Low-power Machine Learning for Visitor Engagement in Museums

Marcus Winter[1][a], Lauren Sweeney[1][b], Katie Mason[1][c] and Phil Blume[2][d]
[1]*School of Architecture, Technology and Engineering, University of Brighton, Brighton, BN2 4GJ, U.K.*
[2]*The Regency Town House, Hove, BN3 1EH, U.K.*

Keywords: Machine Learning, Human Pose Estimation, Embodied Interaction, Visitor Engagement, Museums.

Abstract: Low-power Machine Learning (ML) technologies that process data locally on consumer-level hardware are well suited for interactive applications, however, their potential for audience engagement in museums is largely unexplored. This paper presents a case study using lightweight ML models for human pose estimation and gesture classification to enable visitors' engagement with interactive projections of interior designs. An empirical evaluation found the application is highly engaging and motivates visitors to learn more about the designs. Uncertainty in ML predictions, experienced as tracking inaccuracies, jitter, or gesture recognition problems, have little impact on their positive user experience. The findings warrant future research to explore the potential of low-power ML for visitor engagement in other use cases and heritage contexts.

1 INTRODUCTION

Machine Learning (ML) in cultural heritage is typically concerned with enhancing collections or supporting museum operations. Its potential for audience engagement is largely unexplored, despite the recent emergence of low-power ML technologies (Goel et al., 2020) and open ML model repositories, enabling low-budget development of interactive ML applications that run on consumer-level hardware and avoid privacy issues by processing data locally.

This paper describes a case study developing and evaluating an interactive ML application using these technologies in a heritage context to promote visitors' engagement and learning. Its main contributions are:

- a prototype application for visitor engagement, based on low-power ML technologies;
- an empirical evaluation focusing on visitors' reaction to uncertainty in ML predictions; user experience and visitor engagement.

The following sections contextualise the work, present the prototype application, describe the evaluation study and discuss its findings. The paper concludes by considering limitations of the work and setting out future research directions.

2 RELATED WORK

Museums have long been testbeds for novel technologies, however, their uptake of ML is difficult to assess. French and Villaespesa (2019:102) point out that "the broad language used to describe AI initiatives makes searching for use cases a daunting task". Based on a survey of 61 AI initiatives in museums, they focus their discussion on three application areas: computer vision to enhance collections data, ML for visitor data, and voice assistants for visitor engagement.

Responses to the MAIA survey (Hughes-Noehrer, Jay and Gilmore, 2022:Q8) allow for a similar classification, with 30% describing applications for collections, including enhancing metadata, image tagging, text extraction and OCR; 17% describing applications to support museum operations, including resource planning, ticketing and programming based on visitor data; and 20% describing applications that can be related to visitor engagement such as interactive or personalised experiences, or production of exhibits (other answers mention technologies rather than applications). Contrasting the current 20% of AI initiatives focusing on visitor engagement, 76%

[a] https://orcid.org/0000-0001-6603-325X
[b] https://orcid.org/0000-0001-8312-1550
[c] https://orcid.org/0000-0001-9532-9375
[d] https://orcid.org/0000-0002-3211-3419

Winter, M., Sweeney, L., Mason, K. and Blume, P.
Low-power Machine Learning for Visitor Engagement in Museums.
DOI: 10.5220/0011585600003323
In *Proceedings of the 6th International Conference on Computer-Human Interaction Research and Applications (CHIRA 2022)*, pages 236-243
ISBN: 978-989-758-609-5; ISSN: 2184-3244

of respondents think that AI can enhance the visitor experience on-site, or will be able to in the future (Hughes-Noehrer, Jay and Gilmore, 2022:Q15).

Clusters of past research using ML technologies for visitor engagement include:

- Chatbots offering conversational interfaces to explore collections (e.g., Boiano et al. 2003; Mollica, 2017; Anne Frank House, 2017)
- Robots engaging visitors in the gallery space (e.g., Burgard et al., 1999; Pitch et al. 2011; Del Vacchio, Laddaga and Bifulco, 2020)
- Interfaces to explore ML-enhanced collections (e.g., The Metropolitan Museum of Art, 2022; Harvard Art Museums, 2022a)
- Novel exhibits that would not be possible with conventional technologies (e.g., Tate, 2016; Mihailova, 2021).

