

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

2022

The Creation of Primary Sources Digital Collections in an Academic Library

Keren Barner

University of Haifa, kbarner@univ.haifa.ac.il

Follow this and additional works at: <https://digitalcommons.unl.edu/libphilprac>



Part of the [Arts and Humanities Commons](#), and the [Library and Information Science Commons](#)

Barner, Keren, "The Creation of Primary Sources Digital Collections in an Academic Library" (2022). *Library Philosophy and Practice (e-journal)*. 7535.

<https://digitalcommons.unl.edu/libphilprac/7535>

The creation of primary sources digital collections in an academic library/ Keren Barner

Description of working model & processes in the creation of digital collections at the Younes and Soraya Nazarian Library, University of Haifa

Abstract

This article discusses the creation of primary sources digital collections in an academic library, including implementation, work processes, digitization and providing accessibility of visual materials. These reflect the practical application of the approach developed in The Younes and Soraya Nazarian Library at the University of Haifa regarding the library's position in the field of digital collections. The article denotes the work process model created for these collections, from their reception, through classification, cataloging, scanning and digitization, and to the digital publishing which provides access. Special emphasis is given to practical ways to manage the challenges that arise from handling unique visual materials such as photos, albums, manuscripts, and maps.

Keyword

Digitization, Digital collection, Preservation, Academic library, Digital edition, Scanning

Introduction

This article will deal with the creation of a digital collection model. Implementation, design of work processes, digitization, and accessibility of materials of a visual nature or creation of digital editions of the physical collections. All of these are a practical reflection and translation of the concept that took shape in the Younes and Soraya Nazarian Library, Haifa University (hereinafter...) about the library's place in the field of digital collections. The work process model we created from receiving the materials to their access through digital publishing will be described, with an emphasis on the challenges that arise from handling unique material in terms of their visual nature, such as photos, albums, manuscripts, and maps. visual? digital? The terms and the difference between them must be clarified. One of the main goals of a library is to provide quality service and support to its target audience. In the areas of the digital collections of the Younes and Soraya Nazarian Library, the University of Haifa prefers the use of the term's customer and/or user over audience. First, because they describe in a better way the dynamic and active relations between the library and customers/users not only as information providers but also as information creators of primary materials in digital format. Secondly, these terms specify the identity of the recipients for the benefit of which the library has set itself the goal: a better understanding of the changing and up-to-date information needs of the customers/users and responding to them. The library works to achieve this goal through skilled and professional librarian personnel, using innovative tools and systems. Great emphasis is placed in the work on rich and enriched metadata, and on special and meticulous treatment of the source materials themselves. All these actions are aimed at enabling better access to research, teaching and learning material and more successful retrieval of suitable materials for the customers/users Who is our "user"? Well, the users of an academic library are relatively easy to identify and define, these are the researchers, lecturers, and students. The focus? As an academic library, it is as mentioned to identify and respond to the information needs of these users for research, study, and teaching needs. S.R Ranganathan the librarian and theorist sums it up

nicely in his theory The Five Laws of Library Science (Ranganathan, 1931): 1. Books are for use 2. To each reader his own book 3. Every book has its own reader 4. Save the reader's time 5. The library is an evolving organism The adjustment and changes required Why are they required? For especially light digital collections, by changing the term: book to the term digital representation (representation) and the term "reader" to "user", the five laws preserve the original spirit of things and update them to the present day. As part of the digital collections of the Younes and Soraya Nazarian Library, Haifa University, primary and secondary material files are digitally accessible in a variety of formats: photographs, texts, manuscripts, maps, video, and audio using the PRIMO platform of exlibris, a company specializing in information systems for libraries. Description of the working model on a digital collection. For the benefit of working on digital collections, the library has developed a work model that includes the various work processes that take place from the receipt of the materials to their access to the users, with the participation of librarians and student library employees who deal with the various functions. The various stages will be reviewed with an emphasis on highlighting the challenges arising from the handling of visual materials and other significant points for reference, but first the model itself:

