




Article

# An Interdisciplinary Design of an Interactive Cultural Heritage Visit for In-Situ, Mixed Reality and Affective Experiences

Xabier Olaz , Ricardo Garcia, Amalia Ortiz \*, Sebastián Marichal , Jesús Villadangos, Oscar Ardaiz and Asier Marzo 

UpnaLab, ISC—Institute of Smart Cities, Public University of Navarre, 31006 Pamplona, Spain; xabier.olaz@unavarra.es (X.O.); ricardo.garcia@unavarra.es (R.G.); sebastian.marichal@unavarra.es (S.M.); jesus.villadangos@unavarra.es (J.V.); oscar.ardaiz@unavarra.es (O.A.); asier.marzo@unavarra.es (A.M.)

\* Correspondence: amalia.ortiz@unavarra.es

**Abstract:** Interactive technologies, such as mixed-reality and natural interactions with avatars, can enhance cultural heritage and the experience of visiting a museum. In this paper, we present the design rationale of an interactive experience for a cultural heritage place in the church of Roncesvalles at the beginning of Camino de Santiago. We followed a participatory design with a multidisciplinary team which resulted in the design of a spatial augmented reality system that employs 3D projection mapping and a conversational agent acting as the storyteller. Multiple features were identified as desirable for an interactive experience: interdisciplinary design team; in-situ; mixed reality; interactive digital storytelling; avatar; tangible objects; gestures; emotions and groups. The findings from a workshop are presented for guiding other interactive cultural heritage experiences.

**Keywords:** interaction; interdisciplinary design; interactive digital storytelling; emotional conversational avatars



**Citation:** Olaz, X.; Garcia, R.; Ortiz, A.; Marichal, S.; Villadangos, J.; Ardaiz, O.; Marzo, A. An Interdisciplinary Design of an Interactive Cultural Heritage Visit for In-Situ, Mixed Reality and Affective Experiences. *Multimodal Technol. Interact.* **2022**, *6*, 59. <https://doi.org/10.3390/mti6070059>

Academic Editor: Ester Alba Pagán

Received: 31 May 2022

Accepted: 8 July 2022

Published: 18 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Our aim is to design an interactive tourism experience for The Collegiate Church of Roncesvalles, which is an emblematic spot of the Spanish “Camino de Santiago”. Many tourists and experts in history and cultural heritage (CH) visit the place because of its rich past, full of exciting stories.

This audience heterogeneity makes it difficult to efficiently deliver the content [1]. Interactive technologies can help to contextualize real-life scenarios and the content and therefore to create deep connections with the site [2]. Thus, we decided to develop an interactive system able to engage with the users, adapt content and personalize the experience. At the same time, the project has to respect the aesthetics of the space. The objective of this project is to gather design guidelines for an in-situ interactive experience for a cultural heritage place that emotionally engages with the visitors.

To inform the system design we followed a mixed strategy balancing the expert judgment from an interdisciplinary team and the outcome of a participatory design process. The whole process consisted of two brainstorming sessions and one workshop where user experience (UX) techniques, such as the definition of Persona, Scenarios, User Journeys and role-playing, were applied.

As a result of this process, we designed a prototype where the primary interaction is with an avatar that incorporates storytelling techniques and emotional interaction and narrates local stories associated with historical objects placed around the room. Storytelling is driven through the interaction with real objects and with an emotional conversational avatar. Rather than using VR/AR glasses, the avatar (who represents a miner) is projected on a real wall of the church; the wall simulates a mine using 3D mapping projection. In this way, we take advantage of the existing building and provide a device-less experience, enabling group interaction and collaboration.

In this paper, we present the design process by analyzing the output of each brainstorming and workshop session. We discuss the design rationale of the system and contribute with design guidelines for mixed-reality systems in the CH context.

## 2. Related Work

Incorporating technology in museums and cultural heritage (CH) can improve engagement [3–5], user experience [4,6], presence [7] and cultural awareness [3,4].

We have extracted from previous research the elements that are relevant for designing in-situ immersive experiences where the visitors can move around the environment. These elements are shown in Table 1 and are as follows:

- Interdisciplinary Design Team (IDT): multiple types of expertise are needed to design an emotional and immersive experience for the visitor. We analyze the kind of contribution made by each expert;
- InSitu: we looked for previous research with experiences designed to take place in the real CH spot;
- Mixed Reality (MR): using MR technology to augment the CH real place with multimedia content;
- Interactive digital storytelling (IDS): to generate the narrative guided by the visitors interactions;
- Avatar: a humanoid avatar as a tourist guide;
- Objects: using interactive objects, virtual or real, that are around the environment;
- Gestures: follow the user's position and track their corporal gestures to serve as another input for the interactive system;
- Emotions: taking emotion into account during the design process or introducing affective computing technology;
- Groups: managing the interaction of multiple visitors that come as a group.

**Table 1.** Features of Interactive Cultural Heritage found in existing research.

	IDT	In Situ	MR	IDS	Avatar	Objects	Gesture	Emotion	Groups
Ciolfi & McLoughlin (2012) [2]	✓	✗	✗	✗	✓	✓	✗	✗	✗
Flynn (2013) [8]	✗	✗	✓	✗	✗	✗	✓	✗	✓
Pujol et al., (2013) [9]	✓	✗	✓	✓	✗	✗	✗	✗	✗
Niculescu et al., (2014) [10]	✗	✗	✗	✗	✓	✗	✗	✗	✗
Pietroni & Adami (2014) [11]	✗	✗	✓	✓	✗	✗	✓	✓	✗
Seppala et al., (2016) [12]	✗	✓	✓	✓	✓	✗	✓	✗	✗
Perry et al., (2017) [13]	✓	✗	✗	✗	✓	✓	✗	✓	✓
Vosinakis (2017) [14]	✗	✗	✗	✗	✓	✗	✗	✗	✗
Nisi et al., (2018) [3]	✓	✓	✓	✓	✓	✗	✗	✗	✗
Carrozzino et al., (2018) [4]	✗	✗	✗	✗	✓	✗	✗	✗	✗
Schaper et al., (2018) [5]	✓	✓	✓	✗	✗	✗	✗	✗	✓
Papagiannakis et al., (2018) [15]	✗	✓	✓	✗	✓	✓	✗	✗	✗
Rizvic et al., (2019) [16]	✓	✗	✗	✓	✗	✗	✗	✗	✗
Sylaiou et al., (2020) [7]	✗	✗	✓	✓	✓	✗	✗	✓	✗
Dionisio (2021) [17]	✓	✓	✗	✗	✗	✗	✗	✗	✗
Hammady et al., (2021) [6]	✗	✗	✓	✓	✓	✗	✓	✗	✗
Teston & Munoz (2021) [18]	✗	✗	✓	✗	✓	✓	✗	✗	✗

### 2.1. Mixed-Reality in CH Spaces

The reality–virtuality continuum [19] classifies mixed-reality as the domain between real and virtual environments. Augmented reality (AR) is the part of MR that is closer to real environments, adding information (e.g., visual or auditory) on top of existing real elements.

