http://dx.doi.org/10.14236/ewic/BCSHCI2025.10

'Chattable' Avatars: Using LLMs to Power Visitor Engagement with Historical Persons

Zhuoling Jiang
School of Computer Science
Cardiff University
Cardiff, CF10 5NY, UK
https://zhuoling.space/
JiangZ30@cardiff.ac.uk

Yipeng Qin
School of Computer Science
Cardiff University
Cardiff, CF10 5NY, UK
https://yipengqin.github.io/
QinY16@cardiff.ac.uk

Daniel J. Finnegan
School of Computer Science
Cardiff University
Cardiff, CF10 5NY, UK
https://ps2fino.github.io
finnegand@cardiff.ac.uk

Cultural Heritage institutions such as Galleries, Libraries, Archives, and Museums are tasked with preserving our history and heritage for future generations while engaging new audiences with an appetite for said engagement to be increasingly digital and interactive. To explore how advances in natural language processing, particularly large language models (LLMs), may help GLAMs in their mission, we designed a prototype 'Chattable' avatar, a 3D high-polygon animated character which visitors can talk to and interact with. We report the design of our avatar, and a workshop we conducted with curators and staff from a GLAM institution, to understand the problems, requirements, and opportunities LLMs present in the cultural heritage space. We present results from a qualitative analysis of our workshop highlighting themes such as trust, authority, social experience, and location, finding LLMs may be more suited to deployments focused on non-factual data dissemination. We conclude with implications for GLAMs and suggestions for future research to realise how best to integrate GenerativeAI like LLMs into the GLAM space.

GenerativeAI, Museums and Galleries, Qualitative Research, LLMs, Workshop

1. INTRODUCTION

Galleries, Libraries, Archives, and Museums (GLAMs) are increasingly looking for ways to effectively engage new and younger audiences in an increasingly digital and personalized world. They are faced with factors ranging from site-centric for example, poor WiFi coverage, difficulties installing hardware and sensor technology, to maintenance, for example, staff training, volunteer capacity, and upkeep of software and replacing components. At the same time, GLAMs face increasing budget cuts, forcing them to strategically invest in technology, creating a tension between wanting to innovate and try new things, and the real risk of impact on the organisation introducing new technologies into workflows and exhibitions.

This is not to say that GLAMs are naïve about future change. In contrast, many institutions are aware of the challenges and opportunities brought about by emerging technologies in 3D animation, motion capture, virtual and mixed reality, and AI (Seale 2023). If they were to take advantage of the opportunities, they are often unique spaces to install digital technology as they are primarily public spaces, catering to a wide audience with respect to demographic backgrounds, and focus on historic narratives. It remains an open question how best to integrate digital technology into these spaces (Shehade and Stylianou-Lambert 2020) to offer personalized experiences en masse to cultural heritage site visitors.

Though their mission is history preservation, GLAMs are also future looking, in particular paying attention to advances in emerging technologies in AI. Of particular interest is the broad spectrum of so-called 'GenAI'; generative machine learning models which can produce text, audio, and image/video content from a user prompt. One class of GenAI technology with potential to impact visitor experience are large language models (LLMs). The rapid adoption of these technologies by the IT workforce (Joskowicz and Slomovitz 2024) and the general public (Ronge et al. 2025) is encouraging on the one hand, as this makes them a familiar technology with affordances understood by a wide range of people. However, on the other hand, many issues remain with their use. For example, issues surrounding safety and ethics (Bender et al. 2021), safeguarding content and cybersecurity (Qian et al. 2022), human-centered issues such as privacy and situational awareness and knowing when one is speaking with an AI or a "virtual human" (Ehsan et al. 2021), and issues particular to GLAMs

like volunteer competency training, infrastructure, and familiarity.

In this paper we explore the themes of trust and perceived authority in the context of a national cultural heritage site, understanding not just the interaction paradigms LLMs may provide but the perceived impact on the curatorial process, digital infrastructure, and workforce – both office and site volunteers – of deploying such technology at a site of cultural significance. To operationalize our exploration, we conducted an in-person workshop with a large cultural heritage institution in the UK.

2. RELATED WORK

LLMs have emerged with capabilities far beyond the scope of natural language processing. When appropriately configured they can serve as interactive agents, and with well-crafted prompts, they can realistically mimic the language style of predefined characters Raiaan et al. (2024), and have found widespread application in many fields such as software engineering (Hou et al. 2024), gaming (Gallotta et al. 2024), and healthcare (Maity and Saikia 2025).

2.1. AI Interactive Agents in Museums

Integrating GPT-based conversational agents into a cultural heritage website resulted in improved usability and engagement, especially when contextual question prompts were offered (Geninatti Cossatin et al. 2025). LLMs have also shown to facilitate more fluid and personable interactions compared to rule-based systems (Trichopoulos et al. 2023). However, concerns around misinformation and the factual accuracy of LLM-generated content remain a significant barrier to adoption in GLAMs. Research has suggested strategies for mitigating the damage from factual inaccuracies, by suggesting that historical figure chatbots be framed as narrative or performative agents rather than factual authorities (Padilla Engstrøm and Løvlie 2025), provided their roles were clearly communicated and aligned with visitor expectations (Chen et al. 2025). Our work adopts these strategies as discussion points and possibilities with museum professionals to dive deeper into the organisational and structural barriers that must be overcome to facilitate successful on-site deployment.

2.2. Anthropomorphism and Trust Towards Technology

Care must be taken when deploying AI into public spaces with respect to their form and function, as public perception matters with respect to visitor engagement and interaction. For example, the uncanny valley effect[1] can make visitors uncomfortable and negatively effect their willingness to interact with an AI (Heisler and Becker-Asano 2025). With respect to function, recent work has emphasized the potential of combining human and digital guides to augment, rather than replace, traditional curatorial voices Antoniou et al. (2021). Indeed, museum professionals' attitudes toward AI has been generally open, contingent upon clear roles, technical reliability, and staff inclusion (Cameron et al. 2025).

