Artificial Intelligence for Digital Heritage Innovation

Subjects: Others | Others

Contributor: Sander Münster , Ferdinand Maiwald , Isabella di Lenardo , Juha Henriksson , Antoine Isaac , Manuela Milica Graf , Clemens Beck , Johan Oomen

Artificial intelligence (AI) is a game changer in many fields, including cultural heritage. It supports the planning and preservation of heritage sites and cities, enables the creation of virtual experiences to enrich cultural tourism and engagement, supports research, and increases access and understanding of heritage objects. Despite some impressive examples, the full potential of AI for economic, social, and cultural change is not yet fully visible.

cultural heritage AI agenda

1. Introduction

Digitization is key for protecting, preserving, documenting and opening up European and global cultural heritage (CH) to meet pressing sustainability threats, including environmental ones and increasing social inclusivity. Within the CH sector, economic activities related to digital collections in cultural institutions are a market worth ten bn EUR in 2015 [1]. These developments have been accelerated by the COVID-19 pandemic [2]. Digital technologies can transform the entire value chain model in CH institutions—from capturing and digitizing tangible and intangible heritage and long-term preservation over innovative digital research methods to digital channels allowing people across the globe to interact with digital objects. These channels enable connections to other collections published on the web and accelerate the creation of new artistic works, unearthing new narratives in collections. While all these areas of work could be improved by applying the latest digital technologies, a significant increase is expected during the next few years.

The Strategic Topic Group (STG) Cultural Heritage in Green and Digital Transitions for Inclusive Societies was formed in 2022 within the European Institute of Innovation and Technology's (EIT) Knowledge and Innovation Community for Culture & Creativity and seeks to unlock the potential of CH for the green and digital transitioning of Europe encompassing societal challenges on this key policy topic. The group includes 32 partner organizations in mid-2023 and focuses on four closely connected areas, including (i) upskilling and capacity building; (ii) environmental impact of operations of CH institutions; (iii) increasing outreach and community engagement; and (iv) creation of new business models.

2. Application Fields of AI in CH

In CH, AI is being used in a variety of research areas. These include:

- **Image analysis and restoration**: AI algorithms can analyze and restore old, damaged, or degraded (moving) images, sounds, paintings, and photographs. These algorithms can enhance image quality, remove noise, and even reconstruct missing parts of the artwork, aiding in preserving and restoring cultural artifacts. Examples listed in [3] are the prediction of the painting's style, genre, and artist, the detection of fake artworks by stroke analysis, and the artistic style transfer using adversarial networks to regularize the generation of stylized images." Further research deals with the automatic colorization of images [4] and the restoration of ancient mosaics [5].

- **Object recognition and classification**: AI-powered computer vision techniques enable automatic recognition and classification of cultural objects. By analyzing visual features and patterns, AI algorithms can identify and categorize artifacts, sculptures, and architectural elements [6], facilitating the organization and cataloging of museum collections. Examples are the prediction of color metadata, e.g., for textile objects [7], of technique, timespan, material, and place metadata for European silk fabrics [8], and the recognition and classification of symbols in ancient papyri [9].

- **Translation and transcription**: AI language models are capable of translating. e.g., ancient texts, inscriptions, and manuscripts into modern languages. They can also be used for modern languages by translating metadata or full-text content of heritage objects and related information, making sharing cultural heritage across languages easier. Other models can transcribe handwritten texts, allowing researchers and historians to access and understand historical documents and perform automated analysis (e.g., [10]).

- **Automatic text analysis**: This comprises various approaches [11]. An example is the automatic semantic indexing of pre-structured historical texts, which enables historians to mine large amounts of text and data to gain a deeper understanding of the sources (e.g., [12]); for example, tax lists or registers of letters sent to a historical entity [13].

- **Virtual Reality (VR) and Augmented Reality (AR)**: AI technology supports the creation of immersive VR and AR experiences for CH sites and museums. Visitors can virtually explore ancient ruins, historical sites, or museum exhibitions, interacting with AI-generated virtual characters or objects to enhance their understanding and engagement with the cultural context [14][15].

- **Recommender systems for personalized experiences**: AI algorithms can analyze user preferences, historical data, and contextual information to provide personalized recommendations for CH experiences. Despite the risks of information filtering (e.g., [16]), use is to suggest relevant exhibits, customized tours, or tailored content, AI-powered recommender systems enhance visitor engagement and satisfaction, or—triggered by the advent of large language models (LLMs) such as GPT—dialogue and chatbot systems. Examples are the use of chatbots in museums [17][18] or recommender systems for CH collections (e.g., [19][20]).