A common approach to enhance CH experiences is using AR solutions based on mobile devices [3,20–22] using the technique of Window-on-the-World (WoW) where digital images are overlaid on top of the camera feed of the phone [23]. This approach allows visitors to physically explore the CH site while digital content augments the experience when they see through the device. However, this encourages looking at the phone rather than at the physical space, and focuses on an individual experience [24], constraining the possibilities for group collaboration.

Tangible elements placed in the curated space can mitigate the problem of looking at the phone; for instance, barcodes [25], QR codes [2] or visual markers [3]. This makes visitors pay attention to the physical space before accessing the digital content [3]. Another approach to achieving mobile AR was proposed by [26,27]; they used a see-through head-mounted-display (HMD) and mobile computing equipment to deliver a realistic and interactive immersive experience to the visitors. However, to some extent, this kind of mixed-reality technology also hinders group interaction [25].

Some studies aim at removing the need to carry or wear devices so they propose to project over the physical place [5,11]. Projected augmented reality (AR) offers ways of interacting with an exhibition [28] and improving group experience [29]. For these projection based systems, combining mixed reality with storytelling and gamification is important to achieve user immersion [15].

### 2.2. Interactive Digital Storytelling (IDS)

Using IDS generates a multifaceted experience that encourages conversation and interaction with the visitors [30] by changing from a linear narrative to a more dynamic narrative that, for example, enables the adjustment of the detail and the depth of the presented information without affecting the main educational purpose.

Users appreciated interactive storytelling, empathized with the narrator’s character, and learned information about a cultural monument in an attractive and immersive way [16]. In their work, a website showed elements of a fortress overlooking the city of Sarajevo with an IDS system.

Yasmine’s Adventures [3] mixes storytelling with a physical walk around the area to foster curiosity and willingness to explore the neighborhood [3]. The experience was oriented to individual participants.

For the Etruscanning EU project [11], storytelling and interactive environments are important resources to reconstruct and communicate a cultural context. Their work highlights the importance of interactive storytelling, the definition of adapted content for different audiences and new interaction paradigms to make digital content accessible to all audiences.

CHESS [9] applies content personalization and adaptability to the digital applications of the Acropolis Museum in Athens. This emphasizes the on-site presentation of museum objects through interactive stories personalized to each individual or group of visitors.

Fragments of Laura [3] employs transmedia storytelling to involve tourists in the development of the understanding of cultural heritage. In this case, it uses mobile devices with interactive 360° VR, 2D Motion Comics and Audio Gossips to generate a multimedia story with local awareness and a hypermedia platform. On the other hand, ref. [4] presents three narration alternatives: text panels, audio guides and a virtual guide; the latter improves attention and participation and contributes to better delivery of content and learning.

The use of avatars for the design of an interactive digital narrative [7] gives importance to the social presence, the interrelation between avatars and users and the emotions evoked through different virtual characters that narrate the same story. Thus, the affective potential,

persuasiveness and emotional impact of avatars influence the degree of acceptance of an emotionally charged narrative.

### 2.3. Avatars: Expressive and Realistic Virtual Humans

For an immersive user experience, using a virtual avatar that narrates local stories may guide and make the story more believable to the visitors [7]. Furthermore, a virtual avatar has the data to build on additional historical facts, plus the ability to represent scenes virtually, which helps visitors to understand some concepts and make them more immersive and entertaining [6].

For immersion, having a realistic avatar that resembles a human makes the interaction more believable and relatable for the visitor [31]. These realistic representations make users rate the experience as fairly similar to face-to-face interactions [32]. Avatars can also convey emotions making the interaction more believable [33], and can engage visitors [13].

However, approaching a human-like virtual replica implies some risk, as users may face the uncanny valley [34] when interacting with the avatar. High fidelity characters ended up being considered more attractive, less disturbing, more human and more cheerful than their low fidelity counterparts [35].

Additionally, the real-time capabilities of the avatar are crucial [36], as the visitor needs to receive an almost immediate response for not breaking the immersion.

Some approaches rely on semantics and affective properties to create natural behaviors for creating an empathetic approach where the avatar can listen to the user and react to the speech if it has any mood connotations [33]. Extracting semantic tags from the text to guide the animation could be a solution by dividing them into emotion and intention tags [37]. Other approaches use recurrent neural networks (RNN) to synthesize an output video based on a learned dictionary of personal speech iconic gestures [38].

Face and lipsyncing is a solution present in avatars with bones, action unit and phoneme control [39]. These avatars are often not used in CH environments, as it would need a multidisciplinary team to develop an animation system. However, when they are included [6], they are usually not endowed with emotional capabilities.

### 2.4. Objects

When using tangible objects, the visitors can develop sensory relationships with some of them and value them as physical mementos of their interaction [2]. Some authors allow users to interact with the objects, but these have no real representations and the interaction is performed only with a virtual 3D object [18]. For example, ref. [40] investigates the augmentation of computer generated imagery with a physical replica to increase immersion and awareness; however, the interaction is by means of a laser pointer. Some museum tours have real objects. However, as these are archaeological assets, they must be placed behind security glass, so they use virtual 3D representations to allow the user to look in detail [6].

### 2.5. Gestures

Gestures can be a suitable input from the visitors to control the system, allowing their motion to be tracked and choosing what content they want to access [8]. This gesture detection is usually performed in two ways, by internal tracking of the head or hand position in an AR system [6,11], or by having a physical camera that tracks the user's body or hand gestures [11].

### 2.6. Emotional Interaction

It is well-known that our emotions influence our decision-making [41]. Therefore, it is common for an intelligent virtual tourist guide to have affective capabilities, such as emotions, mood or even a personality [42].

As shown in Table 1, few pieces of research take emotional interaction into account. As is mentioned in [11], there have been significant advancements over the last years in the digitization process with computer graphics techniques and archiving strategies. However,

this does not translate in most virtual museums, as they do not attract the public's attention and involvement: they lack stimulating activities for visitors, narratives metaphors, and emotional impact. In this work, the authors tried to create a deep emotional involvement, using dramatic lighting, storytelling, and sounds. Other studies focus on using emotional content to evoke emotions in the users, for example using virtual reality [13] to connect in a measurably emotional, participatory, interactive and social fashion, or emotional responses caused by virtual humans [7].

### 2.7. Groups Interaction

Although group visits are common in CH environments, the current state-of-the-art in virtual museum experiences tends to focus on individual interactions due to the use of headsets or glasses [6,18]. Other interactions that use mobile apps are also single-user oriented [2,10,16].

However, visitors may vary in age with school groups formed by children that cannot follow instructions for using individual hardware such as glasses or phone apps [5]. In this case, group interaction allows them to participate more easily.