Trust is recognised as a dependent factor in interaction design with AI: both over and under trusting an AI agent can be detrimental to user experience (Desai et al. 2024), and trust can be formed by an agent acting in a way that resonates with a person, making it friendly (Sun and Wang 2025). However, in the context of GLAMs this is further compounded. A fundamental component of curation is 'intellectual authority'; the command over knowledge of the past, or at least the *belief* that such authority is manifest. For example, museums' attraction is the public perception that they are institutions where experts in the field gather to structure and provide access to knowledge, and we trust the information they share is reliable (Longair 2015; Crane 1997). Museums and cultural heritage institutions are often seen as welcoming places with figures of intellectual authority by people inside (i.e., curators, experience designers, volunteers) and outside (i.e., external stakeholders, educational institutions, visitors) and have established trust both within their institutions and the public.

Drawing from this work, and our own motivations to explore how LLMs may be applied in cultural heritage settings, we implemented an LLM powered avatar featuring a high fidelity 3D animated figure, which we call a "Chattable" Avatar (cAv). We focus on museums as a deployment site initially. We have chosen the term 'chattable' as it captures our aim of creating a welcoming and personable character which people can interact: beyond harvesting formal facts and figures from the avatar, people can just talk with it (chat) in an informal way, much like they would with a volunteer at a museum or heritage site.

In summary, we have considered related work in AI based interactive agents, and how anthropomorphism when applied to technology meant for museum deployment can impact end users' perceived trust, to develop our cAv, and use it to support discussion and inspire critical inquiry into the role AI may play in furthering the mission of institutions and organisations in the GLAM space.

[1] Masahiro Mori's original essay was translated from Japanese to English in 2012: https://spectrum.ieee.org/the-uncanny-valley

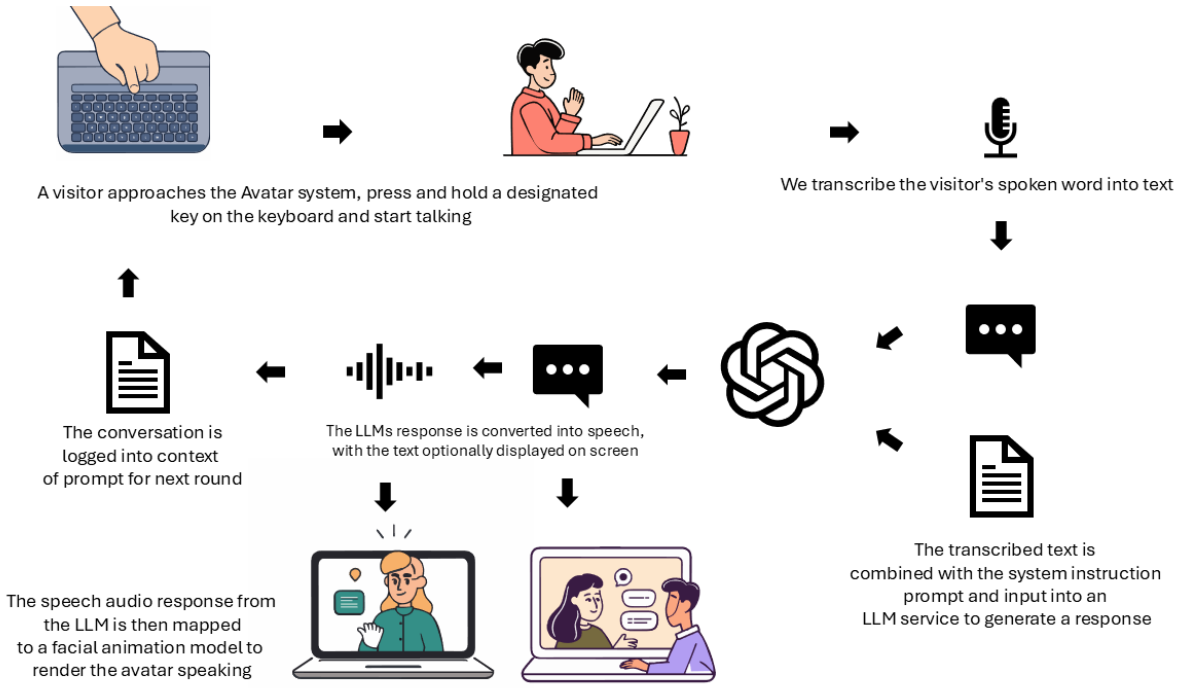

Figure 1: *Visual description of how our system prototype works.*

Figure 2: *A screenshot of our 'Chattable' Avatar running in real time on a standard issue laptop.*

3. PROTOTYPE "CHATTABLE" AVATAR

Our cAv is developed using Unreal Engine 5 [2] and the Epic MetaHuman Plugin [3]. The demo is running on a gaming laptop with a 16" screen, an Nvidia Geforce RTX 3060 GPU, and 12th generation Intel Core i7 CPU. Our avatar was created based on a real person living in the 18th Century who worked as the master of ceremonies at a venue of historical cultural significance. We refer to this person as **RT**. Figure 2 shows our rendered cAv. The choice of clothing, the accent spoken, the decor of the background, the idle animation, and the manner in which the avatar presents itself were curated by cultural heritage professionals. While a critical prerequisite for strong visitor engagement (Hede et al. 2014), this fit our theme of 'chattable-ness' and links with the perceived authority of the cultural heritage institution as a site of reliable knowledge.

Figure 1 describes how our prototype works. We have adopted the MUSETECH model to iteratively design, develop, and evaluate our avatar implementation (Damala et al. 2019a,b). The ability to 'chat' with RT comes from the LLM. Specifically, we used OpenAI's GPT-4 Series model [4] for this version of our prototype cAv. We designed a wrapper workflow around the original text-only GPT-4 chat completion API. We first designed a prompt template [5] based on historical documents provided by a GLAM institution curator collaborating on our project. These documents were outputs from a research project conducted by a GLAM institution based on the real-life person RT. We cannot include the full prompt here as the research and information is under copyright by the GLAM institution. The prompt template features the context and background knowledge, enabling our cAv to 'pretend' and pose itself as RT. The template also features instructions for the LLM to engage the conversation and generate responses that 'sound like' how a human would talk. These context information and instructions are integrated into the system instruction part of the prompt template forming the system context

[2]https://www.unrealengine.com/en-US
[3]https://www.unrealengine.com/en-US/metahuman
[4]https://openai.com/index/gpt-4/
[5]we consider a detailed description of prompt engineering out of scope for this paper, but to interested readers we suggest (White et al. 2023)

and constitute a chat query, which we upload to the LLM using the chat completion API. We maintain a chat log throughout interaction with RT: this allows the LLM to recall points of dialogue between the user and the RT, which in turn are used to produce new responses (Xiao et al. 2024). For example, when asked a question or prompted by a visitor, each word RT speaks is predicted using the chat log, considering all previous words spoken between the visitor and RT.

