

The Copyright Chain in the Digital Curation Process: 'Which Copyright' Project at the Nazarian Library, University of Haifa

Keren Barner

Digital Curation Librarian, University of Haifa

Abstract

The internet era, open access movement, image sharing platforms, and mainly the rise of artificial intelligence (AI) emphasise the importance of providing moderated access and licences for the reuse of collections by memory institutes as a key step in the digital curation process.

Trusting individuals to control their data and its (re)use was an important concern since the early establishment of the digital collections in 1998 at the University of Haifa Library. However, there were no specialised tools for keeping the owner's copy of rights, which led to the use of the note field in the metadata record.

This brief report will describe the experiences in a recent ongoing project to enrich the University of Haifa's digital collections with clear copyright status, with preference to open Creative Commons licences to benefit the user and the owner's experiences.

The steps performed will be explored as part of the workflow designed for the 'Which Copyrights' project 2024–2025. The scope of the project includes approximately 80,000 bibliographic records and more than 200,000 digital representations, covering local cultural heritage data. Within 7 months, over 52,000 records containing approximately 100,000 digital objects have been processed and are available with the copyright's representation notice online. In total, 25% of those records have been updated with CC BY 4.0. The success encouraged the design of the next step, to use a public platform to make the dataset available for the public. The Digital collection of the Younes & Soraya Nazarian Library, University of Haifa offers an open access portal, primarily used by the academic community. Therefore, visibility and usage will be expanded to a broader public, including via non-academic platforms, such as Wikipedia or Wikimedia. By sharing these plans, ideas, and ongoing actions, the library experience could be useful to other institutions that work with digital heritage collections, for the benefit of researchers and the broader public.

Submitted 24 December 2024 ~ Accepted 20 February 2025

Correspondence should be addressed to Keren Barner, kbarner@univ.haifa.ac.il

This report was presented at the International Digital Curation Conference IDCC25, 17-19 February 2025

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: <http://www.ijdc.net/>

Copyright rests with the authors. This work is released under a Creative Commons Attribution License, version 4.0. For details please see <https://creativecommons.org/licenses/by/4.0/>



Introduction

The 'Which Copyright' project, at the Digital collection of the Younes & Soraya Nazarian Library, University of Haifa, addresses the complexities of copyright management within the digital curation process. This initiative highlights the increasing necessity of providing open access and clear licensing for cultural heritage collections, particularly in the context of the internet era, the open access movement, and advancements in artificial intelligence (AI) technologies. By addressing these issues, this project aims to establish a framework that can serve as a model for other Galleries, Libraries, Archives, and Museums (GLAM), emphasising the fundamental principle that good research relies on good data.

The importance of this project is underscored by the common challenges faced by memory institutions in assessing and managing copyrights for extensive volumes of digital items. This project offers a structured approach to copyright management for the Nazarian Library digital collection. In addition, it has the potential to benefit a broader range of GLAM institutions. This collaborative spirit aligns with the understanding that effective digital curation is essential for enhancing accessibility and promoting research initiatives as well as public interest.

Theoretical and Practical Framework

The framework and methodology of the project draw inspiration from "Recommendations Report for Cultural Heritage Image Sharing" Knazook et al. (2023), as well as "A checklist to publish collections as data in GLAM institutions" Candela, et al. (2023). Key elements from both resources combined with in-house experience and expertise, as described in Barner (2022), have been translated into active implementation of clear licensing standards, the promotion of open access through public platforms, and the adoption of ethical sharing practices that adapt to evolving copyright laws.

The following sections present a detailed breakdown of the four chosen points that were adopted from the WorldFAIR recommendations, the steps that were proposed in the workflow, and translated into implementation actions taken at the Nazarian Library for each of them.

FAIR Principles: Findable, Accessible, Interoperable, Reusable

The WorldFAIR project emphasises that cultural heritage data should be made Findable, Accessible, Interoperable, and Reusable (FAIR). This includes ensuring that metadata is machine-readable and available across multiple platforms to improve discoverability and use. The workflow for publishing collections as data recommends the implementation of machine-readable metadata and the use of standardised metadata formats to ensure interoperability across systems. In addition, it advocates for making datasets accessible through APIs to facilitate ease of access by external platforms. At the Nazarian Library, machine-readable metadata has been ensured for all digital images using MARC fields and other international standards for cataloguing. This data is made accessible via our library management system (ALMA) and integrated with discovery platforms, such as PRIMO, ensuring that the collections are findable and interoperable. In addition, access is provided to the datasets via APIs, enabling easier integration and sharing with external repositories.

Clear Licensing Standards: WorldFAIR Recommendation

As the GLAM sector shifts towards viewing collections as data, it is crucial to clarify the copyright and licensing terms for the metadata and underlying data. A clear licence and terms of use should be provided to allow for the reuse of data without restrictions, and a suggested citation should be included for proper attribution.

The workflow suggests applying standardised licences, such as Creative Commons, and embedding clear licensing information within the metadata records. This ensures that users are aware of how they can reuse and cite the datasets.

In line with this recommendation, the Nazarian Library has adopted Creative Commons licences for its digital image collections. In addition, explicit terms of use and a recommended citation format in the metadata to guide users on proper attribution are provided. This approach helps ensure that the datasets can be freely reused while respecting copyright and intellectual property rights.

