A Multimodal XR Framework for Heritage Engagement and Analysis: The Case of Torre del Mar in Borriana (Castellón, Spain)

Original
A Multimodal XR Framework for Heritage Engagement and Analysis: The Case of Torre del Mar in Borriana (Castellón, Spain) / Pupi, Enrico; Spallone, Roberta; Rinascimento, Martina; Gil-Piqueras, Teresa; Rodriguez-Navarro, Pablo. - In: INTERNATIONAL ARCHIVES OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES. - ISSN 2194-9034. - ELETTRONICO. - XLVIII-2/W12-2026:(2026), pp. 399-406. [10.5194/isprs-archives-xlviii-2-w12-2026-399-2026]

Availability:
This version is available at: 11583/3007599 since: 2026-02-13T18:17:33Z

Publisher:
Copernicus

Published
DOI:10.5194/isprs-archives-xlviii-2-w12-2026-399-2026

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-2/W12-2026
11th Intl. Workshop 3D-ARCH "3D Virtual Reconstruction and Visualization of Complex Architectures", 10–12 February 2026, Ancona, Italy

A Multimodal XR Framework for Heritage Engagement and Analysis: The Case of Torre del Mar in Borriana (Castellón, Spain)

Enrico Pupi [1], Roberta Spallone [2], Martina Rinascimento [3], Teresa Gil-Piqueras [4], Pablo Rodriguez-Navarro [5]

[1] Politecnico di Torino, Department of Architecture and Design, Turin, Italy - enrico.pupi@polito.it
[2] Politecnico di Torino, Department of Architecture and Design, Turin, Italy - roberta.spallone@polito.it
[3] Politecnico di Torino, Department of Architecture and Design, Turin, Italy - martina.rinascimento@polito.it
[4] Universitat Politècnica de València. Centro de Investigación PEGASO. Departamento de Expresión Gráfica Arquitectónica, Valencia, Spain - tgil@ega.upv.es
[5] Universitat Politècnica de València. Centro de Investigación PEGASO. Departamento de Expresión Gráfica Arquitectónica, Valencia, Spain - rodriguez@upv.es

Commission II

Keywords: Fortified Heritage, Multi-Source Data, Reality-Virtuality Continuum, eXtended Reality (XR), Digital Fabrication.

Abstract

The transition from Cultural Heritage (CH) digital documentation to dynamic, experiential heritage requires versatile workflows that engage diverse audiences. This paper proposes a scalable, multimodal eXtended Reality (XR) framework for the Torre del Mar (Castellón, Spain) and demonstrates how a single high-fidelity dataset can generate a comprehensive ecosystem of interactive experiences. Starting from a multi-source data fusion of UAV photogrammetry and Terrestrial Laser Scanning (TLS), a master model was developed to drive three distinct interaction pipelines. First, a phygital interface combines a modular, disassemblable 3D-printed replica with model-based Augmented Reality (AR), enabling tangible exploration of constructive details. Second, a collaborative Mixed Reality (MR) environment allows remote experts to co-inhabit the digital space for real-time analysis. Third, a multi-tiered Virtual Reality (VR) strategy optimizes the asset for PC-tethered, standalone, and WebVR platforms, balancing graphical fidelity with accessibility. The results validate a reproducible methodology that transforms technical survey data into an active knowledge system, effectively bridging the gap between scientific preservation and public dissemination while ensuring the long-term valorization of digital heritage assets.

1. Introduction

1.1 Theoretical Framework: The Shift towards Experiential Heritage

The digital documentation of Cultural Heritage (CH) involves managing heterogeneous datasets for knowledge and preservation. Dore and Murphy (2017) defined this systematic collection as essential for providing accurate information on an artefact's survival. Recently, the VIGIE 2020/654 Report (2022) highlighted key quality parameters in 3D digitization. Its predictions have largely materialized: AI algorithms combined with advanced connectivity (5G, cloud computing, IoT) have enhanced data storage and real-time usage. Furthermore, eXtended Reality (XR) is impacting the sector, shifting the focus from digital archiving to public engagement.

Khalil et al. (2020) concentrate on the modeling phase, tracing the evolution from the 1980s digital revolution through CAD to Building Information Modeling (BIM). They emphasize that modern cloud infrastructures and extensive dataset analysis enable the creation of Digital Twins by linking BIM, simulations, and XR.

The pandemic accelerated this shift. While early virtual tours offered passive participation, mobile apps proliferated to replace audio guides. However, spatialized AR models remain underutilized compared to standard multimedia apps, which—previously limited to institutions like the Van Gogh Museum in Amsterdam—are now widespread. The pandemic's impetus has led to new paradigms and a shift from digital documentation to dynamic engagement, one of the most significant outcomes in experiential heritage. This concept, drawing on anthropology, sociology, and cultural studies, is a dynamic process of interactions engaging with tangible and intangible cultural assets through direct, immersive, participatory, and sensory experiences. XR technologies, storytelling, and phygital interactions between physical replicas and digital devices are central media. While often conflated with experiential tourism and public engagement, experiential heritage also involves scholars and professionals who, from immersive and phygital visualization of point clouds, 3D models, and information models, extract information on materials, construction systems, decorative elements, and conservation conditions, add new data, and formulate hypotheses for reconstruction and spatial relationships.

Digital platforms are orienting toward this multilevel use. The Aïoli platform, primarily dedicated to CH scholars and professionals (Abergel et al., 2023), also identifies tourists and broader audiences as potential contributors to the knowledge and conservation of heritage artifacts (http://www.aioli.cloud).

