



# Engaging Museum Visitors with AI-Generated Narration and Gameplay

Wladimir Hettmann<sup>1(✉)</sup>, Matthias Wölfel<sup>1,2</sup>, Marius Butz<sup>1</sup>, Kevin Torner<sup>1</sup>,  
and Janika Finken<sup>1</sup>

<sup>1</sup> Faculty of Computer Science and Business Information Systems, Karlsruhe  
University of Applied Sciences, Karlsruhe, Germany

{hewl1012,matthias.woelfel,buma1060,token1011,fija1013}@h-ka.de

<sup>2</sup> Faculty of Business, Economics and Social Sciences, University of Hohenheim,  
Stuttgart, Germany

**Abstract.** It has become a challenge in recent years to raise interest in museum visits, especially among younger visitors, as the range of alternative entertainment options has become overwhelming and increasingly attractive, interactive, and playful. To re-engage a wide audience with art and cultural heritage, we propose to use artificial intelligence to make the presented artworks more interesting. By using natural language processing to generate a narrative around a selection of individual exhibits and presenting the story as a scavenger hunt, we connect the individual exhibits and make access more playful. The museum visitors are guided through the story by two characters, who also pose challenges to be solved in mini-games. The two characters were chosen as a living being (a puppy) and an embodied agent (a humanoid robot) to indicate whether an utterance is preformulated and fact-based (puppy) or generated and partially made up. By testing the prototype, we could confirm that the generated stories are plausible and exciting, that the participants became more interested in the presented items through the story and mini-games and that the participants could distinguish if the utterance was fact-based or fictional.

**Keywords:** interactive installation · cultural heritage · game narration · natural language processing · engagement · museum

## 1 Introduction

The last few years, in the cultural sector, have been shaped by applying recent developments in *artificial intelligence* (AI) to specific applications. While mass media coverage has made the use of AI to *generate* or *modify* art popular, AI is used in the cultural sector and museums in many other ways—yet not well known to the public. How AI can support museums has been a topic of interest for some time [31] and many use cases have been identified:

For instance, it can be used to provide alternative ways to access information, e.g. through chatbots that can answer questions [10,25], support accessibility, e.g. by automatic speech translation for foreigners or speech synthesis for the blind, to improve visitor experience, e.g. by predicting what’s most interesting through analytics tools, or to understand collections by using machine vision to help recognize and classify images.

One use case, that is gaining more popularity recently, is the use of AI to automatically sort and select items to be presented online, at cultural institutions, or in museums [5]. Curated content is usually clustered following a particular logic which—consciously or subconsciously—influences the way the recipient perceived the individual items and the overall exhibition. This connection, however, is not traceable for many visitors and it stays unclear why and in what order individual items are displayed. For randomly selected items, this connection is completely missing and items appear as randomly selected and placed, leaving museum visitors with uncorrelated content that is hard to consume and remember.

Curation plays an important role in any exhibition. In this process, a curator manages collections of artifacts and artworks by using his or her expertise to select the right artifacts and objects [26,27]. This selection process chooses only artifacts that are appropriate for the theme of the exhibition. Therefore, many objects are not accessible to visitors, and even the existence of many objects remains unknown and hidden in the catacombs of museums. To overcome this limitation we propose to draw, otherwise hidden, items from a museum database and relate them by generating a narration with and around these items in a playful way. The items and related information provided in the database are not altered as we do not want to create new content but add context by natural language processing.

Access to information, to be appealing to a wide audience of different ages and backgrounds, must be easily accessible, user-friendly, of interest, and entertaining. Due to the COVID-19 pandemic, people are more reluctant to touch surfaces, things, or interfaces. Therefore, to present and access the given information and to play mini-games, we used a large screen and touch-free interactions.

## 2 Recent Work

According to Lu, AI, as used by museums, is divided into three key area [18]:

- AI for data analysis  
To segment and cluster large amounts of data to better find, identify and link data or to analyze the behavior of museum visitors.
- AI for restoration of artworks  
Paintings, photographs, etc. may be damaged and need restoration. AI can improve images, e.g. the tool Colourise to color old photographs [17].
- AI to generate art  
AI can simulate artistic works in many forms: paintings, photographs, music,

videos, and writing, such as GPT-3, to generate text content like that written by humans, or DALL-E, to produce images from text labels [23].

