plan to carry out an extensive evaluation with normal visitors, that is, users who did not

participate in the creation of the collection. We shall measure the ability of L2P to engage visitors through the use of AAR and transfer information that can be more easily retained because it was part of an experience. The dataset of works created by the students will serve as a foundation for these experiments. It also represents a valuable resource that could be preserved and expanded through institutional collaborations, potentially developing into a comprehensive repository of audio-augmented cultural heritage.

The project demonstrated significant potential for scalability and adaptation beyond its current implementation. Our initial deployment focused on a specific educational context and selected paintings; however, the underlying methodology and technological infrastructure can be extended to a variety of settings and purposes.

In terms of applicability across different museum contexts, the L2P approach could be tailored to enhance visitor experiences in diverse cultural institutions. This methodology could be effectively utilized, for example, in archaeological museums through sonification of historical artifacts, in science museums to provide auditory representations of natural phenomena, or in ethnographic collections to incorporate culturally relevant soundscapes. Each implementation would necessitate collaboration with domain experts to ensure that sonification choices are appropriate and effective, yet the technical framework remains transferable.

Finally, the experience within the "More Than Words" project showed its potential for individuals with cognitive disabilities; in particular, the possibility of accessing an artwork with a focus on its individual elements seemed of particular relevance for our group. A natural extension of an AAR approach would be to create an experience for users with visual impairments, although it is mandatory to include in the team experts in the field and final users.

References

[1] R. B. R. Dafna Naphtali, "Audio Augmented Reality for Interactive Soundwalks, Sound Art and Music Delivery," in *Foundations in Sound Design for Interactive Media*, 2019.

[2] L. Gayhardt and M. Ackerman, "SOVIA: Sonification of Visual Interactive Art," in *Proceedings of the 12th International Conference on Computational Creativity (ICCC '21)*, Mexico, 2021.

[3] D. Abhraneil, Y. Lee, A. Siddiqui, W. Santos Lages and J. Myounghoon, "Enhancing Art Gallery Visitors' Experiences through Audio Augmented Reality Technology," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2023, Vol. 67(1) 971– 977*, 2023.

[4] N. Orio, A. Micheletti and D. Zilio, "Teaching Computer Skills while Engaging Students: A Gamified Approach to Word Processing," in *GamiLearn'17*, 2017.

[5] B. Manaris, B. Stevens and A. R. Brown, "JythonMusic: An environment for teaching algorithmic music composition, dynamic coding and musical performativity.," in *Journal of Music, Technology & Education (Vol. 9, Issue 1, pp. 33– 56)*, 2016.

[6] B. J. Thenille, B. de Oliveira, G. V. Ferreira, J. R. Sato, C. Feitosa-Santana and P. Vanzella, "The effect of background music on the aesthetic experience of a visual artwork in a naturalistic environment," *Psychology of Music,* vol. 51(I), pp. 16-32, 2023.

[7] J. D. Cho, J. Jeong, J. H. Kim and H. Lee, "Sound Coding Color to Improve Artwork Appreciation by People with Visual Impairments," *Electronics,* vol. 9(11), 2020.

[8] D. Ahmetovic, N. Kwon, U. Oh, C. Bernareggi and S. Mascetti, "Touch Screen Exploration of Visual Artwork for Blind People," in *IW3C2 (International World Wide Web Conference Committee)*, Ljubljana, Slovenia, 2021.

[9] K. Rector, K. Salmon, D. Thornton, N. Joshi and M. Ringel Morris, "Eyes-Free Art: Exploring Proxemic Audio Interfaces For Blind and Low Vision Art Engagement," in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2017.

[10] S. T. Yong and J. G. Peter, "Born Digital: Are They Really Digital Natives?," *International Journal of e-Education, e-Business, e-Management and e-Learning 4,* pp. Yong, Su Ting and Peter J Gates. "Born Digital: Are They Really Digital Natives?" International Journal of e-Education, e-Business, e-Management and e-Learning 4 (2014): 102-105., 2014.

[11] D. Bawden, "Information and digital literacies: a review of concepts," *J. Documentation 57 ,* pp. 218-259, 2001.

[12] Sánchez-Caballé, M. G.-C. Anna and F. M. Esteve-Mon., "The digital competence of university students: a systematic literature review," 2020.

[13] J. Cabero-Almenara, J.-J. Gutiérrez-Castillo, F. D. Guillén-Gámez and A. F. Gaete-Bravo, "Digital Competence of Higher Education Students as a Predictor of Academic Success," *Technology, Knowledge and Learning 28,* pp. 683 - 70, 2022.

[14] J. .. Sparks and I. B. P. M. Katz, "Assessing Digital Information Literacy in Higher Education: A Review of Existing Frameworks and Assessments With Recommendations for Next-Generation Assessment," ETS Research Report Series 16-32, 2016.

[15] M. Warschauer, "The paradoxical future of digital learning.," *Learning Inquiry 1 (2007),* pp. 41-49, 2007.

[16] A. Sadik, "Digital storytelling: a meaningful technology-integrated approach for engaged student learning," *Educational Technology Research and Development 56,* pp. 487-506, 2008.

[17] A. Micheletti, N. Orio and D. Zilio, "What motivates university students to engage in learning computer skills?," in *Italian Research Conference on Digital Library Management Systems*, Bressanone (BZ), 2024.

[18] A. Kozanitis and L. Nenciovici, "Effect of active learning versus traditional lecturing on the learning achievement of college students in humanities and social sciences: a meta-analysis," *Higher Education 86 ,* pp. 1377-1394., 2022.

[19] N. Orio, D. Zilio and M. Andrea, "Teaching Multimedia Skills Through Art Sonification," in *Linking Theory and Practice of Digital Libraries*, Lubiana, 2024.