AI and Image Archives Lit Review
RA 01 Study 2023/24

Authors:
Dharani Persaud, GRA UBC
Jordan Kerr, GRA UBC
Seiji Bessho, GRA UBC
Rachel Paprocki, GRA SJSU
Dr. Jessica Bushey, SJSU

Introduction

Providing access to archival holdings and performing reference services to assist researchers and internal stakeholders in discovering and using archival records is a key activity for archivists, one which has gained increasing importance as technologies and tools have improved to facilitate digitization activities, online collections databases and interactive virtual exhibitions and events. Additionally, users expectations for immediate access to archival holdings, specifically audiovisual materials and images due to publicly available search engines (i.e., Google) and social media platforms (e.g., Instagram, YouTube) has placed significant pressure on archives staff to provide access to these types of archival records sooner, in greater detail and with wider contextual linkages.

Image management is the encompassing term used to describe the processes of storage, organization, access, and preservation of analog, digitized, and born-digital images. Image management is a field of management primed to receive the benefits of AI software and tools. Currently, AI software and tools for image management have been framed within the function of retrieval and access of records in medical archives (Jabarulla & Lee, 2021; Yi et al., 2021); however, limited exploration has been made regarding these tools for public and private archives and their records. The following is organized by the scope and methodology of our survey, three thematic categories, and concerns we see moving forward in the area of AI software and tools for image management in archives.

Scope of the Survey

Our survey of AI software and tools for image management in archives is informed by the work of Colavizza et al. 's (2021) paper “Archives and AI.” Here the authors use AI as “a proxy for Machine Learning (ML), which is the study and development of computer programs that automatically learn from data” (p. 4:3). Our use of AI is additionally inclusive of all Natural Language Processing (NLP) software and tools, like Computer Vision (CV). AI software and tools are thus inclusive of the collection of pre-packaged, customizable, and “black box” software that utilizes one or more of the aforementioned AI processes.

Methodology

With the aim of understanding if and how AI software and tools are being used in the context of archival management, in particular the management of archival images (analogue, digitized and born digital), a literature survey was commenced in Spring 2023. We focused on exploring initiatives and scholarly literature in the field of archives, information science, and digital humanities. We began our investigation with the desire to survey public archives, specifically focusing on public records in national and municipal archives, and AI software and tools directly related to analog, digitized, and born-digital images within the confines of image management. However, we had to reassess these parameters because current published materials - journals and conference materials - did not offer a significant breadth in which to pursue our line of inquiry. Consequently, we opened our criteria beyond archives to include galleries, libraries, private archives, and museums. Furthermore, our initial survey could not find any instance of AI software and tools being used for analog image management, so our investigation was narrowed to digitized and born-digital images.

To identify materials that fit within our scope, specifically relating to the back-end of AI software and tools, we examined AI technology currently used for image management within GLAM (galleries, libraries, archives, and museums) organizations. Additionally, we examined the different functions AI software and tools are being used for image archiving, both within back-end activities like classification and metadata as well as front-end activities like user searching. For this literature review, we searched for terms including archives, digital archives, photoarchive, public records, digital humanities, GLAM organizations, cultural heritage, images, image management, image retrieval, image analysis, photography, metadata, computer vision (CV), explainable AI (XAI), artificial intelligence (AI), deep learning (DL), machine learning (ML), and object detection (OD). We additionally focused on case studies that discussed current GLAM and digital humanities projects using AI software and tools for image management rather than the high-end technical and theoretical applications of said software and tools.

As AI technology can change fairly quickly, we searched for sources that had been published within the last five years, with some exceptions for foundational studies. In addition to published materials, we were interested in ongoing projects, such as EyCon [link needed], which are exploring the application of AI to archival collections and beginning to publish early findings related to their activities. The current lack of projects addressing AI and images collections, specifically in the public records domain is a gap that we have identified through our literature review.