- **Cultural content analysis and interpretation**: AI techniques, such as natural language processing (NLP), are used to analyze large volumes of cultural content, including literature, music, and artwork. This analysis can reveal patterns, themes, and cultural influences, providing valuable insights into historical contexts and artistic movements. Examples are metadata enrichment (e.g., [21][22][23]) and linking to open data sources (e.g., [6]).

- **Heritage digitization and preservation**: AI can be crucial in digitizing cultural artifacts and archives. By automating digitization processes and extracting knowledge, AI speeds up the preservation of CH, allowing researchers and the public to explore and study rare artifacts remotely. Several articles provide an overview of particular technologies, e.g., for 3D acquisition, such as laser scanning [24] or photogrammetry [25], and quantify their use [26]. AI-powered systems can monitor and analyze CH site environmental conditions, helping with early detection of potential threats such as humidity, temperature fluctuations, and structural damage. This real-time monitoring aids in the proactive conservation and protection of cultural landmarks (e.g., [27][28]).

- **Multimodal analysis**: AI is capable of bringing together different sources and types of data. Approaches include text, images [8], 3D models [29], audio [30], and video [31].

- **AI supports or creates artistic expressions**: Applying algorithms that analyze heritage objects (or entire collections) and extract information that either artists and other creators can use to create new works [32] or AI creating "artistic" expressions.

3. Project Examples

To date, there are some impressive examples of the utilization of AI technologies in the field of CH (**Table 1**).

Table 1. Project examples of AI application in CH (all links accessed on 1 December 2023).

	**Art Transfer by Google Arts & Culture** Using AI algorithms, Art Transfer allows users to transform their photos into the style of famous artists such as Van Gogh or Picasso. Link: https://artsandculture.google.com/camera/art-transfer
	**MicroPasts by the British Museum** MicroPasts is a project that combines crowd-sourced data with AI technology. Volunteers contribute by digitizing and tagging images while AI algorithms analyze the data. Link: https://micropasts.org/
	**4Dcity by the University of Jena** This application uses AI to automatically 4D reconstruct past cityscapes from historical cadastre plans and photographs. This 4D model is world-scale and enriched by links to texts and information, e.g., from Wikipedia, and accessible as mobile 4D websites [33]. Link: https://4dcity.org/
	**SCAN4RECO** This EU-funded project combines 3D scanning, robotics, and AI to create digital reconstructions of damaged or destroyed CH objects. Link: https://scan4reco.iti.gr/
	**AI-DA by Aidan Meller Gallery** AI-DA is an AI-powered robot artist developed by Aidan Meller Gallery in the United Kingdom. The robot uses AI algorithms to analyze and interpret human facial expressions, creating drawings and paintings inspired by the emotions it perceives. AI-DA's artworks have been exhibited in galleries across Europe. Link: https://www.ai-darobot.com/

Transkribus by Read Coop SCE
Transkribus is a comprehensive solution for digitization, AI-powered text recognition, transcription, and searching historical documents. A specific emphasis is on handwritten text recognition.
https://readcoop.eu/transkribus/

Transcribathon
The Transcribathon platform is an online crowd-sourcing platform for enriching digitized material from Europeana. It applies the Transkribus handwriting recognition technology to input documents, performs some automatic enrichments (including translation) on the obtained text and metadata, and lets volunteers validate the results.
https://transcribathon.eu/

The Next Rembrandt by ING Bank and Microsoft
This project employed AI algorithms to analyze Rembrandt's works and create a new painting in his style.
https://www.nextrembrandt.com/

Rekrei (formerly Project Mosul)
Rekrei is a crowd-sourcing and AI project aimed at reconstructing CH sites that have been destroyed or damaged. Users can contribute photographs and other data, and AI algorithms help in reconstructing the lost heritage digitally.
https://rekrei.org/

Notre Dame reconstruction
After a fire destroyed parts of the Notre Dame Cathedral in Paris in 2019, a digital twin model was created to experiment—physical anastylosis, reverse engineering, spatiotemporal tracking assets, and operational research—and create a reconstruction hypothesis. The results demonstrate that the proposed modeling method facilitates the formalization and validation of the reconstruction problem and increases solution performance [34].
https://news.cnrs.fr/articles/a-digital-twin-for-notre-dame

Finto AI by the National Library of Finland
Finto AI is a service for automated subject indexing. It can be used to suggest subjects for text in Finnish, Swedish, and English. It currently gives suggestions based on concepts of the General Finnish Ontology, YSO. Link: https://ai.finto.fi

Europeana Translate
This project has trained translation engines on metadata from the common European data space on cultural heritage in order to obtain a service that can translate CH metadata from 22 official EU languages to English, improving the multilingual experience provided to its users. It has been applied to 29 million metadata records so far.
Link: https://pro.europeana.eu/post/europeana-translate-project-brings-together-multilingualism-and-cultural-heritage