While these efforts show that ML already plays an important role in museums, they also suggest a need for research exploring the potential of emergent low-power ML technologies for visitor engagement, which so far have received little attention.

3 HERITAGE CONTEXT

The Regency Town House, built in the 1820s as part of architect Charles Augustine Busby's *Brunswick Estate*, is a museum and heritage centre with a focus on the architecture and social history of Brighton & Hove between the 1780s and 1840s. Among its numerous collections it holds Busby's architectural plans, manuscripts, and aquatints, including an original copy of his 1834 publication, *Collection of Designs for Modern Embellishments*.

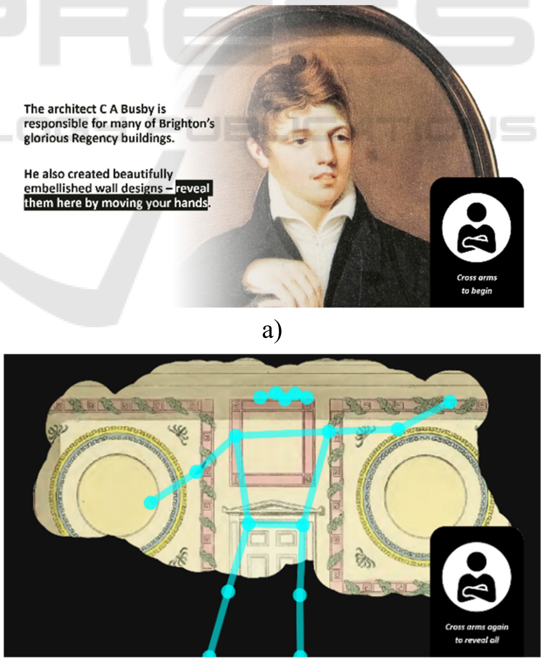
While Busby is mainly known as a designer of buildings, this book features his work on interior designs, which has so far received little attention. To celebrate his interior designs in an authentic setting, the brief was to develop an accessible, immersive, and engaging application that would enable visitors to interact with Busby's colourful and elaborate wall designs in a way that is both educational and fun.

4 PROTOTYPE APPLICATION

The developed prototype uses light-weight ML models for human pose estimation and gesture classification to support embodied interaction (Dourish, 2001) with interactive projections, which has been shown to facilitate meaningful experiences (Tan and Chow, 2017) and lead to high levels of engagement (Lindgren et al., 2016). To satisfy a requirement for cross-platform support and enable other heritage organisations to access the application, the prototype is implemented as a web application. It runs in standards-compliant web browsers and does not require any specialist hardware.

Human pose estimation is used for embodied interaction with designs, turning users' hands into virtual paint brushes to successively reveal designs. A representation of the user's body is shown on screen to support coordination (Figure 1b). The prototype uses the MoveNet pose detection model (Voter and Li, 2021), which provides sufficient performance on most mid-range computers available today.

Gesture recognition is used for control operations. A single trigger gesture is used to start a session from the initial information screen (Figure 1a); progress from reveal action to showing the complete design; and progress from beholding a complete design to loading the next design to reveal. The prototype uses a pre-trained classifier based on body pose data in the COCO format (Lin et al., 2014) and a k-nearest neighbours (k-NN) algorithm. The classifier issues W3C standard Document Object Model (DOM) events for detected gestures, which are listened to by the web application to control the application state.



a)

b)

Figure 1: Interactive wall projection, including (a) initial information screen first encountered by visitors, (b) partly revealed design, with user's body representation and a hint explaining the control gesture in the bottom right corner.