Ultimately, the following themes emerged from our literature review: AI software and tools for image management from a back-end perspective, AI software and tools for image management from a front-end perspective, and AI software and tools for image description and metadata. We conclude with some of the concerns we observed within the analyzed literature. We have chosen to structure the following analysis by archival function within each theme, again drawing on the work of Colavizza et al. (2021). While Colavizza et al. frame their analysis within the Records Continuum model, we have framed ours by archival function. These functions are creation and use, appraisal and acquisition, arrangement and description, retention and preservation, management and administration, and reference and access. Our rationale for this decision parallels that of Colavizza et al. who state this “model offers a holistic, pluralist framework covering all aspects and dimensions of archiving: the creation of the records, their capturing, organizing, and pluralising [*sic*]” (2021, p. 4:2). This literature review is by no means complete; literature reviews of AI technology within the burgeoning fields of GLAM image management will continue to grow and change as more studies emerge.

AI Software and Tools for Image Management (Back End)

AI tools and software for image management are primarily concerned with object detection. In “A Survey of Deep Learning-Based Object Detection,” Jiao et al. (2019) define object detection as “a computer technology related to computer vision and image processing which deals with detecting instances of semantic objects of a certain class” in born-digital and digitized images (p. 128837). These semantic objects include any subject within the image with a distinct identity, like humans and buildings, to provide scene understanding. Object detectors[1] are broken into two categories: one-stage or single-shot detectors and two-stage detectors. The former identifies objects with a singular example per semantic object. Types of single-stage detectors include YOLO[2] (You Only Look Once), SSD (Single Shot Detector), DSSD (Deconvolutional Single Shot Detector), RetinaNet, M2Det, and RefineDet. The latter identifies objects first based on their location then recognition of the semantic object within their location. Types of two-stage detectors include R-CNN, FAST R-CNN, FASTER R-CNN, and MASK R-CNN. Like many AI tools and software, object detectors rely on machine learning and deep learning algorithms to accomplish their output. As Jiao et al. (2019) conclude, object detection is a quickly growing field within AI and deep-learning scholarship, owing to its high accuracy and efficiency. It has increasing applications within and beyond security, military, and healthcare fields – including GLAM (galleries, libraries, archives, and museums) institutions.

Within the last five years, GLAM institutions have increasingly integrated object detection software into digitization and image management projects for digitized historical photos. These projects primarily use single-shot detectors. One example of this technology in practice comes from the Finnish wartime photograph project out of the Finnish Defence Forces’ photograph archive in Finland, published in 2020. Using the single-shot detectors You Only Look Once (YOLOv3) and SSD, nearly 160,000 Finnish World War II photos from 1939-1945 underwent content analysis to detect semantic objects within the digitized photos. YOLO is described as YOLO a “unified AI model for real-time object detection in photographs” through “a single convolutional network simultaneously predicts multiple bounding boxes and class probabilities for those boxes” (Redmon et al., 2016, p. 779). This identification of persons, planes, boats, trains, cars, and animals was used to aid researchers in searching for specific content. Furthermore, the researchers were able to train YOLOv3 and SSD to identify the creators of photos, albeit with 41.1% accuracy (Chumachenko et al, 2020, p. 144192).

[1] Jiao et al. (2019) are explicitly touching on pre-existing domain-specific image object detectors.
[2] This includes YOLOv2 and YOLOv3.

GLAM institutions have also used other computer vision (CV) platforms for image management beyond object detection. Computer vision platforms are a field of AI that enables computers and systems to derive meaningful information from digital and digitized images. The entire process involves image acquisition, screening, analyses, identification, and extraction of information. This technology enables machines to automatically recognize images and describe them with a high degree of accuracy. An example of CV application for digitized archival photographs is the Computer-Aided Metadata generation for Photo Archives Initiative (CAMPI) developed in 2020 at Carnegie Mellon University Library (Corrin et al., 2020). The impetus for the project was to bypass inaccurate or limited labels (and/or tags) linked to archival images and focus on visual similarity searching. The web app CAMPI enabled archivists to tag a photo and then retrieve other images with similar content structure and forms, such as rooms, angles of students facing lecturers, etc. The search retrieves images with visual similarity from within the digitized corpus of the archives. Starting with a "seed image" the app can retrieve similar images that may span across years and throughout collections, enabling the archivist to discover potentially related images that are missing descriptive information. CAMPI also allowed archivists to retrieve a set of similar images and select the best representation in response to a reference request. It also allows for retrospective tagging and the discovery of potentially culturally inappropriate descriptive terminology for archival images.

