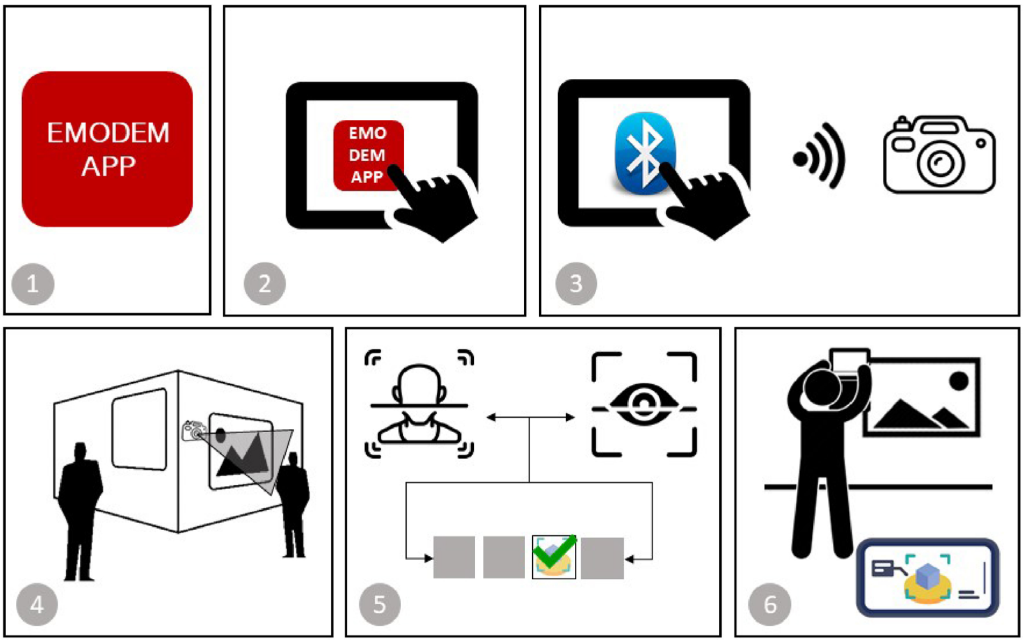
Media Convergence and Museum Education in the EMODEM Project

Paola Puma
Giuseppe Nicastro

Abstract

The interpretation of the museums heritage as an active social element is the basis of the most current cultural institutions projects that envisage forms of documentation, use and dissemination of the cultural heritage increasingly dialogic (the museum community) and dynamics (mixing conversational, experiential, and participative modes).
This scenario is the broad framework of the project conceived to design the EMODEM app based on the convergence of face detection, eye tracking and AR to interface the virtual and the physical space and make the museum experience more visitor-centered, interactive and personalized.
This article integrates the EMODEM research already underway and updates the scientific roadmap according to the progress recently achieved in phases 3 and 4 of the project, presenting the technological innovation that has intervened in the meantime in the project and the advancement of research, currently reached at third field usability testing.

Keywords
museum education, machine learning, face detection, eye tracking, augmented reality.



doi.org/10.3280/oa-845-c206

Introduction

Among the many areas of application of AI to the documentation, use and enhancement of cultural heritage, the main axes of intervention can be summarized in the three main directions: the digitization of collections for catalog optimization, the visitor behavior analysis for predictive purposes, the use of voice technologies as visit assistants [1].
This scenario is the framework of the EMODEM project based on the research started in 2019 on the sample case relating to the monumental complex of Badia a Passignano (Siena, Italy) and then in the Heritage Visual Storytelling Laboratory (HerViSt) transferred to the museum environment for designing an app based on the convergence of face detection, eye tracking and AR to make the museum experience more visitor-centered, interactive and personalized.
This article integrates the EMODEM research already underway [Puma, Nicastro 2021], confirming its general objectives and updates the scientific roadmap according to the progress recently achieved in phases 3 and 4 of the project, presenting the technological innovation that has intervened in the meantime in the project and the advancement of research, which roots on three main strategic axes:
– founding the cultural project on the crosscutting concept of human-centered and personalized use of digital technologies, ICT and AI, pursued through the expansion of the one-way relationship between artwork and visitor in a technologically mediated and personalized relationship between artwork, visitor and museum environment also conceived as a response to the new needs of post-covid museum management and fruition [ICOM 2020; Cicerchia et al. 2021; Charalampos 2022];
– deploying the multidisciplinary scientific set of the project, centered on the relationship between Digital Cultural Heritage and AI, in its scientific links of data documentation of existing reality (Survey), visual representation of digital contents and mixed reality experiences (Drawing) for the dissemination of knowledge [Bekele et al., 2018; Digital Day 2019; Giordano et al. 2021; SCIRES 2021];
– set up the workflow in a clearly and explicitly regulated way with regard to visitor privacy and data governance [UNESCO 2021].