Both MoveNet and the gesture classifier use Tensorflow.js (Smilkov et al., 2019), a JavaScript implementation of the Tensorflow open-source software library for ML (Abadi et al., 2016). The prototype has no other dependencies but uses plain HTML, CSS and JavaScript to ensure standards compliance and minimal overhead.

An admin interface provides functionality to control and customise the application. This includes loading content images, information screens, interaction hints and gesture classifier via URLs, enabling heritage organisations to customise the application for their specific use context and host critical resources in their own web space.

5 EVALUATION

The prototype application was empirically evaluated at the Town House over two days in June 2022, focusing on (i) usability and user experience, (ii) visitors' reaction to uncertainty in ML predictions, and (iii) visitor engagement and learning. The study design was scrutinised by the University of Brighton's ethics board and received a favourable opinion.

5.1 Participants

Participants were recruited through the Town House blog and social media channels to reach their regular audiences. The age distribution in the sample (n=31) is broadly comparable to the average age distribution among heritage audiences in England (DCMS, 2020) with the exception of 35-44 years (6.5%) and 45-54 years (32.3%), which are typically both around 16% (ibid). Participants' frequency of visiting museums and historic buildings (55% up to 5 times per year, 32% 6-15 times, 13% more than 15 times) is higher than the average among heritage audiences in England (ibid). Some participants (10%) say they usually visit on their own, while most visit with friends or family (42%), or say they do both (48%), suggesting that many participants understand museum visits as a social occasion.

5.2 Experimental Setup

The system was deployed in The Regency Town House first floor front dining room (Figure 2), using a notebook computer with Intel Core i7 processor, 16GB RAM and integrated webcam, and a projector with 1920 x 1080 pixels resolution and 6,000 lumen brightness.

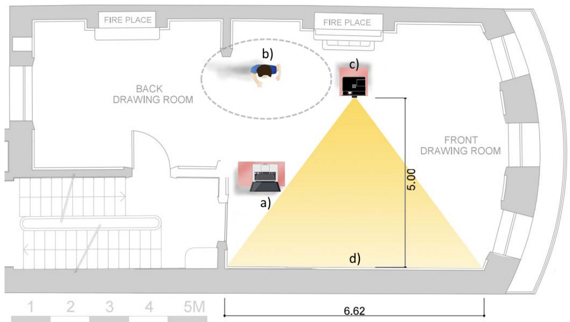


Figure 2: Room layout with positions of (a) computer and webcam, (b) visitor and approximate interaction space, (c) projector and (d) interactive projection.

The projector was fitted with a wide-angle lens to cover the whole sidewall, measuring 6.62m width x 3.72m height excluding bottom skirting, from a distance of 5m. The spatial layout was determined by practical constraints preventing the projector to be mounted at sufficient height for the projection beam to clear visitors interacting in the area in front of the wall. Instead, the projector was mounted on a pedestal at 0.5m height and visitors positioned themselves to the side of it to get a good view of the projection. The computer with integrated webcam was mounted on a small table at 0.65m height facing the user. It was positioned at maximum distance to the user while still clearing the projector beam to not throw a shadow. This resulted in an effective interaction space of approximately 3.4m x 2.2m in which users could operate and their full body was recognised.

5.3 Procedure and Data Collection

Visitors arriving at the Town House were welcomed by a researcher and informed about the context and purpose of the study, before ascending to the first floor dining room. Here Town House staff introduced them to the prototype application, providing information about Busby's interior designs and their relation to the Town House. Participants were also given some initial instructions to get them started with the embodied interaction. A researcher present in the room observed visitors' interaction with the prototype. After visitors were finished with their interaction, they were invited by Town House staff to take part in a short interview about their experience. Interviews were carried out by a researcher in the ground floor reception room. After the interview, participants were asked to complete a short user experience questionnaire.

Observations involved a researcher positioning themselves at the far end of the room to observe visitors' interaction, which typically lasted 5-15 minutes, and take notes recording their initial reactions, level of engagement, learnability and usability issues, and any comments and feedback directed at Town House staff present in the room.