## 3. Design Methodology

A specific methodology for designing interactive tourist experiences has not been found within the related work. We propose a set of tools for designing interactive experiences for cultural heritage places in-situ, emphasizing the visitor's emotions, immersion and freedom to interact with the heritage. As a conclusion extracted from the related work, we propose a profile of experts that should participate in the design process.

### 3.1. Interdisciplinary Team

A common issue in new media research is that it is usually constrained to one scientific field. In Section 2, we found some systems designed by an interdisciplinary team but, except for the design made by [16], the interdisciplinary team was composed of end-users and technical experts.

Interactive digital storytelling needs experts from all relevant fields to develop a new methodology that all target groups can appreciate. As in [16], we specifically address the opposing views of artistic liberty on one side, and technical constraints and historical facts on the other. As described in Section 2, designing an immersive interactive experience in CH involves several features for our research: 3D projection mapping, narration, realistic avatars, images, music, and emotions. Only a multidisciplinary team can combine these notions into an interactive application.

The main objective of the team is to achieve user immersion, the user is the central point. Our multidisciplinary team is formed by experts in designing the communication and interaction with the system to obtain user immersion in the story while taking educational, entertainment, and easy-of-use points of view. The following sections describe expertise areas required for the multidisciplinary development team.

- Computer science. To implement the proposed design with the available hardware and software, the system should run in realtime and be deployed in a computer at the museum. The computer science team must include an expert in human-computer interaction;
- Archaeologists. Interactive tangible objects should be precisely represented so that the tourist learn about the history and art of the place they are visiting;
- Historian. Content must be historically accurate; explanations and dates should be precise;
- Film arts directing. For interactive digital stories, both movies and interactive virtual environments must be taken into account. Once the story is set, the director must design the interaction in these virtual environments so users can experience the story in the best possible way;

- Screenwriter, storytelling for cultural heritage. The story is a core point of the design proposal. The experience starts when the visitors arrive at the place and lasts until they leave, so the characters involved in the story, the storytelling itself and the user interaction should make an immersive experience. Also, the historical facts provided by the historian and existing archaeological knowledge must be translated into an interactive format, so the tourists can learn in an entertaining way;
- Game directing. There should be a gamification process so visitors, both children and adults, can be entertained and are able to interact and complete tasks between short explanations;
- Visual arts, graphics design. Visuals are key for making a first impact on the user. The characters must look real and 3D animations should look natural so the uncanny valley is avoided. Other representations such as 3D environments must look accurate for the story's needs, with special attention paid to textures, 3D models and overall lighting;
- End-users: since we are designing an interactive experience for a cultural heritage place, we defined a tourist as the end user for whom the experience is designed.

### 3.2. Proposed Interactive Technologies

In this section, we propose a set of features to develop an in-situ immersive experience where the visitors are free to touch and explore the CH elements and learn about the history around them:

- In-Situ: in the state-of-the-art, several CH applications have been designed for VMs or Web; however, in this work we focus on applications designed to experience CH in the real place. The technical team must study the environmental restrictions and select a suitable hardware architecture that respects them;
- Mixed Reality (MR): there is a wide range of technologies used to introduce the user to an immersive environment. We focus on technologies that augment reality, leaving the visitor with freedom of movement. Since our aim is for the visitors to feel the experience in-situ, being free to move and interact naturally with real objects (archaeological reproductions and 3D prints) and virtual ones (virtual guide and 3D scans) is crucial. Augmentation will be achieved by projecting the digital content on the physical world employing 3D video mapping. We propose the design of a virtual projection over the real space, given the importance of physical experiences and group interaction. The projection is also supported by the use of tangible objects that are connected with a virtual replica for greater user immersion;
- Interactive digital storytelling (IDS): tourists want to learn about the history and culture of the visited place. The IDS allows the generation of the narrative through the interactive events, giving the tourist the possibility of being the director of his/her own visit;
- Avatar: the storytelling should be performed in an expressive way by a believable avatar;
- Objects: tangible objects will be placed around the space to enable the visitors to directly interact with them. This creates an emotional connection but also allows the system to detect the user's preferences and adapt the storytelling and activities;
- Gestures: to offer an interactive experience the visitor needs to interact with the system. Thanks to the IDS, we will have voice interaction but integrating a body gesture analyzer could offer more information about the visitor's preferences. We will also propose 3D printed replicas of the real objects that can be manipulated by the users as a way of interaction;
- Emotions: the interdisciplinary team must consider the visitor emotions during the design. Technology must be included in order to recognize the visitor emotions and adapt the storytelling to them;
- Groups: the in-situ visits are usually made by groups of people, so the system will need technology to manage group interaction.

### 3.3. Participatory Design

Creating experiences for a cultural heritage persona remains a challenge because no clear methodology currently exists [43]. We propose to use a participatory design method based on the User Journey Map method. Journey maps have user experience deliverables, and they have widespread usage among UX experts [44]. A User Journey is a user-centered design method mainly used in a system's definition or evaluation phases. It shows step by step the user interaction with a system describing their emotions and reactions at each of the points of contact (touchpoints) with the product.

We decided to follow this methodology for several reasons. First of all, a User Journey Map helps us to understand and communicate users' behavior as they progress through a route through interactions as they try to accomplish their goals. The tool is focused on empathy for user needs and understanding their behaviors. Journey maps also provide visualization techniques for effective communication inside multidisciplinary teams [43].

## 4. Workshop

### 4.1. Preliminary Session: Persona, User Goals and Scenario

Each User Journey Map has to reflect the itinerary of a specific person through a given scenario. Therefore, the first step was a preliminary brainstorming session for defining the persona and the task model which explains the main scenario.

This session was performed by seven persons with one expert from each discipline: computer science, archaeologists, historian, film arts director, screenwriter, game directing, visual and end-users.

The session lasted 2 h and was carried out following the approach by [44], who defined a persona using the 3i method: investigate, identify, and imagine.

The first aim of the session was to define the persona. This element will represent a group of users with similar goals. During the first part of the session, the interdisciplinary team defines Mrs. Dubois as our user target. The definition performed of Mrs. Dubois as Persona is detailed in Table 2.