MuseNet by OpenAI
MuseNet composes original music in a wide range of styles and genres. It can create music inspired by different cultural traditions and historical periods, demonstrating the potential of AI in generating new compositions that reflect CH.
Link: https://openai.com/research/musenet

The Hidden Florence by the University of Exeter
The Hidden Florence is an AI-enhanced mobile app that guides visitors through the streets of Florence, Italy, offering insights into the city's rich CH in an engaging way. The app utilizes AI algorithms to provide

18. Bongini, P.; Becattini, F.; Del Bimbo, A. Is GPT-3 All You Need for Visual Question Answering in Cultural Heritage? In Proceedings of the European Conference on Computer Vision; Springer: Cham, Switzerland, 2022; pp. 268–281.

location-based narratives, AR experiences, and interactive storytelling.
Link: https://hiddenflorence.org/

Smartify App by Smartify
Smartify utilizes AI to provide interactive experiences with artworks in museums and galleries. The mobile app uses image recognition to identify artworks, delivering detailed information, audio guides, and curated tours. It is compatible with numerous cultural institutions across Europe and beyond.
Link: https://smartify.org/

Second Canvas App by Madpixel and the Prado Museum
The app uses AI technology to enhance the visitor experience. It provides high-resolution images of artworks, along with interactive features that allow users to explore the details and stories behind the paintings.
Link: https://www.secondcanvas.net/

WAIVE
WAIVE is a smart DJ system utilizing AI to create unique music samples, beats, and loops from the digitized audio archives of the Netherlands Institute for Sound & Vision.
Link: https://www.thunderboomrecords.com/waive

for Designing the Jena Experimental Repository for the DFG 3D Viewer. Appl. Sci. 2023, 13, 9781.

24. Di Stefano, F.; Chiappini, S.; Gorreja, A.; Balestra, M.; Pierdicca, R. Mobile 3D scan LiDAR: A literature review. Geomat. Nat. Hazards Risk 2021, 12, 2387–2429.
25. Remondino, F.; Nocerino, E.; Toschi, I.; Menna, F. A Critical Review of Automated Photogrammetric Processing of Large Datasets. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2017, XLII-2/W5, 591–599.
26. European Commission. Study on Quality in 3D Digitisation of Tangible Cultural Heritage: Mapping Parameters, Formats, Standards, Benchmarks, Methodologies, and Guidelines; VIGIE 2020/654 Final Study Report; European Commission: Brussels, Belgium, 2022.
27. Mishra, M. Machine learning techniques for structural health monitoring of heritage buildings: A state-of-the-art review and case studies. J. Cult. Herit. 2021, 47, 227–245.
28. Tejedor, B.; Lucchi, E.; Bienvenido-Huertas, D.; Nardi, I. Non-destructive techniques (NDT) for the diagnosis of heritage buildings: Traditional procedures and futures perspectives. Energy Build. 2022, 263, 112029.
29. Münster, S.; Bruschke, J.; Hoppe, S.; Maiwald, F.; Niebling, F.; Pattee, A.; Utescher, R.; Zarriess, S. Multimodal AI Support of Source Criticism in the Humanities. In Proceedings of the ADHO DH 2022; Tokyo, Japan, 25–29 July 2022.
30. Ukolov, D. Reviving the Sounds of Sacral Environments: Personalized Real-Time Auralization and Visualization of Location-Based Virtual Acoustic Objects on Mobile Devices; Springer: Cham, Switzerland, 2023; pp. 165–186.
31. Dimitropoulos, K.; Tsalakanidou, F.; Nikolopoulos, S.; Kompatsiaris, I.; Grammalidis, N.; Manitsaris, S.; Denby, B.; Crevier-Buchman, L.; Dupont, S.; Charisis, V.; et al. A Multimodal Approach for the Safeguarding and Transmission of Intangible Cultural Heritage: The Case of i-Treasures. IEEE Intell. Syst. 2018, 33, 3–16.

4. AI Technologies for CH State of the Art

4.1. AI and Images

Historical images hold immense value in documenting our collective heritage. However, analyzing and extracting information from these images manually can be limited, e.g., due to the required effort. Current evolvements in computer visualization are closely coupled to the massive renaissance in ML [35] with the use of convolutional neural networks (CNNs, cf. [36]). There is a large number of computer vision techniques employed in historical image analysis [37][38], including:

- **Content-based image retrieval**: Efficient retrieval and exploration of historical images based on visual similarity and content-based features. However, traditional ML technologies currently require large-scale training data [3][39][40][41], which are only capable of recognizing well-documented and visually distinctive landmark buildings [33] but fail to deal with less distinctive architecture, such as houses of similar style. Even using more advanced ML approaches or combining different algorithms [42] only allows the realization of prototypic scenarios [43][44].
- **Image-based localization**: Connecting images with the 3D world relevant for AR/VR applications requires estimating the original six-degree-of-freedom (6DOF) camera pose. While several methods exist for homogeneous image blocks [45][46], the problem becomes increasingly complex for varying radiometric and geometric conditions, especially relevant for historical photographs [47].
- **Image recognition and classification**: Identifying objects, scenes, or people depicted in historical images using deep learning models, such as CNNs. This field ranges from the detection of WW2 bomb craters in historical aerial images [48], via historical photo content analysis [49], to historical map segmentation [50][51][52].