In our opinion, AI can also help to make access to artwork and cultural information easier and more engaging. Museum visitors can access information easier by using speech recognition and overcoming language barriers through automatic translation.

Audience engagement can take place inside and outside the museum and some approaches can be used on-site but also before or after a visit [16]. For instance, sentiment analysis can be used to analyze and interpret visitor comments to understand what is important to engage them [30].

An outside and personalization use case is the prototype approach called *Storyteller* [1] by The Met, Microsoft, and MIT. This prototype uses AI for speech recognition to track discussions or narratives, which are then matched with artwork from The Met collection and then shared. At the end, when the narrative or discussion is finished, there is the possibility of sharing the generated data and collection via social media, generating an AI museum tour based on the selected artworks, or even creating a print-ready book.

Another use case for engagement and education inside and outside of a museum is *Ping! the museum app* [28]. It uses the familiar Tinder mechanic, allowing users to “date” objects based on their interests. This enables personalized and playful exploration of exhibits. By using the conversation feature, users are allowed to get to know exhibits and create their own collections. With the dialog-based concept the knowledge of exhibits is conveyed in a playful way. The user’s own collections can be viewed through an individual, personalized journey through the museum or even after the museum visit. A level concept enables the unlocking of new objects, which in turn can arouse a user’s curiosity.

### 3 Design Considerations

To provide a unique experience through a customized selection of objects, attention must be paid to the order and way these objects are presented and how they are interconnected. Usually, database entries of cultural assets include a textual description, photos, and other additional material, such as 3D scans in the case of objects. They often contain cross-references to related objects that come from the same site, artist, or art style. As discussed earlier, curating content is an important yet difficult task, in setting up an exhibition. However, neither it is possible to have a curator available for every single collection of objects selected by visitors nor would the information provided in the database be sufficient to solve this task automatically. Therefore, to establish an automated process, one must rely on additional knowledge sources. Language models which have been trained on large amounts of data relate different words—and thus meaning—to one another. The idea we propose in this paper is thus to connect items from the database to one another by correlating them according to a given large language model. But we don’t stop here and also use the language model to provide a narration around these items and to link them to mini-game templates.

### 3.1 Design for Semi-public Spaces

Carr et al. describe in their book *Public Space* essential needs, that public spaces comply with [7]. The first is passive engagement, where people stay around 2–3 meters from the interface. They observe how others are interacting with the interface or have short interactions by themselves. In the active engagement zone (around an arm’s length until 2m away from the interface), people are more attended to the interface and have longer and more focused interaction sessions. Within this active engagement zone, social triangulation (social interaction between strangers) can be observed [19]. In addition, when designing a public display within a museum, it is important to strike a balance between attracting the attention of interested people and not annoying those who are not interested. Potential users must be motivated to interact through challenges, curiosity, freedom, and imagination, without being distracted by intrusive visual stimuli [20]. This balancing act has ultimately led us to a subtle but stylized art style, which is accompanied by interesting characters. The teaser in the form of a game plan arouses the interest of the users and their curiosity increases to see what is behind the fields.

### 3.2 Onboarding

To motivate visitors to interact with the installation potential players when coming closer (recognized through a Kinect sensor) are proactively addressed by two characters who lead the users through the whole application. The two characters were chosen to reflect the hidden processes that are either preformulated and fact-based or generated and partially made up. Text, embedded in a speech bubble, that is preformulated by a human and fact-based are uttered by a living being (a puppy) while text generated by an AI-system and that is partially made up is uttered by an embodied agent (a humanoid robot).

In the onboarding phase, see Fig. 1, the player learns the most important thing that is used to get to the next step and some general information about the installation. As it has been shown that embodied agents that react to the presence and body gesture of their conversation partner have been shown to be perceived as more sympathetic [33] we implemented this mechanism here as well.

### 3.3 Narration

After onboarding the user, the items must be selected, the narration generated, and subsections of the generated narration related to the mini-games. In our prototype, we limited our system to draw a total of five items since we have realized four mini-games and the fifth item extends the possibility of getting more random content. To provide items fitting to the mini-games, there must be one painting, one photograph, and three 3D models.

