Adapting Cultural Heritage Narratives: GPT's Potential in Tailoring Heritage Content for Diverse Audiences

Eleftherios Papachristos
Norwegian University of Science and Tecnology
Gjøvik, Norway
eleftherios.papachristos@ntnu.no

Eike Schneiders
University of Nottingham
Nottingham, UK
eike.schneiders@nottingham.ac.uk

Ole Edward Wattne
Norwegian University of Science and Tecnology
Gjøvik, Norway
ole.wattne@ntnu.no

Abstract

This paper explores the potential of ChatGPT, to adapt cultural heritage narratives for diverse audiences. By analyzing stories from the rurAllure project, we assessed ChatGPT's ability to evaluate and modify content for various demographics, including cultural background, educational level, age, and disabilities. Our findings suggest that while ChatGPT can create an evaluative framework and make adaptations that seem rational on the surface, the process is not without errors and requires careful oversight. The study highlights the promise of LLMs in enhancing personalized storytelling in cultural heritage contexts but underscores the need for further research involving real content creators and audiences to fully understand their effectiveness and utility.

CCS Concepts

• **Human-centered computing** → **User centered design**.

Keywords

Cultural heritage, personalized narratives, large language models (LLM)

1 Introduction

Narratives and stories have long been critical tools for cultural heritage sites, galleries, archives and museums, engaging visitors and creating emotional and intellectual connections [17]. Museums, for example, 'tell stories' through carefully selecting and displaying artefacts, supplemented by explanatory visual and narrative motifs to transform exhibits into immersive narrative experiences [15]. Over the years, technological advancements have enabled the integration of cultural sites and artefacts with virtual elements like reconstructions and digital storytelling, resulting in 'hybrid' approaches that create immersive storytelling experiences in specific locations [3]. These technologies have been used to present historical and cultural narratives immersive and engagingly, reaching broader and more diverse audiences [10].

Traditionally, museum storytelling has often been confined to descriptive, didactic prose, which may not fully engage all visitors, especially those who prefer more dynamic and personalized narratives [16]. However, a significant shift has occurred from an object-centred approach to a more audience-centric perspective in museum display design. This shift acknowledges the importance of empowering visitors to interpret artefacts independently, even beyond the physical confines of cultural institutions [8]. This change is reflected in the current trends in cultural and museum contexts, which are increasingly adopting immersive approaches that are story-led, audience-centred, and multisensory, with a strong focus on participation and engagement [6]. Moreover, extensive research suggests that personalizing storytelling and catering to user differences can greatly enhance visitor experience and engagement [1, 7, 12]. However, creating adaptive and personalised cultural heritage information for visitors is extremely challenging, requiring modelling the user, context, and social aspects [1, 12].

Recent developments in machine learning and large language models (LLMs) like ChatGPT have opened up new possibilities for adapting text and audio, such as cultural heritage content, to meet the needs of diverse audiences. Large language and text-to-image models can generate multimodal output that is, when prompted appropriately, indistinguishable from human-generated content [2]. These advanced models can generate and tailor narratives in real time, offering the potential to personalize storytelling in previously unattainable ways. By adjusting the depth, tone, and focus of information, LLMs could potentially help museums and cultural institutions create more engaging and accessible experiences for visitors with varying interests and backgrounds. However, while LLMs present a valuable tool for enhancing the adaptability of cultural narratives, they are not a silver bullet; their effectiveness depends on careful implementation and ongoing refinement. This paper explores the potential of GPT-based models to adapt existing stories and narratives developed by domain experts, making them more suitable for specific audiences and increasing engagement in cultural heritage contexts.

Research Questions. LLMs have been praised for their transformative potential across various domains, including education [4, 5, 9, 11]. For instance, they can empower learners with disabilities by integrating with speech-to-text or text-to-speech solutions, making learning more accessible for individuals with visual impairments [5]. Additionally, LLMs can provide adaptive feedback and personalized learning content tailored to students' needs and interests [4]. However, despite their impressive capabilities in content adaptation, some studies have highlighted that LLMs still struggle to consistently provide appropriate responses tailored to specific

audiences [13, 14]. This research explores the opportunities and challenges of using LLMs to adapt cultural heritage content to different audiences and enhance engagement. Our guiding research questions are the following:

- To what extent can ChatGPT-4 be used as an analytical tool to evaluate existing stories?
- How effective is ChatGPT-4 in adapting stories to different audiences?