- **Semantic segmentation and object detection:** Locating and recognizing specific objects or regions of interest within historical images using techniques like Faster R-CNN and YOLO, and semantic segmentation to classify parts of images [41][53][54].
- **Image restoration and enhancement:** Repairing and enhancing degraded or damaged historical images through techniques like denoising, inpainting, and super-resolution [55][56].

4.2. AI and Text

Historical texts provide a rich source of information for understanding the past. However, the sheer volume and complexity of historical archives make manual analysis laborious and time-consuming [57]. ML algorithms supported these processes in various ways—from optical character recognition (OCR) to automating the extraction of knowledge and patterns from historical texts [57][58][59]. Approaches include these ML approaches commonly used in historical text analysis:

- **NLP techniques**: Named entity recognition, part-of-speech tagging, sentiment analysis, and topic modeling. The most recent applications of CNNs and Transformer [60] are consistently successful in accurately extracting and reducing the number of errors even with unsupervised pre-training.
- **Text classification algorithms**: Naive Bayes, Support Vector Machines, and Random Forests.
- **Sequence models**: Hidden Markov models, conditional random fields, and recurrent neural networks.

In addition, various preprocessing techniques are used for historical texts to enable their digital processing and respond to challenges such as linguistic variations, archaic vocabulary, and textual degradation:

- **Preprocessing:** Includes character recognition (e.g., OCR), unification, processing of spelling variations and alignment to controlled vocabularies (e.g., [61]).
- **Postprocessing:** Used to check and correct any OCR reading errors via neural network approaches [62].

4.3. AI and Virtual 3D Objects

The application of AI in 3D for CH has gained significant attention in the research community to enhance the analysis, interpretation, and preservation of CH in 3D environments. Here are some key areas of scientific analysis:

- **Object recognition and classification and semantic segmentation:** In 3D/4D reconstruction of CH, ML-based technologies are currently used primarily for specific tasks. This involves AI models to identify specific architectural elements, artifacts, or decorative motifs, to recognize specific objects [39][40][41][63], and to preselect imagery [64][65]. Other tasks include AI-based semantic segmentation techniques to partition 3D models into meaningful regions or components [66].

32. Bocyte, R.; Oomen, J. Content Adaptation, Personalisation and Fine-grained Retrieval: Applying AI to Support Engagement with and Reuse of Archival Content at Scale. In Proceedings of the 12th International Conference on Agents and Artificial Intelligence (ICAART 2020), Valletta, Malta, 22–24 February 2020.
33. Münster, S.; Lehmann, C.; Lazariv, T.; Maiwald, F.; Karsten, S. Toward an Automated Pipeline for a Browser-Based, City-Scale Mobile 4D VR Application Based on Historical Images. In Proceedings of the Research and Education in Urban History in the Age of Digital Libraries; Springer: Cham, Switzerland, 2019; pp. 106–128.
34. Gros, A.; Guillem, A.; De Luca, L.; Baillieul, E.; Duvocelle, B.; Malavergne, O.; Leroux, L.; Zimmer, T. Faceting the post-disaster built heritage reconstruction process within the digital twin framework for Notre-Dame de Paris. Sci. Rep. 2023, 13, 5981.
35. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. Commun. ACM 2017, 60, 84–90.
36. Jahrer, M.; Grabner, M.; Bischof, H. Learned local descriptors for recognition and matching. Proc. Comput. Vis. Winter Workshop 2008, 2008, 39–46.
37. Lang, S.; Ommer, B. Attesting similarity: Supporting the organization and study of art image collections with computer vision. Digit. Scholarsh. Humanit. 2018, 33, 845–856.
38. Rodríguez-Ortega, N. Image processing and computer vision in the field of art history. In The Routledge Companion to Digital Humanities and Art History; Routledge: London, UK, 2020; pp. 338–357.
39. n.b. ArchiMediaL. Enriching and Linking Historical Architectural and Urban Image Collections. Available online: http://archimedial.net/ (accessed on 1 December 2023).
40. Radovic, M.; Adarkwa, O.; Wang, Q.S. Object Recognition in Aerial Images Using Convolutional Neural Networks. J. Imaging 2017, 3, 21.
41. Alger, D.; Allen, B.; Golovinskiy, A. Large-Scale 3D Scene Classification With Multi-View Volumetric CNN. arXiv 2017, preprint. arXiv:1712.09216.
42. Maiwald, F.; Lehmann, C.; Lazariv, T. Fully Automated Pose Estimation of Historical Images in the Context of 4D Geographic Information Systems Utilizing Machine Learning Methods. ISPRS Int. J. Geo-Inf. 2021, 10, 748.
43. Gominski, D.; Poreba, M.; Gouet-Brunet, V.; Chen, L. Challenging Deep Image Descriptors for Retrieval in Heterogeneous Iconographic Collections. In Proceedings of the 1st Workshop on Structuring and Understanding of Multimedia heritAge Contents, Nice, France, 21 October 2019; pp. 31–38.
44. Morelli, L.; Bellavia, F.; Menna, F.; Remondino, F. Photogrammetry now and then—From hand-crafted to deep-learning tie points. ISPRS—Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2022, XLVIII-2/W1-2022, 163–170.