In our target use case, visitors approach RT and strike up a conversation by pressing and holding a designated key on the keyboard and start talking. At this point, the system begins recording from the microphone. Upon release of the key, the spoken word is transcribed into text via an automatic speech recogniser (ASR) service[6]. The transcribed text is integrated into our prompt template creating a bespoke chat query described above. This is sent to an LLM, which generates a text response. The text response is then converted into speech using a text-to-speech (TTS) service provided by the same Epic MetaHuman Plugin, configured with specific voice parameters such as style, accent, speed, and tone. The resulting audio is processed further by the MetaHuman Plugin, synchronizing lip movement to generate facial and animations matching the speech to render the cAv talking in real time.

4. WORKSHOP

We conducted an in-person workshop at the Bath Assembly Rooms, a popular tourist destination. The participants were the authors of this paper (N=3), employees from a cultural heritage institution, all involved in formulating the organisation's digital strategy: curators (N=2), AI governance strategy and IT (N=2), site specific volunteers (N=3); and other academics with expertise in digital heritage, AI, and HCI (N=2) (See Table 1). The workshop took place over 1 day, involving several group-based activities with the objective of creating a cAv experience for the Bath Assembly Rooms. We demonstrated our prototype at this workshop to give participants a hands-on preview of the look and feel of interacting with a cAv.

4.1. Methodology

We deployed a co-design methodology to our workshop, emphasizing activities that involve meaningful stakeholder engagement in our research design to ensure our understanding evolved collectively throughout the day. In our case, colleagues from the GI were actively involved in the research planning phase, where they had explicit participation in planning the workshop activities, and the provocations (below) (Slattery et al. 2020). To stimulate participants and prepare them for discussions, two days prior to the workshop date, all participants were sent a brief 'provocations' slide featuring the following questions aimed at eliciting insights (Bardzell et al. 2012):

- *The best technology provides solutions to challenges and questions we have, rather than relying on novelty. What questions and challenges does your organisation face?*
- *How might we develop visitor literacy around AI?*
- *Where might AI fit into your organisation's mission?*

These provocations were co-produced based on preparatory discussions with the AI and Data Governance team leader from discussions around their digital strategy and what they wanted to focus discussions on at the workshop.

After this, our programme consisted of the following structured activities:

1 Ecosystem mapping exercise focusing on the heritage site experience of a visitor and a curator. The purpose of this activity was to get a first impression from our participants regarding the context of both visitors and curators, the services and resources they interact with when visiting/planning an exhibition. This helped guide our thinking towards the kinds of technology and services that people currently engage in (Abdelnour Nocera et al. 2024), and how cAvs may embed into this ecosystem without disrupting a user's routines and/or their workflows.

2 User persona construction and discussion. Personas are a key part of User Centred Design, particularly in the early stages of development as they help with conceptualizing and ideating new experiences (Salminen et al. 2022). To facilitate and support participants with their user persona creation, we followed best practice and issued templates featuring an image of a person (Salminen et al. 2021), and the following prompts for participants to fill in as much information as possible: occupation; where they live; describe their personality; what challenges they face in their everyday life; and their exposure to digital technology, for example, use everyday, 'early adopter'.

3 After creating their personas, participants took a short break before positioning each of their persona in the centre of an Empathy Map. The purpose of this activity was twofold: 1) to help reach common ground amongst our workshop participants and induce discussion around how our personas might feel in certain contexts (Siricharoen 2021), and 2) to seed

[6]https://openai.com/index/whisper/

Table 1: *Participant Demographics and Roles*

ID	Role	Description
R1	Researcher	Author
R2	Researcher	Author
R3	Researcher	Author
R4	Researcher	Academic faculty member
R5	Researcher	Academic faculty member
C1	Curator	National Cultural Heritage Institution (main office based)
C2	Curator	National Cultural Heritage Institution (site based)
S1	Governance & Strategy	National Cultural Heritage Institution (main office based)
S2	Governance & Strategy	National Cultural Heritage Institution (main office based)
V1	Volunteer	National Cultural Heritage Institution (site based)
V2	Volunteer	National Cultural Heritage Institution (site based)
V3	Volunteer	National Cultural Heritage Institution (site based)

the next activities and prepare participants to make decisions later around the infrastructure and resources that would need to be in place for the successful delivery of a cAv in a GLAM setting. Participants were free to design their persona: the only criteria was that the persona must be interacting in some way with a cAv.

4 Group based hierarchy of needs analysis based on opportunities and challenges each group collectively agrees to prioritize. The purpose was to gauge what curators, IT experts, and academics would prioritize as they worked towards a *practical* installation of a cAv. At the beginning of our workshop day and during refreshments, we invited participants to stick post-it notes capturing their comments on the provocations on a flip chart split in two rows, concerns vs exciting prospects. These were used as reflective artefacts when discussing key opportunities and challenges.

5 Finally, we completed an activity from the AI Museum Planning Toolkit (Murphy and Villaespesa 2020), namely the Stakeholder Mapping exercise. This focused participants' attention on next steps to take on their journey towards deploying a cAv.

5. RESULTS AND OUTPUTS

In the following section we discuss the results from our workshop; outputs from our group activities and discussions. Our a-priori themes of trust and perceived authority were instrumental in designing our workshop activities, from which sub themes arose. We coded the contributions from participants provided on post-it notes, flip charts, maps across all activities completed during our workshop, and included notes taken by a research assistant during the workshop. We conducted a thematic analysis on our data; we take a reflective approach, providing our key (sub)themes in **bold and underlined** and ordered by activity and contextualising in the broader literature as appropriate.