1.2 Research Context: The TOVIVA_project Digital Archive

The TOVIVA project (*Watchtowers and Coastal Defense of the Valencian Littoral: Generation of Metadata and 3D Models for Interpretation and Effective Valorization*) produced the first systematic, metric documentation of the watchtower ensemble of the former Kingdom of Valencia, generating a digital archive designed as a reusable resource for advanced research, including XR systems. Addressing previous gaps in graphic documentation and historiographical errors caused by fragmentary studies, TOVIVA adopted an integrated methodology combining historical analysis with Terrestrial Laser Scanning (TLS) and Structure-from-Motion (SfM) photogrammetry (Rodríguez-Navarro et al., 2015).

The resulting digital archive comprises point clouds, polygonal models, and metadata, documenting towers both as individual artifacts and as elements of a territorial system (https://toviva.blogs.upv.es/descripcion-2/). This multiscalar approach facilitates the analysis of defensive logic and intervisibility, shifting from monument-centered documentation to landscape-oriented analysis. Dissemination relies on an open-access platform and online repositories (https://skfb.ly/oJYrn) featuring optimized models for mobile visualization.

Although TOVIVA predates widespread adoption of immersive technology, its commitment to reusable 3D assets, structured metadata, and scalable formats laid the foundation for subsequent developments.

In this contribution, the TOVIVA digital archive serves as the primary framework for developing a multimodal XR approach for the Torre del Mar of Borriana (Castellón, Spain). In this way, the legacy of TOVIVA continues to advance digital heritage research, aligning with the objectives and methodologies promoted within the 3D-ARCH domain.

1.3 Aim and Objectives: A Scalable Multimodal Framework

While recent scholarship has rigorously examined specific XR modalities, ranging from the immersive efficacy of Virtual Reality (Banfi et al., 2025) to the collaborative potential of Mixed Reality (Chrysanthakopoulou et al., 2025), this research posits a holistic framework that integrates these distinct approaches into a connected ecosystem. By leveraging a high-fidelity master model obtained through multi-source data fusion, the proposed methodology orchestrates a pipeline that enables the adaptation of a single digital asset into three complementary interaction strategies (table 1).

Medium	Audience	Interaction	HW	Feature
Phygital AR	General public	Tangible manipulation and interactive visualization	Smartphone, tablet, physical model	Object tracking and disassembly
Collaborative MR	Heritage experts	Co-presence, sketching, and annotation	HMD (Standalone/ Tethered)	Rapid iteration and remote analysis
Custom MR	Researchers and stakeholders	Tailored data overlay and specific UI	HMD (Standalone)	High-precision tracking and custom interaction
PCVR	Researchers and museums	High-fidelity immersion	HMD and PC	Highest visual quality
Standalone VR	Broad audience	Designed active exploration	HMD	6 DoF hand interaction
Web VR	Remote users	Browser navigation	Smartphone, tablet, PC	Wider accessibility

Table 1. Synoptic overview of the multimodal XR framework. (Editing: E. Pupi)

The first involves engineering a phygital interface in which AR is spatially registered with a modular, disassemblable physical replica, thereby bridging the dichotomy between digital visualization and tactile exploration. A dual-track MR approach complements this: a collaborative environment for rapid remote prototyping and inspection, and a custom-developed application for tailored interactions, enabling specialists to co-inhabit the virtual space for real-time assessment. Concurrently, the framework deploys a multi-tiered VR strategy that adapts to computational constraints, offering high-fidelity PCVR, optimized standalone app-based experiences via HMD, and WebVR for wider accessibility. Beyond these applications, the architecture is designed to accommodate the future integration of Large Language Models (LLMs), enabling static environments to evolve into responsive knowledge systems capable of sustaining historically informed, conversational narratives (Sánchez-Berriel et al., 2025).

2. State of the Art

Current developments in CH indicate a trend shifting from static metric acquisition toward semantic activation, contributing to the creation of dynamic digital ecosystems. The integration of digital modelling and XR is increasingly explored as a means to redefine relational dynamics between artifacts and users, aiming to foster active cognitive processing through multimodal interaction.

2.1 Heritage Digital Documentation

The complexity of historical architecture often necessitates integrated documentation protocols. Research suggests that multi-source data fusion is particularly effective for vertical defensive structures (Baik et al., 2025) and for analyzing geometric changes over decades (Klapa et al., 2025). This approach is also being applied to landscape-scale conservation, where air-ground coordination aims to bridge the gap between settlement patterns and constructive details (Tian et al., 2025).

To address the challenges of massive datasets, recent studies propose automating Scan-to-HBIM workflows. Classification-driven, semi-automated approaches are being tested to streamline the transition from raw scans to semantic models (Salah et al., 2026), while Deep Learning architectures are increasingly being investigated for semantic segmentation of heritage point clouds (Pellis et al., 2025). Furthermore, the potential of Heritage Digital Twins is being expanded through the integration of Knowledge Graphs, which link geometric models with ontological data to support conservation strategies (Hosamo & Mazzetto, 2025).

2.2 XR in Cultural Heritage

XR technologies are being progressively orchestrated within the Reality-Virtuality Continuum, a framework revisited to address the complexity of modern virtual environments (Skarbez et al., 2021). In this context, VR tools are examined for their ability to balance graphical fidelity with accessibility, particularly via WebVR platforms (Banfi et al., 2025). A significant area of experimentation involves Artificial Intelligence, where Large Language Models (LLMs) are proposed to enable virtual agents to sustain historically informed dialogues (Sánchez-Berriel et al., 2025). Concurrently, MR is being developed to facilitate operative telepresence, enabling distributed experts to co-inhabit digital models for collaborative assessment (Chrysanthakopoulou et al., 2025).

2.3 Phygital Interactions: Converging 3D Printing and XR

Phygital approaches aim to bridge the gap between virtuality and materiality by converging digital fabrication with XR. Current research indicates that re-materializing digital data through 3D printing can create Tangible User Interfaces (TUIs) that leverage physical affordance to enhance spatial understanding (Nofal et al., 2025). These methods are also relevant for inclusive design; novel interactive multisensory systems that combine tactile exploration with digital feedback are being validated as tools to broaden accessibility for diverse audiences (Muñoz et al., 2025). Consequently, documentation, semantic modeling, and phygital materialization are increasingly viewed as components of an integrated knowledge system.