2 Methodology

We used professionally written content from the rurAllure project [10] for our investigation. The rurAllure project aimed to promote the cultural heritage of rural areas adjacent to European pilgrimage paths by developing engaging content and applications. Text and audio stories were created to highlight cultural heritage sites along the Saint Olav ways paths in Norway as part of the project's deliverables. These stories, available in both in text and audio formats, were particularly suitable for our evaluation because they are original works produced through meticulous research and are not widely available online, minimizing the likelihood of being included in ChatGPT's training data. This allowed us to assess whether ChatGPT generates inaccurate or fabricated information when asked to modify the stories.

Our primary objective in this exploratory study was to evaluate ChatGPT's capabilities in analyzing existing cultural heritage stories, suggesting improvements, and adapting content for diverse audiences. To achieve this, we selected three distinct stories from the Lake Mjøsa region: "When the Sea Serpents Ruled Lake Mjøsa[1]," "Transport on Lake Mjøsa[2]," and "The Scottish Campaign[3]." We used ChatGPT-4 to evaluate these narratives based on the dimensions that were suggested to be relevant to this context. We then assessed its ability to propose enhancements and tailor the content for audiences with different demographic profiles.

3 ChatGPT as a Tool to Analyse Stories

We explored the potential of ChatGPT as an analytical tool for evaluating existing cultural heritage stories. We crafted a prompt that provided context about the rurAllure project, asking ChatGPT to assume the role of an analyst assessing and enhancing cultural heritage narratives. In a subsequent prompt, we requested that ChatGPT develop a framework for analysis, identifying key storytelling elements or dimensions relevant to the evaluation of these stories.

We then presented ChatGPT with the three aforementioned stories. We asked it to evaluate them across the suggested dimensions using a ten-point scale and explain each rating. ChatGPT identified six evaluative dimensions: 'Historical Accuracy and Depth,' 'Cultural Sensitivity,' 'Engagement and Immersion,' 'Educational Value,' 'Accessibility and Inclusivity,' and 'Sustainability and Impact.' The evaluation results, including ratings and explanations for each story, are detailed in the Appendix (Table 1). At first glance, the analysis seems thorough, with ChatGPT effectively identifying

differences in the narratives and suggesting areas for improvement supported by rational explanations. However, the analysis relies heavily on surface-level criteria. For instance, 'Historical Accuracy' was not judged by verifying historical facts but by simply noting the presence and quantity of references within the text. Similarly, the other dimensions were evaluated mainly based on style and general content structure rather than in-depth contextual analysis. While these assessments may provide value, the actual usefulness of this analysis for content creators remains uncertain. To fully determine its relevance and usefulness, future studies should involve content creators and storytellers focusing on cultural heritage. These studies are crucial for assessing whether ChatGPT's recommendations genuinely enhance the quality of these stories and whether these tools can be effectively integrated into creative processes.

4 ChatGPT for Story Adaptation

To assess ChatGPT's ability to adapt stories for different audiences, we prompted it to modify the three selected stories for various demographics, including Norwegian versus international visitors, blind individuals, and younger versus older age groups. We did this for each story, and we required explanations about what has been changed and why.

In its first attempt to tailor the stories for local and international audiences, ChatGPT significantly shortened the content, omitting entire passages. Even after explicitly instructing the model to retain all information, it continued to condense the stories. Despite this, ChatGPT performed reasonably well in adapting the text for Norwegian audiences by emphasizing shared cultural memories and heritage, using familiar terms and references. For instance, it transformed the original statement, "Mjøsa is Europe's fourth deepest lake," into "Mjøsa is not just any lake—it's a key part of our national landscape and local lore." For international visitors, ChatGPT incorporated well-known global legends, like Scotland's Loch Ness Monster, to make the narrative more relatable and intriguing to those unfamiliar with Norwegian folklore. Additionally, it attempted to provide a geographical context for non-locals. For example, the original line, "The so-called Dølavegen is also known from early times, an ice road that stretched from Eidsvoll in the south to Fåberg in the north," was modified to, "One of the oldest known routes is the Dølavegen, an ice road extending from Eidsvoll in the south to Fåberg in the north, primarily facilitating east-west travel across the lake." However, in doing so, ChatGPT introduced inaccuracies, confusing directions, and misrepresentation of the route's orientation.

When adapting the stories for blind audiences, ChatGPT revised the narratives by emphasizing sensory details beyond the visual, particularly focusing on tactile and auditory elements, while removing colour descriptions to ensure the storytelling remained accessible and engaging. For instance, the original phrase "A green substance flowed out of the carcass, and coloured the water green" was transformed into "The creature's body released a thick, viscous substance that spread through the water." For a younger audience, when prompted to adjust the stories for ten-year-olds, ChatGPT used simpler language, enhanced narrative elements to add mystery and adventure, and made historical and contextual information more accessible. However, its claim to have explained historical

---

references in a way young children could easily understand was not fully realized. For example, instead of clarifying the role of an archbishop, it simply presented the figure as "…a man named Archbishop Walkendorf wrote a letter…," failing to provide the necessary context for young readers.