After the selection step, a link between the keywords of the selected articles is established. Since it has been shown that English models provide better results



**Fig. 1.** User is onboarded and familiarised with the swipe gesture

than German models, the first step is to translate the keywords (e.g. offensive weapon, lance, river find, weapon) using DeepL [2].

On this basis, a connection can be made with the generation process for the story. For this, the natural language model Generative Pre-trained Transformer 3 (GPT-3) [22] must be configured. The configuration includes the selection of the engine *text-davinci-001*, the temperature with the value 0.8, a maximum tokens count of 1500, a top p-value of 1, a frequency penalty of 0.5, and the first-shot mode with a given input content (see Listing 1.1).

**Listing 1.1.** First-shot mode input content as simple string concatenation pseudo code for the NLP model engine *text-davinci-001*

```

"Topic: The Adventure \n Adventure Story with given keywords: "
+ keywords
+ "\n A long time ago ... \n"
+ "Adventure Story with given keywords: "
+ keywords

```

On closer inspection, the first-shot mode is one of the settings that must be considered when creating a story, since the narration and the connection depend on this setting. According to these specifications and settings, a narrative story can be generated (e.g. "The shepherds grabbed their weapons and ran outside to see what was happening. They were shocked to see that it was a group of robbers!"), which is translated back into German with the use of DeepL [2].

In the final process step, the generated story is divided into sub-steps. Each sub-step contains two sentences from the story. This partitioning helps to compensate for the limitation imposed by the tokens in the generation process and to generate enough sub-steps to provide the possibility to assign mini-games.

In a second selection process, each mini-game is assigned to an appropriate sub-step. For this, the fastText [4] word embedding model is used. This gets the sentences of the sub-steps as input and then compares the result of the embedding with the respective keywords of an item. During the comparison, a score is generated for each item sub-step combination and the pair with the highest score (best match) is assigned to each other. The process is iterated until all items are assigned to a sub-step.

In addition, each mini-game of a sub-step contains the corresponding item information. This newly generated content comprises the complete game based on the sub-steps. In a way, this process can be considered a process that a curator or game master would normally perform.

### 3.4 Gameplay

It is known that playful elements can motivate learning and improve cognitive performance. After onboarding, the players followed the generated narration step-by-step and in particular sentences had to solve the tasks given in mini-games. The mini-game not only offers a fun factor but also offers the possibility to improve one's memory performance [21]. Not to be neglected are also the emotions evoked, which can have a positive influence on the memory of the museum objects. In Sect. 5 we describe the mini-game templates we have implemented to demonstrate our system. Please note, that the overall system setup can be easily extended by other game mechanisms and mini-games.

After each mini-game, more information about the exhibit is displayed and the score achieved in the game is announced. At the end of the scavenger hunt, a trophy is displayed (see Fig. 2), depending on the total score achieved. A complete run depends on the visitor's behavior and how fast the mini-games are solved. A usual run takes between 10 and 20 min. The system can be used with a group of visitors or by a single person, it just depends on who is closest to the center of the installation (the so-called sweet spot, see Fig. 1).

## 4 Technical Setup

To provide access to the installation for several visitors at the same time, especially families, we decided to use a large projection area that can be controlled touchless, to not only provide a native and intuitive interface but also to take account of the COVID-19 pandemic (people are more afraid now of touching surfaces). The final application was realized on a large 180° screen with a width of 10.25 m and a height of 2.38 m, which is powered by three overlapping projectors compensating for the non-planar surface and overlap by projection mapping. To accept natural user input, we used an *Azure Kinect* and our self-developed software *Kinetic Space*, which allows us to train and recognize 3D gestures [32].



**Fig. 2.** The scavenger hunt is finished, and the player receives an award in the form of a trophy. As the player has not scored enough points during the mini-games, he does not receive a trophy.

The application is powered by a *React*<sup>1</sup> application, which runs in a full-screen browser window. The data for the game is processed in real-time from a *Flask*<sup>2</sup> backend server.

To realize our application, we applied a set of pre-trained machine learning and deep learning models. These pre-trained models were used to provide text translation, generation, comparison, and instance segmentation. The dataset to finetune the pre-trained models was provided by the *Badisches Landesmuseum Karlsruhe*<sup>3</sup> (dataset not publicly available yet).