3. Case Study and Data Source: The Torre del Mar

3.1. Historical and Territorial Context

The watchtowers of the Valencian littoral emerged during a period of geopolitical instability and rapid military engineering development. While defensive efforts began in the fourteenth century, a strategic shift from maritime to land-based systems occurred under Charles I and Philip II. The military engineer Giovanni Battista Antonelli designed this network, establishing a coherent territorial infrastructure based on visibility and strategic positioning across more than 500 km of coastline.

The Torre del Mar of Borriana (39°52′42″ N, 0°03′13″ W), built between 1553 and 1558, constitutes a paradigmatic case study. Located near a natural landing point and a freshwater source—factors decisive for corsair activity (Melchor Monserrat, 2015)—it ensured maritime surveillance while maintaining visual contact with the tower dels Millars or Almassora to the north and the tower of Moncofa or Biniesma to the south.

Historical research by the Municipal Archaeological Museum of Borriana has solidified the interpretative framework regarding the building's origin and evolution (Melchor Monserrat, 2013; 2015), complemented by typological analyses of coastal fortifications (Celda Cerdán, 2012). Within the TOVIVA project, the tower served as a focal point for digital documentation research (Rodríguez-Navarro et al., 2015) and for early virtualization initiatives (Melchor Monserrat et al., 2016). The present study builds upon this consolidated knowledge to propose an advanced workflow for XR applications.

3.2. Architectural Typology and State of Conservation

The Torre del Mar exemplifies the prismatic square-plan typology (6.30 m per side) with vertical walls devoid of a battered base (*alambor*), characteristic of the transitional period between the late fifteenth and early sixteenth centuries. Its robust geometry, with wall thicknesses ranging from 1.30 to 1.40 m, was designed for structural stability and defense. The tower reaches a height of 10.18 m to ensure long-distance visual control, featuring limited openings consistent with its defensive function.

Internally, the structure is organized into two levels and a flat roof terrace, connected by wooden staircases, initially removable to isolate floors during attacks. The ground floor houses mangers, while the first floor is an open space; a machicolation at the top defends the entrance below.

Following several interventions, the tower is in an excellent state of conservation, retaining high material integrity and legible masonry. Although urban development and apartment blocks partially obscure its seaward visibility, the immediate environment is well-maintained. Despite anthropogenic pressure, the Torre del Mar remains a fundamental landscape landmark for interpreting Borriana's historical defensive system.

3.3. Digital Survey, Data Acquisition, and Processing

3.3.1. Integrated Survey Methodology through UAV Photogrammetry and TLS: The graphic survey aimed to create a high-fidelity digital model by integrating aerial photogrammetry for the exterior envelope and Terrestrial Laser Scanning (TLS) for the interior spaces. This methodology combined active and passive sensors to maximize geometric accuracy. Given the building's height and limited accessibility, UAV photogrammetry was selected to document exterior textures (Rodríguez-Navarro et al., 2016), while TLS captured the overall geometric configuration.

For the exterior, a DJI Mavic 3E drone equipped with a 20-megapixel 4/3-inch CMOS sensor acquired images suitable for both reconstruction and photorealistic texturing. A total of 148 photographs were captured at a resolution of 5280 × 3956 pixels. Flight parameters were optimized for sharpness using a 24 mm equivalent focal length and a fast shutter speed (1/2000 s) to minimize motion blur and ensure homogeneous exposure. The drone’s integrated RTK GPS facilitated automatic image geolocation, streamlining orientation and scaling during processing.

In parallel, a high-resolution survey was performed using a Leica RTC360 Time-of-Flight (ToF) scanner. The strategy involved 26 scans distributed to guarantee optimal point cloud overlap and complete coverage while minimizing shadow zones. A variable-resolution approach was adopted, with low/medium point densities for the interior and medium/high point densities for the exterior (Figure 1).

Figure 1. UAV photogrammetry (left) and terrestrial laser scanning (TLS) data (right). (Editing: E. Pupi)

3.3.2. Data Processing, Optimization, and Model Integration: The transformation of raw data followed a reproducible workflow organized into two parallel pipelines. For photogrammetric reconstruction, 136 of 148 captured images were processed using Agisoft Metashape (Gil-Piqueras et al., 2019). SfM algorithms generated a dense point cloud of approximately 47 million points, cleaned to 45 million. From this, an initial mesh of 15 million polygons was generated and subsequently decimated to 2.5 million to ensure computational sustainability while preserving detail. Photorealistic texturing was achieved through UV mapping and image projection.

Figure 2. The high-fidelity master model of Torre del Mar. The textured polygonal mesh integrates the exterior envelope and interior spaces, serving the multimodal XR ecosystem. (Modelling and editing: E. Pupi)

The TLS dataset processing yielded a dense point cloud of 156 million points (2.8 GB), providing a precise geometric characterization of the interior. The acquisition of overlapping exterior surfaces proved essential for alignment, offering robust shared references (Grussenmeyer et al., 2012). The original TLS cloud (0.001 m resolution) was imported into Metashape, filtered to 115 million points, and triangulated into a high-resolution mesh of 178 million polygons. A further decimation reduced this to 5 million polygons, maintaining metric fidelity, with texturing derived from per-vertex RGB values.

The outcome is a complete, textured, and metrically reliable digital model of the Torre del Mar, integrating UAV-based exterior photogrammetry with TLS interior documentation (Figure 2). This multipurpose asset serves as the primary data source for the XR framework, balancing high geometric detail with optimizations suitable for real-time immersive and hybrid experiences.

4. Methodology: Multimodal Ecosystem Development

The methodological workflow moves beyond linear digitization to establish a generative matrix framework. The high-fidelity digital survey serves as the foundational node for a branching ecosystem of derived assets, each optimized for specific hardware constraints and interaction modalities. The development process focuses on tangible interface engineering for phygital interaction and XR pipeline deployment, ensuring semantic and metric integrity preservation across different digital environments.