State of the Art

The overarching theme of the most recent use of ICT in the field of cultural and museum communication directly concerns EMODEM with regard to three thematic and operational clusters:
1) the design of hardware and data acquisition devices in a specific environment;
2) data processing through machine learning and eye tracking processing algorithms;
3) the production of outcomes dedicated to the application in immersive AR + VR environments outputs.
Topic 1): the infrastructuring of museum environments using hardware and devices specially prepared for the acquisition of user data is dealt with in some recent experiments where the building is mainly detected and is analyzed in some flow components – such as the path accomplished, the number, gender, age class, time and distance of the people who observed each work – or to quantify and register visitors' behavior in terms of monitoring system using sensors and active cameras [Angeloni et al. 2021] – and minor residual attention is paid to the visitor's response in front of the artwork [Ferriani et al. 2021].
Topic 2): a large amount of studies and researches are published with rapidly growing literature on the whole scientific area concerning the AI system engineering, some of whose those are cited for their specific interest in key technical segments for the development of the project proposal, face detection and eye tracking.
Some of the most relevant deep learning approaches to pattern extraction and recognition in visual arts describe how the large availability of digitized painting and drawing collections provides great amounts of data for computer vision studies to be developed in tools to ana-

lyze the sensitivity to a given painting and interpret visual arts. The most impressive advances about image capturing using facial expressions allow to generate detailed visual characterization of the image, and partly to incorporate some aspects related to the observer reaction to the view, as emotions. [Mohamad Nezami et al. 2019].
By machine learning, computer sciences furthermore try to help the arts spreading improving the users 'ability to understand paintings [Lu et al. 2016; Castellano et al. 2021; Duan et al. 2021].
Other studies related to the EMODEM project focus on the eye tracking procedure conducted using the normal webcam. Although most of them are not focusing on museum environment these researches interest our project because they concern the way to process the screen calibration and how to achieve the best trade-off and how to connect questionnaires processing with automatic detection style based on eye tracking technology [Gudi et al. 2020; El Guabassi et al. 2019].
Topic 3): the vast majority of ICT outcomes production in the field of cultural heritage are dedicated to the application in AR, VR outputs and other types of digital environments with variable immersive levels [Bevilacqua et al.; Palestini et al. 2021; Vitali et al. 2021] involving the tactile, vocal, visual and sound dimensions [2].
As is well known, AR solutions allow viewing of digital contents integrated into a real scene; as is well established also in various experiences for museum use, AR allows to enrich the reading of a real scene by superimposing a digital layer capable of containing diversified multimedia outputs. AR therefore presents itself as an enhanced vision of reality and, precisely by virtue of the superimposition of virtual elements to the real world, creates an additional information layer directly connected to a specific physical element of reality. This process thus represents for users the access to critical readings and insights localized about the artwork' world, the final goal of EMODEM.