Working model for digital projects

Process chain



Diagram #1: Working model for digital projects/collections

Admission

The admission process includes various technical aspects related to aspects of organization and transportation. Because these are identical to the treatment of other physical collections that are received by the library, they will not be discussed in this article. Two significant and perhaps slightly different points between receiving a physical collection and receiving a collection intended for digitization that will be discussed are the contributors and the issue of copyright. Who are the donors? Well, a wide variety of types of contributors are behind the projects and digital collections accessed through the library. Among them are private individuals who have personal material documenting historically and socially significant events and places, researchers and the research material collected during their research, small archives of localities such as Nesher, Kibbutz Yagor and others, institutions and departments at the university such as the Department of Theater which delivered and continues to deliver the department's productions as a partner, Jewish Archives Alexandria located in Italy and its materials were donated through the Chair for the Study of the Heritage of Egyptian Judaism named after Yosef Rashlin Berda, the Faculty of Humanities, University of Haifa, the collection of photographs of the photographer Alex Liebek and many others. What they all have in common is that they hold significant material for research, teaching, and study purposes as well as the existence of the material is of value and interest to the public as part of the preservation of cultural heritage. A second crucial point is the reference to the issue of copyrights. A complicated issue that this article does not pretend to cover, and yet we emphasize that when receiving the materials, a careful inquiry was made

on the subject. Are the owners of the materials the legal rights holders of the material? Is it possible to publish it to the public or is it necessary to limit viewing permissions and sometimes even metadata cataloging only and keeping the items in an archival status that does not allow access to the user audience in the hope that the materials will be opened at a future stage when the issue of copyrights will be settled. The coordination of the expectations and the sharing of the joint and separate interests of each of the owners of the parties, that is, the donor, in front of the university library is an important and decisive step regarding the acceptance of the collection for treatment and the determination of the various parameters for the method of treatment, the scope and the way of making the materials accessible. The transfer of the materials is facilitated by the process of digitization in this respect. Donors bring the original materials to the library for care but do not separate from them completely, they receive the materials back after the digitization process and in this way get a high-quality scan of the materials and receive a copy of them without giving up their physical possession and/or their copyrights at the commercial level. The library, for its part, is happy to return the original materials and keep only the digital copies both due to a lack of physical space in the library space, (a problem known to many libraries) and above all the possibility of maximizing the possibilities inherent in digital editions for the benefit of the users. Regarding authorization of use for research, teaching, and study purposes, these are regulated in an official document signed between the donor and the library. Through this document the library allows users to make fair use of the materials. The contact details for the rights holders are entered as part of the record's metadata so that anyone interested can contact the rights holders. Despite the attempt for the series described above, it is not always easy to locate the donors due to age, changes of electronic addresses, telephone numbers and others. The library makes a considerable effort to assist applicants with a large degree of success. An interesting example of a collection whose reception was different is the Ilanot collection, by Prof. J. H. Chajes from the Department of Israeli History at the University of Haifa. The collection, which includes over 500 Kabbalistic verses, was in various libraries and archives around the world, as well as in various private collectors. The work of locating the owners of the rights and their consent to the inclusion and presentation of the items within the digital collection of the University of Haifa Library included contacting each one of the owners of the rights and obtaining a valid legal authorization for this purpose. This Sisyphean activity was done by the researcher himself and his team. In cases where the rights holders were not located or the approval is still in the processing process, they were closed in the system using closing permissions that do not allow viewing by the public except for the metadata data in the record, these were of course created by the library. Of course, this phase also includes aspects of organization and transportation that are identical to the treatment of other physical collections that are taken in by the library, besides, as mentioned, returning the collection to its original owners. .

Classification

It is important to note that the library's digital collections constitute a representative sample of the materials of the original collection and are not a complete coverage of the items in the entire physical collection. The decision on the scope and selection of the collection items to be handled is made according to the advice of a selected academic staff member who serves as an advisor and companion to the selected collection and the librarian responsible for the collection. In some cases, and when necessary, the donor's own opinion is also considered. The sorting process also includes a basic labeling of the collection structure in terms of a super-collection hierarchy and sub-collections according to the different contents and types belonging to the collection. This hierarchy is tested again and again in the process of working on the project until its final accuracy is manifested practically in the stage of

dealing with digital accessibility. That is where the collection was built and accessed on the collection's portal.