4.1 Geometric and Topological Integration

Following spatial alignment of exterior and interior meshes in McNeel Rhinoceros (v. 8.3), the workflow addressed topological discontinuity at transition zones (arrowslits and rooftop hatch). Since physical inaccessibility prevented complete data capture within deep wall penetrations, a reconstruction protocol bridged the shells without occluding voids. Boundary traces of each opening were isolated and processed in Geomagic Wrap (v. 2017.0.0). Instead of standard hole-filling algorithms that seal apertures, an adaptive mesh-generation strategy extrapolated missing geometry based on the curvature and inclination angles of existing mesh traces, ensuring a coherent, watertight topology essential for the manufacturability of physical replicas.

Internal vertical connections required distinct modeling because their geometry was removed during initial TLS point-cloud filtering. To reconstruct these elements with metric fidelity, the original uncleaned dense TLS point cloud was utilized. A subsampled dataset (0.01 m resolution, 5 million points) was imported into Rhinoceros as a spatial reference, enabling precise NURBS-based stair profile reconstruction, subsequently converted to meshes and topologically integrated into the master model.

4.2 Tangible Interface Engineering: The Modular Replica

The translation of the master model into a physical scaled model required an engineering process to ensure both assemblability and digital interoperability. A 1:20 representation scale was selected (approximately 50 cm in height) to retain the high-frequency morphological details captured during the survey. To facilitate the haptic investigation of the tower's constructive logic, the geometry was discretized into twelve modular components using a system of four cutting planes. Two split vertical planes were positioned asymmetrically to reveal internal features, while two horizontal planes were introduced approximately 1 meter above the ground and at the first-floor level to explicate the planimetric layout (figure 3).

Figure 3. The 1:20-scale, 3D-printed, assembled modular replica (left) and a sectioned view (right). (Prototyping and photo: E. Pupi)

The splitting process, executed in McNeel Rhinoceros (v. 8.3), also required preserving the original UV mapping coordinates on the mesh surfaces, a prerequisite for the subsequent 3D object tracking in the AR application. Following segmentation, the mesh volumes were capped using NURBS geometry, subsequently converted to a mesh, and cylindrical housings (Ø 8 mm, thickness 3 mm) were modeled to later accommodate neodymium magnets. This solution replaced traditional interlocking joints, ensuring robust reversibility for repeated assembly cycles. At this stage, the workflow bifurcated: a textured version of the segmented model was retained for the XR pipeline, while the geometry was exported as STL files for manufacturing. Prior to slicing, a topological validation was performed using Autodesk Netfabb Premium (2024) to detect and repair non-manifold edges, overlapping polygons, and extra shells, which are common in complex morphological models. The additive manufacturing process was executed using a cluster of FDM printers (2 Bambu Lab P1S and 1 Bambu Lab H2D). Over approximately 210 hours of cumulative printing, the process consumed 8 kg of PolyMaker Cotton White PLA. This material was selected for its matte surface finish, which minimizes the visual perception of the used 0.2 mm layer height. The components were printed with a 10% infill and 2 perimeters; generated supports were manually removed during post-processing.

4.1 XR Pipeline Implementation

The transition to dynamic environments is orchestrated through a unified development pipeline centered on Unity 6.3 LTS (6000.3.0f1). This engine serves as the integration hub where the optimized 3D geometry, textures, and metadata are synthesized with interaction logic. To ensure scalability across diverse hardware specifications, the project utilizes the Universal Render Pipeline (URP). The pipeline is thus structured to deploy three distinct modalities along the Reality-Virtuality Continuum, each leveraging specific Software Development Kits (SDKs) to address targeted user scenarios.

4.1.1 Model-Based Tracking for Tangible Interaction: The AR implementation of the phygital interface adopts a model-based tracking paradigm, leveraging the specific 3D morphology of the physical replica as a spatial reference anchor. This methodological choice aligns with the definition of Tangible User Interfaces (TUIs) in the architectural domain, where the manipulation of a physical object serves as an intuitive medium for defining relationships between the real and virtual worlds, bridging the gap between static representation and active cognitive engagement (Russo, 2021). By aligning the digital augmentation with the tower's geometry, the system transforms the replica into a scientific mediation tool rather than a mere aesthetic artifact (Balletti & Ballarin, 2019).

To achieve this, the computational workflow in Unity is engineered around a dual-asset architecture that decouples the tracking logic from the visualization layer. The first asset, designated the clay model, is a watertight, geometry-only mesh derived from the master model and optimized to instruct the Vuforia Model Target Generator (MTG). In the target configuration phase, the Advanced Model Target mode was explicitly selected over standard static Guide Views. This mode utilizes deep learning training to enable omnidirectional (360°) recognition, thereby eliminating the need for the user to align the physical object with a pre-set on-screen silhouette. Furthermore, the Motion Hint parameter was set to adaptive, an important optimization for phygital scenarios in which the object is handled by the user rather than remaining static, thereby ensuring tracking stability during manipulation. This approach mirrors recent advancements in heritage visualization where app-based AR leverages specific physical features to anchor interactive digital content, enhancing accessibility and user engagement (Spallone et al., 2025).

Conversely, the visual layer uses a textured overlay model. Precise superimposition is ensured by the clay model acting as an occlusion manager via a custom depth-mask shader that writes exclusively to the Z-buffer. This effectively handles physical occlusions (e.g., the user's hands) and prevents back-face visualization. Finally, photometric consistency is established via ARFoundation, utilizing mobile sensors to match virtual lighting with the real-world environment. Although the preliminary validation focused on the assembled tower, the pipeline is architected to support multi-target tracking, enabling the system to recognize and texturize the 12 components during disassembly independently.