Research Objectives

Enhancing the museum' infrastructure role in educational programs was at the center of the interests of the DigitCH Group, today HerViSt LAB, already in the project "A museum in all senses", where through "talking" replicas of the archaeological finds was created an interactive educational set and in some way the EMODEM project is connected to that experience by using the currently potential of application to CH offered by Big Data analysis.
EMODEM fits into the conceptual framework constituted by the 'AI for society', which is declined in the sense of applying the feasibility of this technological innovation to improve the user experience in the museum by mixing Machine Learning applications and eye tracking with AR.
EMODEM also wants to realize the vision of the museum education framed in the Connected Learning Model, which adopts an socially embedded, interest-driven, and oriented toward educational, economic, or political opportunity approach to education.
Other recently conducted about museum infrastructuring researches have the objective of analyzing visitor' behavior by approaching the topic mainly from the point of view of the environment analysis and develop the issue by extracting data relating to the place where the work is placed. EMODEM' goal instead implements the museum proposal focusing its attention on the visitor and his spontaneous reactions to the artwork by using ML technologies to process the survey for the interpretation and the subsequent selection and proposal of contents consistent with the input received.
The EMODEM project is currently in a second stage of development, the programming of which has also been affected by the restrictions on access to museums deriving from the covid-19 prevention regulations of autumn 2021: after a first test conducted outdoors in Castelnuovo Garfagnana during the spring season 2021, some project parameters were redefined in view of a second applicative experimentation, which was possible to carry out only after several months, at the beginning of 2022, in the museum context of the Murate Art District (MAD) in Florence.

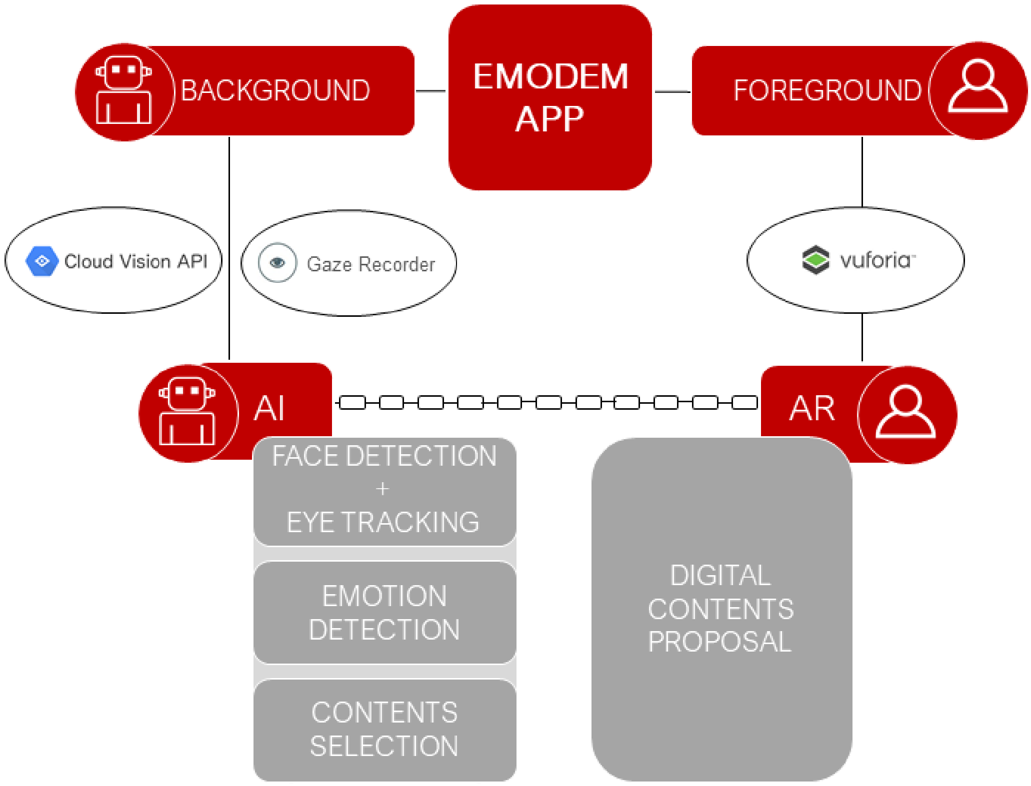


Fig. 1. EMODEM concept.