One clear signal emerging from the ecosystem mapping activity was a dichotomy of perspective. Figure 3 shows two maps created by our workshop participants. **Service Delivery** was a key theme, where visitors take an egocentric perspective and are concerned with the modes of interaction and immediate infrastructure (for example, WiFi) facilitating their visit. For cAvs to be successful, they should integrate with the full journey of the visitor, as C1 says it could offer *"personalised digital interpretation that directly meets (visitor) interest"*. Curators take an allocentric view; they are concerned with the **quality of information** mined from their research. In the context of cAvs, curators were excited to see the capabilities of our cAv: C1 commented *"Its exciting...the ability (of the avatar) to understand questions accurately...and the speed of response"*, with C2 adding *"info(rmation) can go as deep/intellectual as each visitor wants...the depth of content it could share...is a wow factor and fun"*. However, the accuracy of what the avatar says is paramount; it must be authentic and reflect the values of the institution and some participants reflected on the impact this technology may have on volunteers, sceptical of replacing volunteers with automated bots. V1 commented *"Are you (sure you are) hearing facts?"* and *"(we must ensure) it pitches the right amount of info - not too much, but not too little"*.

Next, participants worked together in groups at round tables to create their personas. As our participants were unfamiliar with creating personas and user centred design, we began the session by playing a short video[7] which introduced what personas are and why they are useful. The workshop facilitator (one of the authors of this paper) would circle the room to each table, answering questions participants had regarding user personas generally, though not assisting participants in the task. They also encouraged everyone to give as much detail as possible, leaving no room for assumptions when describing their persona. We present the following personas constructed by our workshop participants:

- *Simon is a tech savvy, business minded, single man. He works hard and makes the most of his spare time socialising with friends. He is interested in the avatar and is enjoying testing it to its limits. He immediately uploads a photo of it to*

[7] https://www.nngroup.com/videos/personas-101/

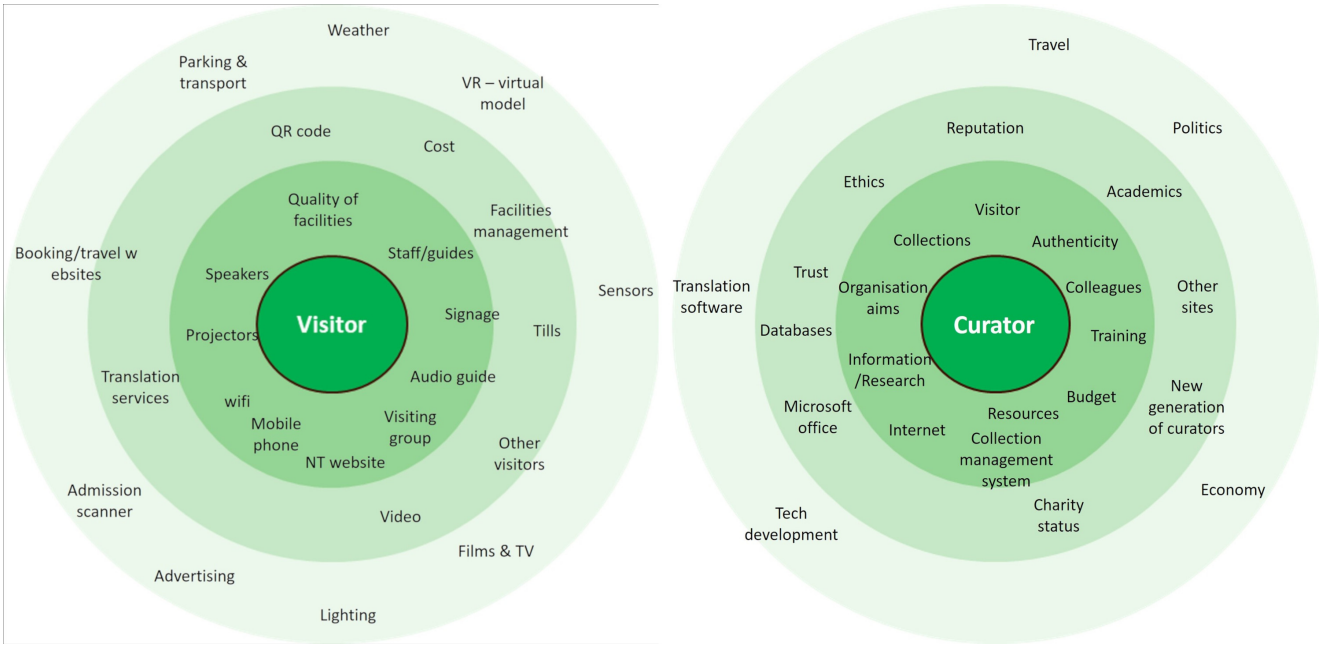

Figure 3: *Results from our ecosystem mapping exercise.*

social media and critiques its interaction. He may be showing off to others but only has a short attention span so does not spend too long with it. He is not the typical visitor and is only visiting the Bath Assembly Rooms as part of a visit with a friend.

- *Aleena is a 30-something female who is visiting the Assembly Rooms with her boyfriend. They live apart and Bath was a meeting place for a romantic weekend away. She is familiar with technology and uses it for work as a customer services manager at a bank. She uses the avatar communally with her boyfriend after initial hesitations about the technology at a heritage site. They are interacting with the avatar in the cafe before they visit the main site.*
- *Sumie-Lee is a recent Business & Media graduate and beauty influencer, who creates content daily for her followers. She is visiting Bath to create content for her social media account. A fan of the television programme Bridgerton, she uses the Assembly Rooms to include content that would make her seem smarter in her posts. She is asking the avatar questions about places to eat in Bath. She spends a long time interacting with the avatar to create the perfect content.*
- *Margaret, a retired teacher is recently widowed. She spends her time helping to care for her two grandchildren and engaging with her church group. Margaret is initially nervous about the avatar as the only technology she usually engages with is sharing photos of her grandchildren on her phone. She enjoys engaging with the avatar, even having a bit of fun by asking it if it can find a husband for her single daughter.*

Participants recognised cAvs may be engaged with as a **communal experience** involving several people, rather than a one-to-one interaction between a person and the avatar. Curiosity, initial hesitation, and social experience were commonly used in discussions. Visitors may approach the cAv in pairs or small groups, similar to how they approach volunteers, so queries on how cAvs can deal with groups scenarios were raised. For example, Aleena and her husband might *"interact cautiously at first" (A1)* while thinking *"Will the AI understand me?" (V2)*. Some participants considered Margaret as someone who would *"watch someone else use if first" (R5)*, and once comfortable would *"ask it lots of questions" (V2)* and feel *"excited to have made it 'work'" (C1)*. As Margaret is a repeat-visitor to the cultural heritage site, one participant questioned if *"(she) would wonder if it (the cAv) would remember me?" (R1)*.