45. Schonberger, J.L.; Hardmeier, H.; Sattler, T.; Pollefeys, M. Comparative Evaluation of Hand-Crafted and Learned Local Features. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 6959–6968.
46. Sarlin, P.-E.; Cadena, C.; Siegwart, R.; Dymczyk, M. From Coarse to Fine: Robust Hierarchical Localization at Large Scale. arXiv 2019, arXiv:1812.03506.
47. Maiwald, F. A Window to the Past through Modern Urban Environments—Developing a Photogrammetric Workflow for the Orientation Parameter Estimation of Historical Images. Ph.D. Thesis, Technische Universität Dresden, Dresden, Germany, 2022.
48. Kruse, C.; Wittich, D.; Rottensteiner, F.; Heipke, C. Generating impact maps from bomb craters automatically detected in aerial wartime images using marked point processes. ISPRS Open J. Photogramm. Remote Sens. 2022, 5, 100017.
49. Chumachenko, K.; Mannisto, A.; Iosifidis, A.; Raitoharju, J. Machine Learning Based Analysis of Finnish World War II Photographers. IEEE Access 2020, 9, 144184–144196.
50. Chazalon, J.; Carlinet, E.; Chen, Y.; Perret, J.; Duménieu, B.; Mallet, C.; Géraud, T.; Nguyen, V.; Nguyen, N.; Baloun, J.; et al. ICDAR 2021 Competition on Historical Map Segmentation; Springer: Cham, Switzerland, 2021; pp. 693–707.
51. Maiwald, F.; Komorowicz, D.; Munir, I.; Beck, C.; Münster, S. Semi-Automatic Generation of Historical Urban 3D Models at a Larger Scale Using Structure-from-Motion, Neural Rendering and Historical Maps; Münster, S., Pattee, A., Kröber, C., Niebling, F., Eds.; Research and Education in Urban History in the Age of Digital Libraries; Springer: Cham, Switzerland, 2023; pp. 107–127.
52. Vaienti, B.; Petitpierre, R.; di Lenardo, I.; Kaplan, F. Machine-Learning-Enhanced Procedural Modeling for 4D Historical Cities Reconstruction. Remote Sens. 2023, 15, 3352.
53. Marmanis, D.; Wegner, J.D.; Galliani, S.; Schindler, K.; Datcu, M.; Stilla, U. Semantic Segmentation of Aerial Images with an Ensemble of CNSS. ISPRS Ann. 2016, 3, 473–480.
54. Hackel, T.; Wegner, J.D.; Schindler, K. Fast semantic segmentation of 3D point clouds with strongly varying density. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. 2016, 3, 177–184.
55. Poterek, Q.; Herrault, P.A.; Skupinski, G.; Sheeren, D. Deep Learning for Automatic Colorization of Legacy Grayscale Aerial Photographs. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2020, 13, 2899–2915.
56. Huang, S.; Jin, X.; Jiang, Q.; Liu, L. Deep learning for image colorization: Current and future prospects. Eng. Appl. Artif. Intell. 2022, 114, 105006.
57. Michel, J.-B.; Shen, Y.K.; Aiden, A.P.; Veres, A.; Gray, M.K.; Team, T.G.B.; Pickett, J.P.; Hoiberg, D.; Clancy, D.; Norvig, P.; et al. Quantitative Analysis of Culture Using Millions of Digitized Books. Science 2011, 331, 176–182.