Study Design

Compared to the three thematic clusters explained in the State of the art chapter, EMODEM approaches topic 1) using low cost devices and open source sw and has its main root in topic 2), where it develops by ML the face detection segment and eye tracking to finalizing the synthesis in AR relating to topic 3 (Fig. 1).
As mentioned above, EMODEM uses image classification of the visitor' face in front of an artwork to analize and read of the reactions induced by the observation of the artwork, using a series of pre-compiled datasets that collect a large number of images relating to emotional states such as surprise, fear, joy, anger.
One of the most evident criticalities that emerged in the first research phases 1 and 2 (carried out in 2019-2020) concerned the actual possibility of the software to interpret natural facial expressions characterized by a reduced "theatricality" and, therefore, by a manifestation spontaneous.
The existing and available databases used by the working group for the alpha test in phase 1 were, in fact, populated by numerous faces images photographed in the very act of emphasizing the emotion felt: in these simulated tests, therefore, the software had been able to correctly recognize all the facial expressions proposed (precisely by virtue of the emphasis that the people photographed had placed in manifesting the different emotional states). On the contrary, the people partecipating to phase 2 test were not provided with any indication of clearly manifest the emotions aroused by the artworks and the input was to behave in a completely natural way, in the attempt to replicate the normal conditions of an artwork exhibited in a museum.
Statistical analysis of the images acquired in phase 2 under these different conditions, however, provided results that discontinuously differ from those of phase 1; taking as an example the recognition of the emotional state indicated as "joy", the overlap between the face detection and questionnaires is between 40% and 80% for 2/3 of the panel [3] and the 1/3

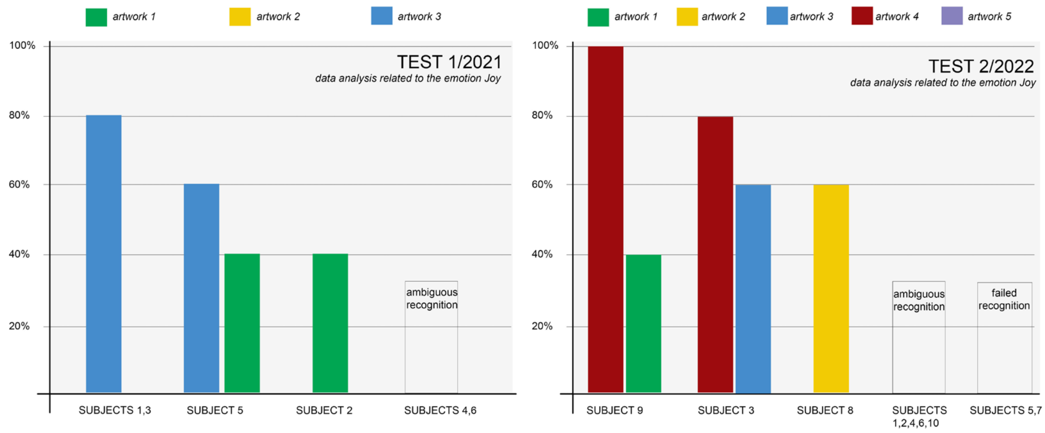


Fig. 2. Statistical analysis of the field tests.

it is cataloged by the sw with a degree of confidence too low to take these results into consideration [4] [Cohen 1988] (Fig. 2).

The results of the tests carried out in phase 2, while presenting these critical issues, nevertheless highlighted some positive aspects considered to be incentives for the continuation of the research. According to the limits of the research in progress, it is in fact necessary to point out that the face detection software used in phase 2 has a standardized training level linked to the choice of use of the open source version (Google Cloud Vision, "ready to use" version).

Despite the critical issues mentioned above, the use of a solution of this type has in any case proved effective due to its ease of use, low cost and effective recognition result of the proposed images, a condition that remains confirmed if we look at the capacity shown by the software to precisely identify common characteristics of the person photographed not directly related to the purposes of this research (sex, presence or absence of glasses, smile, etc.); therefore it is foreseeable that the future implementation of training processes specifically aimed at associating mimic characteristics and emotional states will also increase the accuracy of the application chosen in this specific task.