**Table 2.** Mrs. Dubois definition as Persona.

<b>Behavior</b>	<ul style="list-style-type: none"> <li>• She often travels by car with her partner and children;</li> <li>• She likes to visit places of cultural and historical interest;</li> <li>• Although she is not fluent in Spanish or English, enjoys traveling to other countries, interested in learning about their history and culture;</li> <li>• She likes new technologies and has several gadgets at home;</li> <li>• When she is traveling, usually invests in new experiences that are entertaining and suitable for her children.</li> </ul>
<b>Demographic</b>	<ul style="list-style-type: none"> <li>• Female 40 years old;</li> <li>• Gender: Female;</li> <li>• Nationality: French;</li> <li>• Educational Level: Professional;</li> <li>• Occupation: Works as Engineer;</li> <li>• Family Members: Husband and two daughters;</li> <li>• Language: French;</li> <li>• Civil Status: Married;</li> <li>• Socioeconomic status: Average.</li> </ul>
<b>Objectives</b>	<ul style="list-style-type: none"> <li>• Discover historical and cultural sites of interest in Spain while traveling by car;</li> <li>• Have content adapted to your language;</li> <li>• Experiment with new technologies and find entertainment;</li> <li>• Take away a souvenir of the experience and potentially showing it to her friends through social networks.</li> </ul>

The next step was to define a task model. The task model is a story of what our persona does at each milestone of their journey. To capture the general picture of the experience, we start with a milestone before the persona’s first interaction with our system [44] as is shown in Figure 1. Then, we define four milestones from the beginning to the end with user goals for each milestone. With this information the team was able to specify Mrs. Dubois’s complete scenario as a long story describing her behavior from the beginning to the post-visit (Figure 2).

Milestones	Plan/Research	Discover	Visit	Post Visit
User Goals	<ul style="list-style-type: none"> <li>Get to know local culture</li> <li>Get to know about live experiences</li> </ul>	<ul style="list-style-type: none"> <li>Find a place to visit</li> <li>Get information about what to do there</li> </ul>	<ul style="list-style-type: none"> <li>See and feel the place</li> <li>Listen to the stories</li> <li>Make questions</li> <li>Learn the content</li> <li>Enjoy the visit</li> </ul>	<ul style="list-style-type: none"> <li>Take a memory</li> <li>Share the experience</li> </ul>

Figure 1. Milestones and User Goals of a visitor.

Sra. Dubois	Journey Map: Visit to Roncesvalles with your family					
Stages/ Milestones	PLAN/RESEARCH	DISCOVER	VISIT		POST VISIT	
Touchpoints	Visitor needs to find information about the destination	Visitor select the Roncesvalles visit	The visitor enters the conference room. She sees the environment and the archeological objects. She wants to learn about the place and objects history.		The visitor leaves the room	
Actions	Find in Internet attractions and facilities around Roncesvalles	Access the website where the experience is promoted	Listen to the stories and take pictures. He also asks questions. The human guide tells the avatar that there are little girls in the room!		Share the selfie on your social networks	
Thoughts	"I hope to find something interesting, that we can understand and in turn entertaining for girls"	"Wow! This can be interesting and fun for everyone."	"How great 3D mapping and what interesting stories!!"	"I'm afraid girls will start to get bored of stories!"	"Uff lucky there are games!"	"Hopefully we can go back!"
Emotions						
Pain Points	You feel somewhat anxious to know if you will find something suitable for the whole family	May not find the web or not understanding what is involved	Language can be a barrier	The dynamics of children may be incompatible with that of other visitors	Surely the photo is dark	
Opportunities	Remark on the promotional website that it is an experience for everyone	Good dissemination and multimedia content of the project	Registration on arrival + Adaptive and interactive storytelling	Design experiences with and without games Separate into groups with schedules for children Design playful dynamics suitable for everyone	Take a moment of the experience to take pictures with good lighting	

Figure 2. High level Mrs. Dubois’s Journey Map.

The final step was to determine what information should be included in the diagram:

- Touchpoints: What part of the experience does the persona interact with?
- Actions: What does the visitor do? What information do they look for? What is their context?
- Thoughts: What is the visitor thinking?
- Emotions: What does the visitor feel during the interaction?
- Pain points: What does the visitor want to avoid? What are the obstacles?
- Opportunities: What could we improve to enhance the experience at this point?

At the end of this preliminary session, the multidisciplinary team learned how to define a User Journey Map by creating a high-level Mrs. Dubois’ Journey Map.

#### 4.2. Main Session: User Journey Map Design

This second session aimed at creating a more detailed view of the Visit of Mrs. Dubois’ Journey Map and designing a low-level map for the interactive experience inside the Roncesvalles Place.

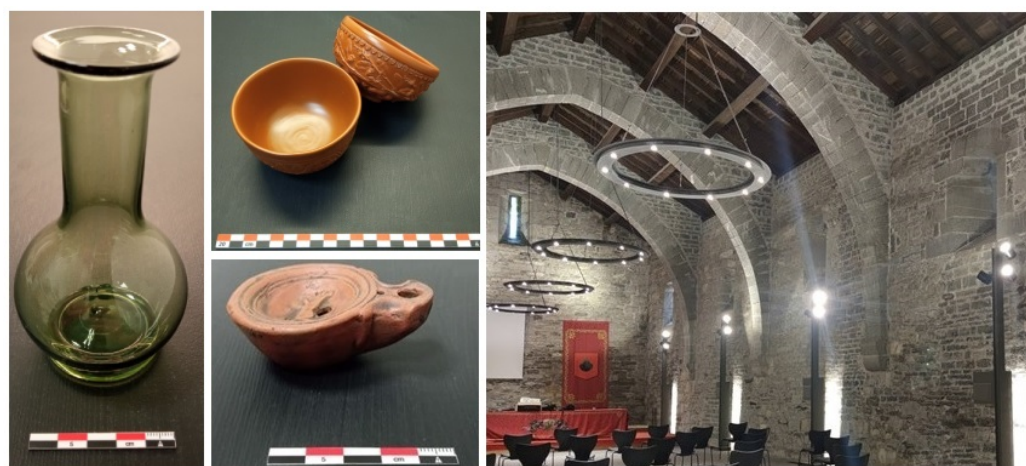


On this occasion, the team was made up of nine people, one history professor and two computer science professors from the Public University of Navarre. Two experts in directing and gamification came from a video game company and two archaeologists came from an archaeological company. Finally, two computer engineers also participated. Each of them took on one or more roles from the profiles defined in the interdisciplinary team. As [16] pointed out, the structure of the multidisciplinary team is inherently flexible as one person can have more than one expertise, and more than one person can share the same expertise. In this session, the specific team structure was composed of five computer scientists, two archaeologists, one historian, one film arts director, one screenwriter, one game director, one visual arts expert and five end-users.

The session lasted 3 h and was divided into three parts—30 min for contextualization, 1 h for brainstorming and 30 min for creating the User Journey Map.

The interdisciplinary team started the session with a summary of the previous sessions and the contextualization of the project. For this purpose, the team was informed about the room where the experience will be installed and the archaeological objects that will be part of the exhibition (Figure 3). Then, the computer science team introduced the interaction toolkit so that non-technical profiles could understand the technological elements that they can use during the design. The toolkit includes the following elements:

- Video Mapping 3D: Explanation of what video mapping 3D is and discussion of the opportunities that this technology can offer to achieve an immersive illusion;
- Avatar: The emotional conversational agent can chat with the tourists, selecting content according to the emotions that are to be transmitted;
- Tangible Objects. Inform that there will be archaeological reproductions in the room. When users interact with them, the system will be able to respond accordingly. We held a discussion about the type of interactions with the objects and their responses;
- Interactive Objects: The system, either through vision or through sensors, will be able to detect when a group of users is around and when they lift the objects;
- Tracking system: Groups of people could be detected; identity can be tracked by comparing the face with a photo taken at the beginning of the experience. The system will also be capable of counting some group gestures as raised hands.