- **Automated music classification** utilizes computer algorithms and ML techniques to automatically categorize music into classes or genres based on features extracted from the music data. Automated music classification has various applications, such as organizing music libraries and archives, and assisting in music research. Music-related classification tasks include mood classification, artist identification, instrument recognition, music annotation, and genre classification. For instance, one study investigates automatic music genre classification model creation using ML [92].
- **Optical Music Recognition (OMR)** research investigates how to computationally read music notation in documents [93]. OMR is a challenging process that differs in difficulty from OCR and handwritten text recognition because of the properties of music notation as a contextual writing system. First, the visual expression of music is very diverse. For instance, the Standard Music Font Layout [94] lists over 2440 recommended characters and several hundred optional glyphs. Second, it is only their configuration—how they are placed and arranged on the staves and with respect to each other—that specifies what notes should be played. The two main goals of OMR are:
  1. **Recovering music notation** and information from the engraving process, i.e., what elements were selected to express the given piece of music and how they were laid out. The output format must be capable of storing music notation, e.g., MusicXML [95] or MEI [96].
  2. **Recovering musical semantics** (i.e., the notes, represented by their pitches, velocities, onsets, and durations). MIDI [97] would be an appropriate output representation for this goal.
- **Automatic Music Transcription (AMT)** is the process of automatically converting audio recordings of music into symbolic representations, such as sheet music (e.g., MusicXML or MEI) or MIDI files. AMT is a very useful tool for music analysis. AMT comprises several subtasks: (multi-)pitch estimation, onset and offset detection, instrument recognition, beat and rhythm tracking, interpretation of expressive timing and dynamics, and score typesetting. Due to the very nature of music signals, which often contain several sound sources that produce one or more concurrent sound events that are meant to be highly correlated over both time and frequency, AMT is still considered a challenging and open problem [98].

4.6. AI and Audiovisual Material

Audiovisual heritage includes various materials such as films, videos, and multimedia content. AI for audiovisual heritage supports various aspects of preserving, analyzing, enhancing, and making accessible audiovisual content of historical and cultural significance. Key areas of application for AI in audiovisual heritage include:

- **Digitization and restoration:** AI assists in digitizing and restoring deteriorating audiovisual materials, improving their quality and preserving their historical significance.
- **Video summaries**: Can speed up the process of finding content in audiovisual archives [99].
- **Content analysis and knowledge extraction:** AI algorithms analyze audio and visual elements within content to identify patterns, objects, scenes, speakers, and other relevant information. It can also help to spot biases and

58. Vidhya, K.A. Text Mining Process, Techniques and Tools: An Overview. Int. J. Inf. Technol. Manag. 2010, 613–622.
59. Ehrmann, M.; Hamdi, A.; Pontes, E.L.; Romanello, M.; Doucet, A. Named entity recognition and classification in historical documents: A survey. ACM Comput. Surv. 2021, 56, 1–47.
60. Rouhou, A.C.; Dhiaf, M.; Kessentini, Y.; Salem, S.B. Transformer-based approach for joint handwriting and named entity recognition in historical document. Pattern Recognit. Lett. 2022, 155, 128–134.
61. Utescher, R.; Patee, A.; Maiwald, F.; Bruschke, J.; Hoppe, S.; Münster, S.; Niebling, F.; Zarrieß, S. Exploring Naming Inventories for Architectural Elements for Use in Multimodal Machine Learning Applications. In Proceedings of the Workshop on Computational Methods in the Humanities 2022, Lausanne, Switzerland, 9–10 June 2022.
62. Drobac, S.; Lindén, K. Optical character recognition with neural networks and post-correction with finite state methods. Int. J. Doc. Anal. Recognit. (IJDAR) 2020, 23, 279–295.
63. Khademi, S.; Mager, T.; Siebes, R. Deep Learning from History. In Proceedings of the Research and Education in Urban History in the Age of Digital Libraries, Dresden, Germany, 10–11 October 2021; Springer: Cham, Switzerland, 2021; pp. 213–233.
64. Münster, S.; Apollonio, F.I.; Bell, P.; Kuroczynski, P.; Di Lenardo, I.; Rinaudo, F.; Tamborrino, R. Digital Cultural Heritage Meets Digital Humanities. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2019, XLII-2/W15, 813–820.
65. Bell, P.; Ommer, B. Computer Vision und Kunstgeschichte—Dialog zweier Bildwissenschaften. In Digital Art History; Kuroczynski, P., Bell, P., Dieckmann, L., Eds.; Heidelberg University Press: Heidelberg, Germany, 2019; pp. 61–78.
66. Russo, M.; Grilli, E.; Remondino, F.; Teruggi, S.; Fassi, F. Machine Learning for Cultural Heritage Classification. In Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain; Franco Angeli: Milan, Italy, 2021.
67. Lowe, D.G. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 2004, 60, 91–110.
68. Knyaz, V.V.; Remondino, F.; Knyaz, V.A. Generative Adversarial Networks for Single Photo 3d Reconstruction. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2019, XLII-2/W9, 403–408.
69. Hermoza, R.; Sipiran, I. 3D Reconstruction of incomplete Archaeological Objects using a Generative Adversarial Network. In Proceedings of the Computer Graphics International 2018, Bintan Island, Indonesia, 11–14 June 2018; pp. 5–11.
70. Nogales Moyano, A.; Delgado Martos, E.; Melchor, Á.; García Tejedor, Á.J. ARQGAN: An evaluation of Generative Adversarial Networks' approaches for automatic virtual restoration of Greek temples. Expert Syst. Appl. 2021, 180, 115092.