Based on these considerations related to the specific field of computer vision applied in EMODEM, it was hypothesized that the only information deduced from the analysis of facial expresssions needed to be integrated by other factors to be included in the input cluster taking into account other data such as body temperature, heartbeat or blood pressure (data easily acquired by sensors installed on the most popular wearable devices such as smartwatches, fit bands, etc.) or, again, data relating to the way in which people observe the images.

Based on this review, the working group therefore decided to strengthen the working methodology by integrating it with a further element of analysis including gaze detection in its workflow. The convergence between face detection, eye tracking and AR in the EMODEM workflow integrates the qualitative methodology of face detection – which detects whether the visitor is interested in the artwork – with a quantitative analysis, which detects the area of the work that the visitor is observing with greater intensity.

The integration of face detection and eye tracking technologies within the EMODEM workflow has therefore determined two distinct results: on the one hand it strengthened the methodology used, on the other hand it reconfigured the ouput-input passage of the WPs, ranging from the representation of the relevant data to the representation of the contents.

The platform for the delivery of educational contents is made up of AR solutions that can be integrated effectively in this specific phase of the EMODEM workflow, allowing to provide the vsitor with personalized information to be enjoyed directly in the vicinity of the artwork he/she is observing for enriching its reading and understanding and triggering any comparisons with further works.

Fig. 3. Project implementation.

Workplan and Research Progress

PHASE 1- task: Alpha testing has been carried out in July 2020, it consists of the selection and testing of the face detection software to be used;
PHASE 2- task: Field usability testing n. 1 has been carried out in May 2021 outdoors as it was held during the restrictions to museums access due the Covid-19 prevention. The outcome consists of the comparison between the face detection and process results of the questionnaires;
PHASE 3- task: Field usability testing n. 2 has been carried out in preparation for PHASE 4. In January 2022 a series of eye tracking surveys were carried out relating to the observation of sample images.
The measurements were conducted in a controlled environment (laboratory) and performed on a panel equivalent (in number, gender and age) to the people involved in Field usability testing n. 1.
The images proposed were also similar in number, type and exposure times. Looking at the technologies currently used in the field of oculometry, it was therefore decided to proceed using the open source "Gaze recorder" eye tracking software using the webcam integrated into a commonly used laptop.
The results related to the ocular fixation point were represented by the software in the form of specific heat maps;
PHASE 4- task: Field usability testing n. 3 has been carried out in February 2022, when it was possible to held the test indoors to verify the potential and criticality of the intermediate platform in a museum environment.
The test was conducted at the Murate Art District-MAD in Florence, a cultural district of creation and residence dedicated to contemporary artistic languages, with similar methods of data delivery and acquisition compared to those followed for field usability testing 1.
The implementation had to face a set of problems consisting of the simultaneous museum activity (visual and sound noise), the outdoor-indoor passage (adequate lighting), the disturbance of the double shooting device (face detection and eye tracking), the need to manage the attendance of participants in safety with respect to the prevention of covid-19 (shots to be carried out without a protective mask) (Fig. 3);
PHASE 5- task: lab technology validation will consist of the DBMS matching the output indications (face detection and eye tracking) with the basket of didactic contents connected to the museum heritage and the AR design of the suggested digital contents. Its implementation is in progress;
PHASE 6- task: Field usability testing n. 4 will consist of experimentation in a museum environment; it's scheduled for September 2022.

Conclusions

The increasingly broad and incisive use of digital technologies, ICTs and AI in heritage sciences promotes a profound rethinking of the notion of heritage, today increasingly understood as the pivot of society and the knowledge economy: an active social factor – of education for inclusion, relationship and promotion of the resilience of the territory – and an important resource with extensive repercussions on the value chain from an economic point of view. The most advanced museology operates in this context, where the commitment to the formation of the museum community prevails over the obsolete model of relationship between cultural institution and visitor which reduces the relationship to the impromptu presence in that single place in that specific fraction of time.
The increasingly accessible application implications of research conducted in the field of AI technologies make affordable this type of management cultural. Museum education can today even more easily use DCH to practice cultural planning that is increasingly concretely committed to outlining audience engagement and audience engagement strategies on which the museum experience is based in advanced and diversified ways.