An example of a collection hierarchy

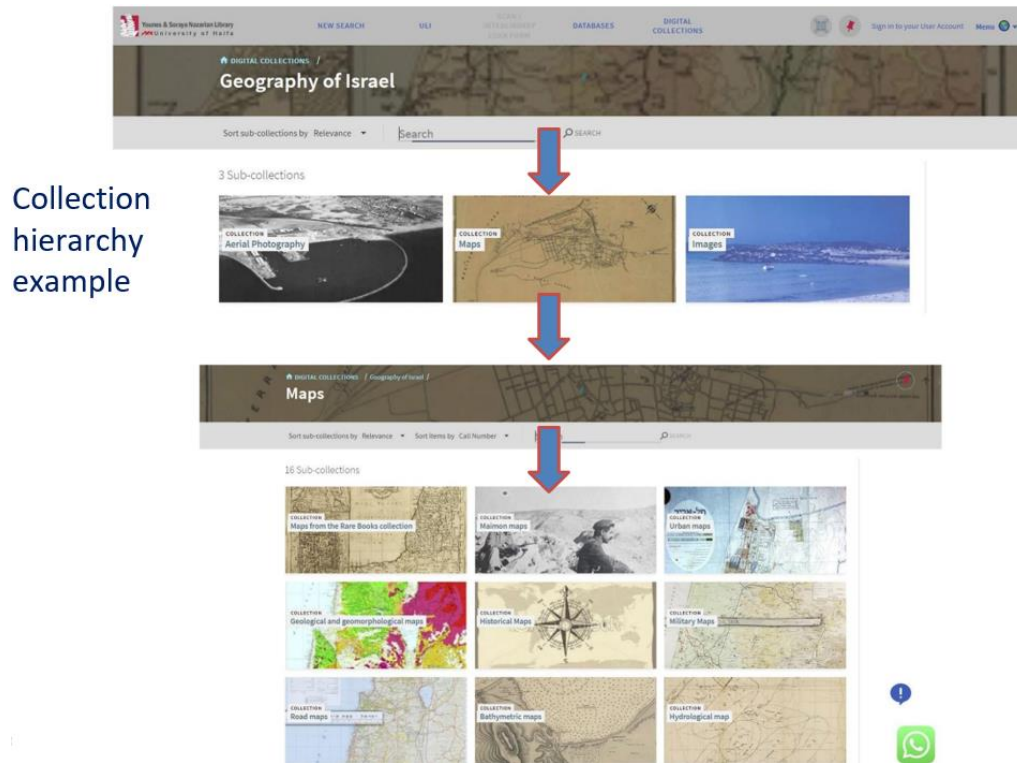


Diagram No. 2: Hierarchy of Collections: Collection on GG of Israel and the ranking of the sub-collections below it.