4.1.2 Multi-tiered VR Environments: The VR ecosystem is articulated into three distinct branches, each designed to address specific trade-offs between graphical fidelity, interaction depth, and accessibility (Banfi et al., 2025). All immersive tiers were validated using the Meta Quest 3 headset to ensure hardware consistency. The first tier focuses on High-Fidelity PCVR, utilized primarily for rapid prototyping and visual validation. This environment was orchestrated using Twinmotion 2025.1.1, leveraging the host computer's GPU to render the master model with real-time global illumination and complex material shaders. This modality allows for immediate, immersive inspection of the architectural digital model without extensive data preprocessing, though interaction is limited to navigation (figure 4).

Figure 4. High-fidelity PCVR experience: immersive visualization of the tower in Twinmotion, featuring real-time global illumination and high-resolution textures. (Prototyping: E. Pupi)

The second tier involves developing a custom Unity application. Unlike the rapid prototyping environment, this tier enables programming advanced interaction logic, semantic enrichment, and informational hotspots that exceed the capabilities of standard visualization software (figure 5). While retaining a high level of geometric detail supported by the headset's Snapdragon XR2 Gen 2 processor, this environment prioritizes the user's active engagement with the heritage asset. To mitigate Visually Induced Motion Sickness (VIMS), a teleportation-based locomotion system was implemented to ensure user comfort during prolonged sessions (Prithul et al., 2021).

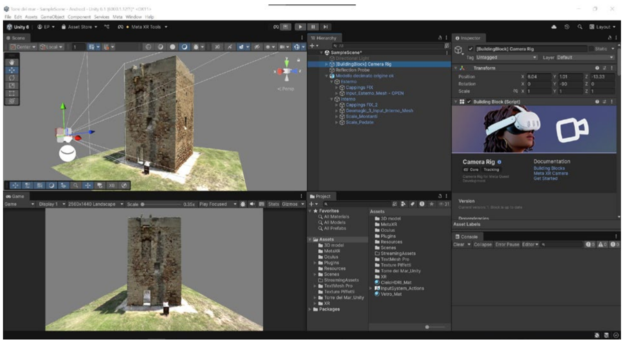

Figure 5. Standalone VR application development: implementation of the custom interaction logic within the Unity engine for the Meta Quest 3 headset. (Prototyping: E. Pupi)

The third tier addresses wider dissemination through WebVR via the Sketchfab platform. This modality imposes a 100 MB storage limit per file on the free plan. Consequently, the master model underwent a targeted optimization pipeline: while the polygonal mesh was further decimated to reduce the vertex count, the high-resolution texture maps were preserved without compression. This choice explicitly prioritizes chromatic and material fidelity over geometric-morphological accuracy, ensuring that the surfaces' perceptual quality remains intact even within a topologically simplified model. Unlike app-based solutions, navigation in this tier is constrained by native platform features.

4.1.3 Collaborative MR for Remote Analysis: The MR module addresses the operative telepresence in heritage science, facilitating a synchronous collaborative environment where interdisciplinary experts can co-inhabit the digital asset. This approach aligns with recent research demonstrating that integrating collaborative and multi-modal interaction methods significantly enhances engagement and cultural learning in virtual heritage contexts (Bekele et al., 2021). To rigorously evaluate the trade-offs between deployment agility and analytical depth, the research juxtaposed two distinct methodological variants, both deployed on Meta Quest 3 headsets to leverage high-fidelity color passthrough and depth sensing.

The first variant utilizes Gravity Sketch, selected not merely as a viewer but for its native identity as an immersive 3D modeling suite. The workflow exploits the platform's cloud-native backend (LandingPad) for rapid asset ingestion, bypassing local compilation. Beyond standard annotation, this environment provides a robust geometric toolkit that enables experts to generate reconstruction hypotheses directly in 3D. A defining technical feature is the fluid manipulation of scale: users can seamlessly transition from a phenomenological 1:1 inspection to a macro-overview, effectively transforming the model into a malleable boundary object for spatial brainstorming (figure 6).

The second variant, a custom application developed in Unity, was engineered to prioritize metric rigor and semantic integration. Networked via Photon Fusion to ensure low-latency synchronization and handle multi-user networking, this application integrates the Meta OpenXR Building Blocks. By utilizing Spatial Anchors, the master model is locked to the user's physical desk, reducing cognitive dissonance. Unlike the commercial alternative, this environment features tailored analytical tools, including dynamic boolean sectioning and semantic hotspots linked to the project's digital archive. This setup allows experts to retrieve historical metadata and technical reports directly within the immersive view, facilitating a comprehensive remote diagnosis supported by multi-modal interaction (Bekele et al., 2021).

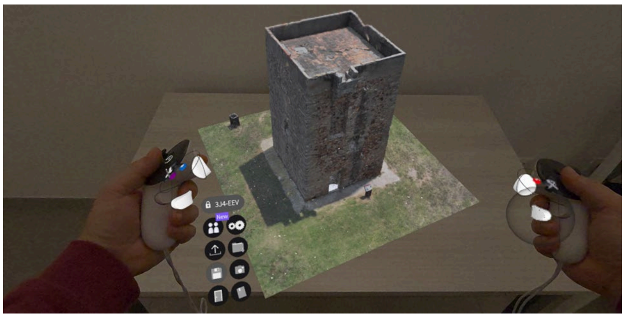

Figure 6. The MR environment using Gravity Sketch enables multi-scalar inspection and 3D annotation directly on the model. (Prototyping: E. Pupi)

4.2 Computational Constraints and Optimization Strategies

While the master model served as a robust foundation, its deployment across the full multimodal spectrum necessitated a secondary optimization phase. The hardware limitations of standalone AR (Unity) on smartphones and tablets, WebVR, and collaborative MR via Gravity Sketch imposed strict limits on polygon counts and texture memory. Consequently, a derived runtime asset was generated through targeted decimation to minimize draw calls without sacrificing chromatic accuracy. This iterative process underscored the critical trade-off between geometric density and frame-rate stability. Although validated on the Torre del Mar pilot, these computational results indicate that a recursive optimization protocol is essential for the broader TOVIVA corpus and requires further tuning to ensure seamless interoperability on constrained hardware while preventing real-time rendering bottlenecks.