#### 4.1 Natural Language Processing

The comparison of keywords and text requires the use of word embeddings. Because of useability and good comparison results, we selected fastText as our tool of choice [4].

By comparing the quality of different models, we found that we could get better results (in terms of more meaningful texts) by automatically translating German text into English, processing it (e.g. using an English GPT-3 model),

<sup>1</sup> <https://reactjs.org/>.

<sup>2</sup> <https://flask.palletsprojects.com/en/2.0.x/>.

<sup>3</sup> <https://www.landmuseum.de/>.

and then retranslating it back to German. To translate content from German to English and back we used the DeepL API [2].

GPT-3 is a deep learning model that has learned the capability to generate human-like textual content. The produced quality of content can be so high that even humans cannot distinguish whether it is machine-generated or human-written content [24]. As GPT-3 is very large we access it through the API [6].

## 4.2 Image Processing

For image processing and analysis, we used the instance segmentation model Mask R-CNN with ResNet50 weights, which extends the approach of Faster R-CNN with another branch for the object mask in addition to the bounding box [13]. To achieve a higher mAP score, and thus reduce the miss matches, we additionally create an inference with the object detector model YOLOv5 [15]. To further improve the accuracy of YOLOv5, we transfer-trained the model with photos generated by neural style transfer [11] and by freezing different layers [29]. While this process did not yield significant improvements in prediction quality for paintings in our test set, we decided to use default YOLOv5 weights. For the final prediction, the results predicted by both models are used.

## 5 Mini-Games

In this section, we describe the mini-games we have used to prove our concept. All mini-games serve as templates where the empty spaces are filled by the respective items from the database.

### 5.1 Mini-Game I: Search for Objects

One mini-game is a “classic search game” within a painting, in which the player must find a list of items. To find objects automatically, we used a Mask-RCNN model with ResNet50 weights. Afterward, an inference with the image detector YOLOv5 is performed and a union set is formed with the results of the Mask-RCNN model to obtain fewer miss matches. For inference, we send the corresponding painting to the aforementioned *Flask* server, which performs the steps described and sends a list of bounding boxes and segmentation masks back to the client. The bounding boxes are then used as a hitbox for the player’s pointing gesture and, as soon as an object has been found, the segmentation mask is placed over it in a slightly transparent way so that the player can recognize the objects already found. When all objects have been found or the maximum number of attempts ( $1.5 * N$  where  $N$  is the number of objects in the picture) has been reached, the game ends and a score is determined depending on the remaining attempts. To mark objects in the painting, the player must point with an outstretched arm and index finger to the place where he assumes the object to be and hold it there for 2s to confirm the entry (see Fig. 3).



**Fig. 3.** The player points to the place in the painting where he suspects an object (white circle represents cursor).

## 5.2 Mini-Game II: Rotate Master and Shadow Match

Two other mini-games use an inertial measurement unit, or IMU, as input to the game. The previously selected 3D objects are displayed as virtual objects and controlled via a physical object. Visualization of the objects in both games is done via separate *Unity WebGL* applications. The applications only act as a view and are controlled by the outer web application. A *Genuino 101* was used to capture the rotations. Among other things, this board has an already integrated inertial measurement unit and an integrated Bluetooth module.

Essentially, the board performs two actions: First, the sensor data from the IMU is computed and then the computed angles are forwarded via Bluetooth. This way the application always gets the current angles of the three rotation axes Yaw, Pitch, and Roll.

Putting it all together: The board calculates the current angles and sends them via Bluetooth to the web application. The web application forwards the rotation angles to the *Unity WebGL* application, which can now update the rotation angles of the virtual object.

The first game is based on the well-known game *Simon Says*. Unlike the giving of arbitrary orders, here rotation directions are given, which then must be transferred to the object in the player's hand. A total of three levels must be completed, with each level containing two rotations. Before each level, a preview of the rotations to be imitated is shown.