contents in terms and track semantic drift in metadata, supporting curators, cataloguers, and others in deciding on potentially updating catalog records [100].

- **Metadata enhancement:** AI enriches metadata for better content organization, search, and context by extracting keywords or using LLMs to organize and enrich metadata records at scale.
- **Transcription and translation:** AI-powered speech-to-text transcription and translation services make audiovisual content more accessible and understandable to wider audiences [101].
- **Partial audio matching**: Supports framing analysis in identifying segments in one source audio file that are identical to segments in another target audio file. Framing analysis can reveal patterns and biases in the way content is being recontextualized in the media to shape public discourse [102].
- **Cross-modal analysis:** AI techniques analyze both audio and visual components of content, facilitating holistic interpretation and understanding.
- **Interactive storytelling and content-generation interfaces:** AI-powered interactive narratives and documentaries engage users with historical events and cultural context. AI can further enhance access by using fine-grained and time-based data extracted by AI systems as a basis for creating "generous interfaces" that allow for the rich exploration of CH collections [103][104] and using conversational speech to provide new ways of interacting with audiovisual collections [105].

71. Microsoft In Culture. See Ancient Olympia brought to life. 2021. Available online: https://unlocked.microsoft.com/ancient-olympia-common-grounds (accessed on 1 December 2023).
72. Mildenhall, B.; Srinivasan, P.P.; Tancik, M.; Barron, J.T.; Ramamoorthi, R.; Ng, R. NeRf: Representing scenes as neural radiance fields for view synthesis. Commun. ACM 2021, 65, 99–106.
73. Srinivasan, P.P.; Deng, B.; Zhang, X.; Tancik, M.; Mildenhall, B.; Barron, J.T. Nerv: Neural reflectance and visibility fields for relighting and view synthesis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2021, Virtue, 19–25 June 2021; pp. 7495–7504.
74. Croce, V.; Caroti, G.; Luca, L.; Piemonte, A.; Véron, P. neural radiance fields (nerf): Review and potential applications to digital cultural heritage. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2023, XLVIII-M-2-2023, 453–460.
75. Kaya, B.; Kumar, S.; Sarno, F.; Ferrari, V.; Gool, L. Neural Radiance Fields Approach to Deep Multi-View Photometric Stereo. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision 2022, Waikoloa, HI, USA, 3–8 January 2022.
76. Murtiyoso, A.; Grussenmeyer, P. initial assessment on the use of state-of-the-art nerf neural network 3d reconstruction for heritage documentation. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2023, XLVIII-M-2-2023, 1113–1118.
77. Vandenabeele, L.; Häcki, M.; Pfister, M. crowd-sourced surveying for building archaeology: The potential of structure from motion (sfm) and neural radiance fields (nerf). Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2023, XLVIII-M-2-2023, 1599–1605.
78. 4dReply. Closing the 4D Real World Reconstruction Loop. Available online: https://cordis.europa.eu/project/id/770784 (accessed on 8 February 2022).
79. Martin-Brualla, R.; Radwan, N.; Sajjadi, M.S.M.; Barron, J.T.; Dosovitskiy, A.; Duckworth, D. NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. arXiv 2021. preprint.
80. Cho, J.; Zala, A.; Bansal, M. DALL-Eval: Probing the Reasoning Skills and Social Biases of Text-to-Image Generative Transformers. CoRR 2022, abs/2202.04053.
81. Li, Z.; Wang, Q.; Cole, F.; Tucker, R.; Snavely, N. DynIBaR: Neural Dynamic Image-Based Rendering. arXiv 2022, arXiv:2211.11082.
82. Uhl, J.H.; Leyk, S.; Chiang, Y.-Y.; Knoblock, C.A. Towards the automated large-scale reconstruction of past road networks from historical maps. Comput. Environ. Urban Syst. 2022, 94, 101794.
83. Liu, C.; Wu, J.; Kohli, P.; Furukawa, Y. Raster-To-Vector: Revisiting Floorplan Transformation. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2195–2203.
84. Oliveira, S.A.; Seguin, B.; Kaplan, F. dhSegment: A Generic Deep-Learning Approach for Document Segmentation. In Proceedings of the 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), Niagara Falls, NY, USA, 5–8 August 2018; pp. 7–12.