Interviews were designed to last 3-5 minutes and included a total of five questions, plus a section on demographic information. Following guidance in Valenzuela and Shrivastava (2008), the researchers took care to make participants feel comfortable and avoided or elucidated technical jargon during the interview. Answers were recorded in a bespoke coding sheet and notes were revised and clarified immediately after the interview while memory was still fresh.

The questionnaire was based on the short form of the UEQ User Experience Questionnaire developed by Schrepp, Hinderks and Thomaschewski (2017). It was filled in by participants after the interview and took 2-3 minutes to complete.

Observation notes, interview notes and completed questionnaires were digitised after the evaluation event to support the data analysis, while the original paper copies were destroyed.

5.4 Data Analysis

Qualitative observation notes and interview notes were transcribed and analysed independently by three researchers using an emergent coding process described in Miles and Huberman (1994). This involved first a data reduction step and then a data visualisation step to identify common themes. Themes identified in the three separate analyses were then discussed and synthesised by all three researchers together, using affinity diagrams as described in Courage and Baxter (2005). This resulted in a set of consolidated themes from these two datasets and allowed for triangulation between observed interaction (observation notes) and self-reported views about the experience (interview notes) to informing the findings.

Quantitative demographic data, as well as two quantitative sub-questions in the interview asking interviewees to rate on a fixed scale the accuracy of the pose estimation and the impact any inaccuracies had on their overall experience, were simply aggregated.

Quantitative data from the short version of the UEQ were analysed using the UEQ Data Analysis Tool developed by Schrepp (2017).

5.5 Findings

5.5.1 Usability and User Experiences

Observations show that most participants had little or no problems interacting with the system. While some were initially unsure how to start and asked staff for instructions, they quickly picked up on how the interaction works and confidently used both body poses and the cross-arms gesture to interact with the system. Overall, the observed behaviour suggests a good degree of usability and learnability in the current design, even though some aspects could be further improved, in particular, the gesture recognition (see discussion of uncertainty below). This is supported by interview data, with almost all respondents reporting that it was clear how the interaction works, even though some qualified this by attributing it to the brief introduction by staff, or observing others before interacting themselves, or saying it took them a while to work out what to do.

Some participants reported that they found it difficult to reach the corners and lower areas of designs, or remarked that the embodied interaction was physically challenging, confirming similar reports in the literature (Hincapié-Ramos et al., 2017; Jang et al., 2017). A few participants linked this to accessibility, pointing out that it was more difficult to use for people at their age, and for others with restricted mobility or dexterity. One suggestion to address these issues was to not only using users' hands as virtual paint brushes but also their feet. Alternative interaction modes were also suggested, for example offering people with mobility issues the opportunity to reveal the projected designs via a touch screen.

Several interview responses suggested alternative ways to provide initial instructions for situations where staff cannot be present (e.g., in writing; short animation). They also commented on the usefulness, design, or timing of the on-screen hints reminding users to cross their arms at certain points. Observations suggest these hints where often unnoticed or ignored by users, with some preferring to continue revealing a design with their hands rather than cutting short the process by crossing their arms. Some responses also questioned the need for an on-screen body representation or suggested alternative representations. All of these comments, however, were in the spirit of further improving the prototype rather than suggesting any fundamental problems in its interaction design.

Regarding their overall user experience, many participants were observed to exhibit behaviour or make comments indicating positive approval, using terms like "amazing" or "extraordinary" or "really cool". A similar picture emerges from interviews, with many participants emphatic about the experience being fun and engaging. Several interviewees suggested that children would appreciate the embodied interaction and some suggested gamifying the experience to make it even more engaging.

Results from the UEQ further support these findings, with particularly positive scores for scales relating to ease of use and stimulation (Table 1). Mean scores of 1.716 (SD=0.928) for pragmatic quality, 1.500 (SD=1.397) for hedonic quality and 1.602 (SD=1.074) overall put the prototype in the top 10% of scores in the UEQ benchmark dataset of 468 studies involving 21,175 people (Schrepp, 2017), which indicates an excellent user experience overall (Figure 3).