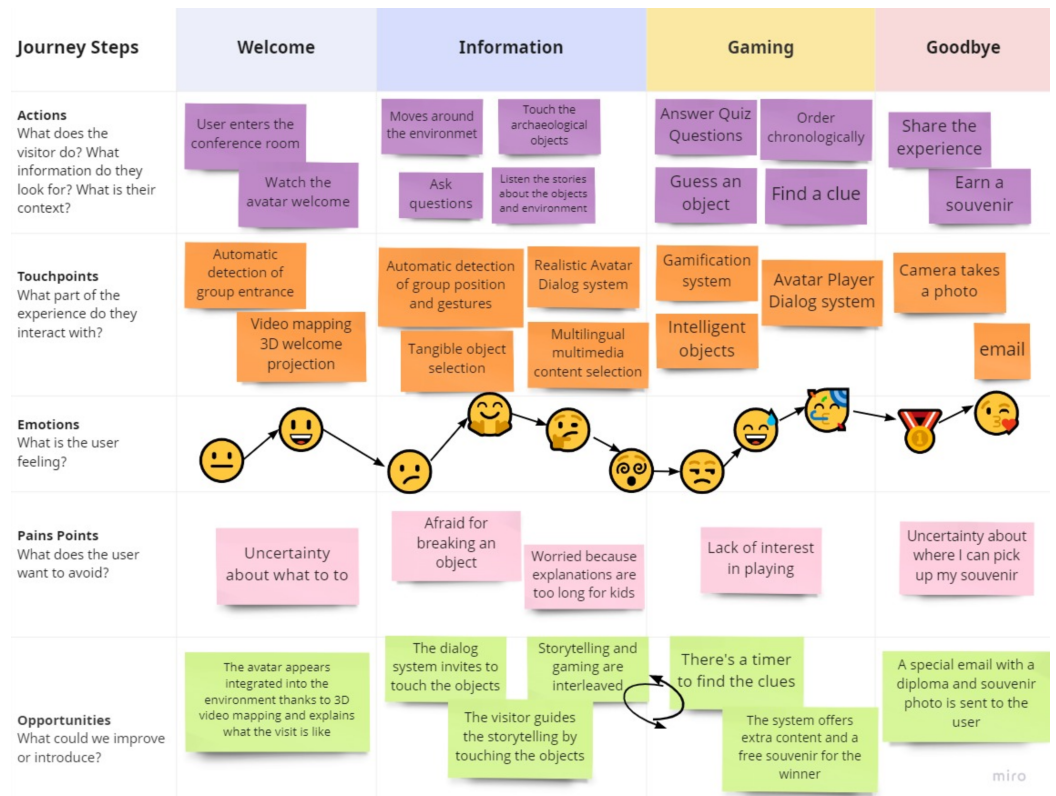
**Figure 3.** Cultural Heritage: Roncesvalles Room and Archaeological objects used as interactive objects.

Then, the interdisciplinary team held a brainstorming session following the User Journey Map Method to design an empathetic and interactive cultural heritage immersive experience. For the brainstorming session the interdisciplinary team used the materials shown in Figure 4.

Finally, the team gathered the generated ideas and designed the final experience in a User Journey Map, obtaining the result shown in Figure 5.



**Figure 4.** The material used during the second brainstorming session for designing an empathetic and interactive cultural heritage immersive experience.



**Figure 5.** Immersive experience in Roncesvalles: User Journey Map.

### 4.3. Workshop Results

We obtained two main contributions. First, a set of guidelines for designing an immersive CH experience defined by an interdisciplinary team of experts. Second, a user-centered design system for developing an immersive CH experience in Roncesvalles. The design is shown by means of a User Journey Map and a simulation of the scene.

#### 4.3.1. Guidelines for Designing an Immersive CH Experience

After completing the workshop, we collected a set of guidelines for designing an immersive and interactive presentation of cultural heritage provided by interdisciplinary experts.

- Experts from CH documentation: Historian and Archaeologists.

- Balance objects' appearance so they invite the users to interact with them. A realistic replica allows visitors to appreciate fine details, yet 3D printed objects encourage users to interact without fearing damaging the object;
- Real books can offer additional information. They can be used as smart objects that enhance the story. In this case, its visualization should be adapted to the group using projections and avatars;
- The information must be correctly structured, avoiding unnecessary material and giving the visitor the opportunity to choose the level of depth of the story he/she wishes to know.
- Experts from arts: graphics design and film directing.
  - For an immersive experience, 3D mapping and creating a seamless transition between the actual wall and the avatar appearance is required;
  - The avatar should be hyper-realistic to give the visitor the illusion of speaking with a guide;
  - For immersion, the scene and camera should be static for allowing spatial coherence;
  - The multimedia content displayed on the walls of the site should always be integrated by 3D mapping to avoid breaking the sense of immersion and the realism of the avatar.
- Screenwriter and storytelling for cultural heritage.
  - The story should not be linear. The visitor should choose what content they want to listen to by asking questions or selecting an object;
  - The IDS information should be structured in different levels of detail, the user being the one who decides how deep to dive into the lore that the object offers;
  - Avatar's sentences should remain short as visitors tend to get bored and disconnect when explanations are long and interaction decreases;
  - The designers should answer the following question: Who should be leading the story explanation? Is the avatar a tool for the real guide or is it the other way around? Will the human guide lead the narrative or is it driven by the interaction with the environment and its objects?
  - The storytelling should follow the segmentation principle where the material is presented in non linear segments so the user can control the narrative.
- Game directing.
  - Intertwine the story with minigames—there may be kids that will lose attention during long explanations;
  - Minigames must be of different types and complexities to attract both the adults and the children;
  - Encourage taking part with a final reward.
- Computer Science.
  - Verify that the interactions defined during the workshop are achievable with the hardware and architecture of the proposed software;
  - The arrangement of the elements inside the room (Projector, Objects, Visitors, Speakers or Cameras) should be designed following the restrictions of the CH.
- End Users.
  - Texts should be translated into different languages;
  - The visitor wants to feel comfortable during the visit and it is important to consider their emotions. For example, in the case of family tourism, the parents usually get worried if the children touch something that might break or if the content is not appropriate for children.

#### 4.3.2. User-Centered Design Experience for Roncesvalles

For designing affective experiences for cultural heritage, it is necessary to consider the emotional journey that the tourist experiences from the beginning of the visit until the end.

