



ReInHerit

Redefining the Future of Cultural Heritage, through a disruptive model of sustainability



www.reinherit.eu



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004545

Project

Project Number	101004545
Project Acronym	ReInHerit
Project Title	Redefining the future of cultural heritage, through a disruptive model of sustainability
Starting Date	01/03/2021
Duration in Months	36
Funding Scheme	Coordination and Support Action
Call (part) Identifier	H2020-SC6-TRANSFORMATIONS-2020
Topic	TRANSFORMATIONS-19-2020 <i>Culture beyond borders – Facilitating innovation and research cooperation between European museums and heritage sites</i>
Website	www.reinherit.eu

Deliverable

Work Package	WP3 - ReInHerit Toolkit
Task	T3.4
Deliverable	D3.8: ReInHerit Toolkit Phase II
Dissemination Level	Public
Type of Deliverable	Report
Leader	UNIFI-MICC
Due Date	
Submission Date	
Keywords	Toolkit, Strategy, Digital Tools, Digital Innovation, AI/CV Interactive Tools, Gamification, CH Management

Version History

Version	Date	Author	Notes
V1.0	02/2023	Marco Bertini, Paolo Mazzanti, Pietro Pala, Filippo Principi	First version submitted to SC for peer-review
V2.0	03/2023	Marco Bertini, Paolo Mazzanti, Reviewers: Panagiotis Charalambous Niki Kyriakou, Chiara Zuanni	Revision
V3.0	09/2023	Marco Bertini, Paolo Mazzanti	Revision - based on the "General Project Review Consolidated Report" No. 2
V3.1	11/2023	Marco Bertini, Paolo Mazzanti, Andrea Oratiou	Revision based on internal review
V4.0	09/2024	Paolo Mazzanti, Marco Bertini (UNIFI - MICC)	Revision - based on the "General Project Review Consolidated Report" No. 3

Acronyms and abbreviations

European Commission	EC
Research Executive Agency	REA
Grant Agreement	GA
Consortium Agreement	CA
Description of Action	DoA
Project Coordinator	PC
Steering Committee	SC
Project Management Team	PMT
Work Package	WP

Disclaimer

This document reflects only the author's view and the Research Executive Agency is not responsible for any use that may be made of the information it contains.



Bank of Cyprus
Cultural Foundation



UNIVERSITÀ
DEGLI STUDI
FIRENZE



UNIVERSITY
of NICOSIA | SCHOOL
OF LAW



MUSEUM
OF CYCLADIC
ART
NICHOLAS AND DOLLY
GOULANDRIS FOUNDATION



CYENS
CENTRE OF EXCELLENCE



Contents:

1. Digital Innovation and Sustainability	8
2. Toolkit Applications	11
3. Toolkit Components	14
3.1 Smart retrieval	14
3.2 Smart video restoration	18
3.3 Strike-a-pose and Face-fit	20
3.3 Smart Lens	23
3.4 Multimedia chatbot	29
3.4.1 Ethical use of AI tools	35
3.4.2 A Participatory Process	38
3.4.3 Co-creative Process	44
3.5 Smart tourism	55
4. Webinars	60
Appendix	63
References	64

Executive Summary

This D3.8 deliverable "*ReInHerit Toolkit Applications Phase (II)*" contains the extended description of the core technologies used in the development of the applications that will be part of a software package of the Toolkit, with a detailed description of all the digital tools developed, complementing the first part introduced in D3.3 "*ReInHerit Toolkit Applications Phase (I)*". This new report describes improvements and updates made in the previous apps illustrated in D3.3, and introduces the other Toolkit apps developed. It will highlight especially the resources produced and published in the Digital Hub of the project, which concern descriptions of the apps, instructions, training webinars, guidelines, prototypes for developing technology-assisted immersive performance and story-based games to achieve the goals concerning the innovation and sustainability themes highlighted in the strategy of D3.2 "*ReInHerit toolkit strategy*" and related to other Tasks and Deliverables of WP3.

1. Digital Innovation and Sustainability

Concerning the development of the Toolkit, in this preliminary section it is important to briefly summarize the **main outcomes** that emerged from the analysis carried out in the previous WP3 Deliverables (in particular D3.1, D3.2 and D3.3) and to also refer to the primary and secondary research carried out in the context of WP2, especially in the ReInHerit focus groups and questionnaires for museums professionals and visitors (reported in D2.3 and D2.4). This research was useful to identify the needs of cultural professionals and different institutions, also comparing them with the needs of users, diversifying them by age, cultural and educational background. Also, it allowed identifying the digital capacities of heritage organizations across Europe and selecting the digital tools to be developed in the Toolkit, considering them in an **innovative and sustainable perspective**. Digital and most innovative tools make the visitor experience more engaging using dynamic solutions to learn, enjoy, discover and interact.¹

From a visitor-experience perspective in museums, innovative technologies, and particularly AI/CV tools, are useful in motivating and creating a stronger relationship between artworks and visitors, especially to interact in a playful approach based on gamification and learning-by-doing techniques. **Young museum visitors** are more likely to interact with digital tools in a museum context and as a consequence they are to be considered the primary target audience that will use the ReInHerit Toolkit. Despite the growing number of museums and cultural institutions adopting a digital strategy, the ReInHerit survey revealed that **smaller museums** still lack the capacity (human resources, skills and funds for development and maintenance) to incorporate digital tools into their daily activities and are in need of upskilling and training on the use of innovative technological tools for heritage management.

The four main pillars (see Fig. 1) that lead the development of the toolkit are shown in the following figure. These elements, introduced in D3.2, are based on visitors preferences and museums' needs. In particular, the Toolkit strategy selected innovative and **interactive tools** able to increase visitor engagement and based on a user-centered approach. Mobile device applications have been developed as first-class targets that makes it easier to follow the **BYOD (Bring-Your-Own-Device)** approach. **AI/CV-based tools** are developed to interact with the performance environment with a playful approach, based on gamification. In order to

¹ More details about innovative and sustainable perspective in relation to our research outcomes have been underlined in "D3.2 The ReInHerit Toolkit Strategy" and "D3.3 The ReInHerit Toolkit Phase I"

adopt a sustainable management perspective, the strategic focus of the toolkit was the development of **open-source code**, so as to facilitate maintenance issues and reuse of applications by different organizations. All the codes, the additional documentation related to the apps, the associated webinars, are accessible in the Digital Hub.

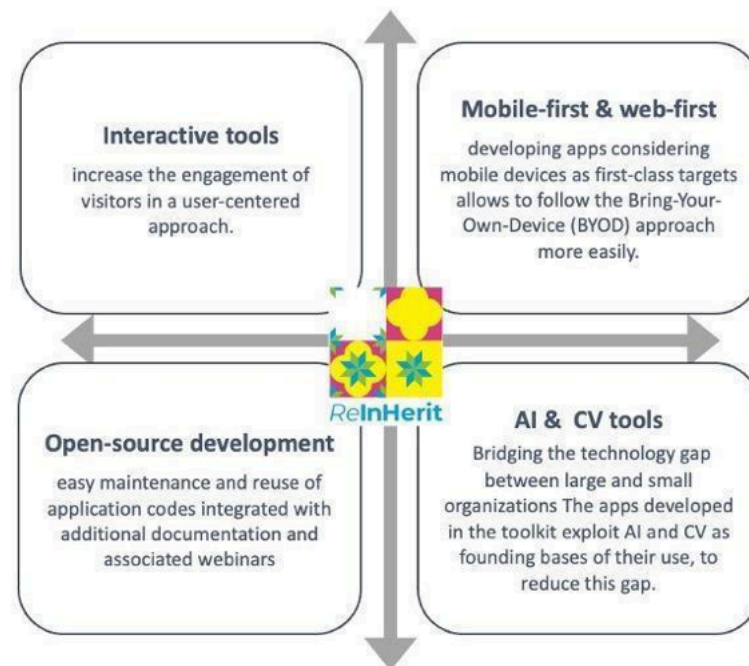


Fig. 1 – ReInHerit Toolkit Strategy

According to focus groups conducted with museum professionals, it is useful and relevant to develop tools in dialogue with visitors. The main goal of digital innovation is to provide not just a tool as a final product, but a collaborative development process, creating a mediation between different disciplinary sectors, and inviting communities into the creation process. The ReInHerit's Toolkit has been designed and tested with a **bottom-up approach**, inviting users to participate in the creation process through workshops and hackathons.

As already outline in the objectives of the Toolkit phase 1 (D3.3 section 1.3) and also In accordance with the strategy of the Toolkit (D3.2 section 5) , it is essential to develop generative artificial intelligence (GenAI) tools following a model that takes into account important critical aspects, such as the scientific accuracy of chatbot results and the ethical implications related to the use of personal and training data. In this regard, the solutions developed will seek to avoid errors and hallucinations by relying on quality content provided by experts, through the development of appropriate prompts capable of directing responses to validated datasets. sectors, including education.

2. Toolkit Applications

The activity of WP3 regarding the development of the ReInHerit toolkit has led to the design of a strategy for the development of the toolkit and the implementation of a [set of innovative applications](#) that employ AI and, in particular, computer vision and NLP.

The selection of the apps has been based on the outcomes of WP2. The development process has been carried out within the technical committee that has been formed within the project and that meets regularly to govern the process; within this framework the specifications of the applications, the design and selection of their characteristics and the revision and control of the whole development process has been carried on. One of the applications has been developed also in collaboration with WP6.

The applications developed are:



[Strike-a-pose](#) and [Face-fit](#): these applications are based on the gamification paradigm, i.e. the idea of using strategies and game dynamics into scenarios that are not a game. AI and CV are used to evaluate how much a user is able to replicate the body pose or facial expression of an artwork, providing back information on artworks and new media based on the user and his interaction with the artwork, for further sharing on social media. Both applications have been demonstrated at the ACM Multimedia 2022 conference, winning the Best Demo Honorable Mention award with the motivation of the “engaging museum experience” they provide.



[Smart retrieval](#): this application is composed of two parts - a system that allows to retrieve multimedia archive materials using a single modality, i.e. the user queries that archive using a sample image or a textual description in natural language, and a multimodal approach where the user combines a sample image and additional text that requires a modification of the visual content in natural language. The tool can be used both by professionals or end users in

web applications. The application has been demonstrated at the IEEE/CVF Computer Vision and Pattern Recognition Conference 2022, the premier annual computer vision conference, winning the Best Demo Honorable Mention award.



Smart restoration: this application, designed for historical video archive professionals, allows to restore analog videos using an innovative neural network that is capable of coping with severe degradations such as tape mistracking. Development of this application has been carried on in collaboration with the Historical Archive of Istituto Luce, which has evaluated the results and provided use cases and test materials. The system has been demonstrated at the ACM Multimedia 2022.



Smart Lens: the goal of this application, co-created with WP6 because of its use for the exhibition of this work-package, is to design a tool for a proactive exploration of artworks, asking visitors to use a smartphone as a sort of magnifying lens to inspect an artifact and using computer vision to identify details selected by the exhibition curators.



Smart tourism: this application is designed for cultural tourism and provides functionalities for landmark recognition using computer vision. The vision system is able to deal also with large monuments, where only a portion of the landmark is visible, using a specific run-time image augmentation process that is combined with a training-time augmentation. The AI and CV components will be exploited in the app for smart tourism planned in WP7.



Multimedia chatbot: this web app allows to design a novel interaction, based on natural language and AI chatbots following the very recent large success of systems such as those popularized by ChatGPT, to ask questions on artworks that address both the visual content and the context of the artwork; this latter aspect is what differentiates the system from other competing approaches. A paper discussing the technology used in the application has been presented in the Vision for Art Workshop of the European Conference of Computer Vision (ECCV) 2022 [Bongini-2022].

All the applications are released on the **ReInHerit Digital Hub**, including source code, documentation and associated webinars that describe the AI and CV technologies used to develop them. For more details on the applications and the development process the more relevant associated deliverables are D3.1 National Surveys, D3.2 Toolkit Strategy and D3.3 Toolkit Phase I. An **overview of the Toolkit** (Fig. 2) has been produced and published on the Digital Hub².

The ReInHerit Toolkit

AI and CV tools
Web applications based on Artificial Intelligence (AI) and Computer Vision (CV) designed to increase visitor engagement in a user-centred and Bring-Your-Own-Device (BYOD) approach.

Open-source development
Easy maintenance and reuse of integrated application codes with additional documentation and associated webinars.

Interactive and Gamification
Interactive tools used to motivate learning and create a stronger relationship between artworks and visitors. Gamification and playful experiences to trigger emotions, inspire creativity and digital learning.

User-Centered Approach
According to ReInHerit's analysis and strategy, the Toolkit's innovative and interactive tools are able to increase visitor engagement and are based on a user-centered approach. Web apps have been developed as first-class targets that make it easier to follow the BYOD approach. To adopt a sustainable management perspective, the strategic goal of the toolkit was the development of open-source code, so as to facilitate the reuse of applications by different organizations.

Make It Your Own!

Digital Tools for Museums and Cultural Heritage

A set of open-source codes, prototypes, instructions, webinars providing resources and guidelines to develop applications for digital interaction in museums and cultural heritage sites.

www.reinherit.eu

Horizon 2020 European Union Funding for Research & Innovation

A Toolkit Developed by MICC

Fig. 2 – Toolkit Summary

² The Toolkit Summary (PDF) can be downloaded by the button at this address <https://reinherit-hub.eu/applications/>. It has also been shared and published in the “European Heritage Hub” as announced [here](#)

3. Toolkit Components

During the second phase of the development of the ReInHerit toolkit the four applications developed within the first phase have been further polished and extended, and three other applications have been developed. The new applications are:

- **Smart Lens** - this is a **web application** that can be used to create an interactive visual guide that recognizes artwork details. The application can be executed on a mobile device like a mobile phone or a tablet. The app can use three different computer vision techniques, each one eliciting a different behavior from the user.
- **Multimedia chatbot** - this is a **web application** that can be **integrated within a museum website** or as a **mobile web** site, and provides an interaction based on natural language processing and chat. Users can either type their questions or interact with speech that is then translated to text. The system can answer questions related to the visual **content** of an artwork or on the **context** of the artwork.
- **Smart tourism** - this is a mobile app that will form the technological base of the smart tourism apps of WP5. It provides functionalities to build smart guides based on visual recognition of landmarks.

The basic technologies and the process used to develop these applications are the same used for the applications reported in D3.3.

Regarding ethical issues of the apps we refer to the annex in D3.2. Further discussion on app co-creation and ethical use of AI tools is provided in Sect. 3.4.1.

3.1 Smart retrieval

A new demonstrator, implemented as a web application, has been developed, showing the text-to-image and image-to-image search capabilities on the Noisy Art dataset [DelChiaro2019]. Currently the system has obtained state-of-the-art results on this dataset [Baldrati-2022c].

These two search modalities complete the text+image to image search available in the previous version of the app. The textual prompt can be expressed in natural language.

The following figures show an example of the starting page, displaying a gallery view of a collection, followed by two examples of image-to-image (i.e. the visual features of a selected image are used to search similar images) and the text-to-image search, where the results are retrieved based on how much they correspond to the textual prompt.

HOME

Search for an artwork inserting a description

Describe the artwork you would like to retrieve

Submit

Or browse through the artwork collection

Fig. 3 – Smart retrieval - gallery view

HOME

Artwork Image

Artwork Info

Title: Annunciation of Fano

Description: Annunciation of Fano is a painting by the Italian Renaissance artist Pietro Perugino, executed around 1488-1490, and housed in the church of Santa Maria Nuova, Fano, central Italy. The work was successful, and a few years later he was commissioned to paint an altarpiece for the same church, the so-called Fano Altarpiece.

Author(s): Pietro Perugino

Related Artworks

Fig. 4 – Smart retrieval - image-to-image search

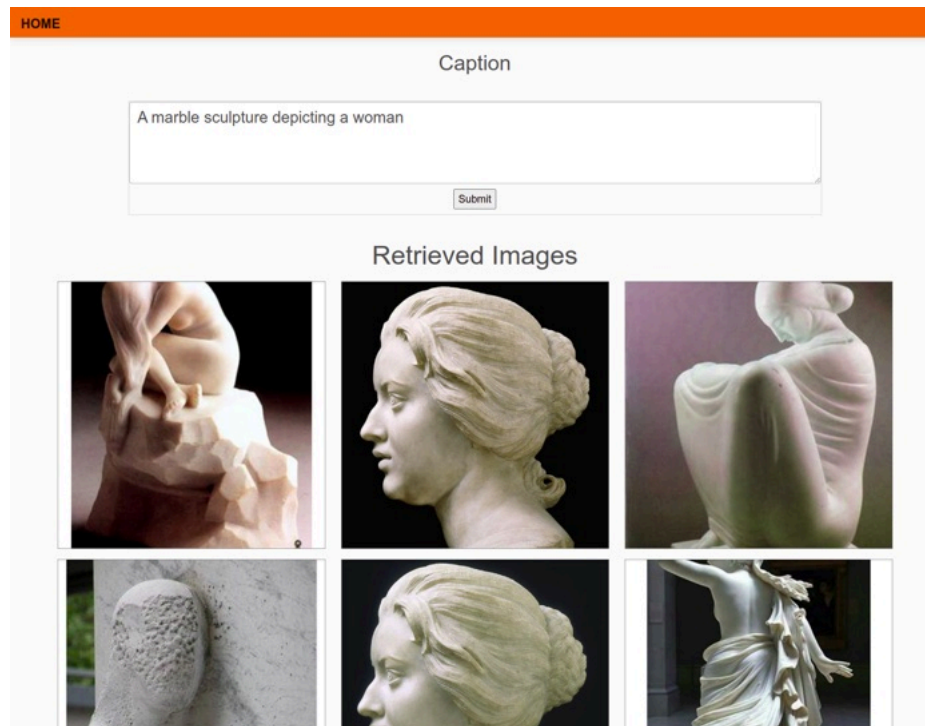


Fig. 5 – Smart retrieval - text-to-image search

The text+image to image search function (i.e. composed image retrieval) is currently being tested in collaboration with the Europeana Fashion Heritage Association³, which has provided a new dataset of historical images related to fashion.