Table 1: UEQ scores and confidence intervals (p=0.05).

Scale		Mean	Conf. Interval	
obstructive	supportive	1.21	0.74	1.69
complicated	easy	2.14	1.73	2.55
inefficient	efficient	1.31	0.89	1.74
confusing	clear	2.17	1.79	2.55
boring	exciting	1.14	0.56	1.73
not interesting	interesting	1.76	1.22	2.30
conventional	inventive	1.79	1.16	2.41
usual	leading edge	1.46	0.89	2.04

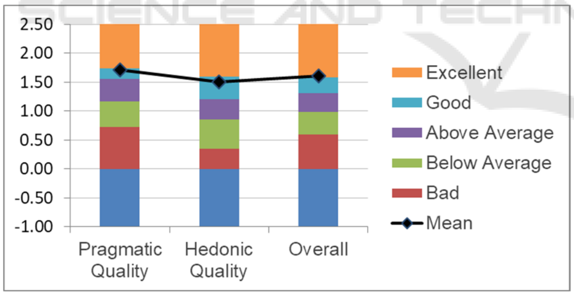


Figure 3: UEQ mean scores against benchmark dataset.

5.5.2 Uncertainty in ML Predictions

Uncertainty in ML predictions manifests itself in the prototype in various ways. In the context of human pose estimation, low confidence scores can lead to some body parts not being rendered on screen. Inaccurate estimations lead to a mismatch between on-screen representation and actual body pose, whereas slight differences in estimations between video frames lead to jitter in the on-screen body representation. In the context of gesture classification, uncertainty can lead to false positives (i.e., a body pose is wrongly classified as a gesture) and false negatives (i.e., a correctly performed gesture is not recognised).

Observations show that the cross-arms gesture was often triggered accidentally (false positives) and, conversely, in some cases did not trigger as expected (false negatives). The former was typically caused by (a) visitors crossing their arms as part of their normal body posture; (b) wrong classification of other body poses involving crossing an arm over the torso; and (c) poor pose estimation due to lighting conditions, loose clothes or shoulder bags obscuring body parts. The latter was typically triggered by visitors performing the cross-arms gesture in a different manner than the classifier was trained to recognise (e.g., several visitors performed an "X" with lower arms crossing diagonally, while the classifier was trained mostly with lower arms crossed almost horizontally). Observations also show tracking inaccuracies and body parts not rendered on screen due to low confidence scores, for example when people performed extreme body poses to explore the limitations of the system. Overall, these aspects had little impact on visitors' engagement. While there were some reactions expressing surprise or even amusement, visitors generally worked through those situations unperturbed and seemed to accept them as part of the experience.

This is supported by interviews, which asked visitors to rate both the application's tracking accuracy and the impact of inaccuracies on their experience, and to expand on their ratings with open comments. Figure 4 shows that participants' ratings tend towards more positive assessments of both aspects. However, while more respondents rate the tracking accuracy as good (52%) rather than perfect (16%), this trend is reversed in how tracking inaccuracies impacted on the experience, with more respondents saying they had no impact (45%) rather than little impact (13%), suggesting a certain level of acceptance of the effects of uncertainty.

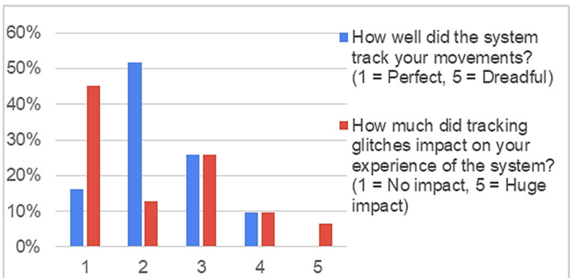


Figure 4: Visitors' ratings of how well the application tracked their movements and how much any tracking inaccuracies impacted on their experience.