An example of the visibility of a digitally accessible collection

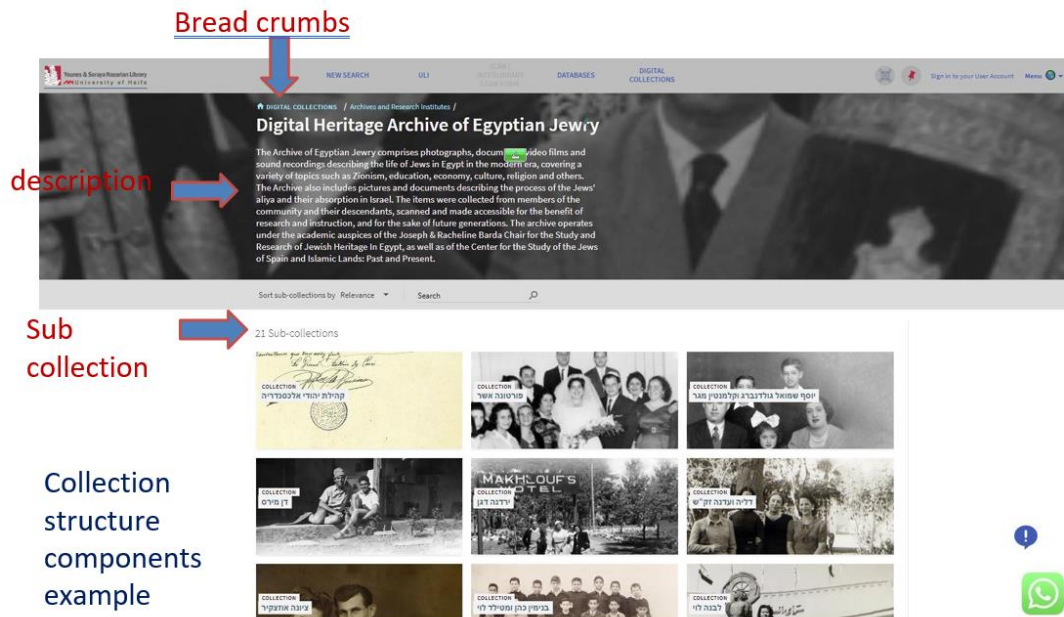


Diagram No. 3: A look at the digital visibility of a super collection and its sub-collections

Scanning and digitizing

The activities related to scanning and digitization are related to each other and therefore appear at the same stage of the work processes although there is a certain separation between these functions. The use of the term digitization in our case refers to the process of converting materials in physical format into digital files and making them accessible in a dedicated web-based system for users. With original materials such as paper, parchment, photo, the digitization process will begin with scanning. The main goals of the digitization processes regarding library materials (Pandey, 2020) concern the optimization of the accessibility of the information items to users in our normal days who expect to access materials remotely via the Internet as well as to improve the preservation level of sensitive and other library materials (Prescott, 2018). High-quality and good scanning of the materials (Malaperdas, 2021) is important not only as a technical step on the way to the digitization and publishing of digital collections, and this is a significant point. Scanning with high standards and quality will provide researchers with materials that can be worked with, treated and manipulated with unique tools such as the digital humanities (Cannelli, 2021), GIS, and more, while relatively low-quality scanning will be well used for teaching and learning needs but not beyond that. Since we have the equipment and skilled manpower for high-quality scanning, we will in most cases prefer high-standard scanning. The world of scanning equipment is well documented and therefore I will mention only a few significant points concerning the creation of digital collections. Two main methods exist in scanners and are manifested in flatbed scanners and overhead scanners equipped with a digital camera. At the Younes and Soraya Nazarian library we use both and use Epson pro750 flatbed scanners and an EDS GAMMA METIS overhead scanner (Metis-EDS-Gamma, 2022) dedicated to scanning cultural preservation materials.

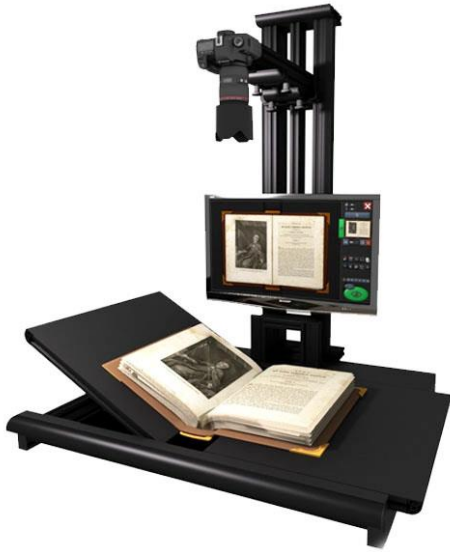


Chart No. 4: EDS GAMMA METIS overhead scanner

Below are some of the outstanding advantages of the scanner not only in terms of quality but also in accordance with library work processes in general and digital collections in particular

- Automatic image naming with user selectable rules (customizable fields, programmable increments, programmable actions, etc.) Automatic or manual exposure and gray balance control (light color temperature compensation)
- Automatic or manual focus control (with Depth Of Field customization)
- Automatic and Manual Crop
- Automatic deskew and curvature correction
- Automatic book center and shape recognition; automatic page split with definable overlap
- Light uniformity correction filter
- Paper color correction filter • Automatic Finger recognition and filter removal
- Anti-Reflection shooting mode (for glossy originals)
- EDS Light Control directly from the EDS software
- Keyboard shortcuts • Image saving in grayscale and color in TIFF, JPEG, JPEG2000, BMP, PNG, PDF, PDF-A, Multipage PDF, Multipage TIFF, TIFF G4 1bit Another notable advantage that is helpful is the control panel on the touch screen that allows even less skilled personnel to use it after dedicated guidance and training. The students employed as scanners all have a background in photography, art and usually also graphics, but naturally without experience in professional scanning and this is reflected in the particularly convenient user weapon. The scanner enables work at a high-quality level relatively easily after appropriate dedicated training through manuals written by us and training tasks that the manpower designated for scanning in the industry goes through.