Another extension of the work has addressed one of the main issues that hampers the development of composed image retrieval systems, i.e. the high effort and cost required for labeling datasets. This issue is particularly relevant in the cultural heritage domain where smaller organizations have difficulty in creating the large datasets required to train the neural networks required to address the task. This work, that has relevant scientific novelty and has been accepted at one of the foremost conferences on computer vision (International Conference on Computer Vision, 2023) [Baldrati-2023], extends Conditioned Image Retrieval (CIR) to a zero-shot context, that does not require a labeled training dataset, thus extending the method to any dataset. The method is still based on the CLIP (Contrastive Language-Image Pre-Training) neural network [Radford-2021] and maps the visual features of the reference image into a pseudo-word token in CLIP token embedding space and integrates it with the relative caption. The method improves on two standard datasets that were used in our previous works (and that were until now the state-of-the-art results per scientific

³ <https://fashionheritage.eu>

literature, [Baldrati-2022a, Baldrati-2022b]), i.e. FashionIQ⁴ and the more general CIRR⁵ dataset.

The motivation for this extension is due to the fact that CIR datasets consist of triplets (Ir, Tr, It) composed of reference image, relative caption, and target image, respectively. Creating a dataset for CIR is expensive as this type of data is not easily available on the internet, and generating it in an automated way is still very challenging. In zero-shot CIR, the aim is to design an approach that manages to combine the reference image and the relative caption without the need for supervised learning on an expensive labeled CIR dataset.

Our method reduces CIR to standard text-to-image retrieval by mapping the reference image into a pseudo-word which is then concatenated with the relative caption. The pseudo-word corresponds to a pseudo-word token residing in CLIP token embedding space. We refer to this mapping process with *textual inversion*, following the terminology introduced in [Gal-2023].

The new demonstrator for Smart Retrieval has been accepted for publication at the foremost international conference on Multimedia, ACM Multimedia 2023, where it will be demonstrated in live sessions and demos [Baldrati-2023b] and at the Deep Multimodal Learning for Information Retrieval workshop associated with ACM MM'23.

Scientific impact:

The new version of **Smart retrieval** has been demoed live at the International Conference on Computer Vision (ICCV), the foremost scientific conference on computer vision and AI applications, over several days and at the ACM Multimedia 2023 conference. The system was used by the attendees of the conference.

A customised version for Fashion Heritage retrieval was developed in collaboration with the The European Fashion Heritage Association (EFHA), an association that brings together more than 50 institutions from 15 European countries. The system was tested on a test dataset of more than 500.000 images provided by the association and the integration of the retrieval system with the website of the association is under development. The EFHA website is an international hub where Galleries, Libraries, Archives and Museums (GLAMs) and brands share their digital heritage assets with the goal of improving the digitisation, access and valorisation of fashion heritage resources.

Associated papers describing the apps:

- A. Baldrati, L. Agnolucci, M. Bertini, A. Del Bimbo, “Zero-Shot Composed Image Retrieval with Textual Inversion”, Proc. of ICCV 2023
The paper has received 58 citations, as of Aug. 2024.
- L. Agnolucci, A. Baldrati, M. Bertini, A. Del Bimbo, “Zero-Shot Image Retrieval with Human Feedback”, Proc. of ACM Multimedia 2023
The paper has received 1 citation, as of Aug. 2024.
- G. Cartella, A. Baldrati, D. Morelli, M. Cornia, M. Bertini, R. Cucchiara, “OpenFashionCLIP: Vision-and-language contrastive learning with open-source

⁴ <https://github.com/XiaoxiaoGuo/fashion-ig>

⁵ <https://www.zheyuanliu.me/CIRR/>

fashion data”, Proc. of International Conference on Image Analysis and Processing (ICIAP) 2023
The paper has received 1 citation, as of Aug. 2024.

3.2 Smart video restoration

The system developed in the first phase has been further improved, reconsidering the design of the neural network used to restore videos. We have designed a novel network based on a Swin-UNet, the architecture used in the first version of the application [Agnolucci-2022], that exploits both neighboring and reference frames by means of a novel Multi-Reference Spatial Feature Fusion block to restore corrupted frames. We propose to employ CLIP for zero-shot artifact detection to select the cleanest frames of each video as references. Finally, we extend a contrastive loss to a multi-frame setting to make the results more perceptually accurate. Both quantitative and qualitative experiments show the effectiveness of our approach when compared with other state-of-the-art methods.



Fig. 6 – Smart Restoration - Old Photos' Restorer

The idea is that even corrupted videos contain some relatively good frames that could be selected based on their quality, and using their visual content provides additional guidance to the neural network that is performing the restoration.

A new scientific paper that describes the novel system has been accepted at one of the foremost conferences on computer vision (IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)) [Agnolucci-2023] and refer to the publication for more details on the design of the neural network. Experimental results, evaluating standard objective visual quality metrics such as SSIM [Wang-2004] and LPIPS [Zhang-2018], show that the new

method outperforms competing approaches such as DeOldify [Deoldify], [Iizuka-2019] and also the very recent method proposed in [Wan-2022].



Fig. 7a – Smart restoration - example of real world video corruption (video courtesy of Archivio Storico Istituto Luce)



Fig. 7b – Smart restoration - example of real world video restoration using the old method (video courtesy of Archivio Storico Istituto Luce)

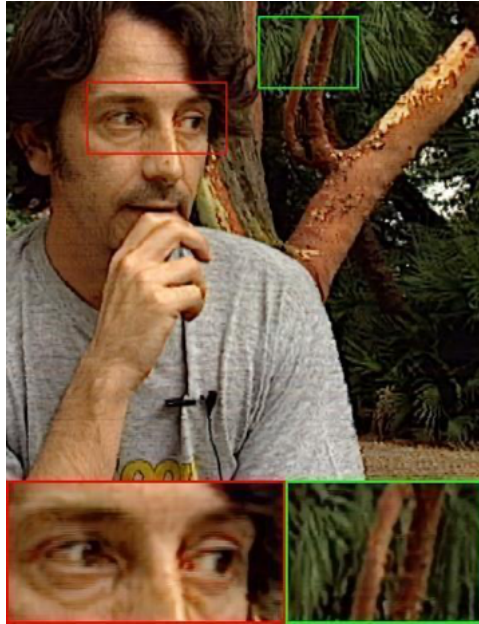


Fig. 7c – Smart restoration - example of real world video restoration using the new method
(video courtesy of Archivio Storico Istituto Luce)

Scientific impact:

The new version of **Smart video restoration** has been demoed live at the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) 2024, one of the most important scientific conferences on computer vision and AI applications, over several days. The system was used by the attendees of the conference. The system has currently the state-of-the-art results for analog video restoration.

Associated papers describing the apps:

- L. Agnolucci, L. Galteri, M. Bertini, A. Del Bimbo, “Reference-based restoration of digitized analog videotapes”, Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2024
The paper has received 1 citation, as of Aug. 2024.
- L. Agnolucci, L. Galteri, M. Bertini, A. Del Bimbo, “ARNIQA: Learning distortion manifold for image quality assessment”, Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2024
The paper has received 1 citation, as of Aug. 2024.

3.3 Strike-a-pose and Face-fit

The apps have been dockerized, i.e. to ease their installation a Docker⁶ file that contains the instructions to automatically create the correct environment for their installation and use has been created.

The initial version of Strike-a-pose that included registration of the user to create personal temporary directories where to store the downloadable videos or to play multi-user challenges has been changed to eliminate this step, effectively avoiding any type of storing personal information. The new download screen asks for the user mail and sends the link to download the video without recording the personal information. (Fig. 8)

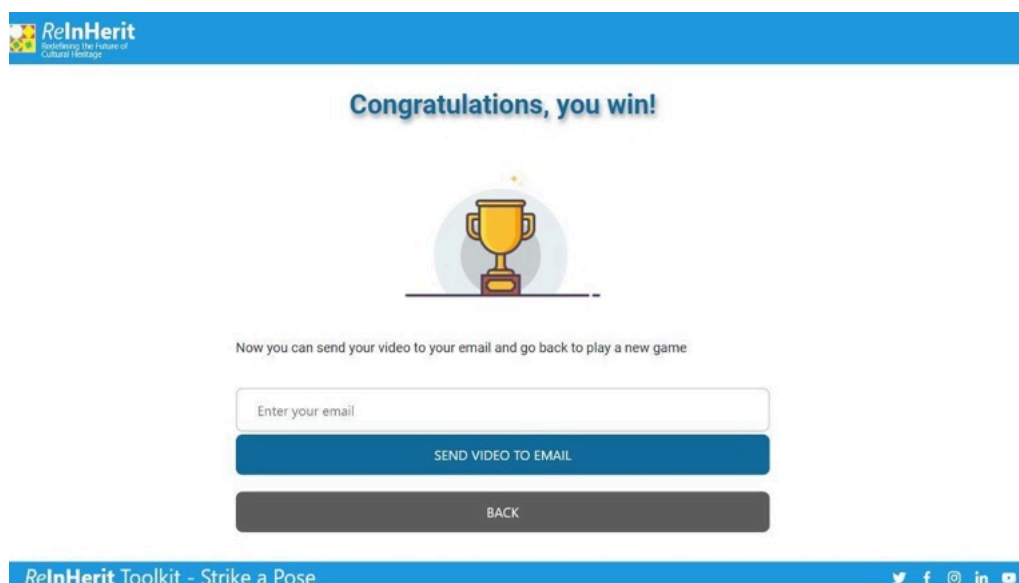


Fig. 8 - Strike A Pose Final Version - download video

The multiplayer version of the application creates anonymous ids to link different players. The interface of the online demo version including Privacy Policy has been updated and simplified (Fig. 9). The online demonstrator has the multiplayer function turned off by default to ease testing the application by inexperienced users.

⁶ <https://www.docker.com>

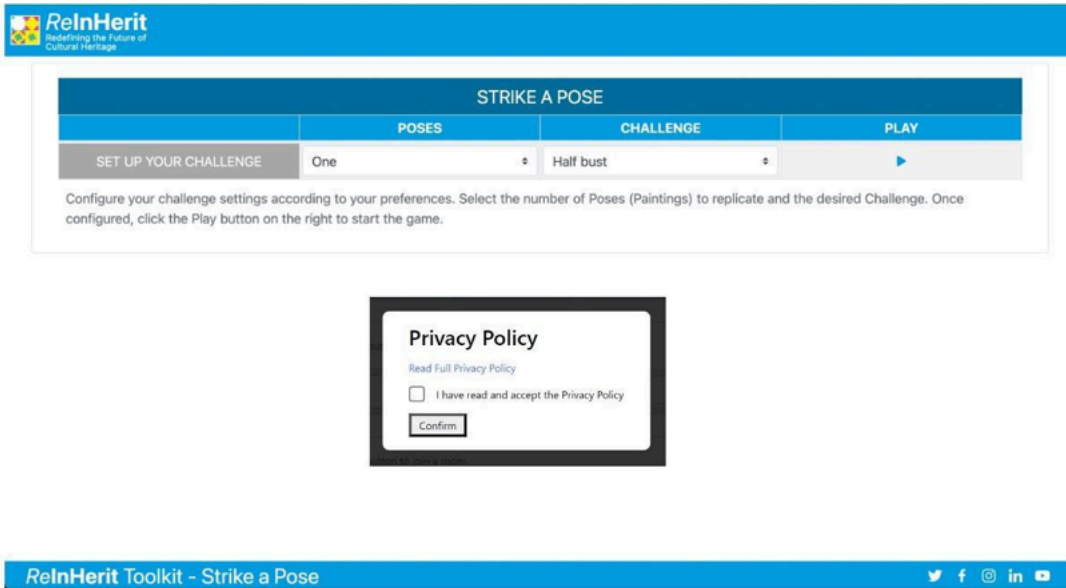


Fig. 9 - Strike A Pose Final Version - Privacy policy

The Face-Fit graphics interface has also been updated and simplified including Privacy Policy. (Fig. 10)

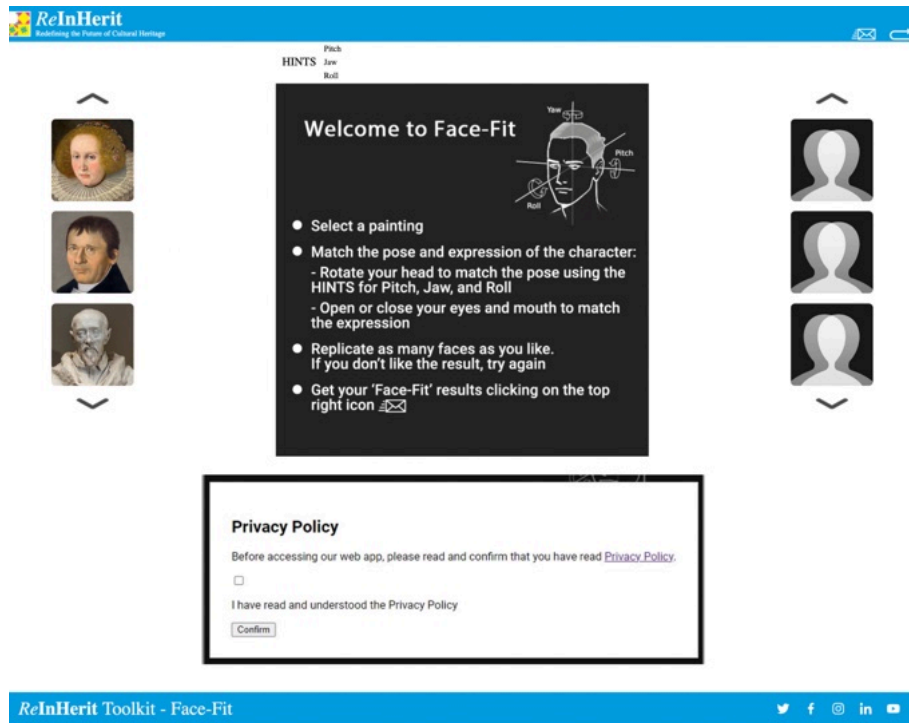


Fig. 10 - Face-Fit Final Version - Privacy policy

Users have the option of receiving the results of their interaction (images/video) via e-mail. The screenshots below are an example of the email received, with related info of the artworks and links for further exploration.

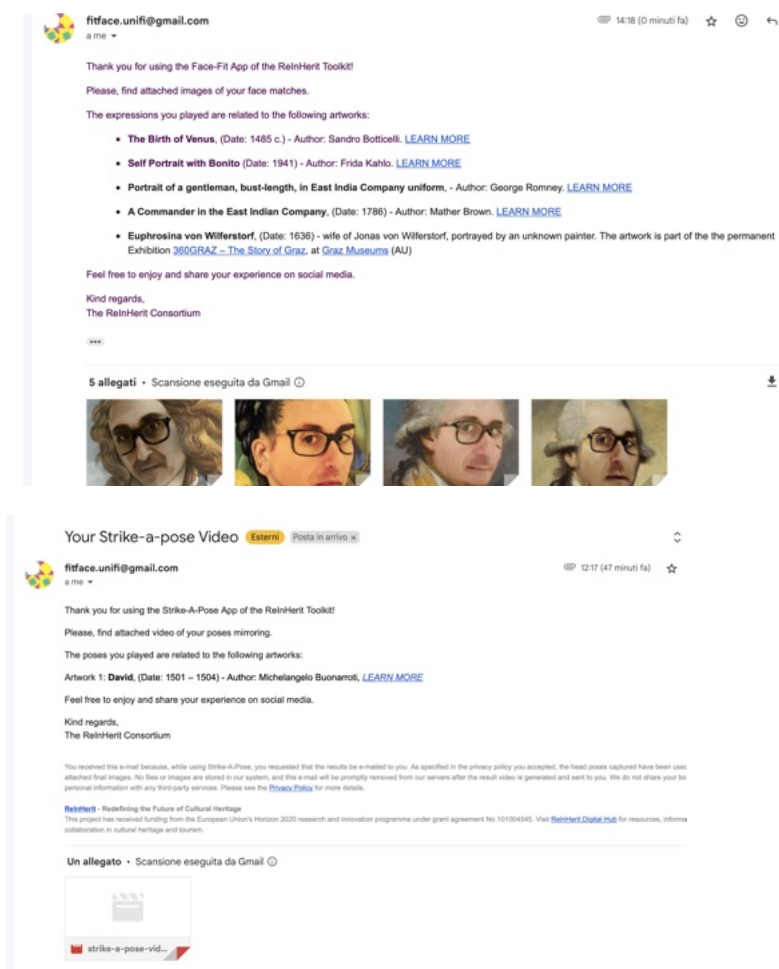


Fig. 11 - Strike to Pose & Face-Fit - interaction results

3.3 Smart Lens

The application interface has been described in D3.7, the following figure shows other examples of how the **details recognized using the computer vision system** implementing the object detection capabilities are shown.