5 Implications and Future Work

The initial exploration of ChatGPT's capabilities in enhancing cultural heritage content and aiding creators in adapting to individual differences and preferences has shown promise. ChatGPT established an evaluation framework, rated existing stories, and provided justifications for its decisions. However, despite the generally sound rationale behind the adaptations, there were instances where the LLM produced erroneous text or overpromised on the modifications it claimed to have made. This highlights the need for careful supervision and suggests that the adaptation process cannot be fully automated or conducted without oversight. This study serves as a preliminary investigation focused on ChatGPT's capabilities, but it is clear that more extensive research is needed. In the future, we intend to explore other LLMs and, more importantly, evaluate the results with actual content creators and real cultural heritage site visitors from different demographics. Our next steps will involve conducting a study where human participants rate the stories based on the same dimensions used by ChatGPT, allowing us to compare the AI-generated scores with human evaluations. Additionally, we will generate stories tailored for specific demographics and collaborate with members of these groups to assess both ChatGPT and human-assigned scores for the adapted narratives. This approach will help us determine three key outcomes: whether the ratings provided by ChatGPT align with human curations, whether ChatGPT can effectively adapt human-written narratives for various demographics, and whether the adapted narratives are preferred by the target groups for which they were intended.

6 Author Biography

Eleftherios Papachristos has a Diploma in Electrical and Computer Engineering and a PhD in Human-Computer Interaction from the University of Patras, Greece. Since 2022, he is an Associate Professor in the Department of Design at the Norwegian University of Science and Technology (NTNU). His research primarily explores the intersection of Human-Computer Interaction and Artificial Intelligence. Currently, he leads the Human-Centered AI Lab within the Department of Design.

Eike Schneiders is a Transitional Assistant Professor at the School of Computer Science, University of Nottingham. His research focuses on Human-Computer and Robot Interaction (HCI/HRI). Eike has done extensive work on the investigation of interactive systems using artificial intelligence (e.g., personal assistant, various types of robots, and LLMs) as well as artist-led research focusing on public engagement. While he focuses on in-the-wild studies, he also conducts a variation of controlled studies (e.g., online or lab-based studies). Most recently Eike has worked on the Cat Royale project (www.eikeschneiders.github.io/catroyale) which sits at the intersection of art, AI, and animal welfare aiming to engage the public in questions surrounding AI and autonomous systems.

Ole Edward Wattne is an assistant professor in information design

at the Norwegian University of Science and Technology (NTNU), Faculty of Architecture and Design, Department of design. His research is focused on wayfinding behaviours and the design of wayfinding systems. He has a Masters degree in Information design from the University of Reading and has worked as an interaction and information designer in the industry for several years before becoming an academic. He teaches and supervises students of graphic and interaction design.

References

[1] Liliana Ardissono, Tsvi Kuflik, and Daniela Petrelli. 2012. Personalization in cultural heritage: the road travelled and the one ahead. *User modeling and user-adapted interaction* 22 (2012), 73–99.

[2] Ömer Aydın and Enis Karaarslan. 2023. Is ChatGPT leading generative AI? What is beyond expectations? *Academic Platform Journal of Engineering and Smart Systems* 11, 3 (2023), 118–134.

[3] Angeliki Chrysanthi, Akrivi Katifori, Maria Vayanou, and Angeliki Antoniou. 2021. Place-based digital storytelling. the interplay between narrative forms and the cultural heritage space. In *International conference on emerging technologies and the digital transformation of museums and heritage sites.* Springer, 127–138.

[4] Wensheng Gan, Zhenlian Qi, Jiayang Wu, and Jerry Chun-Wei Lin. 2023. Large language models in education: Vision and opportunities. In *2023 IEEE international conference on big data (BigData).* IEEE, 4776–4785.

[5] Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, et al. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences* 103 (2023), 102274.

[6] Jenny Kidd. 2018. 'Immersive'heritage encounters. *The Museum Review* 3, 1 (2018).

[7] Panos Kourouthanassis, Costas Boletsis, Cleopatra Bardaki, and Dimitra Chasanidou. 2015. Tourists responses to mobile augmented reality travel guides: The role of emotions on adoption behavior. *Pervasive and Mobile Computing* 18 (2015), 71–87.

[8] Fui-Theng Leow and Eugene Ch'ng. 2021. Analysing narrative engagement with immersive environments: designing audience-centric experiences for cultural heritage learning. *Museum Management and Curatorship* 36, 4 (2021), 342–361.