Encouraging Open Access: WorldFAIR Recommendation

Open access to data should be prioritised, and datasets should be made available through public platforms. The workflow emphasises the importance of using public platforms for data sharing and promoting open access policies. At the Nazarian Library, open access has been embraced by hosting the digital collections on institutional platforms and by actively promoting them on global platforms, such as Wikimedia Commons, Wikipedia, and Wikidata. Users and external collaborators are encouraged to contribute to these platforms, ensuring that the collections are widely available and attributed correctly. This enhances global access to the cultural heritage materials.

Ethical Sharing Practices: WorldFAIR Recommendation

In order to enable ethical sharing practices of digital image sharing, particularly in the context of emerging technologies such as AI we looked for a practical step to promote this important goal. The workflow recommends adding a term of use. The Digital collection of the Younes & Soraya Nazarian Library, University of Haifa chosen an action. It includes regularly review and update its usage terms to ensure that they comply with the latest copyright laws. In addition, appropriate copyright licences are attached to our digital representation, ensuring that users understand the legal parameters of using these materials. Protecting intellectual property is a focus in the context of AI by specifying usage rights and restrictions, helping to safeguard the interests of cultural institutions and creators.

By focusing on these areas, the project seeks to foster an environment where cultural heritage materials could be responsibly shared and reused; therefore, maximising their impact on research and education.

Workflow Action Points

The workflow for processing digital collections at the Nazarian Library consists of 10 key steps. First, the scope of the existing digital collections is assessed to evaluate their extent. Then, the collections are selected for processing, determining which ones will be addressed in this workflow. The permissions are reviewed from collection owners to ensure that existing agreements are up to date. Following this, different copyright licences and usage terms are studied to identify the most appropriate options for each collection. The next step involves exploring copyright options within the digital object management

system to determine how best to integrate licences into the platform. Once this is established, a data sheet is prepared for the selected collections, which includes owner names and contact details. In addition, current contact information for collection owners or their heirs is updated and searched, especially when contact details have changed. After the correct contact information is obtained, the owners or their heirs are contacted to secure updated permissions for the collections. Then, collection records and their digital representations are prepared, ensuring all necessary details are included. Finally, the appropriate licences are added to each digital object, and the records are updated with the new copyright information to reflect the changes. In addition, guidelines are incorporated for 'how to credit' the collection materials within the record metadata. These 10 practical steps helped to perform an efficient workflow.

Outcomes

By the conclusion of the initial phase, the project timeline for 2024–2025 had achieved significant milestones by the end of 2024 (halfway). Over 100,000 digital objects have been processed with appropriate copyright notices, and 52,000 bibliographic records have been updated to include essential copyright metadata. This progress reflects a commitment to transparency and accessibility, with the library successfully reaching over 60% of its goal to process 80,000 bibliographic records within 10 months. In addition, as part of the pilot project for the Wikimedia initiative, the Nazarian Library successfully contributed 220 images to Wikimedia, 100 images to Wikipedia values, 20 files to Wikidata, and 20 more files to the Timna Wikibase project. These images were carefully attributed, ensuring proper citation and adherence to open access principles. This contribution expanded the visibility of the digital collections and fostered greater collaboration with global platforms, enhancing the accessibility and dissemination of cultural heritage materials. The success of this pilot demonstrates a commitment to promoting open access and sharing knowledge via widely used platforms and will soon be expanded.

Summary Insights

Insights gained from the project extend beyond copyright management. They include a deeper understanding of the library's collections and their potential for researchers. The project has prompted the library to standardise its collections, aligning them with international data sharing norms, which enhances their usability and accessibility. In addition, the initiative has invigorated the team's motivation and commitment to pursuing further collaborations and skill development, demonstrating the transformative impact of engaging with copyright issues and digital curation.

In conclusion, the 'Which Copyright' project is demonstrating a cohesive approach to managing digital collections responsibly and ethically.

In addition, the designed theoretical and practical framework and the 10 detailed steps for the workflow action points could provide a valuable model for other GLAM institutions that aim to navigate the complexities of digital curation in an increasingly interconnected world.

Work on the 'Which Copyright' project will continue during 2025, which aims to accomplish the project goals and more.

References

- Barner, K. (2022). The creation of primary sources digital collections in an academic library. *Library Philosophy and Practice (e-journal)*. Retrieved 01.05.2024 from <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=14582&context=libphilprac>
- Candela, G., Gabriëls, N., Chambers, S., Dobрева, M., Ames, S., Ferriter, M., Fitzgerald, N., Harbo, V., Hofmann, K., Holownia, O., et al (2023). A checklist to publish collections as data in GLAM institutions. *Global Knowledge, Memory and Communication*. Advance online publication. DOI: <https://doi.org/10.1108/gkmc-06-2023-0195>
- Knazook, B., Murphy, J., Barner, K., Cassidy, K., Claeysens, S., Cortese, C., Warren, M. et al (2023). *WorldFAIR Project (D13.2) Cultural Heritage Image Sharing Recommendations Report*. Retrieved from <https://zenodo.org/records/7897244>