## Richtige Rotationen: 0 von 2



**Fig. 4.** The *Simon Says* game while rotating the 3D object to the right.

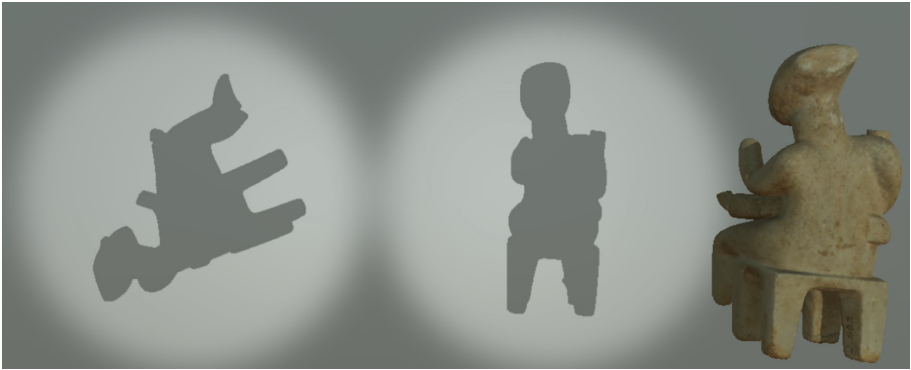
For simplification, only rotations of the pitch and roll axis are considered, only one axis is rotated at a time and the rotation must only reach a certain threshold value to be recognized as correct. An example screenshot is shown in Fig. 4.

In the other game *Shadow Match*, the player tries to rotate the object so that it matches a given virtual shadow. In the view, also shown in Fig. 5, you see two shadows, but on the right, you see the 3D object, which you can rotate. Now you must try to rotate the object so that the right shadow matches the left shadow. Within the *Unity* application, two cameras have been placed in the scene, each pointed at a shadow. To compare the shadows, the two camera images are binarized and a Hamming distance is calculated. Due to the large variety of objects, it is difficult to define a maximum distance that works well for all objects. For simpler objects usually, fewer errors are possible than for more complex objects. This means that different threshold values must be defined.

### 5.3 Mini-Game III: Autumn Floor Projection

Another mini-game developed is based on *Arktos Super Special* from the German TV children's series *Tabaluga Tivi*. The basic idea is to memorize the position of objects randomly scattered on a surface before they are no longer visible. Using one's memories, these objects must be found again. To increase the difficulty level, additional objects have been added that must not be touched. As the description of the game suggests, this game is intended for children and young teenagers.

The text information on the 180° screen informs the user that valuable objects from the museum have been lost. To find them again the player must follow in the footsteps of the dog and the robot. Once there, the lost objects become visible to the user for 7s, but then they are suddenly covered by many colorful leaves due to an autumnal wind - quickly they must be found again within 90s before they get damaged. But watch out! There are many different



**Fig. 5.** Screenshot of the game *Shadow Match*. Left is the target shadow and right is the shadow of the object that is controlled by the player.

animals, such as hedgehogs and marmots, which hibernate in autumn and must not be disturbed.

Therefore, when searching for lost museum objects, do not accidentally step on the animals. When a lost object is discovered, it slowly moves towards the screen, where additional information is displayed. If all the objects can be collected within the time that elapses at the top left of the floor projection, the game is completed successfully, and the player receives full points. However, if the time runs out or a sleeping animal is woken up, then the player only receives the score for the objects already found.

## 6 Conclusion

Preliminary beta tests have demonstrated that the generated narration consisting of 10 to 15 sentences is well perceived and that, most of the time, a consistent story was generated. The story was able to connect the individual items, selected from the database, and individual items fitted well to the single sentence they got assigned to. The use of mini-games with the content of the individual items and gamification supported the story and subjectively made it more entertaining for the beta test group.

The selected different interaction modalities primarily serve for variety to the gameplay and also provide more activity than rigidly playing through console games. By wiping, pointing, and rotating objects, the visitor is inspired to actively participate to progress in the game.

Choosing the scavenger hunt as the game platform from a layering pattern perspective gave us the ability to design and add other mini-games, and benefited from the ability to present the museum's artifacts embedded in a generated narration. The layering pattern was also a design decision in the construction phase of the server/client applications to take advantage of the separation of the generation and usage processes.

## 7 Discussion and Outlook

Besides all the excitement about using AI within the museum context issues of privacy, bias and general awareness should not be ignored. Using AI models to generate a narrative and its subsequent translation raises some ethical concerns about the misinterpretation of the translation or even derogatory words in the content itself. One of the main problems with using today's language models trained with text from the internet and books (such as GPT-3) is the generation of racist, sexist, or otherwise toxic language, as the training data contains such examples and the AI, therefore, tries to mimic this language [12]. This is also the case with AI for translation, where the problem is that it is trained further with each translation, and it is even harder to intervene [9].