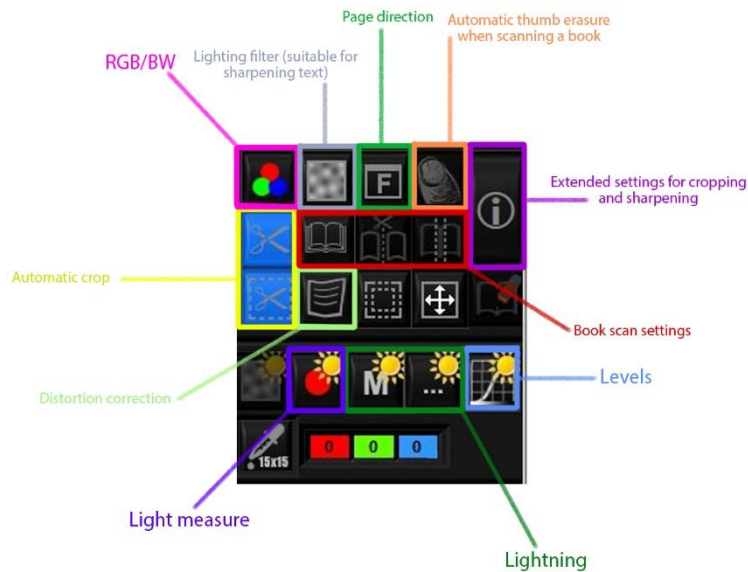


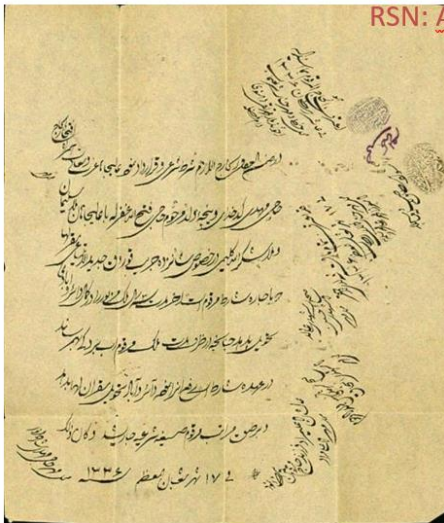
Diagram No. 5: EDS GAMMA METIS touch screen control panel

The three most significant parameters for obtaining a quality scan are resolution, color depth and signal-to-noise ratio (SNR) (Malaperdas, 2021). It is customary to use the term resolution in relation to the number of pixels per inch in digital images expressed in PPI. Our library follows the international standards and tends to adopt the high standards. Today we scan most of our materials at 600 PPI and in some cases, such as slides for example at 1200 PPI. High PPI scanning has advantages and disadvantages that will be discussed later. The depth of color refers to the ability to bring out the variety of tones, shades of the original material. These are represented in bit units. 1bit will represent the material in black and white, 8bit will represent the material in 256 color shades and 24bit will already allow the representation of over a million shades. Even in the case of materials that seem almost monochromatic, such as manuscripts for example, we prefer a high color depth of 24bit that can bring out every detail of the original material. The term signal-to-noise ratio refers to any unwanted component in the image as noise with a general guiding principle that the higher the signal-to-noise ratio, the better the quality of the scan. However, this term requires the highest level of professionalism and here the software component of the overhead scanner auto distortion fix is of high value and particularly effective. So why not raise the resolution indicators and signal-to-noise ratio in advance? Well, there are several reasons for that: The first is the result of extremely large files and the high storage costs of materials, although since cloud storage services we have witnessed a significant drop in prices and therefore this is not the decisive reason especially when the goal is long-term preservation. The options for use by the user are a much more significant reason. Many programs often have difficulty importing and editing particularly large files, alongside the fact that in many cases the resolution is too high, and the signal-to-noise ratio is not proportional. Not only do they have no added value, on the contrary, they add "noise" in the form of lines and dots that are not visible to the human eye in the original material. Examples of scans with different resolutions and color depth - the scan on the right is of higher quality, true to life color and without "noise" expressed in dots and tone changes that can be seen in the example on the left