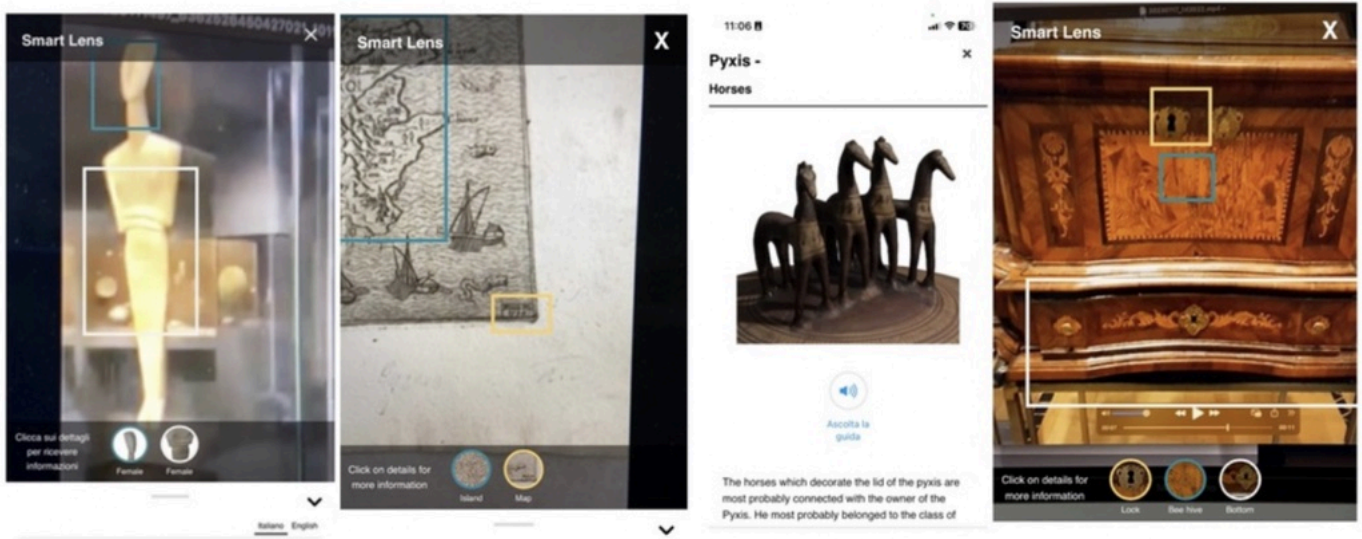


Fig. 12 – Smart Lens - examples of details recognition and information presentation (images courtesy of MCA, Graz Museum, BoCCF)

Additional work has been done to improve the training of the neural networks that implement object detection and image classification capabilities to the app. This is due to the fact that typically these types of networks are typically trained and fine-tuned on large numbers of classes of images and objects; in this context, focusing on artwork details, we have instead to deal with entities that are by definition unique. A set of two backend tools has been developed to augment the training data required to implement the fine tuning of the networks. The system has been implemented in Python and from a small set of images that are taken by the museum curators, it generates a large number of variations altering the perspective of the scene, the position and point of view of the objects, their color and luminance to simulate different types of lenses and image sensors, simulating noise and dirt on the lens, rotating the objects to account for users that employ their device in landscape or portrait orientations; examples of such alterations is shown in the following figure. Since the creation of such augmented datasets is time consuming the code has been parallelized to greatly reduce the time needed to train the networks.

These two applications are components that ease the task of training neural networks for reuse on mobile devices and are part of the ReInHerit Toolkit.



Fig. 13 – Smart Lens - examples of augmentations of the training data used to recognized details of the Cherry Raga (images courtesy of Graz Museum)

To ease the use of the training system by non-expert users, a new web-based application has been developed, integrating also the annotation functionalities required to let curators select the details of artworks from a collection. This system uses JavaScript to implement the annotation interface and to manage the training of the networks, and Python with Flask⁷ and Tensorflow⁸ on the backend to run the augmentation and training tasks. Training is launched as an asynchronous process, so that users can continue to work in the interface.

The **architecture of the system** is the following:

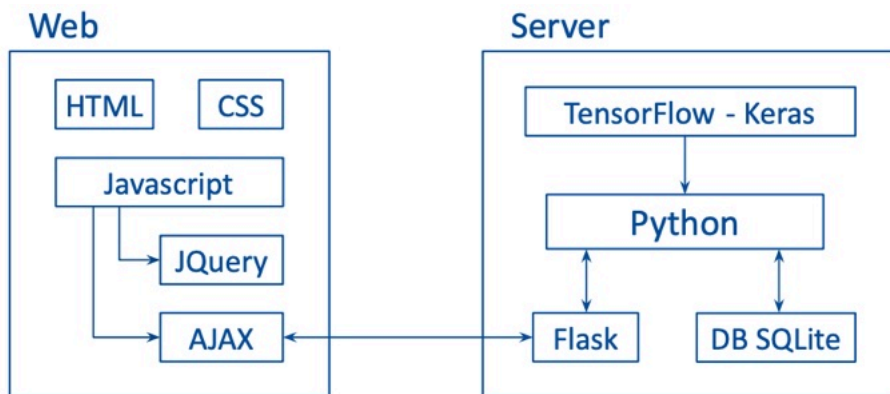


Fig. 14 – Smart Lens - architecture of the neural network training system

The system also allows to run a trained network on a set of images to provide immediate feedback to the user to understand if there is a need for further training.

Examples of the interface of the system are shown in the following images. The system allows users to create their sets of concepts and details (the figure in the following shows the sets of COCO dataset concepts used to test the system).

⁷ <https://flask.palletsprojects.com/>

⁸ <https://www.tensorflow.org>

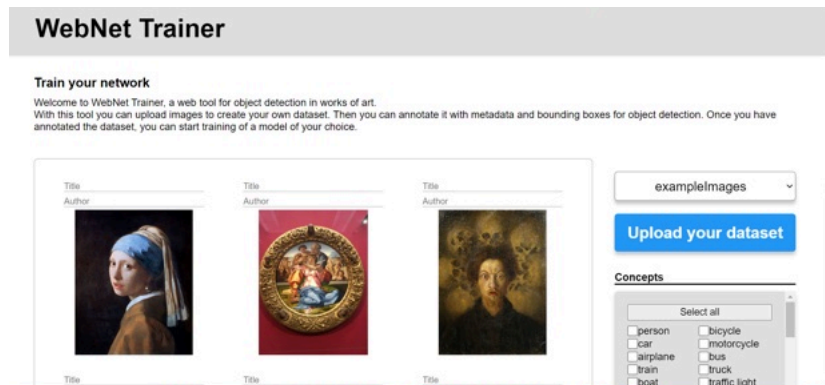


Fig. 15 – Smart Lens - training system gallery interface

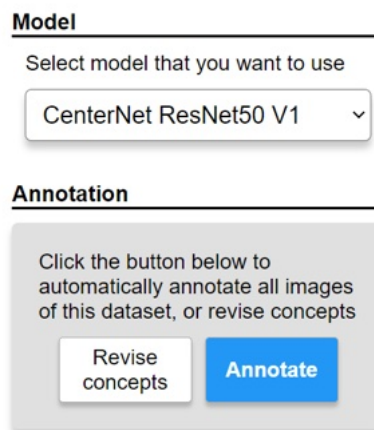


Fig. 16 – Smart Lens -training system - detail of the widget for selection of neural network



Fig. 17- Smart lens - training system | creation of the annotations of details and metadata for an artwork | Examples of automatic (red) and manual (green) annotation of details

To speed up the annotation process it is possible to use a pre-trained network, as shown in the above figure, where manual annotation is highlighted in green and the automatic results are highlighted in orange. The tool can be used, apart from training the network used in Smart Lens, also to train networks for recognition of objects and visual concepts in cultural heritage in general.

Smart Lens web application (Fig. 18) allows museum visitors to create an **interactive visual guide** that recognizes the **details of artworks**. Computer vision is used to recognize which “hotspots” of the artwork are associated with some specific information. This innovative application provides multimedia information about the recognized details, making it an intelligent guide to any collection or set of artworks users are interested in analyzing. Unlike traditional guides that rely on QR codes, Smart Lens requires active and visual engagement from users, making the experience more interactive, engaging and smarter. (Fig. 19)

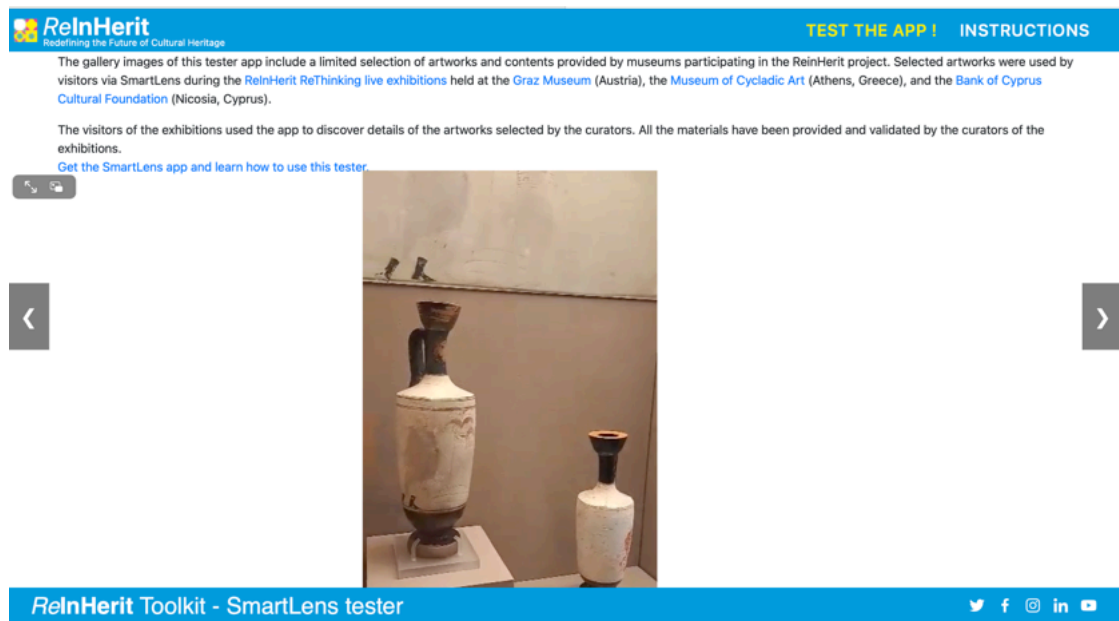


Fig. 18 - Smart Lens Tester

During the **ReInherit Workshop** held at the University of Graz on the 23rd and 24th of May 2022,⁹ MICC visited the permanent collection of the Graz Museum to explore possible applications of the Toolkit with artworks. A number of artworks were identified to be interrogated with Smart Lens (e.g. Fig 20 and 21), i.e. the artworks selected for the WP6 exhibition.

⁹ <https://www.reinherit.eu/news/artificial-intelligence-computer-vision-cultural-heritage-reinherit-workshop-graz/>

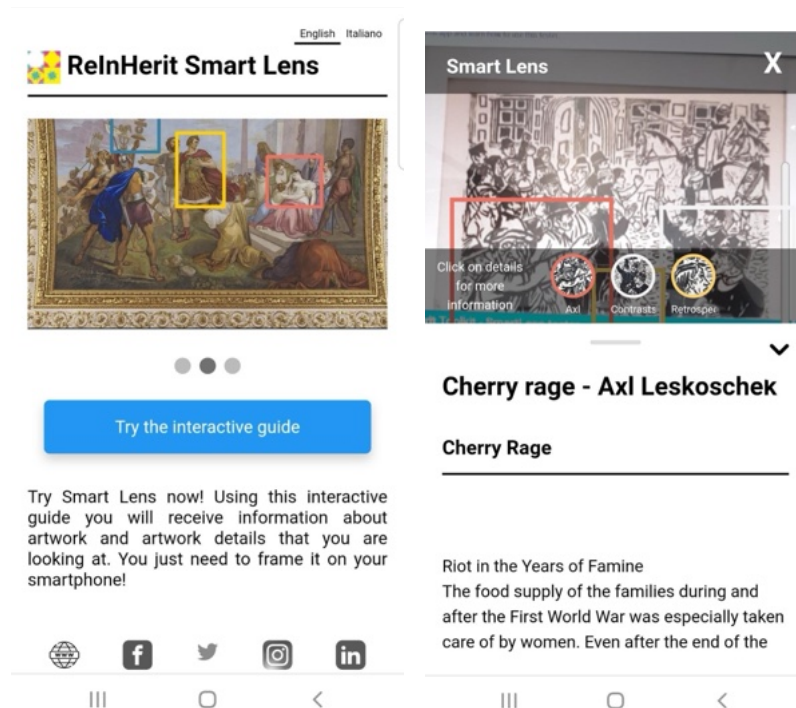


Fig. 19 Smart Lens - Interaction with the App discovering details of the artwork

The textual and descriptive contents of the artwork details were created and provided by the museum curators. Working in cooperation with the curators of all ReinHerit museums, additional artworks were also selected to be interrogated with SmartLens, used by visitors during the ReinHerit *ReThinking* live exhibitions organized at the Graz Museum (Austria), the Museum of Cycladic Art (Athens, Greece), and the Bank of Cyprus Cultural Foundation (Nicosia, Cyprus) in the context of WP6.¹⁰

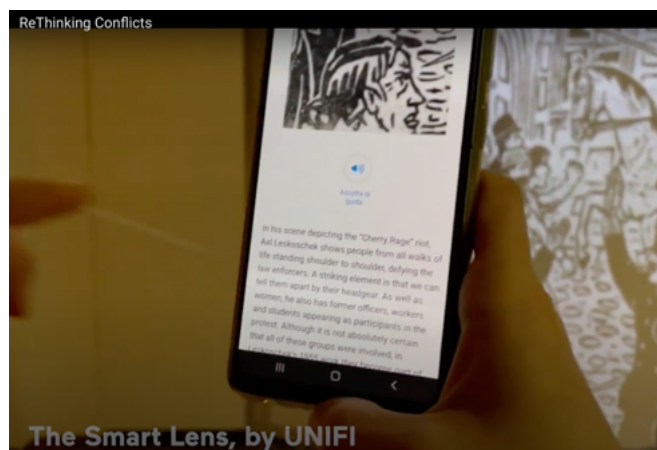


Fig 20 - Video¹¹ screenshot showing Smart Lens used during *ReThinking Conflict* Exhibition, GrazMuseum

¹⁰ <https://reinherit-hub.eu/travellingexhibitions>

¹¹ <https://youtu.be/ySTcIsiZkao?si=GTriywlr33FG1IUv&t=53>

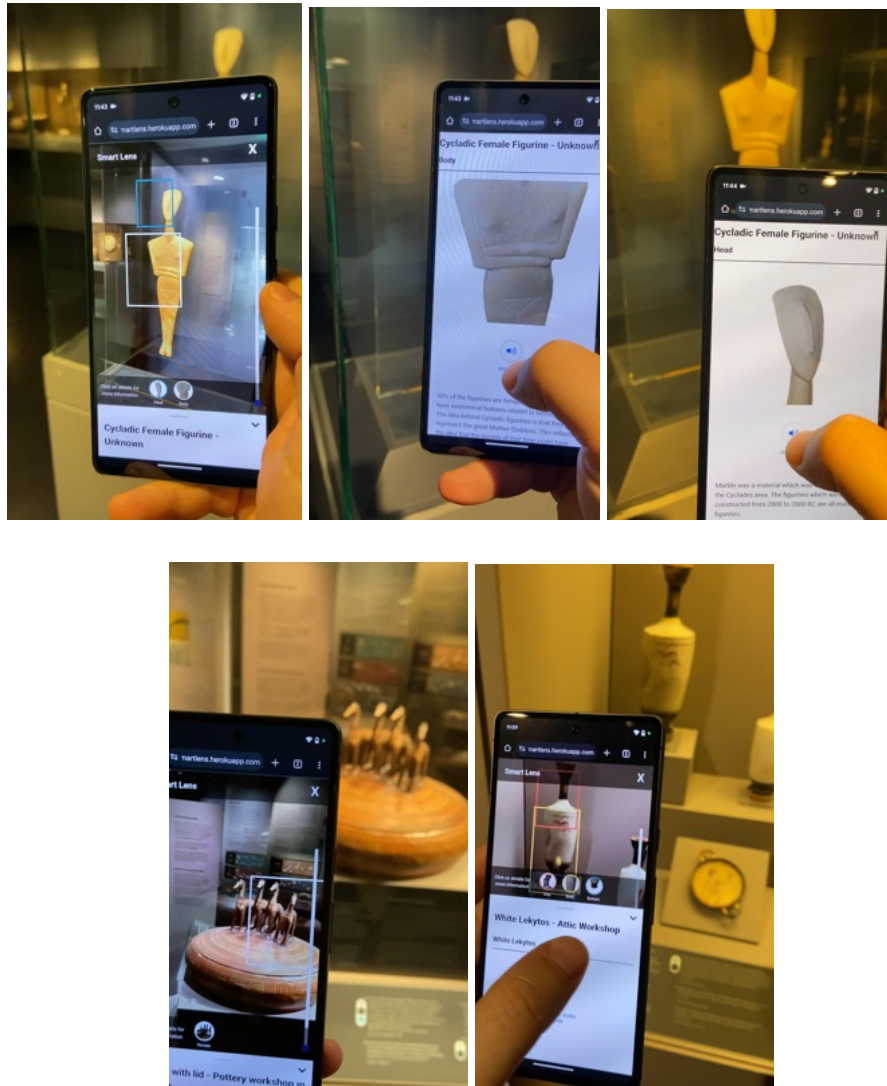


Fig. 21 - Examples of use of the Smart Lens app during ReThinking Conflict Exhibition, Cycladic Museum.

Scientific impact:

Smart Lens has been presented, along with Strike-a-pose and Face-fit, at DRHA2023: Digital Research in the Humanities and Arts 2023 University of Turin, Turin, Italy, September, 2023 Associated papers describing the apps:

- M. Bertini, P. Mazzanti, "Improving Museum Visitors' Engagement using Computer Vision Tools", Proc. of Digital Research in the Humanities and Arts (DRHA), 2023

3.4 Multimedia chatbot

This web-based application implements a chatbot system that can answer questions about visual content of artworks or about their context, e.g. about the author and history of the artwork. The design of this application is motivated by the recent huge interest in chat-based interaction that has been popularized, for example, by ChatGPT.