The only data where we can be sure that the content does not contain derogatory words are the descriptions of the museum objects, which have been carefully checked. For the AI models, precautions are also taken to counteract this, e.g. an AI model is trained with a dataset annotated by users who ensure that no toxic content is created [8]. During translation, a so-called glossary function is integrated, which can be used to influence the translation results and determine how the content should be translated [3].

Another question that needs to be asked when dealing with artificially generated text is how to communicate that it is generated by a machine. Since GPT-3 can generate articles and stories that are difficult if not impossible to distinguish from human content [24], it is important to make users aware of the origin of such content. We address this issue by using the general notion of how humans distinguish what is artificial and what is not by using the narrative of so-called *science-fictional AI*. This involves the general idea of the word *artificial intelligence*, which usually triggers images of a robot or cyborgs [14]. On the contrary, humans imagine by the words *living* or *alive* a being of flesh and blood, which can be an animal or a human being. To put it simply, there are two images, one representing human-created content and the other representing artificially created content. We see this as a potential solution approach for future works, as this concept uses the common human imagination by the distinction of artificial and non-artificial.

The developed solution can work with any kind of data that is accompanied by short textual descriptions, images, or 3D models and therefore can be easily adapted or customized by museums with different scopes and targeted user groups.

## References

1. The Met × Microsoft × MIT. <https://www.metmuseum.org/about-the-met/policies-and-documents/open-access/met-microsoft-mit>
2. DeepL API (2016). <https://www.deepl.com/de/docs-api/introduction/>
3. Individuelle Übersetzungen mit der Glossarfunktion von DeepL (2022). <https://www.deepl.com/de/blog/translate-your-way-with-the-deepl-glossary>

4. Bojanowski, P., Grave, E., Joulin, A., Mikolov, T.: Enriching word vectors with subword information. *Trans. Assoc. Comput. Linguist.* **5**, 135–146 (2017)
5. Bönisch, D.: The curator's machine: clustering of museum collection data through annotation of hidden connection patterns between artworks. *Int. J. Digit. Art Hist.* (5), 5.20–5.35 (2021). <https://doi.org/10.11588/dah.2020.5.75953>
6. Brockman, G., Murati, M., Welinder, P., OpenAI: OpenAI API (2020). <https://openai.com/blog/openai-api/>
7. Carr, S., Francis, M., Rivlin, L., Stone, A.: *Public Space*. Cambridge Series in Environment and Behavior. Cambridge University Press (1992). <https://www.cambridge.org/de/academic/subjects/psychology/social-psychology/public-space?format=PB>
8. Chan, A.: GPT-3 and InstructGPT: technological dystopianism, utopianism, and “Contextual” perspectives in AI ethics and industry (2022). <https://doi.org/10.1007/s43681-022-00148-6>
9. Friedrich, G.: DeepL: so funktioniert die Übersetzung (2021). <https://www.heise.de/tipps-tricks/DeepL-So-funktioniert-die-Uebersetzung-6048772.html>
10. Gaia, G., Boiano, S., Borda, A.: Engaging museum visitors with AI: the case of chatbots. In: Giannini, T., Bowen, J.P. (eds.) *Museums and Digital Culture*. SSCC, pp. 309–329. Springer, Cham (2019). [https://doi.org/10.1007/978-3-319-97457-6\\_15](https://doi.org/10.1007/978-3-319-97457-6_15)
11. Gatys, L.A., Ecker, A.S., Bethge, M.: Image style transfer using convolutional neural networks. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2414–2423 (2016). <https://doi.org/10.1109/CVPR.2016.265>
12. Gehman, S., Gururangan, S., Sap, M., Choi, Y., Smith, N.A.: RealToxicityPrompts: evaluating neural toxic degeneration in language models. *CoRR abs/2009.11462* (2020). <https://arxiv.org/abs/2009.11462>
13. He, K., Gkioxari, G., Dollár, P., Girshick, R.B.: Mask R-CNN. *CoRR abs/1703.06870* (2017). <https://arxiv.org/abs/1703.06870>
14. Hermann, I.: Artificial intelligence in fiction: between narratives and metaphors (2021). <https://doi.org/10.1007/s00146-021-01299-6>
15. Jocher, G., et al.: ultralytics/YOLOv5: v6.1 - TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference (2022). <https://doi.org/10.5281/zenodo.6222936>
16. Lawson, C.A., Cook, M., Dorn, J., Pariso, B.: A STEAM-focused program to facilitate teacher engagement before, during, and after a fieldtrip visit to a children's museum. *J. Museum Educ.* **43**(3), 236–244 (2018). <https://doi.org/10.1080/10598650.2018.1474421>
17. Lim, P.: Bringing black and white photos to life using Colourise.sg - a deep learning colouriser trained with old Singaporean photos (2019). <https://blog.data.gov.sg/bringing-black-and-white-photos-to-life-using-colourise-sg-435ae5cc5036>
18. Lu, F.: What AI is doing for the museum and cultural space today (2021). <https://jingculturecommerce.com/y-lab-art-of-ai-takeaways/>
19. Memarovic, N., Langheinrich, M., Alt, F., Elhart, I., Hosio, S., Rubegni, E.: Using public displays to stimulate passive engagement, active engagement, and discovery in public spaces. In: *Proceedings of the 4th Media Architecture Biennale Conference: Participation*, pp. 55–64 (2012)
20. Müller, J., Alt, F., Michelis, D., Schmidt, A.: Requirements and design space for interactive public displays. In: *Proceedings of the 18th ACM International Conference on Multimedia*, pp. 1285–1294 (2010)