5. UX and UI Design, Accessibility, and Web Integration

To ensure the effective collection, enhancement, and public dissemination of the project's outputs, this study presents a web integration strategy and the redesign of the TOVIVA platform, specifically targeting the digital archive of defensive towers and the primary case study, Torre del Mar.

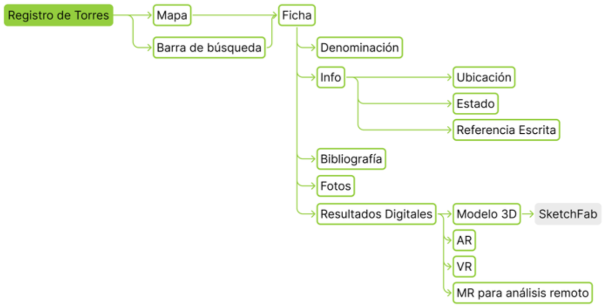

Figure 7. Information Architecture detail of the redesigned webpage, Registro de Torres. (Editing: M. Rinascimento)

A heuristic evaluation based on Nielsen Norman Group's (2012) usability components (learnability, efficiency, memorability, errors, satisfaction) identified friction points, guiding the optimization of human-machine interaction for both specialists and new users.

At the end of the analysis, it was clear what the re-design of the selected webpages should have addressed to achieve usability, memorability, and user satisfaction: flexibility and efficiency of use of the system, findability, visual hierarchy, and consistency.

Consequently, a new Information Architecture was developed to hierarchically cluster content, organizing it for maximum visibility and guiding a clearer user journey (figure 7).

The prototyping phase resulted in a renovated 'Registro de Torres' webpage. The focal point of this interface is an interactive map of the Valencia coastline (figure 8). This map uses dot-like graphics to indicate data density; larger markers indicate a higher concentration of towers at a specific location. This visualization allows users to grasp the richness of the region's fortified heritage instantly. For targeted inquiries, a dual search bar facilitates direct access to specific locations or towers. Interacting with these map markers reveals summary labels that lead users to the dedicated page for the selected asset, such as Torre del Mar. The detail page is structured to balance textual data, such as denomination, conservation state, and bibliographic references, with visual evidence. The layout features a scrollable gallery with different sections showcasing the tower's architecture through recent photography, historical drawings, and survey illustrations, and narrating the process that led to the creation of the tower's 3D model, accessible on the platform SketchFab, and its physical scaled model.

Additionally, the platform conveys the implementation of the tower 3D and physical models for immersive technologies applications, integrating eXtended Reality (XR) experiences designed for both scientific and non-expert audiences. The interface displays frames from the developed AR, VR, and MR experiences, accompanied by explanatory text detailing the technical framework and the evolution of the digital reconstruction (figure 9).

Este mapa recoge las **66 torres defensivas** del siglo XVI-XVII investigadas y digitalizadas en el proyecto

cuanto más grandes son los elementos visuales, más torres hay en la localidad indicada

Figure 8. Registro de Torres webpage prototype. (Prototyping and editing: M. Rinascimento)

Borriana (Castellón)

Torre de la Mar

denominaciónes

Torre de BORRIANA
Antonelli, J. B. (1563) y Manrique Gonzaga, V. (1673)

Torre de BURRIANA
Acuña, J. (1585) y Escolano, G. (1610)

Torre de la MAR (actual)

estado

Restaurada

1ª referencia escrita

S. XVI. Informe de J.B. Antonelli (1563)

XR implementación

A partir del modelo virtual creado, se ha desarrollado un sistema de RA basado en una aplicación para que funcione junto con el modelo físico, en el que el algoritmo de seguimiento de objetos de la aplicación reconoce la forma de cada configuración y le superpone las texturas de alta resolución correspondientes, además activando puntos de información contextuales, creando una experiencia rica e interactiva para dispositivos como móbiles y tabletas

Tecnología Realidad Aumentada RA

Figure 9. Torre del Mar webpage prototype. (Prototyping and editing: M. Rinascimento)

6. System Assessment and Preliminary Validation

6.1 Performance Metrics

The assessment evaluated the framework's computational sustainability, determining where the high-density master model could be deployed versus where an optimized derivative was necessary.

In AR, deployment on mobile devices required an optimized derivative. The computational overhead of Vuforia's tracking algorithms, combined with rendering, made the master model unsuitable. The lighter asset balanced device resources, ensuring the fluid refresh rate essential for stable registration.

In VR, results diverged. Both PCVR and the standalone application successfully handled the master model, confirming that modern chipsets with efficient draw-call management can support high geometric density without inducing VIMS. Conversely, WebVR necessitated the optimized derivative, ensuring accessibility via standard browsers.

MR environments showed a similar dichotomy. The custom application rendered the master model with stable latency despite the passthrough overhead. In contrast, Gravity Sketch required the optimized derivative due to its cloud-synchronization logic.

6.2 Web Summative Evaluations and Future Developments

This design proposal serves as a foundational prototype open to further evolution, particularly regarding User Experience (UX). Future iterations will focus on evaluating the emotional connections and semantic values users derive from the interface. To enhance engagement, an interactive timeline could be added to complement the geospatial map. This feature would allow users to visualize the historical expansion of the Valencian defensive network in time, bridging scientific rigor with public dissemination.

Designed for scalability, the current page layouts function as modular templates. This ensures that as the digital survey expands, new data on other towers can be seamlessly integrated. Finally, a continuous summative evaluation is recommended to validate the platform against user expectations and Peter Morville's (2004) UX facets, ensuring the system remains intuitive and educationally effective.