The application frontend is based on JavaScript, and the interface is based on a reactive design that adapts both to desktop and mobile web browsers, allowing to add this functionality to the existing websites of museums, or to a mobile application to provide a new type of smart guide. To ease this latter type of interaction, queries of the users can be performed using speech recognition, so to avoid typing long queries on the small keyboard of the device.

The backend is implemented in Python, using Flask to provide the REST API to the frontend.

There are two different versions of the backend; one implements a set of three neural networks following the design of the system that we proposed in [Bongini-2020]:

- a neural network classifies the type of the query of the user understanding if it is about the visual content or the context of the artwork;
- a neural network for question answering (QA) uses the contextual information of the artwork, stored as JSON data, to answer questions about the context of the artwork;
- a neural network for visual question answering considers the visual data of the image and the visual description of the artwork, stored as JSON data, to answer questions about the content of the artwork.

The idea of this system is to overcome the limitations of existing visual question answering (VQA) approaches, that take as input an image and a question about the image content and aim to answer correctly to the input question (see following figure). In fact VQA systems are limited in that they:

- are able to answer questions about the image content (visual questions) with a few words;
- are not able to answer questions about the image which involve external information (contextual questions) not inferable from the image content.

However, in the Cultural Heritage domain contextual questions are very frequent (when was the painting depicted?... who is the author?...)

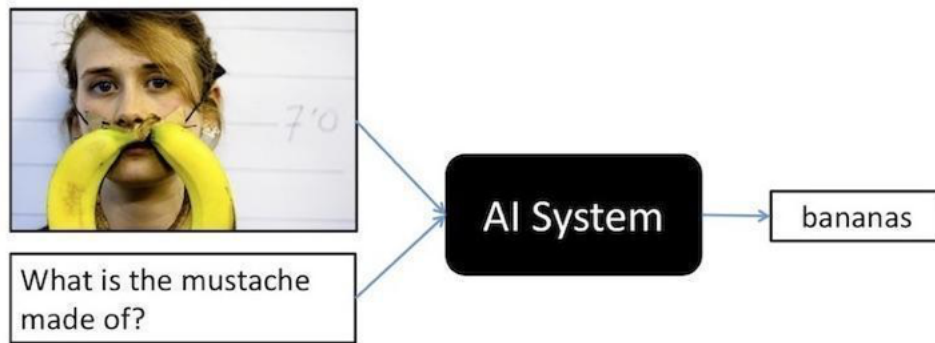


Fig. 22– example of visual question answering (VQA); the user asks questions about the visual content of an image. A neural network associates the question to some part of the image and outputs a (short) text.

The design of the first type of chatbot implemented in the Multimedia chatbot application follows thus the schema represented in the next figure.

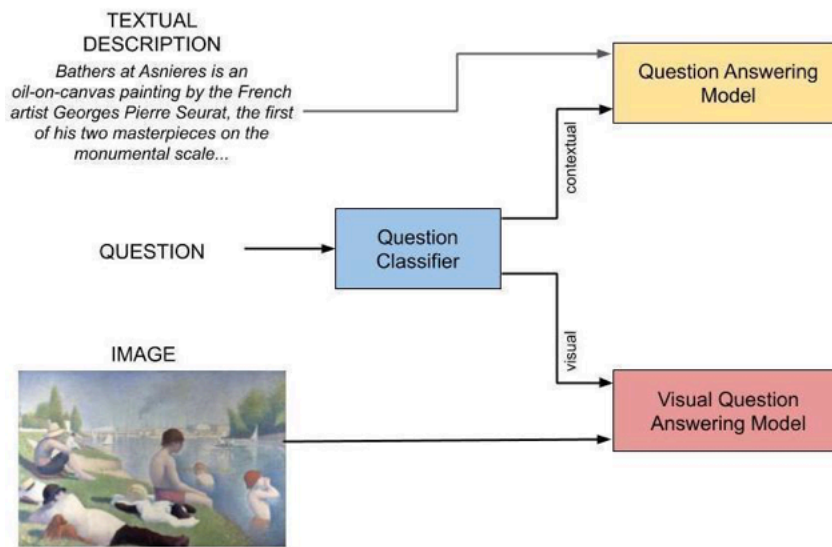


Fig. 23 – multimedia chatbot: example of visual question answering (VQA) for cultural heritage

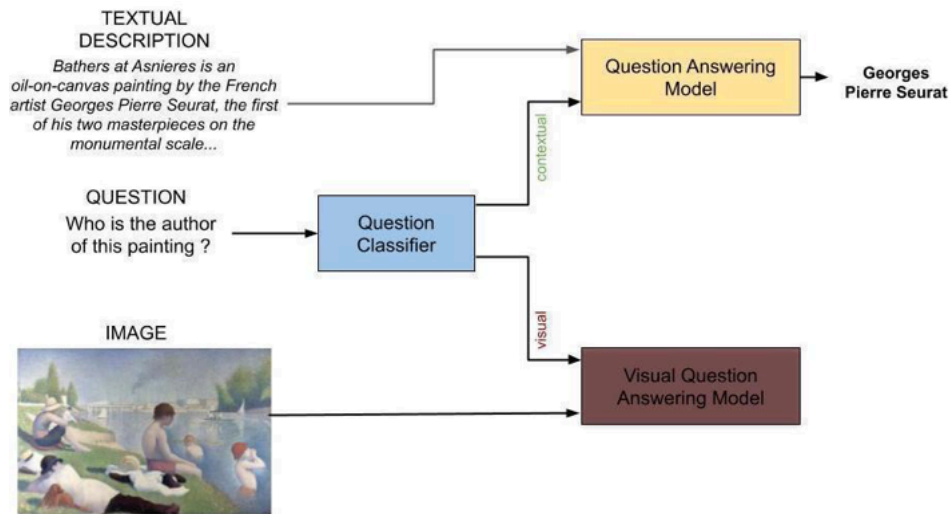


Fig. 24 – multimedia chatbot: example of visual question answering (VQA) for cultural heritage, answering a question related to the context of the artwork

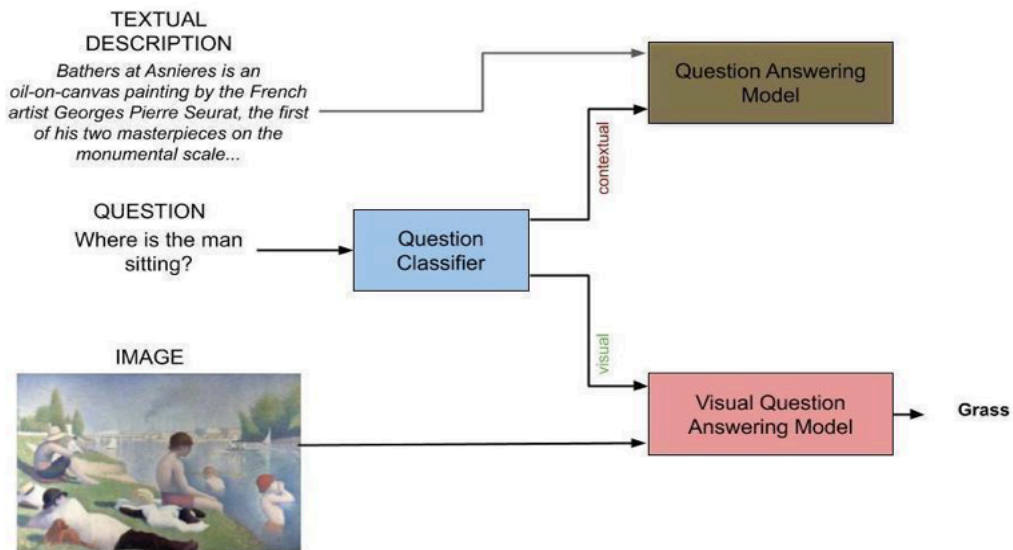


Fig. 25 – multimedia chatbot: example of visual question answering (VQA) for cultural heritage, answering a question related to the content of the artwork

The question classifier network is a network that processes only the textual information given by the question. It has the structure of a Transformer model followed by a classification head and has been trained on questions of both the VQA v2¹² and OK-VQA¹³ datasets.

The VQA network extracts the salient region features of the image using Faster-RCNN [Ren-2015] (pretrained on the Visual Genome Dataset¹⁴). It uses an attention mechanism to

¹² <https://visualqa.org>

¹³ <https://okvqa.allenai.org>

¹⁴ <https://homes.cs.washington.edu/~ranjay/visualgenome/index.html>

filter the image regions according to the input question and has been trained on examples of VQA v2 dataset.

The QA network uses an attention mechanism to find the answer to the question in the text. It has the structure of the Transformer models and has been trained on Stanford Question Answering Dataset (SQuAD) dataset¹⁵.

The following figures show some qualitative results obtained by the system, highlighting in red the mistakes.



QUESTION CLASSIFIER	QA MODEL	VQA MODEL
Who is the author of this painting? contextual		
What is in the background? visual		
When this painting was depicted? contextual		
Where is the painting now? visual		
How many people are in the image? visual		
When this painting was depicted? contextual		
What is she wearing on her head? contextual		
What is she wearing on her head? visual		
Who is the portrayed in this painting? contextual		
What is hanging on the wall? visual	<p>Who is the author of this painting? george william joy</p> <p>When this painting was depicted? 1895</p> <p>What Joy done when he was warking on this painting? borrowed bus from a company</p> <p>What was the early omnibus? horsedrawn carriage drawn along set route picking up ...</p> <p>Who mainly used this form of transport? middle classes</p>	<p>How many clocks are in the figure? two</p> <p>Is there a tree trunk? yes</p> <p>What is the color of the sand? brown</p> <p>Are we at the beach? yes</p> <p>What is in the background on the left? surfboard</p>

Fig. 26 – multimedia chatbot: qualitative results of the question classifier, question answering network and visual question answering network.

In addition to this first type of chatbot system, following the emergence and success of neural networks based on GPT architectures and training, a second chatbot engine has been added to the backend, using a GPT-based neural network. A paper describing the use of this type of neural network has been published in a workshop dedicated to applications of AI and computer vision to cultural heritage in one of the foremost conferences on computer vision (European Conference on Computer Vision 2022) [Bongini-2022].

This second system allows to create longer answers compared to the first approach, thus creating a more natural interaction. To cope with the fact that GPT-based systems tend to create texts that sound plausible only from a linguistic point of view, but that are not based on actual knowledge, the system has been designed using prompt-engineering techniques that force the GPT model to adhere to the contextual and content-based information provided by the JSON files that contain actual details and information related to the artwork.

¹⁵ <https://rajpurkar.github.io/SQuAD-explorer/>

Examples of the interface of the multimedia chatbot for mobile devices are shown in D3.7, in the following are shown examples of the web application (pilot version) designed for PC browsers. (Fig. 27-31)

The multimedia chatbot application is dockerized to ease its installation.



Fig. 27 – VIOLA Multimedia Chatbot: landing page

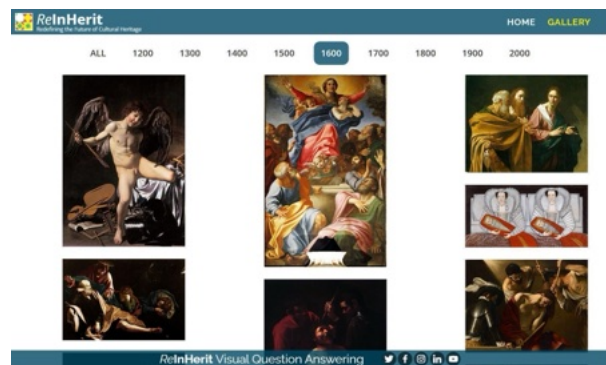


Fig. 28– Gallery view - version 1

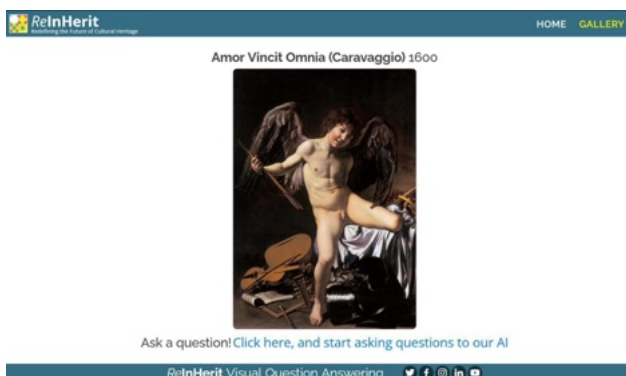


Fig. 29 – Image selection context of an artwork

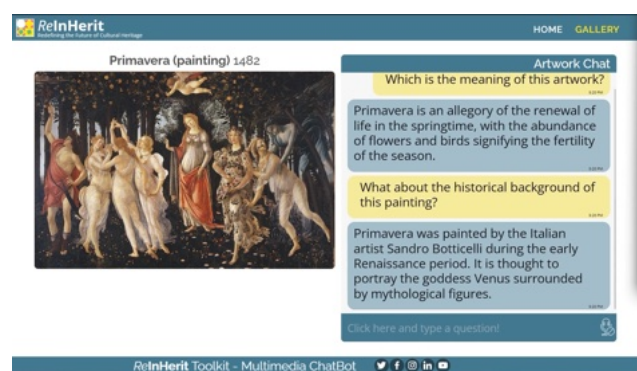


Fig. 30 – Answering questions related to the

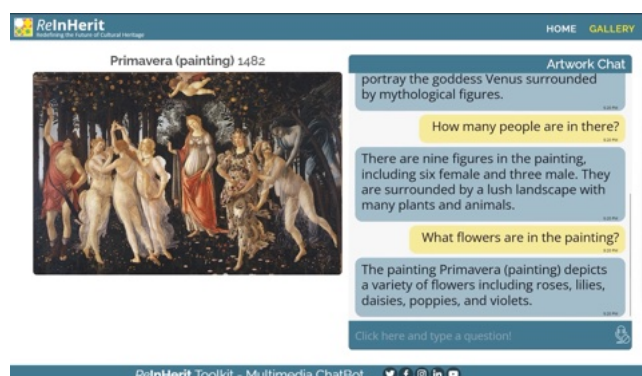


Fig. 31– Answering questions related to the visual content of the painting

The images and texts used in this first version of the chatbot, and shown in the above images, are from a standard dataset used within the Computer Vision community to assess the effectiveness of Visual Question Answering (VQA) systems, thus are not aligned with common

practices in the Cultural Heritage community. For this reason some extensions and changes to the system were implemented, and are described in the following sections.

3.4.1 Ethical use of AI tools

Concerning the multimedia chatbot VIOLA and based on the ReInHerit Ethics Card, data accuracy, ethical issues and the handling of personal data for user interaction have been better defined and specified¹⁶. In the **first pilot version**, data has been imported using a dataset (called ArtPedia - <https://aimagelab.ing.unimore.it/imagelab/page.asp?IdPage=35>) of images and texts from Wikipedia that is commonly used in the computer vision community to test Visual Question Answering (VQA) systems. The choice of this dataset was motivated by its common use as a benchmark in the computer Vision community. This version is only a demonstrative test Alpha phase with an extensive image gallery.

A **second version 2.0** of the VIOLA Multimedia Chatbot was implemented. The number of images will be significantly reduced to showcase only a gallery of images (Fig. 32) related to a limited selection of artworks provided by the Museums involved in the ReinHerit Project (GrazMuseum, Cycladic Museum, Bank of Cyprus Foundation), published online for the [Digital Collection](#)

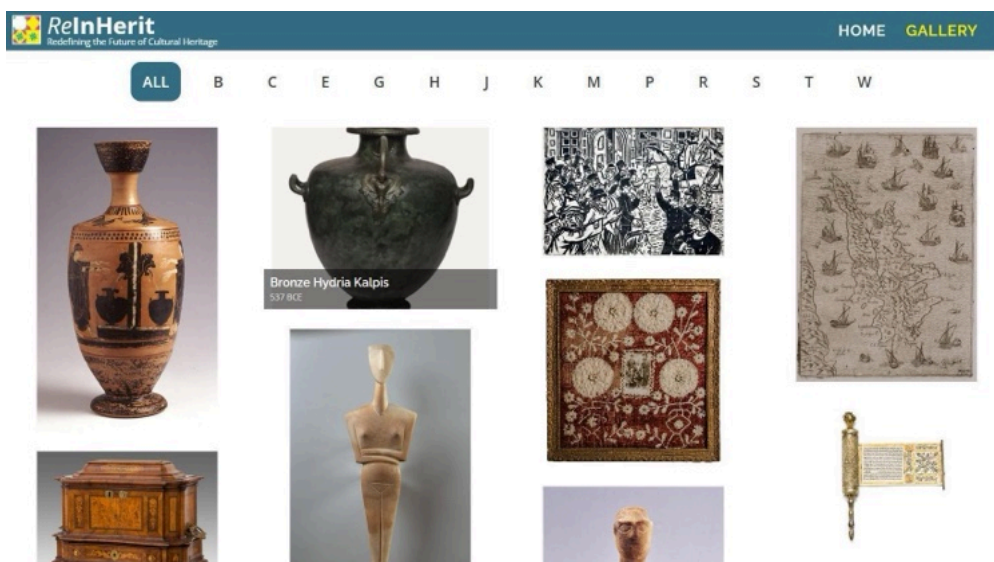



Fig.32 VIOLA Multimedia Chatbot v. 2.0 - The Gallery View

In order to include qualified content in accordance with scientific standards, a temporary “Google Artwork Addition Form” was created (Fig. 33), selecting a data structure for each artwork based on the [Object ID - International Council of Museums \(icom.museum\)](#). We

¹⁶ “Ethics annex “ D3.2 “ReInHerit Toolkit Strategy”

asked the museum curators and quality content experts of the ReInHerit Consortium to fill in the fields of the Form (Title, Subject, Measurement, Type of object, Date, Maker, Materials and techniques, Description, Location, Additional Links). Cooperation with museums in content management and their contribution is useful to collect information about the artworks. Related data helped to validate the chatbot's results, ensuring that the system's answers are in line with scientific accuracy.