Open answers describe issues as experienced by participants. While some of these relate to uncertainty in ML predictions, others are caused by ML inference latency, or are the result of interaction design decision aimed at mitigating some of the challenges of embodied interaction:

- Some participants remarked on the body representation sometimes being inaccurate or out of proportion or not at scale. This is only partly caused by uncertainty in ML predictions, as the prototype also makes purposeful adjustments that enable all participants to reach the top of the screen regardless of their body height or distance from the camera.
- Some respondents remarked on jitters in their on-screen body representation, which is caused by slight variations in key point predictions between video frames.
- Some respondents remarked on lag or delay, which is caused by latency in ML inference, and is invers proportional to the processing power of the machine it runs on.

Overall, the data shows that several aspects of the user experience are affected by ML uncertainty or latency, however, it also suggests a high level of acceptance and willingness to work around any issues. Some interviewees even suggested that glitches made the experience more interesting as they added an element of unpredictability.

5.5.3 Visitor Engagement and Learning

Observations show how several visitors methodically reveal designs, often over prolonged periods of time, engaging with one design after an other, directing enquiries about specific design features and colours to staff during their interaction, and talking about designs and the differences between them with staff and other visitors after their interaction. While most participants are clearly excited about the technology, they also show their appreciation for the designs they reveal, suggesting that the prototype overall manages to "preserve the primacy of the object and aesthetic encounter" (vom Lehn and Heath, 2003, p.3) and raises visitors' interest in the designs, motivating them to enquire and learn about them.

Interview data strongly supports this, with overwhelmingly positive answers when asked whether the prototype was an engaging way to learn about Busby's interior designs. Several participants expressed their excitement at the vivid colours of the designs and at seeing them projected at scale in an authentic environment. While there were two responses saying they would prefer static images or written materials, most expressed their satisfaction at the interactive and immersive nature of the prototype, with some explicitly stating that it was more fun to reveal the designs rather than just observing them.

Several participants pointed out the value of additional narrative delivered by staff, either as part of the introduction, or commenting about specific designs, or answering questions by participants during or after their interaction. This reinforces the notion that the interaction awakens interest and motivates visitors to learn about the designs by asking staff for more information. It also indicates that the prototype can be part of a wider engagement and learning strategy in museums involving staff or other experts who can provide additional information and support a more conversational form of learning.

6 DISCUSSION

The empirical evaluation suggests good usability and learnability of the developed prototype This was supported by participants receiving initial instructions by staff, the direct mapping between users' body pose and on-screen representation, and on-screen hints when to use the cross-arms gesture. Feedback suggests a need to improve the gesture recognition and explore the possibility of users customising or hiding their body representation.

While the presence of staff clearly added value to the experience, this might not always be feasible, especially in smaller museums with resource constraints. Future work should explore how the prototype can be further developed for use cases with no staff present. Feedback suggesting to gamify the experience, and observations of visitors' persistence in trying to reveal - or "complete" - all areas of a design, indicate a natural challenge to build on.

Uncertainty in ML predictions, perceived as tracking inaccuracies, jitter in the on-screen body representation, and gesture recognition problems, had little impact on participants' overall user experience, who generally showed a high level of tolerance, simply trying again when something did not work as expected. As this behaviour is likely to be influenced by staff readily providing hints and explanations, future work should explore if this tolerance holds without staff being present.

Despite these issues, the UEQ results suggest an excellent overall user experience against benchmark data (Schrepp, 2017; Schrepp, Hinderks and Thomaschewski, 2017). The high score for pragmatic quality supports findings indicating good usability, while the high score for hedonic quality supports findings that participants enjoyed the experience. They add further support to literature on the engaging qualities of gesture-based interfaces (van Beurden, Ijsselsteijn and de Kort, 2012) and show that ML uncertainty does not diminish their appeal.

From a museum perspective, key questions include whether the application supports visitors' engagement and learning, and whether its use of ML is a feasible alternative to specialist hardware. The results show that the application is highly engaging and motivates visitors to ask questions about designs and learn more about Busby's work. Participants had no privacy concerns about being observed by a camera, and while they noticed the effects of uncertainty in ML predictions, these had little impact on their positive experience.