21. Mullins, J.K., Sabherwal, R.: Gamification: a cognitive-emotional view. **106**, 304–314 (2020). <https://doi.org/10.1016/j.jbusres.2018.09.023>, <https://linkinghub.elsevier.com/retrieve/pii/S0148296318304739>
22. Radford, A., Narasimhan, K.: Improving language understanding by generative pre-training (2018)
23. Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., Lee, H.: Generative adversarial text to image synthesis (2016). <https://doi.org/10.48550/ARXIV.1605.05396>, <https://arxiv.org/abs/1605.05396>
24. Sagar, R.: OpenAI Releases GPT-3, The Largest Model So Far, 31 July 2020
25. Schaffler, S., Ruß, A., Sasse, M.L., Schubotz, L., Gustke, O.: Questions and answers: important steps to Let AI chatbots answer questions in the museum. In: Wölfel, M., Bernhardt, J., Thiel, S. (eds.) ArtsIT 2021. LNICST, vol. 422, pp. 346–358. Springer, Cham (2021). [https://doi.org/10.1007/978-3-030-95531-1\\_24](https://doi.org/10.1007/978-3-030-95531-1_24)
26. Staff, M.: Museum curator job guide: how to become a museum curator (2022). <https://www.masterclass.com/articles/museum-curator-job-guide#4-types-of-museum-curators>
27. Staff, T.: Museum curator (2021). <https://theartcareerproject.com/careers/museum-curator/>
28. Thomas Lilge, C.S.: Ping! Die Museumsapp. Stiftung Humboldt Forum im Berliner Schloss (2020). <https://www.museum4punkt0.de/ergebnis/ping-die-museumsapp/>
29. Torrey, L., Shavlik, J.: Transfer Learning. Handbook of Research on Machine Learning Applications (2009). <https://doi.org/10.4018/978-1-60566-766-9.ch011>
30. Villaespesa, E.: Diving into the museum’s social media stream. Analysis of the Visitor Experience in 140 Characters. <https://mw2013.museumsandtheweb.com/paper/diving-into-the-museums-social-media-stream/>
31. Villaespesa, E., French, A.: AI, visitor experience, and museum operations: a closer look at the possible, pp. 101–113 (2019)
32. Wölfel, M.: Kinetic space: 3D-gestenerkennung für dich und mich. Konturen Ausgabe **31**, 58–63 (2012)
33. Wölfel, M., Purps, C.F., Percifull, N.: Enabling embodied conversational agents to respond to nonverbal behavior of the communication partner. In: Kurosu, M. (ed.) HCI 2022. LNCS, vol. 13304, pp. 591–604. Springer, Cham (2022). [https://doi.org/10.1007/978-3-031-05412-9\\_40](https://doi.org/10.1007/978-3-031-05412-9_40)