[9] Mengqi Liu and Faten M'hiri. 2024. Beyond Traditional Teaching: Large Language Models as Simulated Teaching Assistants in Computer Science. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1.* 743–749.

[10] Martín López-Nores, Tsvi Kuflik, Manolis Wallace, and Yannick Naudet. 2019. Preface to the UMUAI special issue on personalized delivery of cultural heritage content: perspectives on 7 years of progress in the field. *User Modeling and User-Adapted Interaction* 29 (2019), 1–7.

[11] Eleftherios Papachristos, Yavuz Inal, Carlos Vicient Monllaó, Eivind Arnstein Johansen, and Mari Hermansen. 2024. Integrating AI into Design Ideation: Assessing ChatGPT's Role in Human-Centered Design Education. *Authorea Preprints* (2024).

[12] Laia Pujol, Akrivi Katifori, Maria Vayanou, Maria Roussou, Manos Karvounis, Marialena Kyriakidi, Stamatia Eleftheratou, and Yannis Ioannidis. 2013. From personalization to adaptivity–Creating immersive visits through interactive digital storytelling at the Acropolis Museum. In *Workshop proceedings of the 9th International conference on intelligent environments.* IOS Press, 541–554.

[13] Riaz Qureshi, Daniel Shaughnessy, Kayden AR Gill, Karen A Robinson, Tianjing Li, and Eitan Agai. 2023. Are ChatGPT and large language models "the answer" to bringing us closer to systematic review automation? *Systematic Reviews* 12, 1 (2023), 72.

[14] Donya Rooein, Amanda Cercas Curry, and Dirk Hovy. 2023. Know Your Audience: Do LLMs Adapt to Different Age and Education Levels? *arXiv preprint arXiv:2312.02065* (2023).

[15] Maria Roussou. 2004. Learning by doing and learning through play: an exploration of interactivity in virtual environments for children. *Computers in Entertainment (CIE)* 2, 1 (2004), 10–10.

[16] Maria Roussou, Francesco Ripanti, and Katerina Servi. 2017. ENGAGING VISITORS OF ARCHAEOLOGICAL SITES THROUGH" EMOTIVE" STORYTELLING EXPERIENCES: A PILOT AT THE ANCIENT AGORA OF ATHENS. *Archeologia e Calcolatori* (2017).

[17] Ektor Vrettakis, Vassilis Kourtis, Akrivi Katifori, Manos Karvounis, Christos Lougiakis, and Yannis Ioannidis. 2019. Narralive–Creating and experiencing mobile digital storytelling in cultural heritage. *Digital Applications in Archaeology and Cultural Heritage* 15 (2019), e00114.

A   Evaluation Dimensions used for the three narratives.

	Sea Serpent in Lake Mjøsa		Transport on Lake Mjøsa		The Scottish Campaign	
Dimension	Score	Justification	Score	Justification	Score	Justification
Historical Accuracy and Depth	8	Well-researched with historical references, although the blending of fact and folklore may be confusing.	8	Provides historical data and detailed accounts, but could benefit from more sources to verify claims.	7	Offers historical details and context, though some claims lack confirmation, affecting overall credibility.
Cultural Sensitivity	7	Respects local folklore and uses it as a narrative framework, but lacks broader cultural perspectives.	7	Respects the historical significance but lacks mention of how these histories interact with local cultures today.	6	Attempts to recount events neutrally but could improve in presenting both Scottish and Norwegian perspectives more equitably.
Engagement and Immersion	9	Highly engaging with vivid descriptions and a compelling mix of suspense and intrigue.	6	Factual and informative but lacks the narrative flair that would make the history more vivid and engaging.	7	Provides a compelling narrative with vivid combat descriptions, though some parts may need more emotional depth to fully engage readers.
Educational Value	8	Offers significant insights into the role of folklore in history and cultural identity.	9	High educational content, effectively providing a historical overview of transportation on Lake Mjøsa.	8	Educative in illustrating mercenary use in European wars and regional military strategies.
Accessibility and Inclusivity	8	Uses accessible language and clear narrative style, suitable for a wide audience.	7	Language is straightforward and accessible, though some terms might need clarification for broader audiences.	7	Generally accessible in language and presentation but could benefit from clearer explanations of certain historical terms and contexts.
Sustainability and Impact	5	Enhances cultural understanding, though direct promotion of conservation or community benefit is limited.	5	Discusses historical usage but does not connect this to current sustainability efforts or impacts on the community.	5	Mainly historical recounting with limited discussion on its modern-day relevance or lessons for contemporary conflict resolution.

Table 1: Comaperative table based on ChatGPTs evaluation dimension for the three stories