The **location** of a cAv was considered important: C2 claimed *"we could put it in a café which may reach non-fee-paying customers, offering an engagement opportunity, versus those who have crossed the pay barrier and are already invested in the exhibition or site"*. R5 considered the spectator's perspective *"...a sense of enjoyment and fun comes from the interactions, adding to the visitor's experience"* which is something *"Sumie-Lee may want to capture (this) to share with her social media followers" (R4)*. Finally, the cAv was also considered to have potential as an effective gateway experience: R4 said *"Simon may be familiar*

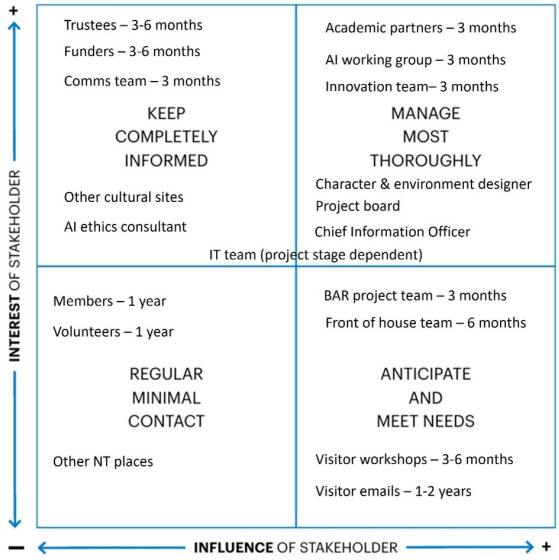

Figure 4: *Digitized depiction of the flipchart our workshop participants used to complete the stakeholders map.*

with technology although not necessarily AI" and *"he might want to see how far he can stretch it, break it"*.

After the lunch break, we continued with a round-table discussion, using the provocations sent in advance of the workshop (See Section 4.1). Table 2 summarizes the main discussion points during this activity. Participants were drawn towards opportunities and challenges surrounding the visitor experience, but also technical issues and the need for volunteer training and system maintenance. For our participants, the unique selling point of a cAv would be drawing in new audiences and improving the experience for established visitors. Bringing **history to life** and providing a new, **engaging way of learning**, especially for younger audiences, were the two prime goals discussed. Our participants were excited and saw technology as an opportunity for augmenting visitor interactivity in creative ways. Hesitations surrounded the practicality of implementing a cAv such as cost, technology support, and managing the experience if volunteers and/or other staff at the site are busy.

Again, the information's historical accuracy was viewed as important; however, how it manages sensitive topics in a respectful manner (**sensitivity and respect**) was vital to it's adoption in our participant's opinions. Reflecting on the personas from earlier, V2 commented *"She (Margaret) might feel nervous at first, and would expect a certain attitude and mannerism in its tone and what it says"*. Asking questions to a historical character posed consideration as to how this is managed in line with typical GLAM institutional values. C1 summarised it well *"The characters should be inspired by history...(yet) we have a level of responsibility."*.

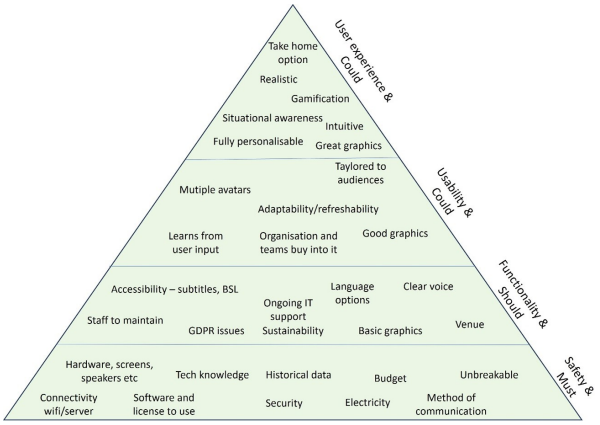

Figure 5: *Digitized depiction of the flipchart our workshop participants used to complete the hierarchy of needs exercise.*

Although accuracy in historical content is needed, just as important is *how* content is relayed to visitors, and that it should be done so with sensitivity and in a way that acknowledges what happened in the past whilst highlighting that some of the actions from historical characters were not appropriate. **Biases** were raised as an issue, particularly how a cAv might be trained to be aware of biases. How to "train it" in ethics and diversity was also raised. For example, C2 said it was *"vital to ensure that it is tested with a diverse audience to ensure inclusivity to all before it is introduced"*. Finally, internally evaluating a cAv before public display would be of value, such as allowing staff or visitors in a private, closed environment to provide feedback regarding their interactions. Having real time data for the most asked questions and how many people interacted with it would also enable ongoing evaluation of the content.

After the round-table discussion, our participants came together to decide how best to sort the key points discussed at our workshop following a hierarchy of needs structure. Figure 5 shows how our participants ranked importance of factors when creating, facilitating, and promoting cAvs. Confirming the sentiment across earlier activities, aspects of safeguarding, security, resources, ethics, and functional requirements were all considered the foundations which must be met. For our participants, to facilitate interaction with a cAv and ensuring integrity, the correctness of the knowledge imbued into a cAv regarding the topic of interest/historical site of significance where it is installed is essential (our theme of quality of information).