As outlined in Prokop et al.'s (2021) article "AI and the Digitized Photoarchive," the Frick Art Reference Library (FARL) in New York used a researcher-created AI algorithm to read analog metadata to automatically classify artwork based on visual elements to reduce their digitization backlog. Modeled off the Witts library photoarchive, FARL's photoarchive applied automatic image classifiers to photographs to capture and complete descriptive metadata for end users. Similarly, Kim and Lee's (2022) "Emotions and Colors in a Design Archiving System" sought to curate digitized artwork collections for users based on a metadata and design work data, the latter schema being an expression of intangible emotions. Using an algorithm of the authors' creation, they associated each pixel of digitized artwork with a color. Each color was prescribed a specific emotion based on art history classifications and added as an additional metadata field to the artwork. This retrieval system allowed users to search the collection by either color or emotion to see a specifically curated online gallery. This project focused on adding metadata not image similarity. Drawing from art history and color theory, the authors developed a pixel-by-pixel AI program breaking down the color scheme of digitized works of art into its most basic form. Using an algorithm of their creation this color pixel was treated as metadata and processed via AI to create a curated image retrieval system based on emotions attributed to the identified colors. Artwork that displayed predominantly cool colors and tones, like dark purple and black, were classified as sad, while light colors and tones, like red and orange, were classified as bright and happy. Users could then search the CV platform and determine what curated digitized artwork they wanted to view based on color and/or emotion.

Additionally, Mohanty et al. 's (2019) "Photo Sleuth," used the CV software Photo Sleuth for facial recognition of unknown United States of America Civil War portraits. Photo Sleuth is “a webbased platform that combines crowdsourced human expertise and automated face recognition to support Civil War portrait identification” (p. 547). As a crowd-sourced project, Photo Sleuth relied heavily on front-end user participation to determine the accuracy and validity of AI-determined facial recognition. Despite this reliance, the authors reported users were very reliable in their added metadata and identification. As Mohanty et al. state: Photo Sleuth “does not depend on a training set. Instead, it exploits the strengths of existing face recognition algorithms in a hybrid pipeline by integrating additional relevant information from visual clues in a photograph into the search process to enhance accuracy” (p. 548). Ultimately the project was quite successful in identifying, tagging, and adding photos to provide names to unnamed portraits.

User Experience (Front-End AI Use) (DP)

There are several recent studies and articles that focus on the user experience and AI tools for image collections (Angelova, et al, 2021; Arnold & Tilton, 2020). Angelova, et al. (2021) developed a study exploring a CV-based search platform for users to search and discover visual collections. The results of this study showed that a useful CV search allowed users to both look for a specific image in a directed search as well as conduct a more exploratory one. Users essentially used visuals instead of text as search criteria. This design enabled cross-collection image linking within and between UK image repositories, similar to what EyCon is attempting to do. However, their own literature review shows that format and quality of images varies from place to place, as well as metadata quantity and type, which remains a barrier to intra-collection image-linking.

Arnold and Tilton (2020) also use CV techniques to increase user discovery. However, their use of CV relates to detecting regions of images, which they describe as “stuff”: sky, water, trees, etc. They apply this to colour photographs from Farm Security Administration-Office of War Information Collection at the US Library of Congress by doing object annotations, people annotations, and captions. However, they found that object annotations were only right about 70% of the time, which means manual validation would have to also be employed. Captions were only correct 30% of the time and therefore did not seem like the best way to produce captions. Annotating people, however, had a high success rate, which means there is some possibility for creating a structured language to boost discoverability and access for users.