7. Discussion and Future Outlook

7.1 Scalability and Long-term Preservation Strategy

The modular architecture of the proposed framework demonstrates high scalability, establishing the master model as a systematically applicable generative matrix across the entire TOVIVA defensive network. This approach enhances long-term preservation strategies by shifting the focus towards dynamic multimodal dissemination. Valorizing the heritage asset across complementary technological tiers, the framework ensures that the historical and morphological values of the Torre del Mar permeate a diverse audience. Consequently, the digital ecosystem acts as an active agent of cultural resilience: it guarantees the persistence of historical memory through widespread accessibility, fostering a collective consciousness that constitutes the ultimate prerequisite for the physical safeguarding of built heritage.

7.2 Towards AI-Driven Heritage Narratives

A pivotal trajectory involves integrating LLMs and Retrieval-Augmented Generation (RAG) to ground conversational agents within the TOVIVA scientific corpus (Sánchez-Berriel et al., 2025), transforming the virtual assistant into a dynamic comparative tool. Furthermore, adopting Agentic AI paradigms (Pavlidis, 2025) could evolve the system into a cognitive twin. By embedding semantically structured risk memory, the digital model shifts from a passive repository to an active entity capable of reasoning about the asset's temporal trajectory and acting as a collaborator in anticipatory preservation.

8. Conclusions

This research validates a multimodal framework transcending the specific case of Torre del Mar, offering a replicable protocol applicable to the entire coastal defensive network documented by the TOVIVA project. A defining outcome is demonstrating the value of legacy data: high-fidelity metric surveys acquired over a decade ago have been successfully reactivated, proving that rigorous digital documentation constitutes a future-proof investment capable of supporting visualization technologies that were nascent or non-existent at the time of acquisition. By leveraging the increasing ubiquity of consumer-grade hardware, the framework democratizes access to architectural information, effectively translating technical data into experiential heritage. The integration of physical fabrication, immersive environments, and web accessibility establishes a scalable paradigm in which the digital asset serves as a dynamic interface, ensuring the historical memory of the Mediterranean defensive system remains a shared, living, and queryable resource for future generations.

Acknowledgements

The research on Valencian towers is directed by Pablo Rodriguez and Teresa Gil, as part of the TOVIVA project. This paper was written by E. Pupi (par. 2, 4, 6.1 and 7.2), R. Spallone (par. 1.1), M. Rinascimento (par. 5 and 6.2), and Pablo Rodriguez and Teresa Gil (par. 1.2 and 3). The authors wrote paragraphs 1.3, 7.1, and 8 together.

References

Abergel, V., Manuel, A., Pamart, A., Cao, I., De Luca, L. 2023: Aïoli: A reality-based 3D annotation cloud platform for the collaborative documentation of cultural heritage artefacts. *Digital Applications in Archaeology and Cultural Heritage*, 30, 2023, e00285. https://doi.org/10.1016/j.daach.2023.e00285

Aïoli, http://www.aioli.cloud, accessed 06/01/2026.

Baik, A., Al-marwai, M., Kateb, A., Al-Harthi, M., 2025. Digital documentation of Saudi built heritage: Al-Arfa' Castle. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-M-9-2025, 87–94. https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-87-2025

Balletti, C., Ballarin, M., 2019. An application of integrated 3D technologies for replicas in cultural heritage. *ISPRS Int. J. Geo-Inf.*, 8(6), 285. https://doi.org/10.3390/ijgi8060285

Banfi, F., Cazzani, A., Liu, W., 2025. XR technologies for enhanced cultural engagement: From HBIM to a comparative analysis of VR and WebVR development tools. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-M-9-2025, 95–102. https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-95-2025

Bekele, M.K., Champion, E., McMeekin, D.A., Rahaman, H., 2021. The influence of collaborative and multi-modal mixed reality: Cultural learning in virtual heritage. *Multimodal Technol. Interact.*, 5(12), 79 https://doi.org/10.3390/mti5120079

Celda Cerdán, V. (2012). Torres de vigía y defensa del litoral valenciano (III): de la Torre de Burriana a la Torre de El Puig. Retrieved from https://riunet.upv.es/handle/10251/17562

Chrysanthakopoulou, A., Chrysikopoulos, T., Arvanitis, G., Moustakas, K., 2025. Reimagining historical exploration: Multi-user mixed reality systems for cultural heritage sites. *Applied Sciences*, 15(5), 2854. https://doi.org/10.3390/app15052854

Dore, C., Murphy, M., 2017. Current state of the art Historic Building Information Modelling. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W5, 185–192. https://doi.org/10.5194/isprs-archives-XLII-2-W5-185-2017

Gil Piqueras, T., Pérez-Vila, A. & Rodríguez Navarro, P. (2019) Fotogrametría multi-image mediante SFM. Revisión del software disponible. In: Lloréns Corraliza, S., Rincón Millán, M. D. & Martín Pastor, A. (eds.) *Avances en Expresión Gráfica Aplicada a la Edificación*. Valencia, Tirant Humanidades, pp. 645–659.

Grussenmeyer, P., Alby, E., Landes, T., Koehl, M., Guillemin, S., Hullo, J.F., Assali, P., Smigiel, E., 2012. Recording approach of heritage sites based on merging point clouds from high resolution photogrammetry and terrestrial laser scanning. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIX-B5, 553–558. https://doi.org/10.5194/isprsarchives-XXXIX-B5-553-2012

Hosamo, H., Mazzetto, S., 2025. Integrating knowledge graphs and digital twins for heritage building conservation. *Buildings*, 15(1), 16. https://doi.org/10.3390/buildings15010016

Khalil, A., Stravoravdis, S., Backes, D., 2021. Categorisation of building data in the digital documentation of heritage buildings. *Applied Geomatics,* 13, 29–54. https://doi.org/10.1007/s12518-020-00322-7

Klapa, P., Żygadło, A., Pepe, M., 2025. 3D heritage reconstruction through HBIM and multi-source data fusion: Geometric change analysis across decades. *Applied Sciences*, 15(16), 8929. https://doi.org/10.3390/app15168929

Melchor Monserrat, J.M., 2013. Diez años del Servicio Municipal de Arqueología de Burriana (2003-2013). *Magnífico Ayuntamiento de Burriana*, Burriana.

