

Case Study: The Endangered Archives Programme's use of AI tools in evaluating Jacques Toussele's Cameroonian photography archives

Kaila Fewster¹

Educational applications: This case study illustrates the uses of facial recognition and pattern matching technologies in assessing a photographic archive, which highlights the practical implementation of AI-powered tools, like computer vision and deep learning, in archival research. It also demonstrates how AI might be used for other archival practices like arranging and describing materials with photographic or non-textual collections. Additionally, this case emphasizes the importance of collaboration with professionals and experts from all types of disciplines when it comes to working AI tools on archival projects or integrating these technologies into broader archival and records management practices.

Educational topics: AI for non-textual records (photographs), types of AI/ML for photographs, AI for collection management and access, collaboration in archival AI projects².

About: This case study is part of a series of learning materials developed by InterPARES Trust AI³ researchers and educators to train archival professionals and students to effectively leverage artificial intelligence in their archival work. The final draft was completed on November 2nd, 2023. It has a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International BY-NC-SA 4.0 license, which requires that reusers give credit to the creator. It allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, for noncommercial purposes only. If others modify or adapt the material, they must license the modified material under identical terms.⁴

This case study explores how a project team in the Endangered Archives Programme, in collaboration with the Toussele family, used pattern-matching and facial recognition AI technology to evaluate the completeness of Cameroonian photographer Jacques Toussele's photographic archives held at the British Library. The goal was to estimate the percentage of Toussele's total photographic output held in a single collection. While there were no archivists on the project team, the tools used in this project have pragmatic applications for organizing and annotating poorly documented archives.

¹ InterPARES Trust AI Graduate Academic Assistant, University of British Columbia.

² Educational applications map to a Body of Knowledge proposed by InterPARES researchers for AI/ML for the archival professionals.

https://docs.google.com/document/d/1UsjkkkGeSJrgCDJGASCAy5q0Uo_ZkQpzi_Ch8XUcqYw/edit?usp=sharing

³ This case study is an outcome of InterPARES Trust AI, an international research partnership led by Drs. Luciana Duranti and Muhammad Abdul-Mageed, University of British Columbia, and funded by the Social Sciences and Humanities Research Council of Canada (SSHRC). <https://interparestrustai.org/>

⁴ Case Study: The Endangered Archives Programme's use of AI tools in evaluating Jacques Toussele's Cameroonian photography archives © 2024 by Fewster, Kaila is licensed under CC BY-NC-SA 4.0. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-sa/4.0/>

Jacques Toussele was born in 1939 in Cameroon and was an active studio photographer between 1960 and the 1990s. The archive contains roughly 45,000 images mostly scanned from negatives, but also including around 2300 scans of prints, while the Toussele family retains the originals. The project team chose two different technological approaches to determine the completeness of the archive. Both methods were applied when testing for completion of the negatives collection by determining if a scanned print's original negative was also held in the archive. This was accomplished using *Cameroon Image Search*, developed by the Visual Geometry Group at Oxford, that uses Computer Vision and Deep Learning technology to enable facial recognition and pattern-matching in images. This allowed for the project team to search for an individual's face and retrieve all other images containing a similar face regardless of age, pose or lighting, using the facial recognition modality. Additionally, using the pattern-matching modality, images taken in the same studio, with the similar clothes or the same props could be identified. This was done by selecting an input feature, like a background or floor pattern, from an initial image to locate other similar images. Using these two methods, a random sample of 105 prints was compared against the original negatives and 32, or roughly 30%, were identified. The rest of the collection's completeness was determined using non-computational methods and demonstrated that the collection has high rates of loss, and likely only holds close to 30% photographs of Toussele ever took. Overall, this case study illustrates that while archives are incomplete, AI technology like facial recognition and pattern-matching can be valuable tools for systematically assessing archival collections and their exhaustiveness and extend the possibility for access points.

Potential Discussion Questions:

1. How can institutions balance the use of new technologies like AI with traditional methods of archival research and preservation, and what skills do archivists and records managers need to work in that balance?
2. What are the potential ethical concerns related to using facial recognition technology in archival research, and how should these concerns be addressed, especially with sensitive or private images like those in personal fonds?
3. How can collaborations between technological developers and archivists be improved to ensure that AI tools are designed with the practical needs of archival research in mind?

References

University of Oxford. (n.d.). *Visual Geometry Group Software*. Retrieved November 2, 2023, from <https://www.robots.ox.ac.uk/~vgg/software/>

Zeitlyn, D., Coto, E., & Zisserman, A. (2021). Bounding an archiving: Assessing the relative completeness of the Jacques Toussele archive using pattern-matching and face-recognition. *Visual Studies*, 1–25. <https://doi.org/10.1080/1472586X.2021.1991238>

Zeitlyn, D. (2015, December 15). *Endangered archives blog: Using face recognition to find an EAP Christmas Card*. https://britishlibrary.typepad.co.uk/endangeredarchives/2015/12/face-recognition-camera-photography.html?_ga=2.85497041.95921504.1661987499-244054761.1661987499.

Zeitlyn, D. (2017, June 9). *Archiving a Cameroonian photographic studio*. Endangered Archives Programme. <https://eap.bl.uk/project/eap054>