ReInHerit | VIOLA Multimedia Chatbot - Artwork addition form

In order to include qualified content in accordance with scientific standards, we selected a data structure for each artwork based on the [Object ID - International Council of Museums \(icom.museum\)](https://icom.museum)

We kindly ask museum curators and quality content experts to contribute the necessary information about the artworks in your museum. Your data will help us validate the chatbot's results, ensuring that the system's answers are in line with scientific accuracy. Thank you for your collaboration!

Fig. 33 Temporary Google Artwork Addition Form

The chatbot works this way: the questions of the users are encapsulated within a set of instructions that force the chatbot to answer only using the knowledge about the artwork provided with this form, to avoid the hallucinations typical of chatbots by GenAI tools like ChatGPT. This solution will seek to avoid errors by relying on quality content provided by experts, through the development of **appropriate prompts** capable of directing responses to the validated dataset collected. The more the textual description the better the chatbot will answer the questions of the users.

In place of the Google Form, a final **backend CMS interface** for museum curators (Fig. 34) has been developed, in which museum curators can add and edit the artwork data, also with the possibility to upload a code via a JSON file. (Fig. 35)

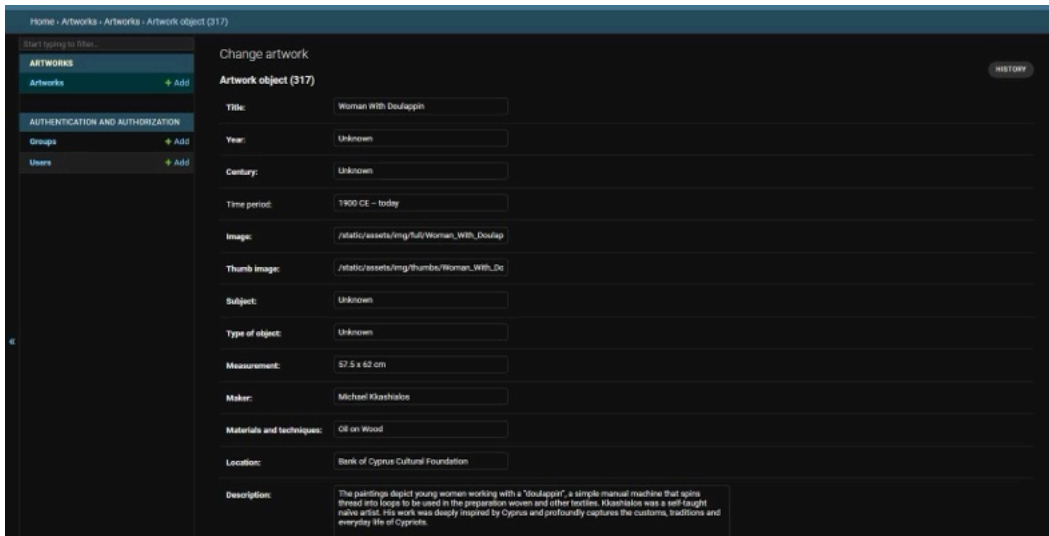


Fig. 34 Backend interface for Manual input of data

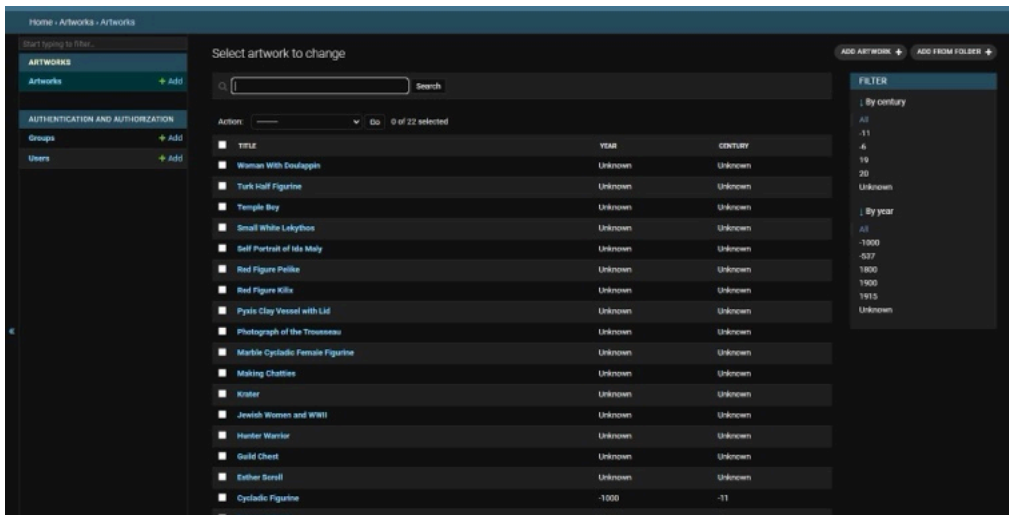
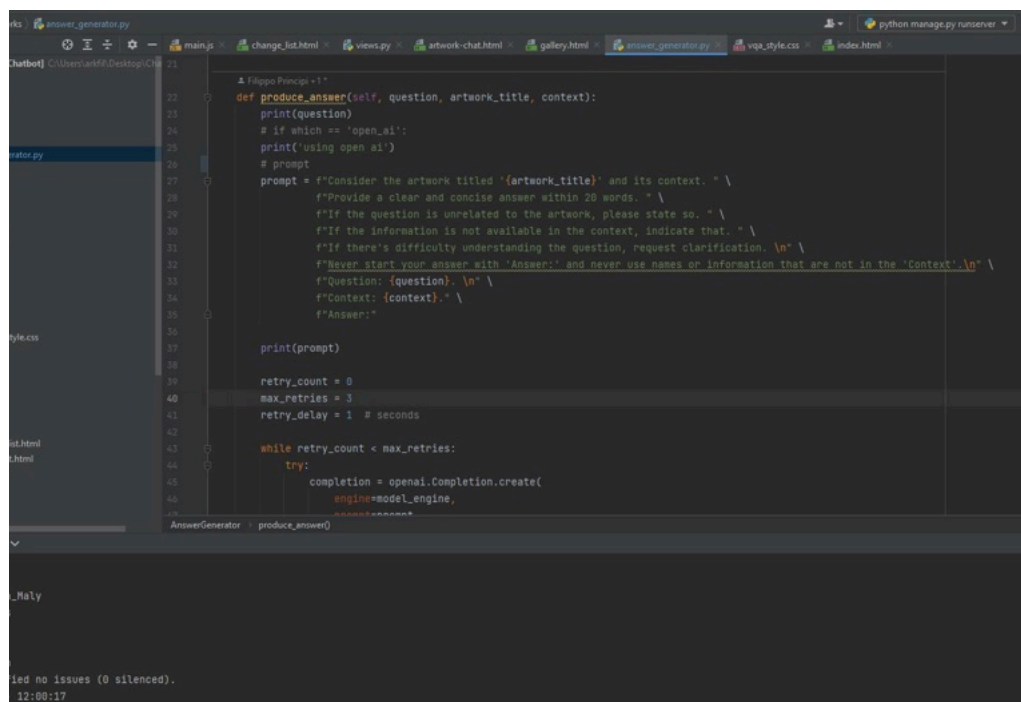


Fig. 35 Uploading data via a JSON file

VIOLA Chatbot was updated with a prompt that computes the answer from the contextual knowledge associated with the artwork; the prompt instructs the chatbot to answer using only this information to guarantee that hallucinations are not introduced. (Fig 36) This approach is related to a new technique called **Retrieval Augmented Generation (RAG)**, in which the system outputs answers based on knowledge from the questioner's request and a database of existing information (in this case the documentation of the artwork provided by a domain expert as a curator); the role of the chatbot is thus to understand the question and compose the answer based on data prepared by an expert; in this context the LLM is used as a sort of query tool that allows to search information in a database using natural language. This approach has additional benefits: LLMs may not be trained on specific artworks that are lesser known and available in the training datasets of the LLMs, thus more prone to hallucinations. Being able to update and adapt the descriptions used by the chatbot allows to reduce costs needed to implement fine-tuning of the models, another possible approach to

reduce hallucinations. The chat functionality based on ChatGPT will be validated and controlled based on the content provided by the Museums, ensuring validated and accurate responses.¹⁷



```
def produce_answer(self, question, artwork_title, context):
    print(question)
    # if which == 'open_ai':
    print('using open ai')
    # prompt
    prompt = f"Consider the artwork titled '{artwork_title}' and its context. " \
            f"Provide a clean and concise answer within 20 words. " \
            f"If the question is unrelated to the artwork, please state so. " \
            f"If the information is not available in the context, indicate that. " \
            f"If there's difficulty understanding the question, request clarification. \n" \
            f"Never start your answer with 'Answer:' and never use names or information that are not in the 'Context'." \
            f"Question: {question}. \n" \
            f"Context: {context}." \
            f"Answer:"
    print(prompt)
    retry_count = 0
    max_retries = 3
    retry_delay = 1 # seconds
    while retry_count < max_retries:
        try:
            completion = openai.Completion.create(
                engine=model_engine,
                prompt=prompt)
            return completion.choices[0].text
```

Fig. 36 Prompt Code for validated and accurate data

3.4.2 A Participatory Process

Following the **testing** of the chatbot conducted with users and experts, results showed that answers were sometimes missing or incorrect. These tests highlighted the need for a more dynamic and participatory process for updating content, particularly the "artwork description" field. To address this, a dynamic and participatory process was implemented using the CMS for content management, involving developers, curators, and users. The goal was to enable the VIOLA Chatbot to accurately and scientifically answer as many questions as possible based on user interactions.

In this process, the primary goal of VIOLA "**prompt engineering code**" is to prevent incorrect answers. The **textual data describing artworks**, provided by museum curators, mainly helps to prevent hallucinations or incorrect responses. Secondly, it is also important to minimize the number of unanswered questions.

¹⁷ "Please refer to D3.2 - section 6 Ethics Annex

The chatbot's **interaction mechanism** is based on **dialogue** and is designed to cater to diverse audiences, addressing not only historical questions primarily posed by experts or curators but also the general curiosities and stories about the artworks from everyday users. During the testing workshops conducted as part of the dissemination activities of the ReInHerit Project, we observed that users generally start the conversation with questions driven by curiosity or related to visual content rather than historical or scientific inquiries. As the conversation progresses, the questions become more detailed and related to the historical-scientific context, influenced by the answers received during the interaction.

Therefore, answers to questions about **curiosities** and stories related to the artwork motivate users to continue interacting with the chatbot. Users progressively learn new insights into the historical and cultural context of the artwork, inspired by the responses that may prompt further questions and explanations on **historical, cultural, and scientific details** and aspects. Designed as a tool to support museum visits, usable during or before/after the visit, the chatbot should not only provide correct answers but also engage and interest the user.

There will always be new and some unanswered questions, but we have followed a strategy to reduce this number. The priorities in design and development are:

- 1) **Prevent incorrect answers.**
- 2) **Provide as many answers as possible regarding both curiosities and its historical, scientific, and cultural context.**

To achieve these goals, we developed “**prompt engineering** to generate and process correct answers and collaborate more closely with museum curators to create descriptive texts of artworks containing accurate and detailed “**user-oriented scientific information**”. By using prompts that work with detailed, curator-validated descriptive texts, we ensure the information provided is correct, scientific, and responsive to user curiosity. To enable this, we also created two technical functions:

- **CMS admin interface** for museum curators to log in, manage, and edit all descriptive content related to the artworks. (Fig. 37)
- System to track and store all **unanswered questions**, creating an exportable list for museum curators. This list is useful to understand visitors' curiosities and missing descriptive information. (Fig. 38)

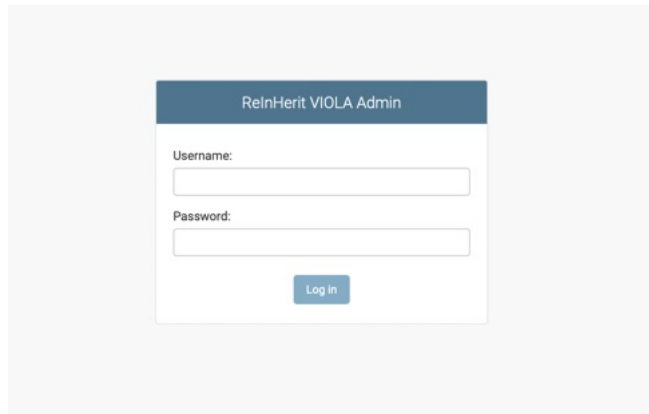


Fig. 37- CMS - New VIOLA Admin interface

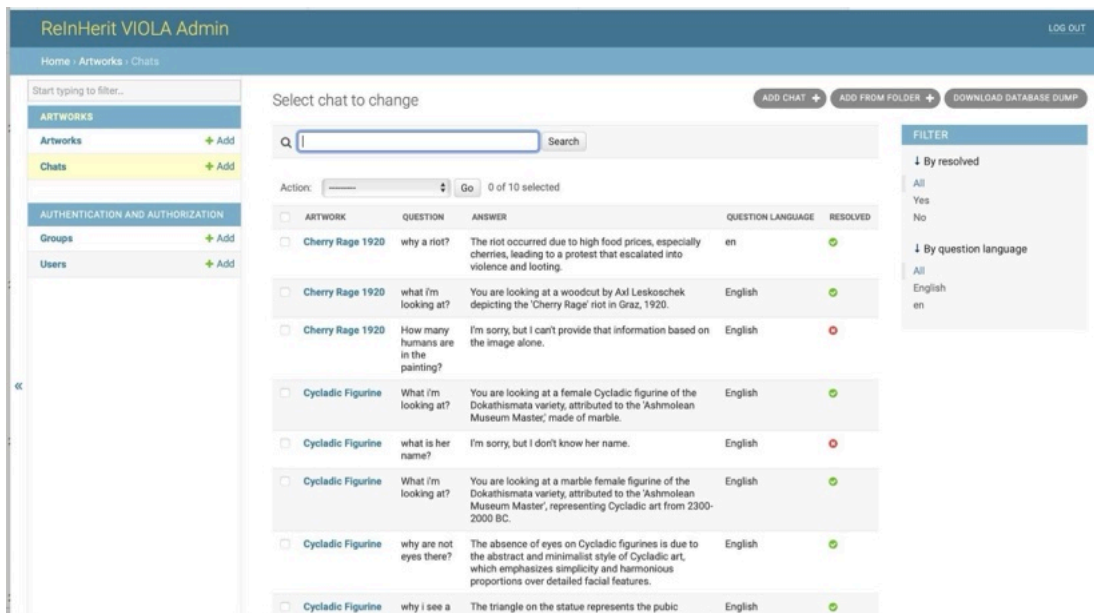


Fig. 38- ADMIN CMS New version with the unanswered questions list

This interdisciplinary working method highlighted the importance of closer collaboration between developers and museum curators as a best practice, ensuring the continuous updating of artwork-related content. In this way, the content is not static like in catalogue records but dynamic and curiosity-based. It is open to dialogue and promotes a human-centered definition of the museum, following a **collaborative and participatory approach**. (Fig. 39)

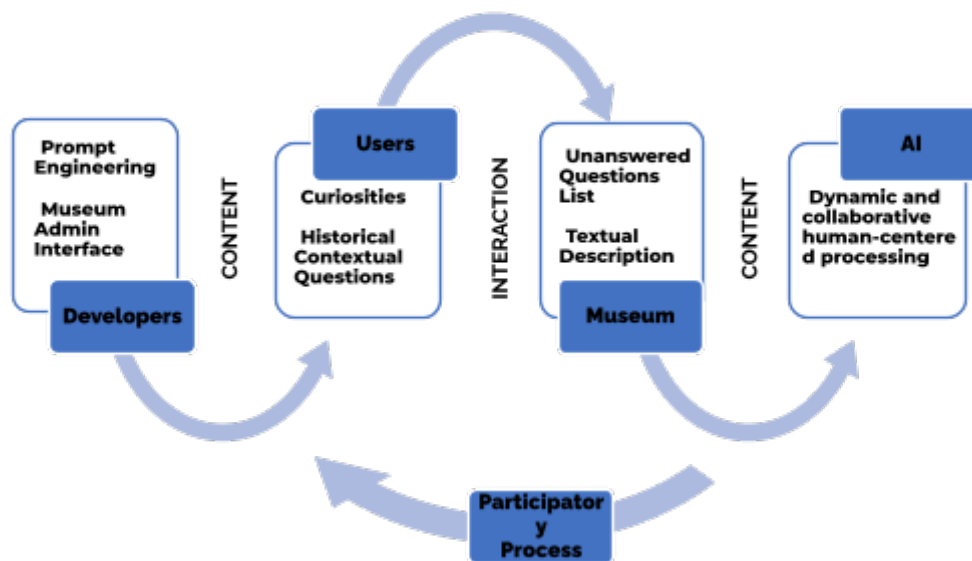


Fig. 39 - Participatory Process to generate “user-oriented” and quality contents

The third **final version** of the online Gallery has been updated and contains a selection of artworks from ReInHerit Consortium museums. These artworks include more detailed descriptions integrated with additional information validated by museum curators, specifically from the Cycladic Museum, GrazMuseum, and the Bank of Cyprus Cultural Foundation. Before entering the gallery, users are required to read and confirm the Privacy Policy, where it is stated that unanswered questions are collected as part of a collaborative approach to improve the descriptive content of artworks. Unanswered questions and answers will be kept anonymously and not shared with third-party services. (Fig. 40)

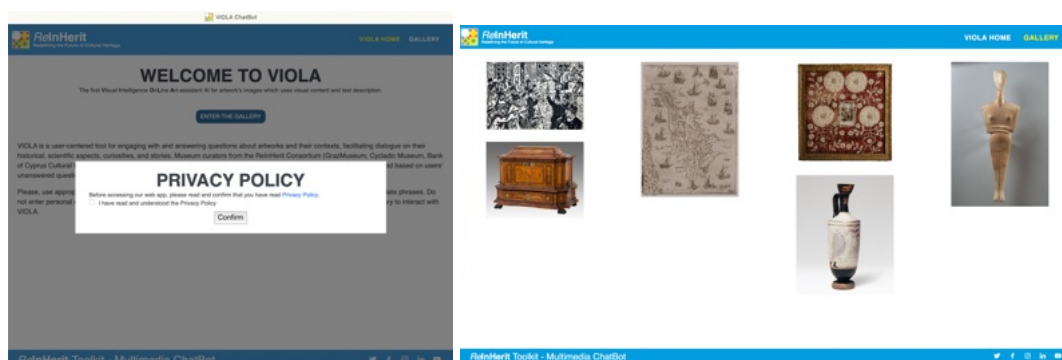






Fig.40 - Final Version VIOLA Home Page, Privacy Policy and Final Gallery

The following artworks were selected by the museum curators:

MUSEUM	Museum of Cycladic Art 	Graz Museum 	Bank of Cyprus Cultural Foundation 
ARTWORK	Cycladic Figurine	Cherry Rage	Cocoon Frame
ARTWORK	Small White Lekythos	Guild Chest	Cipro Map

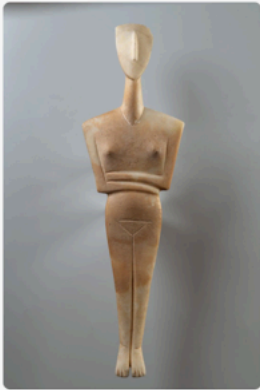
Museum curators have been involved to edit and add “**descriptive** section”.