Description & Metadata (DP)

Metadata is an important focus of research in AI and images collections, with emphasis on studies involving facial recognition (Bakker et al., 2020; Proctor & Marciano, 2021; Milleville et al., 2023) and a study on object detection algorithms (Aske & Giardinetti, 2023). Many studies mention the time-consuming and labor-intensive activity of adding metadata and description to images and Bakker, et al. relate this to the fact that while time-consuming, it is often very important for end-user engagement. There is also a lack of adequate metadata across the archival and digital collections world (Bakker et al., 2020) and as a result the ability to accurately use AI for speedier metadata processing would greatly improve the field. Bakker et al. (2020) analyzed facial recognition applications OpenCV, Face++, and AmazonAWS to determine the most effective ones for producing metadata for digital collections. Images consisted of public figures and city and municipal officials between the 1920s and 1990s, which did raise some ethical concerns as discussed in the ethics section below. Bakker et al. (2020) found that Amazon AWS was strongest, with 96.9% confidence percentage but was the slowest of the applications. They also found that while in-house trained models were faster, they needed larger datasets for the same quality of results compared to cloud-based datasets. However, cloud-based platforms are definitely useful for easy-to-start environments where the collections are not as large. Their findings reveal the potential decrease in time spent processing a photograph collection, which provides relevance of using facial recognition software for archival collections. Faces in new photographic acquisitions can use existing collections metadata for faster and less labour-intensive processing. They also mention the issue of institutional memory where an archivist may be in an institution for a long time and therefore easily recognises faces in collections, but that knowledge is lost when the archivist leaves.

A very recent paper by Milleville et al. (2023) details testing of a generic image enrichment pipeline based on facial recognition in the FAME (facial recognition as a tool for metadata creation) project. They first built a reference set of known persons, but as this had to be compiled manually it was one of the most time-consuming steps. They found that they needed about 3-5 images per person for it to work, and that it was important to include images in the reference set from a similar time period as the collection. This system worked by identifying matching face embeddings from the reference set with the archival collections using cosine similarity. They used InsightFace's pre-trained model, and later compared the accuracy of that model to FaceNet, noting that the recognition accuracy for InsightFace was higher. Even with no fine-tuning of the pretrained model, over 93% of the person's predictions were correct.

Proctor and Marciano's (2021) study at Spelman College photo archives tested a computational processing workflow that integrated textual and photographic resources for better contextual metadata. Linked data is highly valuable for collections especially when trying to integrate other voices and authorities in a collection, but it is very time-consuming and labor-intensive so automating it is desirable. The framework they used was designed to work with existing item-level metadata standards, and their process started off with conducting a statistical analysis of the whole photo collection using OpenRefine. Their study found that this framework's performance shows promising results for the future of automated Linking. There is a potential for this approach to help build a more streamlined approach for archivists to integrate deeper text-based contextual information into photos.

As noted earlier, the EyCon Project, led by Dr. Lise Jaillant, seeks to connect and analyze digitization efforts across institutions. They hope to increase the discoverability, usability, and accessibility of overlooked and scattered archival materials relating to armed conflicts up to 1918. EyCon has two goals: One is to aggregate photographic material from separate repositories into a thematic collection focusing on non-European wars. The other is developing AI techniques that aid in data enrichment for large collections of photographs. They hope to use AI tools to compare and search images, create subgenres and identify anomalies, and more easily visualize a large collection using *image embedding, topic modeling, and clustering*. They also hope to use this project to *re-train existing datasets* that involve conflict photographs of this same time period (https://eycon.hypotheses.org/le-projet-en-quelques-mots/english).