We have designed an interactive experience located at Roncesvalles. Roncesvalles is a very important town within the Camino de Santiago because it is the point where the roads converge. Roncesvalles is a suitable place to make an in-situ exhibition given its geographical location and rich historical content. In this work, the archaeologists and historians designed an exhibition in the context of Roman times, since the Romans built a road that passed through the town and its history is based on legends. The exhibition was designed to be installed in the conference room at “Royal Collegiate Church of Roncesvalles’s Santa María”, shown in Figure 3.

As represented in the User Journey Map created in the workshop, the experience is organized into four steps—Welcome, Information, Gaming and Goodbye (Figure 5).

At the beginning, the tour group enters the room and the system automatically detects the entrance to start a wall projection with 3D video mapping. The group will notice how suddenly the wall breaks and falls apart. Behind the wall, a mine is revealed and a virtual archaeologist greets the group of visitors. The projected mine is a photogrammetry of a real mine found in the Roncesvalles area. The video is an opportunity to alleviate the visitor’s uncertainty and explain how to start the visit in an immersive way.

After the welcome step, the group can move around the room to observe the environment and the replicas of archaeological objects placed in the room. The objects selected for this first prototype were a lacrimarie, a bowl and an oil lamp shown in Figure 3. They are related to the historical context selected for the visit and were found at the archaeological excavations around Roncesvalles. The system automatically detects if the tour group is around an object to trigger the avatar storytelling related to that object as is simulated in Figure 6.

When the avatar begins to tell stories related to the objects, adult visitors may think that these stories will be boring for the children and they will lose interest. However, parents are relieved when they discover that the avatar can detect children in the room and proposes activities for the children, who can play a series of games. During the design workshop the team decide to mixed storytelling with gaming in order to entertain the whole family, both adults and children. So, together with the adults, the children play mini-games that follow the gamification proposed by the multidisciplinary team, with playable activities adapted for different audiences. The mini-games designed for this experience are related to the tangible interactive objects in the room and with the storytelling explained by the avatar. For example, one mini-game is to find an object hidden in the room following the clues given during the storytelling. Another example, after listening to the avatar’s narration, the tourists must order the objects chronologically. For designing the whole set of mini-games, theories such as the flow concept provided by Csikszentmihalyi [45] should also be taken into account. Therefore, the set of mini-games should present activities with enough difficulty and involving some risk.

Finally, at the goodbye step, the avatar surprises the visitors by telling them that she has taken some photos of the activities and that she can send them as a gift to their e-mails so that they can share the experience on social networks.



**Figure 6.** Simulation of visitors listening to the projected Avatar during the visit.

## 5. Proposed Framework

In this article, we present a design for implementing an interactive on-site experience. The proposed framework is shown in Figure 7. This design considers two inputs (visual and auditory) through which the information reaches the virtual avatar. The auditory input, in this case coming from a microphone, extracts the multimodal information perceived from the voice of the visitor, passing it to the Text module, which transcribes the audio content and the Emotional Audio Analyzer module, which detects the emotional tone of the voice.

The visual input is captured using two Kinect v2 cameras, which track the position of the visitors around the area. They also detect the interaction of the group with objects in the environment, the individual gestures and the emotion on facial expressions.

This information is sent to the interaction manager module. The information is classified, structured and sent to the dialogue manager module, which is in charge of deciding what to perform and the response it should give, as well as how it affects the avatar state.

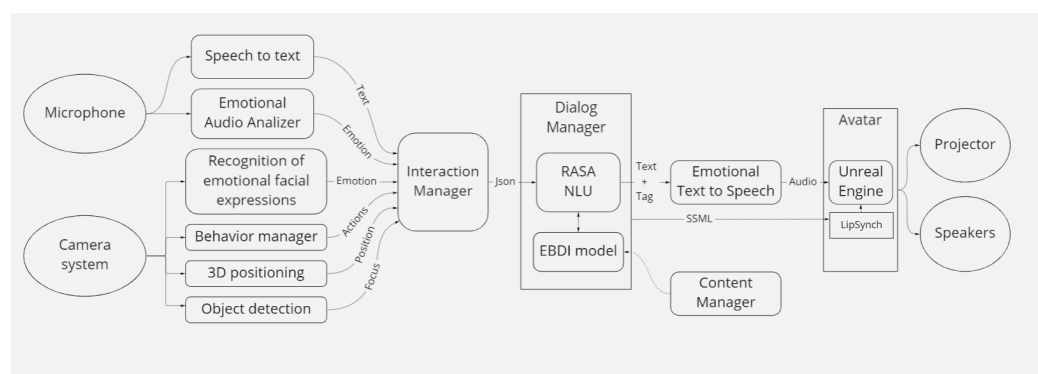
The NLU module detects the conversational intention of the visitors and the emotional polarity of the sentence. At the same time, the personalized action based on the EBDI model establishes the beliefs related to the incoming information. Finally, it offers plans according to the situation, using the content manager and thus offering a plain text and an SSML (Speech Synthesis Markup Language) file as a response. The markup text response goes, on the one hand, to the Emotional Text-to-Speech module to generate the response audio and, on the other hand, to the avatar module to generate the expressive animations. Finally, the system mixes the rendered avatar with the audio and then projects it onto the wall. The audio is played through the speakers.

Regarding the avatar module, there are several options for generating a believable avatar; however, in the current state-of-the-art, the Epic Games Unreal engine offers realistic avatar support called Metahumans [46]. As it is a customizable engine, it allows us to build up the proposed avatar module that will take the generated text and convert it into final face and body animations for the avatar.

The proposed avatar module takes the SSML input from the dialog manager and reads the visemes so it can generate a lipsync animation from it. Parallel to this, the lipsync plays along with the generated audio, so it can both have a voice and move the lips while reading the input.

As the dialog manager offers a conversational agent and an emotional EBDI model, it allows the export of emotional tags that will endow the avatar with emotional reactions and an empathy system to follow and react to not only the content of the user's message but its intention. These tags will generate different animations on the avatar based on emotions, intensity, and semantic features to emphasize essential terms or to be able to point out exciting objects following the story's needs.

The framework is designed to implement a sufficiently scalable prototype since user inputs are considered and the storytelling as well as the avatar animation are generated in real time from text. We note that the group sizes are limited during the regular exhibition and this limitation can also be used to ensure that the interactive experience works properly.



**Figure 7.** Proposed Framework for implementing an interactive on-site experience.

## 6. Conclusions

Our work designed interactive experiences for cultural heritage. We informed the system design following a participatory design approach combined with UX techniques. The whole design process was enriched by an interdisciplinary team who addressed previously suggested challenges [3] regarding the storytelling content, mixed-reality environment and the architecture of the physical space. Specifically, we contribute by summarizing the process in terms of design guidelines and recommendations for future developments.