Acknowledgments

The editorial and scientific responsibility of the chapters is recognized to: Paola Puma for Introduction, Research objectives, Study design, Conclusions; Giuseppe Nicastro for State of the art, Workplan and research progress; both the authors Paola Puma and Giuseppe Nicastro for the figure and References.

Notes

[1] See the Smart Archive Search project conducted at the Polo '900 cultural center in Turin, where smart tools are used to relate the data present in the various archives of the hub; the research conducted at the Louvre Museum aimed at analyzing the behavior of visitors regarding their stay inside the museum or their movements within the complex; the "The Voice of Art" project, in which visitors can dialogue through the use of an APP with seven works from the Pinacoteca do Estado de São Paulo, or the audio guide of the Egyptian Museum of Turin which uses the voice of a artificial intelligence to guide in the halls of the Invisible Archeology exhibition.

[2] For their area of direct reference to the topic of EMODEM we report some EU funded projects: TIME MACHINE project https://cordis.europa.eu/project/id/820323; INCEPTION project, https://cordis.europa.eu/project/id/665220; EMOTIVE project, https://cordis.europa.eu/project/id/727188; GIFT project, https://cordis.europa.eu/project/id/727040.

[3] Taking in account that a prevalent qualitative methodology was used, the number of panelist is in the usual range for this kind of test. the recruitment was carried out with incremental numbers (n. 6/10 people in phase 2/4) and, addressing the problems due the covid-19 prevention occurred all along the research, no inclusion / exclusion criteria were adopted. For the detailed description of the test delivery methods see Puma, Nicastro 2021, p. 317-318.

[4] There are cases in which both sadness and amazement have been classified as Joy. In some cases the software did not produce any valid response, returning non-qualifying results for any of the acquired images and in particular for one of the panelists the failure was total.

References

Angeloni Renato, Pierdicca Roberto, Paolanti Marina, Mancini Adriano, Tonelli Andrea (2021). Measuring and evaluating visitors' behaviors inside museums: the CO.ME. project. In *SCIRES-IT – SCIentific RESearch and Information Technology*, 11 (1), 2021, pp. 167-178.

Bekele Mafkereseb Kassahun, Pierdicca Roberto, Frontoni Emanuele, Malinverni Eva Savina, Gain James (2018). A Survey of Augmented, Virtual, and Mixed Reality for Cultural Heritage. In *ACM Journal on Computing and Cultural Heritage*, 11 (2), 2018, pp. 1-36.

Bevilacqua Marco Giorgio, Fedeli Anthony, Capriuoli Federico, Gioli Antonella, Monteleone Cosimo, Piemonte Andrea (2021). Immersive Technologies for the Museum of the Charterhouse of Calci. In Giordano Andrea, Russo Michele, Spallone Roberta (eds.) *REPRESENTATION CHALLENGES. Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain*. Milano: FrancoAngeli, pp. 167-172.

Castellano Giovanna, Vessio Gennaro (2021). Deep learning approaches to pattern extraction and recognition in paintings and drawings: an overview. In *Neural Computing and Applications*, 33 (19), 2021, pp. 12263-12282.

Charalampos Dimoulas (2022). Cultural Heritage Storytelling, Engagement and Management in the Era of Big Data and the Semantic Web. In *Sustainability*, 14 (812), 2022, pp. 1-229.

Cicerchia Annalisa, Miedico Cristina (2021). *Musei In_visibili. Visioni di futuro per i musei italiani per il dopo emergenza Covid-19*. Roma: Fondazione Scuola Beni Attività Culturali.

Clini Paolo, Quattrini Ramona (2021). Digital heritage. Common good. In *SCIRES-IT – SCIentific RESearch and Information Technology*, 11 (1),2021, pp. 1-208.