Figure 4 summarises the key stakeholders, frequency of communication and the interest of the stakeholder in a cAv. In line with the toolkit's intentions (Murphy and Villaespesa 2020), after our mapping of stakeholders, discussions at our workshop emphasized how the influence and interest of stakeholders will evolve over time. For example, at the co-design and planning stage (i.e., now), a member of the IT team would have

Opportunities	Challenges
Visitor Engagement	
Drive visits indoors	Historical (In)accuracy
Representing famous figures	Explaining to visitors how to use it
Bring characters to life	How to ensure age-appropriate responses from the LLM
Present different perspectives	Technology e.g. wifi, sound, general kit
Empower visitors	Artificial voice may disengage some visitors
Educating children/help school visits	Cost
Ability to display real historical events	How it handles sensitive subjects for example, slavery
Great social media interest	Final form of avatar e.g. projection, screen
Accessibility	
Personalisation of experience	May alienate or disengage current audiences
Visitors not feeling judged when asking it questions	Safeguarding, disclosing personal data
Bring in more visitors	Competing noise of avatar with visitors
Improved accessibility – languages, learning needs	Volunteers may feel threatened by it
Security and Organisation Related	
Support visitor services if short on staff	Managing experience if busy
Updates with new research	Expertise required for ongoing maintenance

Table 2: *Opportunities and Challenges raised by our workshop participants. Themes around visitor engagement, security and safeguarding, and accessibility were highlights from our discussions.*

minimal interest and influence (the centre of Figure 4). However, once a cAv is launched, the IT department would become a team of vital importance and it would be prudent of any institution to ensure they were in regular contact with the design team. Any Board of Trustees an institution may have would likely be notified at the earliest possibility, particularly considering the potential legal, ethical, and accessibility issues cAvs may raise. They would maintain a close eye on development and outputs from future workshops at general meetings for example.

Communications and public relations teams will receive immediate and regular updates following workshops. Less immediate contact with institution volunteers, members, and the public is expected, however they would be engaged once development enters the final stages, with deployment imminent. Finally, our participants suggested a big issue for many GIs revolves around visitor perceptions of GLAMs being places of fact, featuring well researched experiences visitors can 'trust'. By introducing forms of Generative AI models i.e., LLMs which offer possibilities of unscripted visitor experiences, to the GLAM space one removes the content control from the organisation. This in turn fundamentally deregulates the personally curated experience to one where checks and balances are less understood.

6. DISCUSSION

Our workshop focused on key challenges for GLAM institutions looking to use cAvs for engaging visitors, but also identified several key topics participants had not considered before. For example, the need to engage not just visitors but their partners (as with our persona Aleena) *beyond* the place of cultural consumption like in a café, or how to engage visitors who come to the place for reasons other than cultural heritage engagement (as with our persona Sumie). Another example was that though well aware of privacy concerns, our participants had not considered the *negative* impact of zero data retention on the visitor experience, for example how visitors may form bonds with the cAv only for it to forget who they are upon their next visit like with our persona Margaret. It is known that aspects of system design negatively impacts the trust a user has in that system (Weinberger and Felt 2016), but its less clear how this loss of trust would translate between the avatar and the institution, as the avatar may be seen as an institutional representative similar to a volunteer. This could potentially have knock on effects to the perceived authority of the public in the institution, particularly as LLMs are known to generate inaccurate or even fictitious responses (Döbler et al. 2024). There is clearly a need to balance safeguarding and privacy of visitors' sensitive data with a compelling personalised user experience.

Our hierarchy of needs, round-table discussions, and stakeholder mapping conveyed a clear body of work around infrastructure: what is needed to make cAvs a reality, and a core prerequisite for our sub themes of engagement and learning in novel ways. Recent work exploring acceptance and adoption of robots in museums identifies three phases of adoption; pre use, initial use, and sustained use. The authors reflect on how professionals would adapt their use of the robot to suit their specific needs and context and turning barriers into opportunities (Cameron et al. 2025). We expect the same dynamic when adopting cAvs as both curators and volunteers emphasized service delivery, yet success here hinges not just on the confidence in staff and volunteers to operate the technology but the possibility of improving *their* experience, not just the visitors'.

With respect to themes of social experience and engagement of visitors reflect that of recent literature in the area of Generative AI and chatbots (Chen et al. 2025; Quinto Lima et al. 2025; Lin and Hu 2025). Visitors may query the avatar, but if or when they recognise it has provided the wrong information the experience could lead to negatively impacting their trust in not just the avatar, but the institution too. By making the avatar's shortcomings transparent, this could be an opportunity for critical engagement, particularly in a social setting. GLAMs could even adopt cAvs in a way that visitors are encouraged to feedback on their interaction with the avatar. This could be used to encourage further critical thinking and engagement in an already informal learning environment. Our workshop participants also highlighted the importance of inclusivity, managing and mitigating bias and the sensitivity of responses from an avatar. cAvs offer a highly personalised experience, and can adapt on-the-fly based on participants' queries, and its (non)anthropomorphic appearance could even be controlled by the visitor. This could have positive impact on different demographics, benefitting neuro-diverse and other individuals who may benefit from sensitively-aware and conversational learning paradigms, drawing on suggestions from (Hall et al. 2024) regarding designing for independence. Individuals are free to engage the avatar however they like: consider our personas Sumie-Lee and Margaret, and Aleena and how they would each engage the avatar in completely different ways, with different motivations and intentions, and alone or with others.

A concern from workshop participants regarding cAvs replacing volunteer staff arose during our ecosystem mapping and again during roundtable discussions focused on challenges. Though we (researchers) acknowledge these concerns are warranted and could sympathise, we stressed the importance of appropriation to the other participants: that use of an LLM need not, by itself, result in the *substitution* of human volunteer interaction. This is a perspective held by our participants, in line with previous research (Antoniou et al. 2021), and is a belief we posit may be widely held in the cultural heritage space. On the contrary, cAvs could be used as cooperating curators, a tool for professionals to use while creating an exhibition, or theatrical agents, as part of a show with human actors to create an augmented reality experience and encourage active interpretation (Alsford and Parry 1991; Jackson and Leahy 2005). Participants at our workshop, though optimistic and excited about the possibilities of LLMs and cAvs to revolutionize visitor experiences, maintained a healthy cautionary stance on their application, due rightly or wrongly to feelings of uncertainty this technology may create.

While our study provides valuable insights into organisational perspectives on deploying Chattable Avatars (cAvs) within GLAM settings, it does not include museum visitors—the ultimate end users—as participants. This was an intentional focus: our aim was to understand the organisational challenges, concerns, and opportunities that shape early-stage planning and governance. However, we acknowledge that the absence of visitor perspectives limits our ability to assess how cAvs might be received and used in real-world settings. Future work will address this by involving end users through field studies and in-situ evaluations, complementing institutional insights with direct observations of engagement, trust formation, and usability. Integrating both perspectives will be essential to designing socially attuned and trustworthy AI experiences in cultural heritage contexts.