A recent EyCon paper by Aske and Giardinetti (2023) discusses the possibility of automating contextual metadata. Different institutions approach metadata creation in different ways creating a disparity in what kind of metadata exists. This is a drawback when researchers want to cross-examine photographs in potentially related collections at different institutions. EyCon has created its own training database extracted from different mediums as well as testing object detection algorithms to categorize images and enrich descriptive content metadata. Their methodology for using pre-trained object detection is to first create specific classes that correspond to objects being detected. They then annotate part of the collection body with all elements needed as well as identify location to automate recognition. Training is then semi-supervised to allow AI to learn from algorithms. They are also looking at AI layout analysis that can detect and extract text blocks in publications, which is particularly useful for newspapers where text may relate to corresponding photos. They are also using CV to do similarity searches that are not reliant on existing metadata. They work solely by digital visual matching and extracting features of each image as vectors.

A recent EyCon workshop explored the application of a multi-modal approach to historical research that would utilize a visual similarity tool to quickly find instances of the same image in different publications, different formats, and in different archives, then utilize a text processor to analyze and compare attributions such as the photographer, location and/or date to support an action of suggesting attributions for the images (Dentler 2023 https://eycon.hypotheses.org/1539). This approach utilizes visual similarity algorithms and textual similarity algorithms. The latter incorporates probabilistic predictions for word order to produce results. From a user perspective, the EyCon application would provide the option to search solely on the basis of visual similarity, solely on the basis of text, or with a combined approach. A user could query the database using an image, or enter a text query, or provide both an image file and additional text associated with the image. Issues raised specific to metadata linked to images in archival collections are standards for image metadata. EyCon is promoting the use of the International Image Interoperability Framework (IIIF) standard. Additionally, the suggestions provided to researchers are AI-generated and there is a desire to highlight that the suggestions are not coming from an archivist or subject area expert so that researchers stay informed as to the AI construct they are working within. Lastly, without an explanation of the AI tools and AI-generated suggestions and correlations, there is a future risk of perpetuating hidden collections and limiting what researchers are aware of and ask for.

Digital Curation (DP)

Eiler, Graf, and Dorner (2018) conducted a study testing artificial neural networks (ANN) and convolutional neural networks (CNN) to pre-classify and organize historic photographs. Their sample set was from Atelier Seidel, an archive and museum for photographs in South Bohemia. They found that ANN had high accuracy of classification, which means they would be able to reduce classification uncertainty through a stable and standardized classification scheme. Additionally, ANN-based classification is accurately replicable given there is already a training dataset available. They also found that classification can be a recurring process where the training of a network can result in continuously improving classification quality and speed.

Wevers and Smits (2020) discuss the introduction of convolutional neural networks (CNN) in 2012 as a development of computer vision (CV) that enables researchers "to take the analysis of large numbers of images a step further. They can be used to explore the content (what is represented) and the style (how it is represented) of images." (p.195) This article explains how neural networks are applied to computer vision tasks and the challenge that variation in images pose to computer vision algorithms. "A CNN can recognize all these different 'possibilities' of a cat, by looking at multiple lower-level features that uniquely predict the presence of a cat in all its variations" (Wevers & Smits 2020, p.196). It is the combination of lower-level features that point to specific high-level features that enable it to predict spatial relationships, such as a cat's face. The CNN's architecture consists of multiple layers of convolutions and during training a neural network can learn the optimal configuration of convolutions to predict the correct label for an annotated image. This article provides a step-by-step explanation of how CNN was applied to a large corpus of Dutch Newspapers containing photographs and illustrations. The article explains how CNN can be used in the back-end to group visual content according to image identification and similarity. In the front-end, CNN requires a classificatory system based on both textual and visual input to retrieve accurate images based on researcher questions. Ideally, classification "should be done using the input of domain experts, and perhaps related to the research question at hand." (Wevers and Smits 2020, p.201). The article concludes with the assertion that CNNs provide an opportunity to bypass textual captions and titles for historical images and focus solely on the visual content, which can then be used for additional description. "CNNs are a new technique to query visual content of digital archives, without having to rely on textual elements" (Wevers and Smit 2020, p.204). This article highlights the benefits of computer science techniques for humanities research.