The findings provide useful insights informing future development and research. While not allowing for extrapolation to other use cases or contexts, they give an indication of the potential of low-power ML as an enabling technology for visitor engagement and provide a snapshot of related user experience issues.

7 LIMITATIONS

Aiming for high ecological validity, the evaluation study took place in the intended target environment and involved participants recruited via The Regency Town House's social media channels. The sample size and composition, choice of methods and rigorous data analysis ensure high internal validity, however, the bespoke nature of the prototype and the evaluation environment make it problematic to generalise findings to other applications and environments. As such, no recommendations or design guidelines are offered for low-power ML applications for visitor engagement.

8 CONCLUSIONS

This paper describes a prototype application using human pose estimation and gesture recognition for visitor engagement with interior designs in a heritage setting. Unlike other applications involving embodied interaction, it does not require specialist hardware but uses pre-trained ML models and runs on a mid-range computer with a webcam, putting it into reach for smaller museums with limited budgets and development capabilities.

The prototype uses low-power ML technologies, which are particularly well suited for interactive applications as they process data locally rather than transmitting to a server, reducing latency and preserving visitors' privacy.

An empirical evaluation in the intended target environment found it usable, learnable and offering an excellent overall user experience. Besides engaging visitors of all ages, it motivated them to ask questions about the interior designs they revealed and to learn more about them in informal conversations with staff. Uncertainty in ML predictions, perceived by visitors as tracking inaccuracies, jitter in their on-screen representation and gesture recognition issues, had little impact on their positive experience.

The findings indicate that low-power ML holds great promise for visitor engagement in heritage contexts and warrant future research to explore this potential. This includes developing designs that can run unsupervised in the gallery space, without staff being present to provide information and assistance, and exploring how other ML capabilities can support visitor engagement and learning in museums.

ACKNOWLEDGEMENTS

This research was supported by the Ignite funding scheme of the Community University Partnership Programme (CUPP) at the University of Brighton. We would like to thank visitors to The Regency Town House for taking part in the evaluation and sharing their valuable views and feedback.

REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X. (2016). TensorFlow: A System for Large-Scale Machine Learning. Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16). arXiv:1605.08695.

Anne Frank House (2017). Anne Frank House launches bot for Messenger. Available https://www.annefrank.org/en/about-us/news-and-press/news/2017/3/21/anne-frank-house-launches-bot-messenger/. Retrieved 6 August 2022.

Boiano S, Gaia G, Caldarini M (2003) Make your museum talk: natural language interfaces for cultural institutions. Museums and the Web 2003. Available https://www.museumsandtheweb.com/mw2003/papers/gaia/gaia.html. Retrieved 6 August 2022.

Burgard, W., Cremers, A. B., Fox, D., Hähnel, D., Lakemeyer, G., Schulz, D., Steiner, W., and Thrun, S. (1999). Experiences with an interactive museum tour-guide robot. Artificial Intelligence, 114(1–2), 3–55

Courage, C. and Baxter, K., 2005. Understanding your users: A practical guide to user requirements methods, tools, and techniques. Gulf Professional Publishing.

DCMS (2020). Taking Part 2019/20: Cross-sectional survey. Technical Report. Available https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/916246/Taking_Part_Technical_Report_2019_20.pdf. Retrieved 27 July 2022.

Del Vacchio, E., Laddaga, C., & Bifulco, F. (2020). Social robots as a tool to involve student in museum edutainment programs. In Proceedings of the 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 476-481.

Dourish, P. (2001). Where the action is: the foundations of embodied interaction. MIT Press, Cambridge, Mass.

French, A. and Villaespesa, E. (2019). AI, visitor experience, and museum operations: a closer look at the possible. In Humanizing the Digital: Un-proceedings of the MCN 2018 Conference, 101-113.

Gaia, G., Boiano, S., and Borda, A. (2019). Engaging museum visitors with AI: The case of chatbots. Museums and digital culture, 309-329. Springer.