7. CONCLUSIONS AND FUTURE WORK

Our work has implications for LLMs in museums. Our workshop highlighted the broad and diverse visitor demographic that would come to visit a site of cultural heritage. These demographics are not all necessarily interested in facts and figures; we posit that one way to mitigate the risks of LLMs disseminating false information – a major risk to trust and authority – is to shift the focus towards entertainment, nudging visitors to contact a human volunteer. Previous work encourages us not to focus only on the end user experience of factual content but instead promoting discovery (Prentice et al. 1998; Ch'ng et al. 2023). We see LLMs as a step towards promoting discovery; instead of positioning the AI as a repository of information and expecting the public to trust it, we suggest pivoting towards a model that deploys LLMs as 'partner detectives', agents capable of engaging in dyadic interactions to promote curiosity, nudging visitors towards factual sources like recognised research repositories.

Future work may investigate how operating a cAv at a specific cultural heritage space may impact resources,

and ensuring it stays compliant with internal and external policies across visitor, site, and cultural heritage professional perspectives. We suggest a field study, where visitors will have the opportunity to 'speak' with an avatar about anything they wish. Of particular interest is sampling **what** people talk about and **how** they engage in a dialogue with a cAv.

Additionally, we recommend more workshops, with a focus on data governance, ethics, and technological infrastructure necessary to scale up cAvs for deployment across several sites under a large GLAM institution's care. This will enable future research exploring contextual factors of cAvs, how visitor's engage with them, and what impact they have on visitors' interpretation of at-site heritage experiences, but also how they may be integrated into online systems for remote heritage consumption. Finally we encourage further study on the changing ecosystem surrounding volunteer staff at a cultural heritage institution that adopts cAvs to understand the support they would need, skills to develop, and exploring the interaction between volunteer and visitor mediated with and through a cAv.

8. ACKNOWLEDGMENTS

This research was funded by Cardiff University's AHRC Impact Acceleration Account. We thank the School of Computer Science and Informatics at Cardiff University. We thank all our workshop participants, and the Bath Assembly Rooms for providing the venue to host our workshop.

REFERENCES

Abdelnour Nocera, J., L. Nielsen, and C. Li (2024, October). Introducing and Evaluating Service Design Tools in Organisations. In *Adjunct Proceedings of the 2024 Nordic Conference on Human-Computer Interaction*, NordiCHI '24 Adjunct, New York, NY, USA, pp. 1–3. Association for Computing Machinery.

Alsford, S. and D. Parry (1991, March). Interpretive theatre: A role in museums? *Museum Management and Curatorship*. Publisher: Taylor & Francis Group.

Antoniou, A., M. Vayanou, A. Katifori, A. Chrysanthi, F. Cheilitsi, and Y. Ioannidis (2021, December). "Real Change Comes from Within!": Towards a Symbiosis of Human and Digital Guides in the Museum. *J. Comput. Cult. Herit. 15*(1), 15:1–15:19.

Bardzell, S., J. Bardzell, J. Forlizzi, J. Zimmerman, and J. Antanitis (2012, June). Critical design and critical theory: the challenge of designing for provocation. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, New York, NY, USA, pp. 288–297. Association for Computing Machinery.

Bender, E. M., T. Gebru, A. McMillan-Major, and S. Shmitchell (2021, March). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, New York, NY, USA, pp. 610–623. Association for Computing Machinery.

Cameron, H. R., G. Reyes-Cruz, A.-M. Piskopani, P. Barnard, A. Boudouraki, P. Caleb-Solly, S. D. Castle-Green, J. E. Fischer, R. Hyde, A. Kucukyilmaz, and H. A. Maior (2025, April). Acceptability, Acceptance and Adoption of Telepresence Robots in Museums: The Museum Professionals' Perspectives. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, New York, NY, USA, pp. 1–18. Association for Computing Machinery.

Chen, B., R. Wen, S. Tan, and Y. Li (2025, April). Exploring User Preferences for Museum Guides: The Role of Chatbots in Shaping Interactive Experiences. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '25, New York, NY, USA, pp. 1–8. Association for Computing Machinery.

Ch'ng, E., S. Cai, P. Feng, and D. Cheng (2023, April). Social Augmented Reality: Communicating via Cultural Heritage. *J. Comput. Cult. Herit. 16*(2), 29:1–29:26.

Crane, S. A. (1997). Memory, Distortion, and History in the Museum. *History and Theory 36*(4), 44–63. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/0018-2656.00030.

Damala, A., I. Ruthven, and E. Hornecker (2019a, January). The MUSETECH Companion : Navigating the Matrix. Num Pages: 934339 Place: Glasgow Publisher: University of Strathclyde.

Damala, A., I. Ruthven, and E. Hornecker (2019b, February). The MUSETECH Model: A Comprehensive Evaluation Framework for Museum Technology. *J. Comput. Cult. Herit. 12*(1), 7:1–7:22.

Desai, S., C. Z. Wei, J. Sin, M. Dubiel, N. Zargham, S. Ahire, M. Porcheron, A. Kuzminykh, M. Lee, H. Candello, J. E. Fischer, C. Munteanu, and B. R. Cowan (2024, May). CUI@CHI 2024: Building Trust in CUIs—From Design to Deployment. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '24, New York, NY, USA, pp. 1–7. Association for Computing Machinery.

Döbler, M., R. Mahendravarman, A. Moskvina, and N. Saef (2024, March). Can I trust You? LLMs as conversational agents. In A. Deshpande, E. Hwang, V. Murahari, J. S. Park, D. Yang, A. Sabharwal, K. Narasimhan, and A. Kalyan (Eds.), *Proceedings of the 1st Workshop on Personalization of Generative AI Systems (PERSONALIZE 2024)*, St. Julians, Malta, pp. 71–75. Association for Computational Linguistics.

Ehsan, U., Q. V. Liao, M. Muller, M. O. Riedl, and J. D. Weisz (2021, May). Expanding Explainability: Towards Social Transparency in AI systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, pp. 1–19. Association for Computing Machinery.

Gallotta, R., G. Todd, M. Zammit, S. Earle, A. Liapis, J. Togelius, and G. N. Yannakakis (2024). Large language models and games: A survey and roadmap. *IEEE Transactions on Games*, 1–18.