A series of screenshots as examples of artwork chat (Fig. 41):

 **ReInHerit**
Redefining the Future of Cultural Heritage
VIOLA HOME **GALLERY**

Cycladic Figurine
2300-2000 BC

Maker: Unknown
Material: Marble
Measurement: H. 39.1 cm
Location: Museum of Cycladic Art
Artwork Description Validated by: Cycladic Museum - 4 June 2024



Artwork Chat

Ask me something about this artwork!
3:04 PM

Why is this artwork unclothed and naked?
3:04 PM

The figurine is unclothed to emphasize the human form and possibly to symbolize fertility and reproduction, as indicated by its female characteristics.
3:04 PM






Does the position of the arms have a specific meaning?
3:05 PM

The folded arms of female figurines denote a passive position.
3:05 PM

Tell me the story of the Cycladic female figurines
3:05 PM

The Cycladic female figurines are masterpieces of Cycladic Art, primarily made of marble. They often depict women with folded arms, emphasizing features like breasts, pubic triangle, and slightly swollen abdomen, indicating fertility and reproduction. These figurines, dating back to 2300-2000 BC, are characterized by their simplicity, harmonious proportions, and angular contours. They were likely created by skilled male marble sculptors and are thought to represent pregnant women or fertility deities. The figurines were mostly found in graves, suggesting a possible association with funerary rituals. The Cycladic Art, including these figurines, has significantly influenced modern artists.
3:05 PM

ReInHerit Toolkit - VIOLA ChatBot

Cherry Rage 1920
1955 CE

Maker: Axl Leskoschek
Material: paper, woodcut
Measurement: 44,5 x 31,5cm
Location: Graz Museum
Artwork Description Validated by: Graz Museum - 4 June 2024



Artwork Chat

Hi! Nice to meet you!
2:51 PM

Ask me something about this artwork!
2:51 PM

what is represented in this artwork?
2:52 PM

The artwork represents the Cherry Rage riot in Graz, 1920, showing protesters clashing with police.
2:52 PM

Why Cherry?
2:52 PM

The term 'Cherry' refers to the high prices of cherries that sparked the protest in Graz in 1920.
2:52 PM

Tell me the story depicted
2:53 PM

The artwork depicts the Cherry Rage riot in Graz, 1920, where women protested high food prices, leading to violence, looting, and police intervention, resulting in deaths and injuries.
2:53 PM

Cipro Map
1570 CE

Maker: Nicolò NELLI
Material: Loose-sheet map
Measurement: 19cm x 13,6cm
Location: Bank of Cyprus Cultural Foundation
Artwork Description Validated by: Bank of Cyprus Cultural Foundation - 4 June 2024



Artwork Chat

Ask me something about this artwork!
2:56 PM

What is drawn on this map?
2:57 PM

The map shows Cyprus surrounded by 19 warships, 18 with sails and 1 without sail.
2:57 PM

What historical event is represented by the author?
2:57 PM

The Fourth Ottoman-Venetian War (1570-1573) over Cyprus.
2:57 PM

tell me about the Fourth Ottoman-Venetian War
2:58 PM

The Fourth Ottoman-Venetian War (1570-1573) was a conflict over Cyprus, resulting in Ottoman occupation until 1878.
2:58 PM

Can you tell me anything about the author of the work?
2:59 PM

Nicolò Nelli was a Venetian engraver, mapmaker, and publisher active between 1562-1573, influenced by Marco Dente.
2:59 PM

Fig. 41 - VIOLA Multimedia Chatbot Final Version - Answering questions

3.4.3 Co-creative Process

The VIOLA chatbot web app has also been tested and studied during the [AI&XR Summer School held in Matera in July 2023](#) (Fig. 42). A multidisciplinary group of young Ph.D. students worked on the theme "Smart Interaction and Digital Contents," tested the Viola Chatbot and SmartLens apps, using open-source codes from Digital Hub and adapted them to artworks from some local museums. This co-creation process has revealed new potential technological developments for the apps and user interaction scenarios.



Fig. 42 - AI&XR Summer School held in Matera

During Summer School MICC - University of Florence presented a practical lecture on 'Innovative and sustainable approaches for user engagement and digital interaction with cultural heritage'. In order to adopt a sustainable and user-centered approach, the aim was to share the relevant results of the playful approach and user engagement studies conducted within the ReInHerit H2020 research project. Cutting-edge applications based on artificial intelligence developed by the MICC were presented to inspire the students' project proposals. Participants were invited to use, test, and explore the Toolkit in a collaborative and interdisciplinary approach, linking technological and cultural sectors.



Fig. 43 - Hackathon working Group, AI&XR Summer School Matera

During the week, international speakers and famous experts debated and engaged with international students and researchers with different educational content and skills. Young PhD students worked in an interdisciplinary way on the ReinHerit web apps Multimedia Chatbot and SmartLens using open-source codes shared by the Digital Hub. This co-creation process added new technological developments for apps and user interaction scenarios, improving engagement, inclusivity and new design features¹⁸. (Fig. 43)



Fig. 44 - Experts and Professionals, AI&XR Summer School Matera

The hackathon involved experts and professionals, (Fig. 44) including Dr. Luca Melchionna an Italian journalist and entrepreneur. He worked for 11 years in the MART - Museo di arte moderna e contemporanea di Trento e Rovereto (Italy) - press office, and later collaborated with Palazzo Grassi-Pinault Foundation, Fondazione Fitzcarraldo, TSM Trento and contributed to the debate on audience development and the management of technological innovations in the museum sector. He is one of the founders of "Machineria - Stories that Work", a company that develops content for the cultural sector - and which has just released a chatbot

¹⁸ ReinHerit Digital Hub <https://reinherit-hub.eu/summerschool/>

(NeroBot) for the Colosseum Archaeological Park (Rome, IT)¹⁹. He was involved as a supervisor for the group work during the hackathon to provide his advice on the functionalities and potential of VIOLA Chatbot version 1. He tested its functionality and usability, moving from specific questions about paintings to more general topics. He tried to use the prototype to produce texts addressed to different audiences, leaving insights for the hackathon also using the Digital Hub Forum. (Fig. 45)

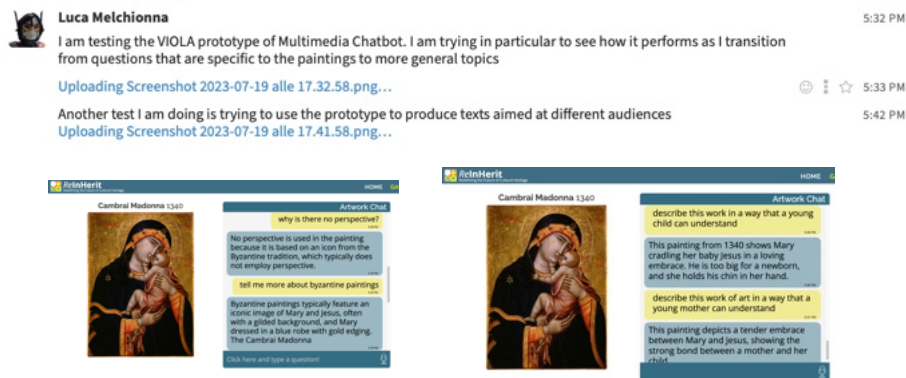


Fig. 45 - Interaction on Digital Hub Chat / Forum, AI&XR Summer School Matera

AIFOCUS²⁰ proposal aims to merge "VIOLA" chatbot and SmartLens to create an interactive art experience, offering detailed artwork insights via conversation. (Fig. 46)



Fig. 46 - AIFOCUS Hackathon Proposal

Users explore surroundings through cameras, engaging for artwork insights, including multimedia contents. The goal is to promote knowledge among young people and stimulate

¹⁹ <https://colosseo.it/en/>

²⁰ <https://reinherit-hub.eu/summerschool/314ed627-6f30-4d43-9428-4e55aee28066>

As in the SmartLens app, users can scan with the smartphone camera for details to be examined. When users launch the webapp, they can explore their surroundings through the smartphone camera and discover details of the artwork. They can initiate a conversation with the chatbot, which provides precise and detailed answers related to the analyzed details. (Fig. 49)

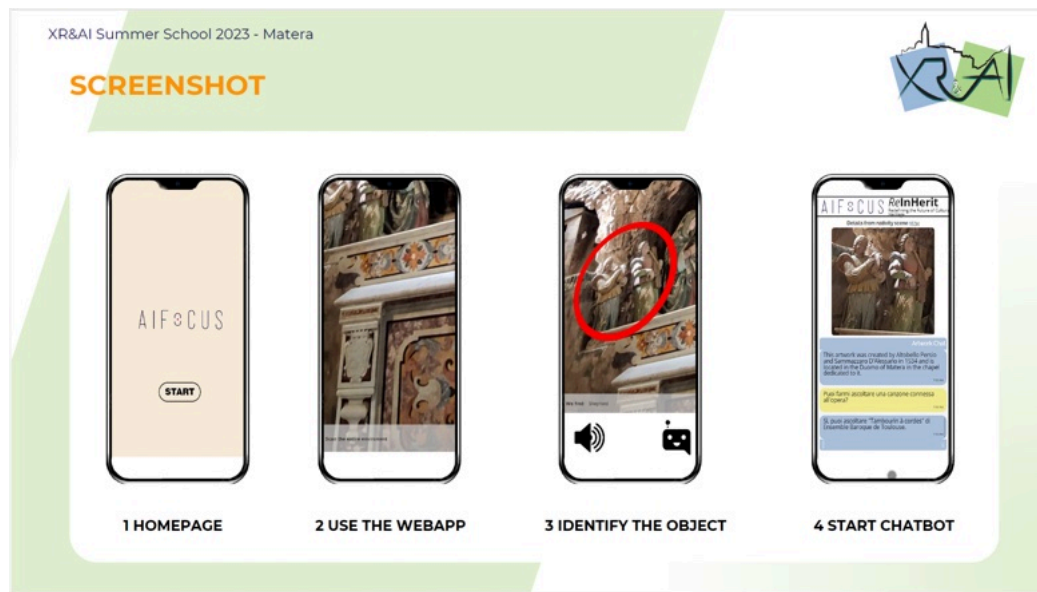


Fig. 49 - AIFOCUS Hackathon Proposal Screenshots

The project also highlighted the possibility of integrating additional multimedia content, such as audio and video related to the observed detail.

The webapp offers users an engaging and dynamic experience to learn about the details of a work of art, allowing them to directly identify what interests them. Its main objective is to promote knowledge among young people and stimulate their curiosity about the art around them, making learning an engaging and accessible experience for all.

Youtube short video: https://youtu.be/THX_2PNmJ9I?si=JN27NSn28G5Dpuf9

Following the discussion we had with some experts on AI and the use of chatbots in museums and after the co-creation process carried out during the hackathon, we will try to add more guidance on transparency on training data and the monitoring of scientific accuracy for VIOLA Multimedia Chatbot, SmartLens as well as Strike-Pose and Face-Fit. We will explore usability, accuracy and related Ethical issues of these apps with experts from the leading working group on AI and museums (“Museums+AI Network” Goldsmiths University in London and Pratt Institute in New York). As already outlined in Deliverable 3.2 of the Toolkit Strategy (p. 22-23), the experience and reference to this important Network is useful for testing and

updating the ReInHerit Toolkit following an interdisciplinary, sustainable and up-to-date approach on the impact of AI in the museum world.

Along these lines, MICC has been invited by Giuliano Gaia²¹ to participate in “Museums+AI Workshop”,²² in Milan (9 October 2023) on ethical AI (Fig. 50), presenting some case studies and tools on AI in museums and produced in the context of the ReInHerit Toolkit. The collaboration with the IULM AI Lab and the Museums+AI Network at Goldsmiths University in London and the Pratt Institute in New York demonstrates the importance of bringing together experts from various fields to discuss and advance the ethical implementation of AI in cultural institutions. The presence of Dr. Oonagh Murphy and the presentation of the Toolkit developed by the Museums+AI Network further highlight the importance of ethical considerations in AI applications. The testing of various applications through the working groups will provide valuable insights and feedback to refine AI practices in museums. Overall, this workshop represents a valuable opportunity for partnership, knowledge exchange and collective efforts to ensure that AI in museums meets ethical guidelines and enhances the visitor experience in a responsible manner.

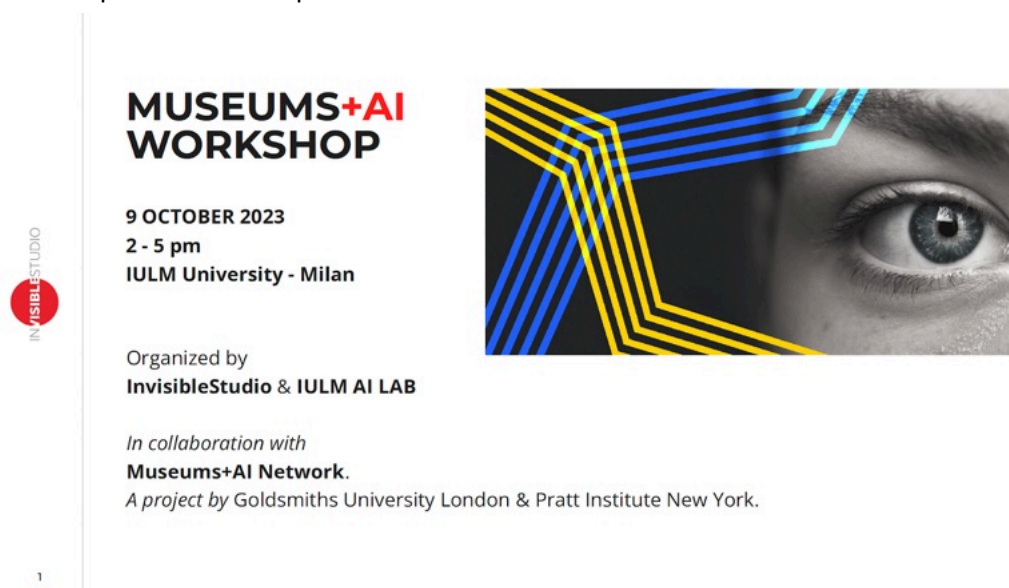


Fig. 50 - Museums+AI Workshop, Milan 9 October 2023

Topic of the Workshop : "The recent explosion in generative AI means that museums have to quickly adapt to the new landscape and understand how to make the most of it ethically." To help museums in doing so InvisibleStudio together with two leading

²¹ Giuliano Gaia, a pioneer in the field of digital communications for museums in Italy /UK, and co-founder together with Stefania Boiano of the cultural company Invisible Studio. He talked about how visitors can enrich their experience using digital technology in museums and how the importance of a human-centered design is crucial to create better technology experiences in museums, in [Museums Up - ReInherit's Podcast sixth episode: "Digital communication with audiences in museums"](#)

²² <https://www.invisiblestudio.net/portfolio/museums-ai-workshop>

institutions, the Goldsmiths University in London and IULM AI Lab in Milan, designed the workshop focused on navigating the ethical landscape of AI in museums.