Han et al. (2022)'s study on the Frick Art Reference Library's Photoarchive developed a deep learning framework specifically tailored to the Photoarchive's collection. AI annotated images using existing headings within the Photoarchive's classification system. This study shows the benefits of collaborating between cultural heritage preservationists and AI experts, much like the EyCon project. Additionally, it shows a possibility for deep neural networks to be adapted to classify images via "specialized, hierarchical, multilabel classification systems" (p.57), which only require small changes. This allows for better performance by using pre-trained models. However, the accuracy of the label depends on how many training images are tagged with that label.

In contrast to other initiatives, the Sherratt and Bagnall's (2019) study used pre-trained OpenCV and Python scripts to identify portraits and photographs of faces at the National Archives of Australia in specific relation to White Australia policy. Unlike previously mentioned initiatives that focus on archival collections held at private institutions, art galleries and museums, this study focuses on the archival holdings of a public institution. They used screen-scrapers and OpenCV to essentially create their own framework, and is an interesting example of a sort of boot-legged process, unlike other studies in this lit review that were conducted by official organizations. This experiment brought to light non-white faces within the archives that had previously been overlooked. The authors also noted that difficulties working with RecordSearch were made easier by using OpenCV. Another paper written by Wang et al.'s (2008) details an attempt to use automatic annotation for historical images by linking to existing meta tags from contemporary images. This paper is more specifically about images within historical books where text is easily analyzed but images are left out, not standalone photographs. This is an older study but highlights the importance of using a mix of color, shape, and textures within a linking algorithm to produce proper links to tags. This was a smaller dataset, but an interesting preliminary experiment nonetheless.

We also looked at the LUSTRE project [https://lustre-network.net/], which aims to bring policymakers together with GLAMs professionals, computer scientists, and digital humanists in an effort to address AI and governmental records, especially digital ones. Their talks so far have focused on using AI for organizing email content, visual text query building, transparency in recordkeeping, and ethics, among others. However, there is little mention of specifically using AI for image records, whether digitized or born-digital.

Concerns

AI software and tools for image management in archives have limitations. Firstly, object detection software does not have a mechanism to process analog images. As a result, archival image collections without digital descriptions or digitized images will remain hidden from AI tools employed by archivists and end-users for search and retrieval. It should be made clear that archival institutions will need to invest in digitization infrastructure and activities in order to transform analog collections into data that can be used for training AI and/or data that can be curated by AI tools. The term "data" instead of records has been used purposefully to highlight the approach taken by current AI activities. It is the absence of context that should raise alarms for archivists. Absences within the archival holdings may also be accentuated by AI that is trained by what already exists, thus creating a confirmation bias for researchers of only finding what is there - instead of being able to discover the gaps in holdings and the silences of communities not represented.

CV platforms rely on metadata (sourced either from image content or the textual descriptions linked to the images) to perform search and retrieval processes. Secondly, AI software for digitized and born-digital images possess a threat to the archival bond. As Colavizza et al. (2021) argue in "Archives and AI," AI is increasingly beneficial for the organizing and accessing of digitized and born-digital archival materials; however, it is insensitive to "archival hierarchy and context" (p. 8). By removing images from their wider contextualization for identification or retrieval through AI, the software and tools reviewed for image management could cause a confirmation bias rather than "seeking the unknown or surfacing gaps and absences in the data" (Colavizza et al., 2021, p. 8). Unless some other method of "recontextualization" is provided so that the archival bonds between records of the same activity and creator can be made explicit, AI tools for retrieval and curation run the risk of treating archival records as data. Colavizza et al., provide a thorough (if not comprehensive) literature survey on AI and Archives, drawing upon articles and workshops in the archives and information science field and digital humanities from 2015 - 2021.