Goel, A., Tung, C., Lu, Y. H., and Thiruvathukal, G. K. (2020). A survey of methods for low-power deep learning and computer vision. In 6th World Forum on Internet of Things (WF-IoT), 1-6. IEEE.

Harvard Art Museums (2022). AI Explorer: Explore how a computer sees art. Available https://ai.harvardartmuseums.org/about. Retrieved 11 August 2022.

Hincapié-Ramos, J. D., Guo, X., Moghadasian, P., and Irani, P. (2014). Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1063-1072. ACM.

Hughes-Noehrer, L., Jay, C. and Gilmore, A. (2022). Museums and AI Applications (MAIA) Survey. University of Manchester. Dataset. https://doi.org/10.48420/19298588.v1

Jang, S., Stuerzlinger, W., Ambike, S., and Ramani, K. (2017). Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 3328-3339. ACM.

Lee, L., Okerlund, J., Maher, M. L., & Farina, T. (2020, July). Embodied interaction design to promote creative social engagement for older adults. In International Conference on Human-Computer Interaction, 164-183. Springer, Cham.

Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., and Dollár, P. (2014). Microsoft COCO: Common objects in context. In Proceedings of the European conference on computer vision (ECCV), 740-755. Springer.

Lindgren, R., Tscholl, M., Wang, S., and Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. Computers & Education, 95, 174-187.

Mihailova, M. (2021). To dally with Dalí: Deepfake (Inter) faces in the art museum. Convergence, 27(4), 882-898.

Miles, M.B., and Huberman, A.M. 1984. Qualitative Data Analysis. Newbury Park, CA: Sage.

Mollica, J. (2017). Send Me SFMOMA. Available https://www.sfmoma.org/read/send-me-sfmoma/. Retrieved 6 August 2022.

Pitsch, K., Wrede, S., Seele, J. C., and Süssenbach, L. (2011). Attitude of german museum visitors towards an interactive art guide robot. In Proceedings of the 6th international conference on Human-robot interaction, 227-228. ACM.

Schrepp (2017). UEQ Data Analysis Tool. Available https://www.ueq-online.org/Material/Short_UEQ_Data_Analysis_Tool.xlsx. Retrieved 26 July 2022.

Schrepp, M., Hinderks, A., and Thomaschewski, J. (2017): Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). IJIMAI, 4 (6), 103–108.

Smilkov, D., Thorat, N., Assogba, Y., Nicholson, C., Kreeger, N., Yu, P., Cai, S., Nielsen, E., Soegel, D., Bileschi, S. and Terry, M. (2019). Tensorflow. js: Machine learning for the web and beyond. Proc. of Machine Learning and Systems, 1, 309-321.

Tan, L., and Chow, K. K. (2017). Facilitating meaningful experience with ambient media: an embodied engagement model. In Proceedings of the 5th International Symposium of Chinese CHI, 36-46.

Tate (2016). Can a machine make us look afresh at great art through the lens of today's world? IK Prize 2016: Recognition. Available https://www.tate.org.uk/whats-on/tate-britain/exhibition/ik-prize-2016-recognition. Retrieved 11 August 2022.

The Metropolitan Museum of Art (2022) The Met Art Explorer. Available https://art-explorer.azurewebsites.net/search. Retrieved 11 August 2022.

van Beurden, M.H., Ijsselsteijn, W.A., de Kort, Y.A. (2012). User Experience of Gesture Based Interfaces: A Comparison with Traditional Interaction Methods on Pragmatic and Hedonic Qualities. LNCS, vol 7206, 36-47. Springer, Berlin.

Voter, R. and Li, N. (2021). Next-Generation Pose Detection with MoveNet and TensorFlow.js. Available https://blog.tensorflow.org/2021/05/next-generation-pose-detection-with-movenet-and-tensorflowjs.html. Retrieved 26 July 2022.

Winter M., Jackson P. (2020) Flatpack ML: How to Support Designers in Creating a New Generation of Customizable Machine Learning Applications. LNCS vol 12201, 175-193. Springer Nature.