The Hosts & Workshop Facilitators were:

- **Oonagh Murphy** Goldsmiths University AI+Museums Network.
- **Guido Di Fraia** AI Lab IULM University
- **Giuliano Gaia** InvisibleStudio LTD IULM University

Italian Case Histories:

- **Paolo Mazzanti & Marco Bertini** MICC University of Florence / H2020 ReinHerit Consortium "Increase User Engagement and motivate learning in Museums with Computer Vision and AI - The ReInHerit Toolkit"
- **Luca Melchionna** Machineria "AI and attempts not to reinvent the wheel. Lessons learned in the recent past"
- **Francesco Ronchi** Synesthesia "AI experiments at the Museo Nazionale del Cinema of Turin IT"



Fig. 51 - Museums+AI Workshop, Milan 9 October 2023

This hands-on event brings together museums, university researchers and private companies to explore the responsible integration of AI technologies in cultural institutions. Paolo Mazzanti and Marco Bertini MICC presented a case history at Museums+AI Workshop²³ (Fig. 51). The presentation of these applications through the working groups provided valuable insights and feedback to improve AI practices in museums, and was an opportunity to discuss and address the topics related to Implications of using AI, described in section 6.1 of D3.4. ReInHerit Apps are part of Italian case studies included in the Italian translation of "AI: Museum Planning Toolkit"²⁴ edited by Dr Giuliano Gaia InvisibleStudio, in collaboration with Dr. Oonagh Murphy co-leader of the project AI+Museums Network. (Fig. 52).



Fig. 52 - Museums+AI Workshop, ITA Version

On September 6th 2023, as part of the Summer School DigitArch. Digital Collection Curator for Archaeology²⁵ - Marco Bertini and Paolo Mazzanti (MICC Unifi) gave a lecture titled "Emotions, interactions and Digitals _ the museum experience" presenting innovative approaches for digital learning experience and emotional engagement, based on AI/CV tools (Fig. 53). The Summer School offers a unique high-level international training aimed at putting into practice the digital and technological skills of students and professionals working in the cultural heritage sector, with a special focus on the archaeological field, in line with the professional profile of the Digital Collection Curator. Focused on the use of ICT for communication and enhancement of cultural sites. During the visit to the Civic Museums of Reggio Emilia IT with Dr. Georgia Cantoni (Communication Manager, Web and New Media Area and Curator of the Ethnographic Collections of the Civic Museums of Reggio Emilia), the

²³ <https://reinherit-hub.eu/news/5c0b413f-51f1-40c6-9524-9dc0dc87be3b>

²⁴ https://themuseumai.network/wp-content/uploads/2024/04/20190317_museums-and-ai-toolkit_web_ita_v2-1.pdf

²⁵ <https://reinherit-hub.eu/news/db96b296-45f2-4f8c-8bc7-5d3f5dca61d1>

participants were invited to identify works of art to be used with the ReInHerit H2020 toolkit. And to be used as a working proposal for the creation of textual materials during the two-day Editathon with experts from Wikimedia Italy and museum curators.



Fig. 53 Summer School DigitArch 2023

MICC presented and demonstrated the interactive applications of the ReInHerit toolkit during the Macerata Humanities Festival,²⁶ organized by the University of Macerata, in two different sessions on Friday 29th and Saturday 30th 2023 from 15:00 to 18:00. During these events, the Smart Lens, VIOLA chatbot, Strike-a-pose and Face-fit were presented and made available in an interactive installation (Fig. 54). The event was free and open to the citizenry, and aims to show how art, human and social sciences are the key to a better understanding of contemporary society and to the valorisation of the cultural heritage of the territory. Special and welcome guests were the families with children who enjoyed and tested the Toolkit apps with fun and interest. (Fig. 55)



Fig. 54- Poster ReInHerit toolkit - Macerata Humanities Festival 2023

²⁶ <https://reinherit-hub.eu/news/8b82d931-2b8f-4ce5-a7aa-a33b38d83031>



Fig. 55- Demo and Testing ReInHerit toolkit - Macerata Humanities Festival 2023

Additionally, we are working with curators of the following small and medium-sized museums to test and add more artworks: [Fondazione Musei Senesi](#) IT, [Giara Gipsoteca di Arte Antica e Antiquarium University of Pisa](#) IT, [Musei Civici di Reggio Emilia](#) IT, [Museo Capitolare Diocesano Foligno](#) IT. MICC organised lectures and workshops in small and medium-sized national museums to test and work on improving the Chatbot and other Toolkit applications, following a co-creative and bottom-up approach.²⁷ (Fig. 56)

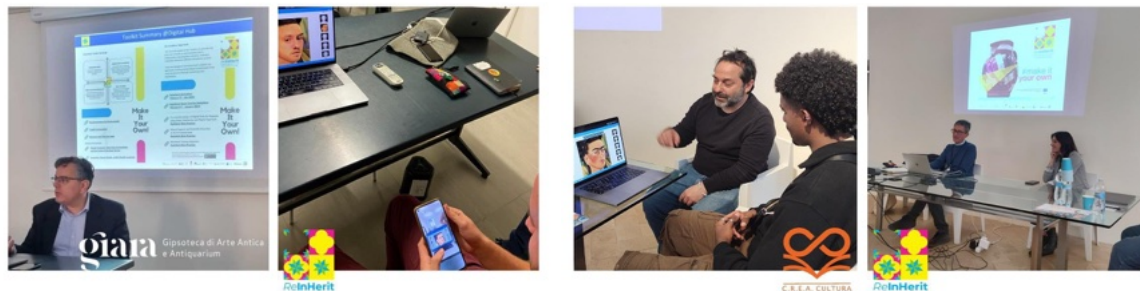


Fig. 56 Lectures and Workshops - testing the Apps

MICC tested the Multimedia VIOLA Chatbot and Face Fit apps with museum experts and a large group of young students who interacted with a selection of artworks. The images and related content have been selected working in collaboration with the curators of the Museo Capitolare Diocesano (Foligno IT) and GIARA Gipsoteca (Pisa IT) in order to be added to online App Galleries of the ReInHerit Digital Hub.

MICC researchers participated in the 15th ACM Multimedia Systems Conference held from 15-18 April, 2024 in Bari, Italy with a DEMO, titled "Context-aware Chatbot using MLLMs for Cultural Heritage"(Fig. 57). MMSys 2024 provides a forum to leading researchers from academia and industry to present and share their latest findings in multimedia systems. Based on work conducted as part of the ReInHerit Toolkit's VIOLA Multimedia Chatbot, the demo presented to Multimedia experts the innovative approach that integrates contextual descriptions of artworks to improve the answering of visual questions, addressing a crucial

²⁷ [Workshop - CREA Cultura Festival 2024, Foligno IT](#) (March 22, 2024) and Lecture on "Reinherit Toolkit AI-based Apps for Digital Engagement" at GiARA Gipsoteca Arte Antica e Antiquarium, UNIPI (May 9, 2024)

need in cultural heritage interpretation. It highlights the need for chatbots that function with scientific accuracy, free of hallucinations, to provide reliable information in the cultural domain. A great opportunity to present the results achieved to the scientific community, introducing the system architecture, the chatbot dataset, design considerations, discussing possible developments and challenges, following the chatbot test carried out with the community of prominent researchers from around the world attending the Conference.²⁸



Fig. 57 - VIOLA Chatbot Demo at MMSys 2024

Considering NEMO's shared focus and recommendations on Artificial Intelligence that museums are at the forefront of an AI-driven society²⁹, we engaged the NEMO Network by participating in the third NEMO MeetUp on June 12, 2024. During this event we presented the ReInHerit Toolkit, a great opportunity to share project results with associate members of this important European museum network.³⁰ The Toolkit was also presented to interdisciplinary experts attending the Culture Research Observatory Action Workshop, July 11-12, 2024 at Monash University Prato and Centro Pecci Prato (Italy) organized by the Institute for Digital Culture (University of Leicester) in collaboration with the Monash University European Research Foundation.³¹ A Toolkit introduction was included in "Book of Abstracts from the international conference *"Quest: How to make digital work for, not against learning"* held online June 6, 2024, Bratislava 2024.³² In the context of the International Council of Museums, the Toolkit was presented to ICOM National Working Group "Multimedia and Emerging Technologies" during the May 13, 2024 Webinar entitled *"Emotional Experiences at the Museum. Engaging, interacting and connecting even with AI-based tools."*³³ Our Paper titled *"ReInHerit AI Toolkit: Enhancing Museum User Interaction and addressing Ethical Aspects"* has been accepted at the ICOM AVICOM (Audiovisual, New Technologies and Social Media) Annual Conference 2024 - *"What's New from the World of Digital Museums"* - Turin, Italy on October 17-18.³⁴

²⁸ <https://reinherit-hub.eu/news/3dc34ed3-2680-493c-9bb2-be00965f5732>

²⁹ <https://www.ne-mo.org/news-events/article/nemo-presents-3-recommendations-addressing-the-development-of-ai-technology-in-museums>

³⁰ <https://reinherit-hub.eu/news/c29fbf53-b2a1-4c95-af66-33045ee945d9>

³¹ <https://reinherit-hub.eu/news/e636543a-cc97-4987-a433-2f10c6f6dac5>

³² <https://doi.org/10.31577/2024.9788097443467> (ENG Version)

³³ <https://reinherit-hub.eu/news/c375e2e1-b76a-42a8-b188-7ccf4bf69b87>

³⁴ <https://avicom.mini.icom.museum/welcome-to-turin/>

Scientific impact:

VIOLA Multimedia chatbot was demoed live during the ACM Multimedia Systems 2024 conference, one of the most important conferences on multimedia.

VIOLA has also been presented, along with Strike-a-pose and Face-fit, at the International Conference Extended Intelligence for Cultural Engagement (ExICE), Bologna, Italy, March 2023.

Associated papers describing the app:

- P. K. Rachabatuni, F. Principi, P. Mazzanti., M. Bertini, “Context-aware chatbot using MLLMs for Cultural Heritage”, In Proc. of the 15th ACM Multimedia Systems Conference (MMSys), 2024.
- F. Becattini, P. Bongini, L. Bulla, A. Del Bimbo, Ludovica Marinucci, Misael Mongiovì, and Valentina Presutti. “VISCOUNTH: A Large-scale Multilingual Visual Question Answering Dataset for Cultural Heritage”, in ACM Transactions on Multimedia Computing, Communications, and Applications, 2023
The paper has received 6 citations, as of Aug. 2024.
- P. Bongini, F. Becattini, A. Del Bimbo, “Is GPT-3 All You Need for Visual Question Answering in Cultural Heritage?”, In Proc. of European Conference on Computer Vision (ECCV) Workshops, 2022
The paper has received 20 citations, as of Aug. 2024.
- P. Mazzanti, M. Bertini, A. Baldrati, L. Agnolucci, A. Ferracani, F. Principi, P. Bongini, “The ReInHerit Toolkit: Artificial Intelligence and a people-centered approach for cultural engagement”, Proc. of Extended Intelligence for Cultural Engagement (ExICE), 2023

3.5 Smart tourism

The smart tourism app is a native Android app that provides the basic functionalities for the smart tourism app of WP5. It has been developed as an Android app to overcome the computational limitations of web-based systems, that are still not feature complete when using TensorFlowJS³⁵ with respect to TensorFlow Lite³⁶. Despite this it works also on low and mid-level devices, since it has been designed to use a set of very fast neural network that can be executed in real-time (based on variations of the MobileNet V3 architecture [Howard-2019]).

The system uses computer vision to recognize landmarks and monuments, providing multimedia information, and can be personalized to create different tours. The end-user

³⁵ <https://www.tensorflow.org/js>

³⁶ <https://www.tensorflow.org/lite>

application is complemented by a set of backend tools that create the set of images needed to recognize landmarks, performing an augmentation that is used both at training time and at test time; these tools are to be used by the creators of the smart guides. The vision system implements a content-based image retrieval (CBIR), using techniques that are completely different, because of the nature of the task that is addressing landmarks, from that of the Smart Lens app.

In fact, when considering the recognition of landmarks, a common case is that the user does not frame with the lens the whole object, and that the object itself has many different and possibly diverse views. To this end when creating the guide the curators must use a large number of images. To cope with the variability of the point of view of the users each image is further split in different parts and zoomed in and out versions are created, as shown in the following figures. All these augmentations result in the creation of a very large dataset of images representing the landmarks.



Fig. 58– image splitting augmentation



Fig. 59– image zooming augmentation

To cope with the large number of images, these are indexed using FAISS³⁷, an indexing library that allows to perform approximate nearest neighbor retrieval, thus reducing the number of actual image comparisons that are needed to determine the landmark that is framed.

³⁷ <https://github.com/facebookresearch/faiss>

This need is exacerbated by the fact that we implement also a test-time augmentation, i.e. the image that is taken with the mobile phone is used to generate two zoomed versions, and each image is further split into 3x3. This test-time augmentation is used to improve the performance of the CBIR system in terms of accuracy. For each split of each image obtained during the use of the application is computed a descriptor using the MobileNet network, and using FAISS a nearest neighbor image of the database is retrieved and ranked in terms of visual similarity, selecting the 3 most similar images, as shown in the following figure. A K-NN classifier is used to recognize the landmark based on these retrieved images.



Fig. 60 – smart tourism app: landmark recognition

To test the performance of the computer vision system we have experimented its performance in terms of precision on the challenging Paris Revisited dataset³⁸, developed by the University of Oxford. This dataset is composed of 5000 images of landmarks; we used $\frac{2}{3}$ for training and $\frac{1}{3}$ for test.



Fig. 61 – revisited Paris dataset: for each query image is provide a set of good (dark green), medium (light green) and hard (yellow) images that are associated. A CBIR system should provide the results in these categories in this order.

Considering the computational costs of the variants of MobileNet V3 tested, the smaller version allows to compute the visual features in 0.006 seconds, making it suitable for low-end

³⁸ <https://github.com/filipradenovic/revisitop>

devices that do not have much computational capabilities. This at the cost of losing 5 percent points in terms of accuracy. Conversely, mid and high-end devices can use a medium and a large network version, that require 0.023 and 0.038 seconds. Thanks to the use of FAISS, the search of similar images takes a negligible amount of time (0.004 seconds). Once the set of similar images has been obtained the result of the K-NN classifier is instantaneous. The accuracy of the large network is ~76%, while the medium version obtains ~73% and the smaller one 68%. Using the test-time augmentation improves the performance by 6 points, thus reaching up to 82%, showing the benefit of this technique.

The following figure (Fig. 62) shows screenshots of the pilot first version of the application, with additional debug information in the first two images, and the recognition of a landmark and then associated information. More details on the **latest version** of the Smart Tourism App and the collaborative³⁹ process of co-creating and managing museum content are described in the section “Application description” at pag. 12 of D7.9 “Free Smart Tourism Apps Report”. The **co-creative process** took place during the ReinHerit Smart Tourism Hackathon in Nicosia, January 2024.⁴⁰

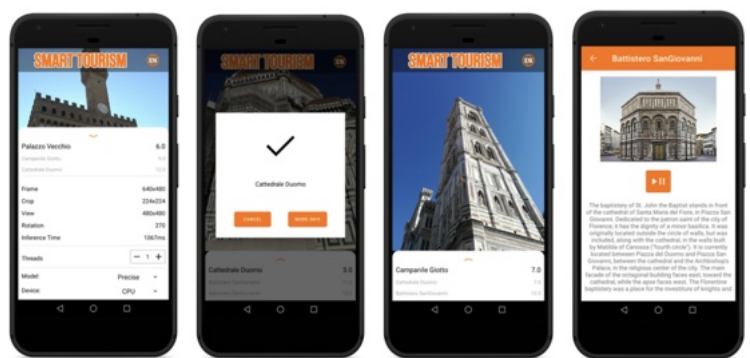


Fig. 62 – First Pilot Version of Smart Tourism app | views of the application: debug information to test the capability to differentiate between visually similar landmarks; debug info on camera setup and neural network inference (using CPU instead of GPU); examples of landmark recognition and information on the recognized landmark.