Although beyond the scope of this review, AI software and tools for image management also present significant ethical concerns for archives owing to how they learn. For example, Bakker et al.'s (2020) analysis of facial recognition applications OpenCV, Face++, and AmazonAWS used open-source photos with high circulation to train their software; however, this dataset lacked gender, race, and age diversity because it was composed primarily of middle-aged white men. Additionally, Proctor and Marciano (2021) found that the facial recognition systems they used (Distant Viewing Toolkit (DVT) and OpenCV) had to be tested for baseline accuracy because most of the image makeup of this collection consisted of young Black women. As archives and their holdings are a source of privilege and power, it is necessary to not reproduce structures found in analog materials in the digital and digitized realm. Using insensitive datasets to teach CV and object identification software could result in reduced retrieval and identification of BIPOC and other minority materials in favor of historically privileged materials. AI is not neutral, as it is directly impacted by human biases during the learning phase.

There are different levels of risk associated with AI algorithms to improve work processes in archives and generative AI that is unsupervised and provides results directly to the public. These must be considered by archivists in the early stages of planning and integration of AI tools. Utilizing a risk management framework prior to implementation of AI tools into archival processes and workflows is necessary to inform both internal and external stakeholders of the known risks, potential issues and possible solutions. As transparency and accountability are paramount for public institutions, it is reasonable that a risk assessment be undertaken before AI technologies and tools are utilized in both cultural heritage institutions and public archives responsible for preserving and providing access to historical records.

Conclusion

In conclusion, our survey of AI software and tools for image management in GLAM organizations from a back-end and front-end perspective found gaps that require additional exploration. Further research into how archivists are interacting with and using the aforementioned AI software and tools beyond their technical application is required. This includes how personnel are being trained to navigate these new tools and how these tools function for staff. Namely, do they reliably aid in image retrieval? Do they affect the daily activities of archivists? And do they affect record activities? This includes how AI software and tools for image management are considering and following the trust ontology laid out within contemporary diplomatic analysis through InterPARES AI. Additionally, the majority of the case studies examined for this survey were limited to a particular collection or goal, there is no discussion of how such technologies and tools can be applied beyond the scope of their specific projects. Further research into how AI software and tools are being broadly applied to image management is required. Funding and cost are also a significant point of concern. As all projects discussed relied on previously digitized or born-digital collections and out-of-the-box or customized AI software and tools, the question of how feasible these tools are for everyday image management within GLAM organizations is raised. Moreover, these out-of-the-box and black-box services fall victim to human subjectivities and errors that directly affect training sets, accuracy, and output. AI software and tools for image management within GLAM organizations would benefit from increased intra-institutional work, like that done by EyCon. Expecting future users to search across institutional collections without the structured and linked data to support such functionality will cause problems as AI software and tools advance - and as digitization and born-digital images continue to grow in an archival collection

References

Angelova, L., Ogden, B., Craig, J., Chandrapal, H., & Manandhar, D. (2021). *Deep discoveries: A towards a national collection foundation project final report*. https://doi.org/10.5281/zenodo.5710412

Arnold, T., & Tilton, L. (2020). Enriching historic photography with structured data using image region segmentation. *Proceedings of the 1st International Workshop on Artificial Intelligence for Historical Image Enrichment and Access*, 1–10. https://aclanthology.org/2020.ai4hi-1.1

Aske, K., & Giardinetti, M. (2023). (Mis)matching Metadata: Improving Accessibility in Digital Visual Archives through the EyCon Project. *Journal on Computing and Cultural Heritage*. https://doi.org/10.1145/3594726

Bakker, R., Rowan, K., Hu, L., Guan, B., Liu, P., Li, Z., He, R., & Monge, C. (2020). AI for archives: Using facial recognition to enhance metadata. *Works of the FIU Libraries*. https://digitalcommons.fiu.edu/glworks/93

Colavizza, G., Blanke, T., Jeurgens, C., & Noordegraaf, J. 2021. Archives and AI: An overview of current debates and future perspectives. J. Comput. Cult. Herit. 15, 1, Article 4 (December 2021), 15 pages. https://doi.org/10.1145/3479010

Chumachenko, K., Mannisto, A., Iosifidis, A., & Raitoharju, J. (2020). Machine learning based analysis of Finnish world war II photographers. IEEE Access, 8, 144184-144196. https://doi.org/10.1109/ACCESS.2020.3014458.