Resolution: 300 ppi VS 600 ppi

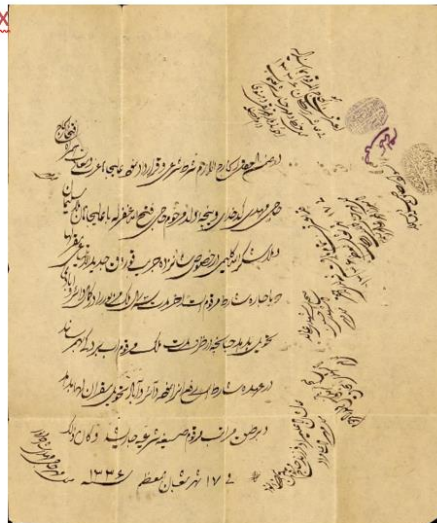
Color depth: 8 bit VS 24 bit

RSN: Autofix



color distortion and noise

of Haifa, Nazarian Library - All rights reserved



accurate coloring and less noise

Chart No. 6: Comparison of scans in resolutions, color depth, signal-to-noise ratio

To conclude this phase, it was said that from the work experience, it seems that two significant points for a high-quality scan: suitable equipment and competent and skilled personnel. The digitization processes include, as mentioned, the conversion of materials from a physical format to a digital format, which is an intelligent file date used to prepare them for entry into a digital object management system. These are done in the Younes and Soraya Nazarian Library using the ALMA-DIGITAL EXLIBRIS system (ALMA manage digital resources, 2022), a URM framework for managing digital resources. The management of the digital collections in the system includes the construction of the collection tree, a hierarchy of super collections and sub-collections, loading processes that combine using marc xml between the records and the content files to create digital representations with instances for the various formats tif, jpg, mp4, pdf, (Brenner, 2021), creating representative thumbnails (figurines) to the collections, the introduction of verbal descriptive content in English and Hebrew.

Cataloging and metadata

The cataloging and metadata phase is a significant and well-known phase in the world of librarianship and informatics and not only in the context of digital collections and visual materials of course. Many challenges arise in the cataloging phase due to the uniqueness of different types of visual materials, only a few of them that are relevant to the work processes in digital projects will be briefly mentioned here. Cataloging should include the following elements and through them we will try to clarify and illuminate the considerations and challenges in handling visual materials: zip code, cataloging rules, mandatory fields, known/unknown details Postal code - the guideline is to set a distinct logical postal code for each record. In the case of visual items, manuscripts and/or parts of books chosen to be represented in the collection (photos, maps, or Kabbalistic charts for example in Prof. J. H. Chajes collection of Kabbalistic Ilanot) is it a single item or a group of items that should be

represented in one record that includes several images or each image separately? If the item is part of an existing work, how will we make sure that the context of the work will be clearly documented in the metadata? How to create the contexts between the different items we curate into a collection with common content from different sources and more. There is no single answer to all of these questions and a lot of thought, professionalism and familiarity with the various fields and the possibilities of using them is required and perhaps the most important and significant thing in the context of the work processes on digital collections that must be carefully considered at the beginning of the work is what the desired end result is, how we would like the digital collection to look to the user and to examine how it is possible to catalog the item in such a way that this intention will be fulfilled.