The final application (Fig. 63) is designed for cultural smart tourism and provides functionalities for landmark recognition using computer vision and personalized tour recommendations based on user preferences in terms of types of tours (designed by the guide creators, e.g. historical buildings or museums) and properties of the locations (e.g. guided tours, free entrances, kid-friendly...). The ReinHerit app lets tourism organizations create localized apps and types of visits, adapting them to the preferences of the visitors. It’s a tool to implement Smart Tourism, i.e. facilitating access to tourism and hospitality products, services, experiences. (Cultural) Tourism landmarks are organized in categories, e.g. types of landmarks or because they belong to a type of visit. The choice is up to the tourism organization. Each landmark is associated with properties that may be of interest for the

³⁹ <https://reinherit-hub.eu/news/df5ec792-13ae-4e91-86fe-f756ee8a5a3c>

⁴⁰ <https://reinherit-hub.eu/hackathonnicosia/>

visitors or that characterize a visit. “Smart Tourism” Youtube short video <https://www.youtube.com/watch?v=Mlfbmtp1LKK>

App enables innovative **Active Exploration** using responsive, user-centred CV/AI solutions:

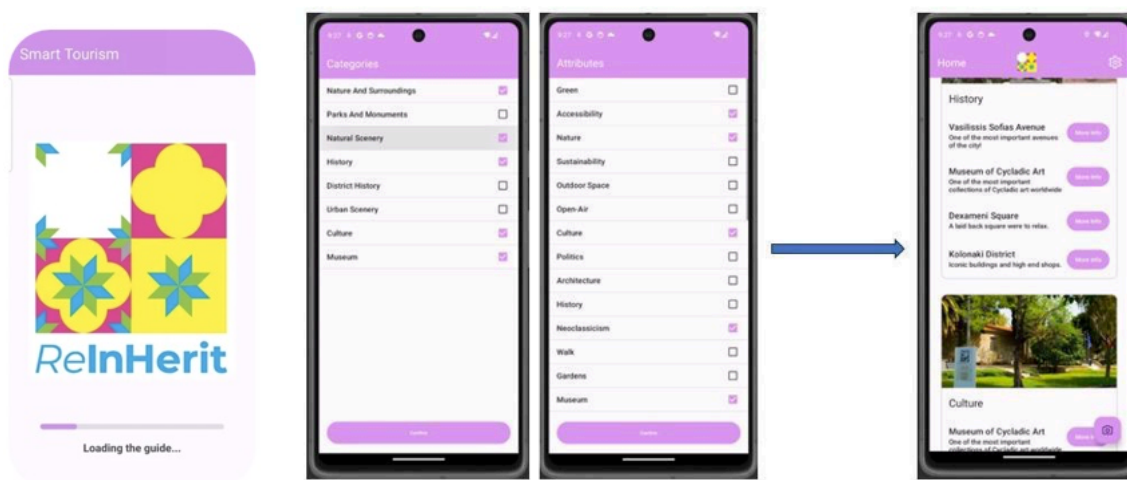
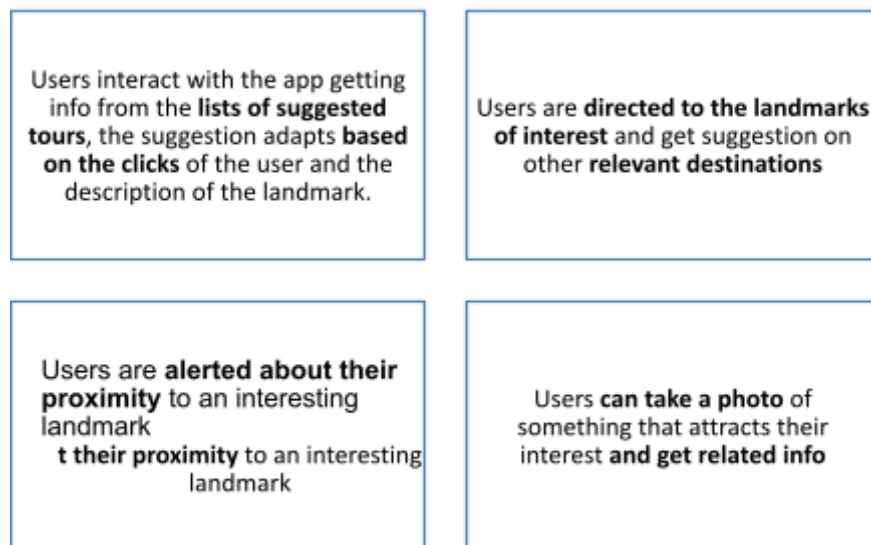


Fig. 63 - Final version of Smart Tourism app | Active Exploration - Users interact with the app getting info from the lists of suggested tours, the suggestion adapts based on the clicks of the user and the description of the landmark. Users are directed to the landmarks of interest and get suggestions on other relevant destinations.

User data and privacy have been taken into the utmost consideration in the implementation of the Smart Tourism app. The recommendations for landmarks and visits are computed directly on the device. This process takes into account several factors, including user interactions with landmark descriptions, the similarity of different landmark descriptions, explicit user preferences, GPS data, and camera access for recognizing landmarks. Importantly, all this information is stored locally on the device and is not logged externally,

ensuring that user data remains private and secure. This approach ensures that users benefit from personalized recommendations while maintaining full control over their personal information.⁴¹

4. Webinars

In addition to the application development the activity has regarded the implementation of a series of **webinars**, whose production will continue in the following months.

Each Consortium partner organized webinars based on a set of themes, identified by the WP2 analysis, and indicated in D7.5 “Dissemination and Exploitation Plan 2nd Version”.

Macro Topics for the Digital Hub Webinars’ section are **TECHNOLOGY, TOURISM, MANAGEMENT**, covering themes such as:

- Digital and emerging technology
- IPR Intellectual Property Rights
- Traveling Digital Exhibition planning
- Immersive Performance
- Gamification
- Digital Hub
- Cultural Tourism
- Conservation and preservation
- Inclusive Museum
- Sustainability and Heritage
- Co-creation with museums
- etc.

Webinars started in November 2022 and will end in October 2023.

As of February 2023, 8 webinars have been delivered with an average participation of 60 registrants per webinar from different European countries.

All the video recording of the **webinars** and the materials presented are made available on the Digital Hub. <https://reinherit-hub.eu/webinars> (Fig 64). Documentation of all webinars are included in “D3.9 Training Curriculum and Syllabi”.

⁴¹ More info on “Ethics Annex “ section 6 - D3.2 “ReInHerit Toolkit Strategy”

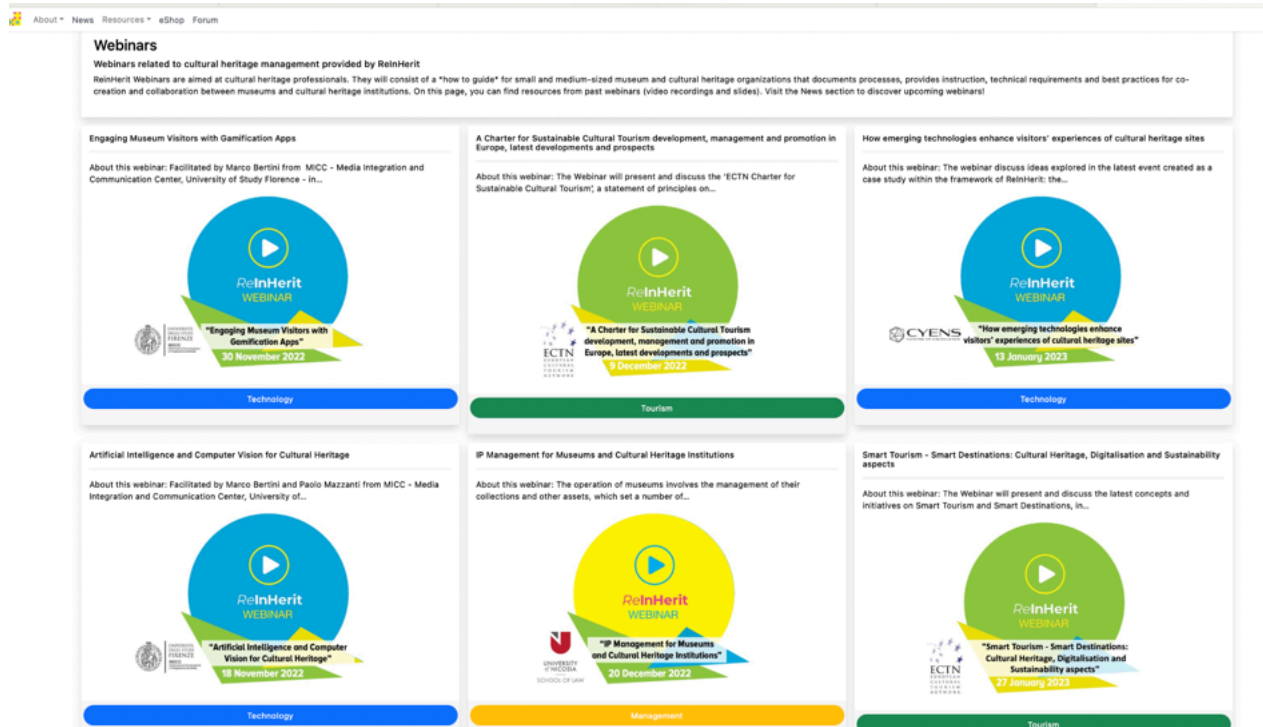


Fig.64- Webinars Resources on Digital Hub (Feb. 2023)

Below is the temporary list of webinars with dates and topics:

- Webinar n. 1 – Technology / Digital and emerging technology
18 november 2022 “Artificial Intelligence and Computer Vision for Cultural Heritage”
MICC-UNIFI, Marco Bertini / Paolo Mazzanti
- Webinar n. 2 - Technology / Digital and emerging technology
30 November 2022 “Engaging Museum Visitors with Gamification Apps”
MICC-UNIFI, Marco Bertini
- Webinar n. 3 / Tourism / Cultural Tourism
9 December 2022 “A Charter for Sustainable Cultural Tourism development,
management and promotion in Europe, latest developments and prospects” ECTN
Manos Vougioukas / Jordi Tresserras
- Webinar n. 4 / Management / IP Management for Museums
20 December 2022 – “IP Management for Museums and Cultural Heritage
Institutions” UNIC, Irini Stamatoudi
- Webinar n. 5 / Technology / Emerging Technologies in cultural heritage sites
13 January 2023 – “How emerging technologies blend with cultural heritage sites”
CYENS, Kleanthis Neokleous/Andreas Papapetrou
- Webinar n. 6 / Tourism / Smart Tourism Destinations
Smart Tourism - “Smart Destinations: Cultural Heritage, Digitalisation and

Sustainability aspects” ECTN, Manos Vougioukas / Nasos Hadjigeorgiou (54 registrations)

- Webinar n. 7 / Management / Sustainability and Cultural Heritage
10 February 2023 “How to Manage Museums in a Sustainable way?”
ARTE SOSTENIBILE, Jaime López
- Webinar n. 8 / Management / IP Management for Museums
24 February 2023 "A Sustainable CH Management Plan for Museums" UNIC
Konstantinos Roussos.

Additional 13 webinars are planned addressing the following topics:

- 3D Scans for Collection Objects
- Collaboration platforms in Cultural Heritage
- Digital and emerging technology skills
- Co-creation in museums
- Inclusive Museums
- Storytelling and Soft Skills
- Preservation and AI
- Green museums
- Gamification

Appendix

List of **meetings and workshops** associated with the production of this deliverable.

1. Workshop “Artificial Intelligence and Computer Vision for Cultural Heritage” Graz - AU, in collaboration with MICC-Unifi (WP3), UniGraz (WP4) and GrazMuseum (WP6) - (23-24 May 2022)
2. Smart Lens app meetings - WP6 partners (23 September 2022, 16 January 2023, 19 January 2023)
3. Hackathon planning meeting (26 January 2023) - with BOCCF and MateraHub partners
4. Innovation workshop - all ReInHerit partners (14 February 2023)
5. Smart retrieval app meeting for application within European Fashion (15 February 2023) - with European Fashion Heritage Association Managing Director
6. Regular meetings of the tech committee are taking place every month. In collaboration with MICC-Unifi, UniGraz, Bank of Cyprus Cultural Foundation, CYENS
7. ReInHerit Study Visit and Consortium Meeting in Brussels, 8-12 May 2023 Brussels
8. ReInHerit mobility activities, visit and meeting with BoCCF in Florence at MICC Media Integration and Communication Center of the University of Florence, 13 July 2023
9. Poster/Demo on "ReInHerit Toolkit: Artificial Intelligence and a People-centered Approach for Cultural Engagement" at "ExICE - Extended Intelligence for Cultural Engagement" 15 March 2023 Bologna Italy
10. "ExICE – Extended Intelligence for Cultural Engagement" Conference, Bologna, Italy. This conference was organized as part of the SPICE Project under the Horizon 2020 research and innovation program (Grant Agreement N. 870811).
11. ReInHerit Hackathon | XR&AI Summer School 2023 Matera IT 17-22 July 2023
12. Summer School DigitArch. Digital Collection Curator for Archaeology, Reggio Emilia IT, 6 September 2023.
13. Magma Fest, 24-25 November 2023. Matera, Italy.
14. ReInHerit Toolkit - Macerata Humanities Festival 29-30 September 2023 Macerata IT
15. ReInHerit Toolkit - Museums+AI Workshop IULM AI LAB Milan IT, 9th October 2023 Milan IT
16. ReInHerit Smart Tourism Hackathon, 20-21 January 2024, Nicosia, Cypro
17. ReInHerit Toolkit Workshop - CREA Cultura Festival 2024, 22 March 2024 Foligno IT
18. Lecture on "ReInHerit Toolkit AI-based Apps for Digital Engagement" at GiARA Gipsoteca UNIPI 9 May 2024, Pisa IT
19. DEMO titled titled "Context-aware Chatbot using MLLMs for Cultural Heritage”, at ACM Multimedia Systems Conference 2024 15-18 April 2024 in Bari, Italy
20. Introducing ReInHerit to NEMO Network Members - Third NEMO Meet-Up, Online 12 June 2024
21. Culture Research Observatory Action Workshop, at Monash University Prato and Centro Pecci Prato (Italy) organized by the Institute for Digital Culture (University of Leicester) in collaboration with the Monash University European Research Foundation, July 11-12, 2024, Prato, IT.

References

- DoA, Part A / B
- D2.1 “Focus Group Report”
- D2.2 “State of the Art Report - Literature Review”
- D2.3 Questionnaires Report
- D2.4 Focus Groups Report Phase II
- D2.5 “CH Management Guidelines”
- D2.6 “A sustainable model of CH management state of the art report”
- D3.1 “National Surveys Report”
- D3.2 “Toolkit Strategy”
- D3.3 “ReInHerit Toolkit phase I”
- D3.4 “Consolidated Report on ICT in CH Management”
- D3.7 “Demonstrator Mobile Applications - final version report”
- D4.1 “Requirements Analysis Report”
- D4.2 “Digital hub”
- [Agnolucci-2022] L. Agnolucci, L. Galteri, M. Bertini and A. Del Bimbo. “Restoration of Analog Videos Using Swin-UNet” In Proceedings of the 30th ACM International Conference on Multimedia (ACM MM), 2022
- [Agnolucci 2023] L. Agnolucci, L. Galteri, M. Bertini, A. Del Bimbo, “Reference-based Restoration of Digitized Analog Videotapes”, Proc. of WACV 2024
- [Baldrati-2022a] A. Baldrati, M. Bertini, T. Uricchio, and A. Del Bimbo. Conditioned and composed image retrieval combining and partially fine-tuning clip-based features. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pages 4959–4968, June 2022.
- [Baldrati-2022b] A. Baldrati, M. Bertini, T. Uricchio, and A. Del Bimbo. “Effective conditioned and composed image retrieval combining clip-based features.” In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 21466–21474, June 2022.
- [Baldrati-2022c] A. Baldrati, M. Bertini, T. Uricchio and A. Del Bimbo. “Exploiting CLIP-based Multi-modal Approach for Artwork Classification and Retrieval”. In Proceedings of Heri-Tech - The Future of Heritage and Science and Technologies Conference, Springer, 2022
- [Bladrati-2023] A. Baldrati, L. Agnolucci, M. Bertini, A. Del Bimbo, “Zero-Shot Composed Image Retrieval with Textual Inversion”, Proc. of ICCV, 2023
- [Baldrati-2023b] A. Baldrati, L. Agnolucci, M. Bertini, A. Del Bimbo, “Zero-Shot Image Retrieval with Human Feedback”, Proc. of ACM Multimedia, 2023
- [Bongini-2020] P. Bongini, F. Becattini, A.D. Bagdanov, A. Del Bimbo, “Visual question answering for cultural heritage”, IOP Conference Series: Materials Science and Engineering, 2020
- [Bongini-2022] P. Bongini, F. Becattini, A. Del Bimbo, ““Is GPT-3 all you need for Visual Question Answering in Cultural Heritage?””, Proc. of European Conference on Computer Vision workshops, 2022
- [DelChiaro2019] R. Del Chiaro, A. D. Bagdanov, and A. Del Bimbo. “Noisyart: A dataset for webly-supervised artwork recognition.” In VISIGRAPP (4: VISAPP), pages 467–475, 2019.

- [Deoldify] 2018. DeOldify. <https://github.com/jantic/DeOldify>
- [Donadio-2022] M.G. Donadio, F. Principi, A. Ferracani, M. Bertini and A. Del Bimbo. "Engaging Museum Visitors with Gamification of Body and Facial Expressions" In Proceedings of the 30th ACM International Conference on Multimedia (ACM MM), 2022
- [Gal-2023] R. Gal, Y. Alaluf, Y. Atzmon, O. Patashnik, A. H Bermano, G. Chechik, and D. Cohen-Or, "An image is worth one word: Personalizing text-to-image generation using textual inversion", Proc. of International Conference on Learning Representation, 2023
- [Howard-2019] Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., ... & Adam, H. "Searching for MobileNetV3" In Proceedings of the IEEE/CVF international conference on computer vision, 2019
- [Iizuka-2019] Satoshi Iizuka and Edgar Simo-Serra. 2019. "DeepRemaster: Temporal Source-Reference Attention Networks for Comprehensive Video Enhancement." ACM Transactions on Graphics, vol. 38, n. 6, 2019
- [NEMO 2023] [Digital Learning and Education in Museums - Innovative Approaches and Insights](#)" Report 2023, NEMO - The Network of European Organizations
- [Nielsen-1993] J. Nielsen. 1993. Iterative user-interface design. Computer 26, 11 (1993), 32–41.
- [Radford-2021] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever. "Learning transferable visual models from natural language supervision," 2021
- [Ren-2015] Ren, S., He, K., Girshick, R., Sun, J. "Faster R-CNN: Towards real-time object detection with region proposal networks" In Proc. NIPS, 2015
- [Selvaraju2019] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Ba- tra. Grad-cam: Visual explanations from deep networks via gradient-based localization. International Journal of Computer Vision, 128(2): 336359, Oct 2019
- [UNESCO 2023] UNESCO 2023 "Guidance for generative AI in education and research" <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- [Villaespesa Murphy 2020] Elena Villaespesa, Oonagh Murphy, "AI: A Museum Planning Toolkit" - The Museums + AI Network - Goldsmiths, University of London New Cross London SE14 6NW <https://themuseumsai.network/toolkit/>
- [Wan-2022] ZiyuWan, Bo Zhang, Dongdong Chen, and Jing Liao. 2022. Bringing Old Films Back to Life. CVPR (2022).
- [Wang-2004] ZhouWang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. 2004. "Image quality assessment: from error visibility to structural similarity." IEEE Transactions on Image Processing (TIP) 13, 4 (2004), 600612.
- [Zhang-2018] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. "The unreasonable effectiveness of deep features as a perceptual metric." In Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018