**Author Contributions:** Conceptualization, X.O., R.G., A.O., S.M. and J.V.; methodology, A.O. and S.M.; software, X.O. and R.G.; validation, J.V., O.A. and A.M.; resources, A.O., J.V., O.A. and A.M.; writing—original draft preparation, X.O., R.G., S.M. and A.O.; writing—review and editing X.O., R.G., S.M., A.O., J.V., O.A. and A.M.; supervision, A.M.; funding acquisition, A.O. and J.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the R & D projects of the Government of Navarra under grant agreement No 0011-1365-2021-000063.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank Carlos Zuza and Nicolás Zuazúa from the archaeological company Gabinete TRAMA for their contributions for the designing of Storytelling and Tangible objects. We also thank to Javier Larumbe from Nexxyo Labs company for his contribution to arts and gamification design. Finally, we would like to thank Iñigo Mugueta of Medieval History & Historical Heritage, for his contributions to Storytelling documentation and game design.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

CH	Cultural Heritage
VM	Virtual Museums
VR	Virtual Reality
MR	Mixed Reality
AR	Augmented Reality
IDS	Interactive Digital Storytelling
IDT	Interdisciplinary Design Team

## References

1. Raptis, G.E.; Fidas, C.; Katsini, C.; Avouris, N. A cognition-centered personalization framework for cultural-heritage content. *User Model. User Adapt. Interact.* **2019**, *29*, 9–65. [[CrossRef](#)]
2. Ciolfi, L.; Mcloughlin, M. Designing for meaningful visitor engagement at a living history museum. In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense through Design, Copenhagen, Denmark, 14–17 October 2012; pp. 69–78. [[CrossRef](#)]
3. Dionisio, M.; Barreto, M.; Nunes, N.; Nisi, V. A Mixed Reality neighborhood tour: Understanding visitor experience and perceptions. *Entertain. Comput.* **2018**, *27*, 89–100. [[CrossRef](#)]
4. Carrozzino, M.; Colombo, M.; Tecchia, F.; Evangelista, C.; Bergamasco, M. Comparing different storytelling approaches for virtual guides in digital immersive museums. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Cham, Switzerland, 2018; pp. 292–302. [[CrossRef](#)]
5. Schaper, M.M.; Santos, M.; Malinverni, L.; Berro, J.Z.; Pares, N. Learning about the past through situatedness, embodied exploration and digital augmentation of cultural heritage sites. *Int. J. Hum. Comput. Stud.* **2018**, *114*, 36–50. [[CrossRef](#)]
6. Hammady, R.; Ma, M.; AL-Kalha, Z.; Strathearn, C. A framework for constructing and evaluating the role of MR as a holographic virtual guide in museums. *Virtual Real.* **2021**, *25*, 895–918. [[CrossRef](#)]
7. Sylaiou, S.; Kasapakis, V.; Gavalas, D.; Dzardanova, E. Avatars as storytellers: Affective narratives in virtual museums. *Pers. Ubiquitous Comput.* **2020**, *24*, 829–841. [[CrossRef](#)]
8. Flynn, B. v-Embodiment for Cultural Heritage. In Proceedings of the 2013 Digital Heritage International Congress (DigitalHeritage), Marseille, France, 28 October–1 November 2013; pp. 347–354.
9. Pujol, L.; Katifori, A.; Vayanou, M.; Roussou, M.; Karvounis, M.; Kyriakidi, M.; Eleftheratou, S.; Ioannidis, Y. From personalization to adaptivity. Creating immersive visits through interactive digital storytelling at the Acropolis Museum. In *Museums as Intelligent Environments Workshop (MasIE), Proceedings of the 9th International Conference on Intelligent Environments, Athens, Greece, 16–17 July 2013*; IOS Press: Amsterdam, The Netherlands, 2013; pp. 541–554. [[CrossRef](#)]
10. Niculescu, A.I.; Yeo, K.H.; D’Haro, L.F.; Kim, S.; Jiang, R.; Banchs, R.E. Design and evaluation of a conversational agent for the touristic domain. In Proceedings of the Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific, Siem Reap, Cambodia, 9–12 December 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 1–10. [[CrossRef](#)]
11. Pietroni, E.; Adami, A. Interacting with virtual reconstructions in museums: The etruscanning project. *J. Comput. Cult. Herit.* **2014**, *7*, 1–29. [[CrossRef](#)]
12. Seppälä, K.; Heimo, O.I.; Korkalainen, T.; Pääskylä, J.; Latvala, J.; Helle, S.; Härkänen, L.; Jokela, S.; Järvenpää, L.; Saukko, E.; et al. Examining User Experience in an Augmented Reality Adventure Game: Case Luostarinmäki Handicrafts Museum. In Proceedings of the IFIP International Conference on Human Choice and Computers, Salford, UK, 7–9 September 2016; pp. 257–276. [[CrossRef](#)]
13. Perry, S.; Economou, M.; Young, H.; Roussou, M.; Pujol, L. Moving beyond the virtual museum: Engaging visitors emotionally. In Proceedings of the 2017 23rd International Conference on Virtual Systems and Multimedia, VSMM 2017, Dublin, Ireland, 31 October–4 November 2017; pp. 1–8. [[CrossRef](#)]
14. Vosinakis, S. The use of digital characters in interactive applications for cultural heritage. In *Applying Innovative Technologies in Heritage Science*; IGI Global: Hershey, PA, USA, 2020; pp. 109–137. [[CrossRef](#)]
15. Papagiannakis, G.; Geronikoulakis, E.; Pateraki, M.; López-Menchero, V.M.; Tsioumas, M.; Sylaiou, S.; Liarokapis, F.; Grammatikopoulou, A.; Dimitropoulos, K.; Grammalidis, N.; et al. *Mixed Reality, Gamified Presence, and Storytelling for Virtual Museums*; Springer International Publishing: Cham, Switzerland, 2018; pp. 1–13. [[CrossRef](#)]
16. Rizvic, S.; Boskovic, D.; Okanovic, V.; Sljivo, S.; Zukic, M. Interactive digital storytelling: Bringing cultural heritage in a classroom. *J. Comput. Educ.* **2019**, *6*, 143–166. [[CrossRef](#)]
17. Dionisio, M.; Nisi, V. Leveraging Transmedia storytelling to engage tourists in the understanding of the destination’s local heritage. *Multimed. Tools Appl.* **2021**, *80*, 34813–34841. [[CrossRef](#)]
18. Teston, A.; Munoz, A. Digital Avatars as Humanized Museum Guides in the Convergence of Extended Reality. In Proceedings of the MW21 Conference, Washington, DC, USA, 5–9 April 2021.