Cohen Jacob (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum.

Cooperation on advancing digitisation of Cultural Heritage (2019).
https://ec.europa.eu/digital-single-market/en/news/eu-member-states-sign-cooperate-digitising-cultural-heritage
(03 January 2022).

Duan Yingjing, Zhang Jie, Gu Xiaoqing (2021). A Novel Paradigm to Design Personalized Derived Images of Art Paintings Using an Intelligent Emotional Analysis Model. In *Frontiers in psychology*, 12, 2021, pp. 1-14.

El Guabassi Inssaf, Bousalem Zakaria, Al Achhab Mohammed, Jellouli Ismail, EL Mohajir Badr (2019). Identifying learning style through eye tracking technology in adaptive learning systems. In *International Journal of Electrical and Computer Engineering*, 9 (5), 2019. pp. 4408-4416.

Ferriani Stefano, Marghella Giuseppe, Moretti Lorenzo, Pagnutti Simonetta, Scipinotti Riccardo, Carta Aurelio (2021). Il sistema Shareart, IOT e Reti neurali per valutare il gradimento delle opere d'arte durante le visite museali. Sperimentazione presso le collezioni comunali d'arte di Bologna.
https://iris.enea.it/handle/20.500.12079/56301 (03 January 2022).

Giordano Andrea, Russo Michele, Spallone Roberta (eds.) (2021). *Representation Challenges. Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain*. Milano: FrancoAngeli.

Gudi Amogh, Li Xin, van Gemert Jan (2020). Efficiency in Real-Time Webcam Gaze Tracking. In Bartoli Adrien, Fusiello Andrea (eds.). *Computer Vision – ECCV 2020*. Cham: Springer.

ICOM International Council of Museums (2020). Museums, museum professionals and COVID-19: follow-up survey.
https://icom.museum/wp-content/uploads/2020/11/FINAL-EN_Follow-up-survey.pdf (03 January 2022).

Lu Xin, Sawant Neela, Newman Michelle, Adams Reginald, Wang James, Li Jia (2016). Identifying Emotions Aroused from Paintings. In: Hua Gang, Jégou Hervé (eds.). *Computer Vision – ECCV 2016*. Cham: Springer.

Nezami Mohamad Omid, Dras Mark, Wan Stephen, Paris Cecil (2019). Image Captioning using Facial Expression and Attention. https://arxiv.org/abs/1908.02923v2 (03 January 2022).

Palestini Caterina, Basso Alessandro (2021). Evolutionary Time Lines, Hypothesis of an AI+AR-Based Virtual Museum. In Giordano Andrea, Russo Michele, Spallone Roberta (eds.). *Representation Challenges. Augmented Reality an Artificial Intelligence in Cultural Heritage and Innovative Design Domain*. Milano: FrancoAngeli, pp. 253-259.

Puma Paola, Nicastro Giuseppe (2021). The Emotion Detection Tools. In The Museum Education EmoDeM Project. In Giordano Andrea, Russo Michele, Spallone Roberta (eds.). *Representation Challenges. Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain*. Milano: FrancoAngeli, pp. 315-320.

Vitali Marco, Bertola Giulia, Natta Fabrizio, Ronco Francesca (2021). AI+AR: Cultural Heritage, Museum Institutions, Plastic Models and Prototyping. A State of Art. In Giordano Andrea, Russo Michele, Spallone Roberta (eds.). *Representation Challenges. Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain*. Milano: FrancoAngeli, pp. 59-65.

UNESCO (2021). Recommendation on the ethics of artificial intelligence.
https://unesdoc.unesco.org/ark:/48223/pf0000380455?1=null&queryId=369f0115-0a2a-44b1-be22-4cec74cb9bf5
(03 January 2022)

Authors
Paola Puma, Dept. of Architecture, University of Florence, paola.puma@unifi.it
Giuseppe Nicastro, Dept. of Architecture, University of Florence, giuseppe.nicastro@unifi.it

 Isbn 9788835141945