Geninatti Cossatin, A., N. Mauro, F. Ferrero, and L. Ardissono (2025). Tell me more: Integrating llms in a cultural heritage website for advanced information exploration support. In *Information Technology & Tourism*, Volume 27, pp. 385–416.

Hall, K., P. Arora, R. Lowy, and J. G. Kim (2024, May). Designing for Strengths: Opportunities to Support Neurodiversity in the Workplace. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA, pp. 1–14. Association for Computing Machinery.

Hede, A.-M., R. Garma, A. Josiassen, and M. Thyne (2014, January). Perceived authenticity of the visitor experience in museums. *European Journal of Marketing 48*(7/8), 1395–1412. Publisher: Emerald Group Publishing Limited.

Heisler, M. and C. Becker-Asano (2025). Conversations with andrea: Visitors' opinions on android robots in a museum. *arXiv preprint arXiv:2506.22466*.

Hou, X., Y. Zhao, Y. Liu, Z. Yang, K. Wang, L. Li, X. Luo, D. Lo, J. Grundy, and H. Wang (2024, December). Large language models for software engineering: A systematic literature review. *ACM Trans. Softw. Eng. Methodol. 33*(8).

Jackson, A. and H. R. Leahy (2005, November). 'Seeing it for real . . . ?'—Authenticity, theatre and learning in museums This article draws on the combined efforts of the research team: Anthony Jackson, Helen Rees Leahy, Paul Johnson (Research Assistant, Centre for Applied Theatre Research, Manchester University) and Verity Walker (museum consultant and director of 'Interpret-action'). *Research in Drama Education: The Journal of Applied Theatre and Performance 10*(3), 303–325. Publisher: Routledge _eprint: https://doi.org/10.1080/13569780500275956.

Joskowicz, J. and D. Slomovitz (2024, February). Engineers' Perspectives on the Use of Generative Artificial Intelligence Tools in the Workplace. *IEEE Engineering Management Review 52*(1), 258–267.

Lin, W. and X. Hu (2025, April). GenAI-Supported Creative Learning in Digital Museum Education: A Case Study of Maritime Art Painting Creation. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '25, New York, NY, USA, pp. 1–8. Association for Computing Machinery.

Longair, S. (2015, January). Cultures of Curating: The Limits of Authority. *Museum History Journal 8*(1), 1–7. Publisher: Routledge _eprint: https://doi.org/10.1179/1936981614Z.00000000043.

Maity, S. and M. J. Saikia (2025). Large language models in healthcare and medical applications: A review. *Bioengineering 12*(6).

Murphy, O. and E. Villaespesa (2020, January). AI: A Museum Planning Toolkit. Report, Goldsmiths, University of London, London. ISBN: 9781913380212 Num Pages: 24.

Padilla Engstrøm, M. and A. S. Løvlie (2025). Using a large language model as design material for an interactive museum installation. *arXiv preprint arXiv:2503.22345*. Work-in-Progress paper accepted at DIS 2025.

Prentice, R., S. Guerin, and S. McGugan (1998, February). Visitor learning at a heritage attraction: a case study of *Discovery* as a media product. *Tourism Management 19*(1), 5–23.

Qian, J., L. Dong, Y. Shen, F. Wei, and W. Chen (2022, February). Controllable Natural Language Generation with Contrastive Prefixes. arXiv:2202.13257 [cs].

Quinto Lima, S., G. Buraglia, W. Kam-Kwai, and J. Roberts (2025, April). Data Bias Recognition in Museum Settings: Framework Development and Contributing Factors. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, New York, NY, USA, pp. 1–15. Association for Computing Machinery.

Raiaan, M. A. K., M. S. H. Mukta, K. Fatema, N. M. Fahad, S. Sakib, M. M. J. Mim, J. Ahmad, M. E. Ali, and S. Azam (2024). A review on large language models: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access 12*, 26839–26874.

Ronge, R., M. Maier, and B. Rathgeber (2025, March). Towards a Definition of Generative Artificial Intelligence. *Philosophy & Technology 38*(1), 31.

Salminen, J., S.-G. Jung, J. M. Santos, A. Mohamed Sayed Kamel, and B. J. Jansen (2021, May). Picturing It!: The Effect of Image Styles on User Perceptions of Personas. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, pp. 1–16. Association for Computing Machinery.

Salminen, J., K. Wenyun Guan, S.-G. Jung, and B. Jansen (2022, April). Use Cases for Design Personas: A Systematic Review and New Frontiers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, pp. 1–21. Association for Computing Machinery.

Seale, N. (2023, October). Education, Entertainment, and Engagement in Museums in the Digital Age. In *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY Companion '23, New York, NY, USA, pp. 326–329. Association for Computing Machinery.

Shehade, M. and T. Stylianou-Lambert (2020, January). Virtual Reality in Museums: Exploring the Experiences of Museum Professionals. *Applied Sciences 10*(11), 4031. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute.

Siricharoen, W. V. (2021). Using Empathy Mapping in Design Thinking Process for Personas Discovering. In P. C. Vinh and A. Rakib (Eds.), *Context-Aware Systems and Applications, and Nature of Computation and Communication*, Cham, pp. 182–191. Springer International Publishing.

Slattery, P., A. K. Saeri, and P. Bragge (2020, February). Research co-design in health: a rapid overview of reviews. *Health Research Policy and Systems 18*(1), 17.

Sun, Y. and T. Wang (2025). Be friendly, not friends: How llm sycophancy shapes user trust.

Trichopoulos, G., M. Konstantakis, M. Loupis, E. Loukis, and D. Michail (2023). Crafting a museum guide using chatgpt4. *Big Data and Cognitive Computing 7*(3), 148.

Weinberger, J. and A. P. Felt (2016). A week to remember: The impact of browser warning storage policies. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, pp. 15–25.

White, J., Q. Fu, S. Hays, M. Sandborn, C. Olea, H. Gilbert, A. Elnashar, J. Spencer-Smith, and D. C. Schmidt (2023, February). A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT. arXiv:2302.11382 [cs].

Xiao, G., J. Tang, J. Zuo, J. Guo, S. Yang, H. Tang, Y. Fu, and S. Han (2024, October). DuoAttention: Efficient Long-Context LLM Inference with Retrieval and Streaming Heads. arXiv:2410.10819 [cs].