19. Milgram, P.; Takemura, H.; Utsumi, A.; Kishino, F. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and Telepresence Technologies*; SPIE: Bellingham, DC, USA, 1995; Volume 2351, pp. 282–292.
20. Lanir, J.; Wecker, A.J.; Kuflik, T.; Felberbaum, Y. Shared mobile displays: An exploratory study of their use in a museum setting. *Pers. Ubiquitous Comput.* **2016**, *20*, 635–651. [[CrossRef](#)]
21. Egglestone, S.; Roussou, M.; Brundell, P.; Chaffardon, C.; Kourtis, V.; Koleva, B.; Benford, S. Indoors and outdoors: Designing mobile experiences for Cité de l'espace. In Proceedings of the NODEM 2013 Conference, Stockholm, Sweden, 1–4 December 2013.
22. Kenderdine, S.; Chan, L.; Shaw, J. Pure Land. *J. Comput. Cult. Herit.* **2014**, *7*, 1–15. [[CrossRef](#)]
23. Milgram, P.; Kishino, F. A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Inf. Syst.* **1994**, *77*, 1321–1329.
24. Betsworth, L.; Bowen, H.; Robinson, S.; Jones, M. Performative technologies for heritage site regeneration. *Pers. Ubiquitous Comput.* **2014**, *18*, 1631–1650. [[CrossRef](#)]
25. O'Hara, K.; Kindberg, T.; Glancy, M.; Baptista, L.; Sukumaran, B.; Kahana, G.; Rowbotham, J. Collecting and sharing location-based content on mobile phones in a zoo visitor experience. *Comput. Support. Coop. Work.* **2007**, *16*, 11–44. [[CrossRef](#)]
26. Vlahakis, V.; Ioannidis, M.; Karigiannis, J.; Tsoiros, M.; Gounaris, M.; Stricker, D.; Gleue, T.; Daehne, P.; Almeida, L. Archeoguide: An augmented reality guide for archaeology sites. *IEEE Comput. Graph. Appl.* **2002**, *22*, 52–60. [[CrossRef](#)]
27. Papagiannakis, G.; Ponder, M.; Molet, T.; Kshirsagar, S.; Cordier, F.; Magnenat-Thalmann, N. LIFEPLUS: Revival of Life in Ancient Pompeii. VSM2002-Invited Paper. 2002. Available online: <https://infoscience.epfl.ch/record/100872> (accessed on 30 May 2022).
28. Plecher, D.A.; Ulschmid, A.; Kaiser, T.; Klinker, G. Projective Augmented Reality in a Museum: Development and Evaluation of an Interactive Application. In Proceedings of the ICAT-EGVE 2020—International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments, Orlando, FL, USA, 2–4 December 2020; Argelaguet, F., McMahan, R., Sugimoto, M., Eds.; The Eurographics Association: Geneva, Switzerland, 2020. [[CrossRef](#)]
29. Grammenos, D.; Michel, D.; Zabulis, X.; Argyros, A.A. *PaperView*; ACM: New York, NY, USA, 2010; pp. 57–60. [[CrossRef](#)]
30. Wyman, B.; Smith, S.; Meyers, D.; Godfrey, M. Digital Storytelling in Museums: Observations and Best Practices. *Curator Mus. J.* **2011**, *54*, 461–468. [[CrossRef](#)]
31. Karuzaki, E.; Partarakis, N.; Patsiouras, N.; Zidianakis, E.; Katzourakis, A.; Pattakos, A.; Kaplanidi, D.; Baka, E.; Cadi, N.; Magnenat-Thalmann, N.; et al. Realistic Virtual Humans for Cultural Heritage Applications. *Heritage* **2021**, *4*, 4148–4171. [[CrossRef](#)]
32. Rogers, S.L.; Broadbent, R.; Brown, J.; Fraser, A.; Speelman, C.P. Realistic Motion Avatars are the Future for Social Interaction in Virtual Reality. *Front. Virtual Real.* **2022**, *2*, 750729. [[CrossRef](#)]
33. Yalçın, Ö.N. *Modeling Empathy in Embodied Conversational Agents*; ACM: New York, NY, USA, 2018; pp. 546–550. [[CrossRef](#)]
34. Mori, M.; MacDorman, K.; Kageki, N. The Uncanny Valley [From the Field]. *IEEE Robot. Autom. Mag.* **2012**, *19*, 98–100. [[CrossRef](#)]
35. Higgins, D.; Egan, D.; Fribourg, R.; Cowan, B.; McDonnell, R. *Ascending from the Valley: Can State-of-the-Art Photorealism Avoid the Uncanny?* ACM: New York, NY, USA, 2021; pp. 1–5. [[CrossRef](#)]
36. Rebol, M.; Gütl, C.; Pietroszek, K. *Real-Time Gesture Animation Generation from Speech for Virtual Human Interaction*; ACM: New York, NY, USA, 2021; pp. 1–4. [[CrossRef](#)]
37. Hu, L.; Qi, J.; Zhang, B.; Pan, P.; Xu, Y. *Text-Driven 3D Avatar Animation with Emotional and Expressive Behaviors*; ACM: New York, NY, USA, 2021; pp. 2816–2818. [[CrossRef](#)]
38. Liao, M.; Zhang, S.; Wang, P.; Zhu, H.; Zuo, X.; Yang, R. Speech2Video Synthesis with 3D Skeleton Regularization and Expressive Body Poses. In Proceedings of the Asian Conference on Computer Vision, Singapore, 20–23 May 2021; pp. 308–323. [[CrossRef](#)]
39. Aneja, D.; McDuff, D.; Shah, S. *A High-Fidelity Open Embodied Avatar with Lip Syncing and Expression Capabilities*; ACM: New York, NY, USA, 2019; pp. 69–73. [[CrossRef](#)]
40. Bugalia, N.; Kumar, S.; Kalra, P.; Choudhary, S. *Mixed Reality Based Interaction System for Digital Heritage*; ACM: New York, NY, USA, 2016; pp. 31–37. [[CrossRef](#)]
41. Hu, J.; Guan, C.; Wang, M.; Lin, F. Model of emotional agent. In *Pacific Rim International Workshop on Multi-Agents*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Berlin/Heidelberg, Germany, 2006; pp. 534–539. [[CrossRef](#)]
42. Sánchez-López, Y.; Cerezo, E. Designing emotional BDI agents: Good practices and open questions. *Knowl. Eng. Rev.* **2019**, *34*, e26. [[CrossRef](#)]
43. Subhi, N.N.A.; Bell, D. Hybrid user experience (UX) modelling in a heritage setting. In Proceedings of the Winter Simulation Conference, Gothenburg, Sweden, 9–12 December 2018; pp. 1298–1309. [[CrossRef](#)]
44. Szabo, P.W. *User Experience Mapping*; Packt: Birmingham, UK, 2017; pp. 1–352. [[CrossRef](#)]
45. Csikszentmihalyi, M. *Flow: The Psychology of Optimal Experience*; Harper Perennial: New York, NY, USA, 1991.
46. Wang, S.; Liu, H.; Zhang, X. *Design and Implementation of Realistic Rendering and Immersive Experience System Based on Unreal Engine4*; ACM: New York, NY, USA, 2020; pp. 11–15. [[CrossRef](#)].