85. Ignjatić, J.; Bajic, B.; Rikalovic, A.; Culibrk, D. Deep Learning for Historical Cadastral Maps Digitization: Overview, Challenges and Potential. In Proceedings of the 26th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision in co-operation with EUROGRAPHICS Association 2018, Delft, The Netherlands, 16–20 April 2018.
86. Kartta Labs. 2023. Available online: https://github.com/kartta-labs (accessed on 1 December 2023).
87. Petitpierre, R.; Kaplan, F.; di Lenardo, I. Generic Semantic Segmentation of Historical Maps. In CEUR Workshop Proceedings; CEUR-WS: Aachen, Germany, 2021.
88. Petitpierre, R. Neural networks for semantic segmentation of historical city maps: Cross-cultural performance and the impact of figurative diversity. arXiv 2020, preprint. arXiv:2101.12478.
89. Available online: https://ismir.net (accessed on 1 December 2023).
90. MIRtoolbox. Available online: https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox (accessed on 1 December 2023).
91. Mignot, R.; Peeters, G. An Analysis of the Effect of Data Augmentation Methods: Experiments for a Musical Genre Classification Task. Trans. Int. Soc. Music. Inf. Retr. 2019, 2, 97–110.
92. Tulisalmi-Eskola, J. Automatic Music Genre Classification—Supervised Learning Approach. Master's Thesis, Metropolia University of Applied Sciences, Helsinki, Finland, 2022.
93. Calvo-Zaragoza, J.; Jr, J.H.; Pacha, A. Understanding Optical Music Recognition. ACM Comput. Surv. 2020, 53, 1–35.
94. Standard Music Font Layout. Available online: https://w3c.github.io/smufl/latest/index.html (accessed on 1 December 2023).
95. Music XML. Available online: https://www.musicxml.com (accessed on 1 December 2023).
96. Music Encoding Initiative. Available online: https://music-encoding.org (accessed on 1 December 2023).
97. Official Midi Specifications. Available online: https://www.midi.org/specifications (accessed on 1 December 2023).
98. Benetos, E.; Dixon, S.; Duan, Z.; Ewert, S. Automatic Music Transcription: An Overview. IEEE Signal Process. Mag. 2019, 36, 20–30.
99. Tsalakanidou, F. Deliverable 2.3—AI Technologies and Applications in Media: State of Play, Foresight, and Research Directions; AI4Media Project (Grant Agreement No 951911); 2022. Available online: https://www.ai4media.eu/wp-content/uploads/2022/03/AI4Media_D2.3_Roadmap_final.pdf (accessed on 1 December 2023).
100. Ferrara, A.; Montanelli, S.; Ruskov, M. Detecting the Semantic Shift of Values in Cultural Heritage Document Collections. Ceur Workshop Proc. 2022, 3286, 35–43.

101. Van Noord, N.; Olesen, C.; Ordelman, R.; Noordegraaf, J. Automatic Annotations and Enrichments for Audiovisual Archives. In Proceedings of the 13th International Conference on Agents and Artificial Intelligence—Volume 1: ARTIDIGH, Online, 4–6 February 2021; SciTePress: Setúbal, Portugal, 2021; pp. 633–640, ISBN 978-989-758-484-8; ISSN 2184-433X.
102. Kemenade, P.v.; Bocyte, R.; Oomen, J. You've Been Framed—Partial Audio Matching Functionality to Support Framing Analysis. In DARIAH Annual Event 2023: Cultural Heritage Data as Humanities Research Data? Zenodo: Budapest, Hungary, 2023; Available online: https://zenodo.org/communities/dariahannualevent2023chdata-hrdata/ (accessed on 1 December 2023).
103. Wigham, M.; Melgar Estrada, L.; Ordelman, R.J.F. Jupyter Notebooks for Generous Archive Interfaces. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; Song, Y., Liu, B., Lee, K., Abe, N., Pu, C., Qiao, M., Ahmed, N., Kossmann, D., Saltz, J., Tang, J., et al., Eds.; pp. 2766–2774.
104. Piet, N. Beyond Search; Netherlands Institute for Sound & Vision: Hilversum, The Netherlands, 2023.
105. Beelen, T.; Velner, E.; Ordelman, R.; Truong, K.; Evers, V.; Huibers, T. Designing conversational robots with children during the pandemic. arXiv 2022, preprint. arXiv:2205.11300.