Eiler, F., Graf, S., & Dorner, W. (2018). Artificial intelligence and the automatic classification of historical photographs. *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality*, 852–856. https://doi.org/10.1145/3284179.3284324

Han, X. Y., Papyan, V., Prokop, E., Donoho, D. L., & Johnson, C. R. (2022). Chapter 1: Artificial intelligence and discovering the digitized photoarchive. In L. Jaillant (Ed.), *Digital Humanities Research* (1st ed., Vol. 2, pp. 29–60). Bielefeld University Press, Verlag. https://doi.org/10.14361/9783839455845-002

Jabarulla, M. Y., & Lee, H. (2021). Blockchain-based distributed patient-centric image management system. *Applied Sciences, 11*(1), 1-20. https://doi.org/10.3390/app11010196

Jiao, L., Zhang, F., Liu, F., Yang, S., Li, L., Feng, Z., & Qu, R. (2019). A survey of deep learning-based object detection. IEEE Access, 7, 128837-128868. https://doi.org/10.1109/ACCESS.2019.2939201

Kim, H. J., & Lee, H. (2022). Emotions and colors in a design archiving system: Applying AI technology for museums. *Applied Sciences*, *12*(5), 2467. https://doi.org/10.3390/app12052467

Milleville, K., Broeck, A. V. D., Vanderperren, N., Vissers, R., Priem, M., Van De Weghe, N., & Verstockt, S. (2023). Enriching image archives via facial recognition. *Journal on Computing and Cultural Heritage*, 3606704. https://doi.org/10.1145/3606704

Mohanty, V., Thames, D., Mehta, S., & Luther, K. (2019). Photo sleuth: Combining human expertise and face recognition to identify historical portraits. *IUI '19: Proceedings of the 24th International Conference on Intelligent User Interfaces*, 547-557. https://doi.org/10.1145/3301275.3302301

Niwata, A., & Watanave, H. (2019). “Rebooting memories”: Creating “flow” and inheriting memories from colorized photographs. *SIGGRAPH ASIA Art Gallery/Art Papers*, 1–12. https://doi.org/10.1145/3354918.3361904

Proctor, J., & Marciano, R. (2021). An AI-assisted framework for rapid conversion of descriptive photo metadata into linked data. *2021 IEEE International Conference on Big Data (Big Data)*, 2255–2261. https://doi.org/10.1109/BigData52589.2021.9671715

Prokop, E., Han, X. Y., Papyan, V., Donoho, D. L., & Johnson, C. R. (2021). AI and the digitized photoarchive: Promoting access and discoverability. *Art Documentation, 40*(1), 1-20. https://doi.org/10.1086/714604

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779-788. https://doi.org/10.1109/CVPR.2016.91.

Sherratt, T., & Bagnall, K. (2019). The people inside. In K. Kee & T. Compeau (Eds.), *Seeing the Past with Computers* (pp. 11–31). University of Michigan Press. https://www.jstor.org/stable/j.ctvnjbdr0.4

The EyCon project (eng.). (n.d.). *EyCon*. Retrieved April 24, 2023, from https://eycon.hypotheses.org/le-projet-en-quelques-mots/english

Wang, X., Ye, L., Keogh, E., & Shelton, C. (2008). Annotating historical archives of images. *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries*, 341–350. https://doi.org/10.1145/1378889.1378948

Wevers, M. & Smits T (2020). The visual digital turn: Using neural networks to study historical images. *Digital Scholarship in the Humanities 35* (1), 194–207. https://doi.org/10.1093/llc/fqy085

Yi, T., Pan, I., Collins, S., Chen, F., Cueto, R., Hsieh, B., Hsieh, C., Smith, J. L., Yang, L., Liao, W., Merck, L. H., Bai, H., & Merck, D. (2021). DICOM image Analysis and archive (DIANA): An open-source system for clinical AI applications. Journal of Digital Imaging, 34(6), 1405-1413. https://doi.org/10.1007/s10278-021-00488-5