Melchor Monserrat, J. M. (2015). Datos históricos sobre la Torre del Mar (Burriana - Castellón). In: Rodríguez-Navarro, P. (Ed.), *Defensive Architecture of the Mediterranean. XV to XVIII centuries*. Vol I. Editorial Universitat Politècnica de València, pp. 113–116. http://dx.doi.org/10.4995/FORTMED2015.2015.1678

Melchor Monserrat, J. M., Martínez, J., Bonafe, C., Cabrera, A. (2016). La virtualización en el Museo Arqueológico de Burriana (Castellón – España). *Proceedings of the 8th International Congress on Archaeology, Computer Graphics, Cultural Heritage and Innovation 'ARQUEOLÓGICA 2.0'*, pp. 78–83. https://doi.org/10.4995/ARQUEOLOGICA8.2016.3036

Morville, P. (2004). *User experience design*. Semantic Studios. http://semanticstudios.com/user_experience_design/, accessed 14/01/2026.

Muñoz, A., Climent-Ferrer, J.J., Martí-Testón, A., Solanes, J.E., Gracia, L., 2025. Enhancing cultural heritage engagement with novel interactive extended-reality multisensory system. *Electronics*, 14(10), 2039. https://doi.org/10.3390/electronics14102039

Nielsen, J. (2012, January 3). *Usability 101: Introduction to Usability*. Nielsen Norman Group. https://www.nngroup.com/articles/usability-101-introduction-to-usability/, accessed 14/01/2026.

Nofal, E., De Marco, R., Voinea, A.E., Rejeb, A., 2025. Developing digital skills for architectural heritage: A didactic approach to 3D mapping and phygital storytelling. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-M-9-2025, 1089–1095. https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-1089-2025

Pavlidis, G., 2025. Agentic AI for cultural heritage: Embedding risk memory in semantic digital twins. *Computers*, 14(7), 266. https://doi.org/10.3390/computers14070266

Pellis, E., Masiero, A., Betti, M., Tucci, G., Grussenmeyer, P., 2025. A deep learning multiview approach for the semantic segmentation of heritage building point clouds. *Int. J. Archit. Herit.*, 19(12), 3117–3139. https://doi.org/10.1080/15583058.2025.2485242

Prithul, A., Adhanom, I.B., Folmer, E., 2021. Teleportation in virtual reality; A mini-review. *Front. Virtual Real.*, 2, 730792. https://doi.org/10.3389/frvir.2021.730792

Rodríguez-Navarro, P., Gil Piqueras, T., Verdiani, G. (2016). *Drones para el levantamiento arquitectónico. Aplicación para la documentación de las torres del litoral valenciano*, en: El arquitecto, de la tradición al siglo XXI: docencia e investigación en expresión gráfica arquitectónica: 16 Congreso Internacional de Expresión Gráfica Arquitectónica. Echeverría Valiente, E. y Castaño Perea, E., eds. Alcalá de Henares: Fundación General de la Universidad de Alcalá, 2016. ISBN 978-84-88754-39-4. Vol. 2

Rodríguez-Navarro, P., Verdiani, G., Gil Piqueras, T. (2015). Comprehensive methodology for documenting the defense towers of the Valencian Coast (Spain). In: Rodríguez-Navarro, P. (Ed.), *Defensive Architecture of the Mediterranean. XV to XVIII centuries.* Vol I. Editorial Universitat Politècnica de València, pp. 321-328. http://dx.doi.org/10.4995/FORTMED2015.2015.1741

Russo, M., 2021. AR in the architecture domain: State of the art. *Applied Sciences*, 11(15), 6800. https://doi.org/10.3390/app11156800

Salah, R., Géczy, N., Ajtayné Károlyfi, K., 2026. Architectural heritage digitization: A classification-driven semi-automated scan-to-HBIM workflow. *Buildings*, 16(1), 21. https://doi.org/10.3390/buildings16010021

Sánchez-Berriel, I., Pérez-Nava, F., Pérez-Rosario, L. (2025). Natural Interaction in Virtual Heritage: Enhancing User Experience with Large Language Models. *Electronics*, 14(12), 2478. https://doi.org/10.3390/electronics14122478

Skarbez, R., Smith, M., Whitton, M.C., 2021. Revisiting Milgram and Kishino's reality-virtuality continuum. *Front. Virtual Real.*, 2, 647997. https://doi.org/10.3389/frvir.2021.647997

Spallone, R., Teolato, C., Russo, M., Vitali, M., Palma, V., Pupi, E., Rinascimento, M., 2025. Visualising Piffetti's library in Villa Della Regina Museum: An interdisciplinary digital project for knowledge accessibility. In: Bienert, A., Emenlauer-Blömers, E., Lengyel, D. (Eds.), *EVA Berlin 2025. Electronic Media and Visual Arts*, arthistoricum.net, Heidelberg, 143–152. https://doi.org/10.11588/arthistoricum.1568.c24085

Tian, M., Zheng, C., Lou, N., 2025. Research on the integrated application of multi-scale 3D digital technologies in village heritage conservation. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-M-9-2025, 1499–1506. https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-1499-2025

VIGIE 2020/654. 2022. Study on quality in 3D digitisation of tangible cultural heritage: mapping parameters, formats, standards, benchmarks, methodologies, and guidelines. https://digital-strategy.ec.europa.eu/en/library/study-quality-3d-digitisation-tangible-cultural-heritage, accessed 06/01/2026.