Known/unknown details - in non-library materials many data may be missing (author's name, year of photograph, identity of people and places recorded) and legality must be decided in the way these gaps are handled. In addition to the expertise and adaptation of a cataloger to the type of content of the collection, we try to attach academic advice from a faculty member who is an expert in the field to each collection and/or digital project to try and complete these gaps even in the identification phase prior to the cataloging. The cataloger reflects the remaining gaps in the record according to the cataloging laws and local laws established for this purpose. Cataloging rules - as with any other library item, you must choose the accepted cataloging language at the institution and work according to it. In our case MARC. In unique materials such as manuscripts, for example, challenges arise such as the existence of different editions that are similar but not identical in different libraries around the world, in archives and with private collectors, without a known author, without a known year, with different names and difficulties from copyright areas. All of these should be reflected in the catalog under its strict rules. One of the ways to be flexible is to develop dedicated local fields for the more complicated issues. As mentioned, while maintaining the catalog rules and consistency that does not conflict with the other items in the library catalog. Mandatory fields - each catalog record has mandatory fields that are necessary for its existence. One such field is the header field. One of the significant challenges in digital image type materials is the title. Unlike a book published in an organized publishing house, even if it has a well-known title and a large number of printings and editions, in digital images it is usually a unique item that does not exist in any other database/collection, and its visual content forces the cataloger to "invent" a title, also he must make sure that it is worded in the manner of the creator The connection to the collection to which the item belongs but is also meaningful when retrieved from the catalog as a single item. The collaboration between researchers and catalogers is extremely important in the handling of research digital projects as well as in the handling of digital collections in general because they contain primary materials that require both extensive general knowledge and academic expertise in the narrow fields. A suitable example of this is the Ilanot Kabbalah project of Prof. J. H. Chajes, a researcher and faculty member of the University of Haifa. In his article "Reading the maps of the divinity of the Kabbalists" (Hayat, 2015) Prof. describes and explains that "in Kabbalistic literature there appear tree-like diagrams known as Ilanat. These Ilanat served as decorative additions to the research books published in recent decades but were hardly discussed in the books themselves - and this even though the Ilanat constitute an important Kabbalistic literary genre in its own right. This genre, which has been neglected in research until now, is at the center of my work. The Ilanat diagrams present the Deity as it was understood by various Kabbalists from the 13th century onwards. Because Kabbalists tended to imagine the Deity in a rather mechanical, technical, and complex way, these diagrams were essential for various uses in Kabbalistic practice: for learning and

be searched and retrieved through the collection's portal and/or through the library catalog, both access points are embedded on the Younes and Soraya Nazarian Library website. Chart No. 8: Digital Collections Portal, screenshot from the website of the Younes and Soraya Nazarian Library, University of Haifa Each item in a digital collection includes rich metadata data and representations containing the scan files of the primary material itself in an appropriate format where the terms of use are of course while maintaining copyright laws. We offer three types of formats according to the type of material: tif, jpg, pdf Tif - heavy and high-quality format, suitable for long-term preservation, for high-quality printing, as well as for use with advanced DH tools Jpg - a relatively easy format, convenient for online use for a wide variety of uses PDF - accepted format for materials defined as documents containing written material of various types (handwritten, printed...). Some of these materials are also OCRed based on professional consideration of added value to the item and other considerations. All digital collections (Digital collections, 2022) are accessed through the two forms contained in the exlibris PRIMO platform iif viewer and alma viewer These days we are working on introducing local improvements in the weapon for the convenience of the user. Most of the digital collections are open access and therefore accessible to all except items whose usage permission has been restricted due to copyright issues.

Future aspects, summary, and conclusions

The author wanted to share her experience within the Younes and Soraya Nazarian Library, Haifa University in the field of building digital collections of original materials of a visual nature and hope that this will be of interest and help to libraries, archives and any interested party engaged in and/or intending to engage in the field. The guiding concept as an academic library sees the library as a body that is not satisfied with the role of providing information but also a producer of primary information in its digital form (Barner, 2011) and this is the guiding rationale for a large part of the digital collection tree offered to the public of researchers and students and the general public within the digital collections portal of the Younes and Soraya Nazarian Library, Haifa University. In the article, a work model was presented for working on digital projects, the meaning of which goes beyond the detailed description of the processes, is the definition of an orderly work process in which there are many control stations, contributed by those with various expertise in the library whether in the field of copyrights, scanning, cataloging, sorting, library information systems and more. Much emphasis is placed in this professional article on the scanning stage not only as a technical stage but as a significant stage in the preparation of the material not only for the immediate term but also for the long term and for the exploitation of advanced technological and analysis possibilities from the fields of digital humanities for example, on the possibilities of accessibility in different and many formats to exploit the possibilities of use inherent in them. The three significant insights from the work experience and from the professional literature on working on digital collections are the need and advantage of close cooperation with researchers and listening to their needs (Costis, 2017), searching for ways of doing things that are characterized by flexibility and creativity and constantly examining what is being done in the world and in Israel in similar and different projects to receive new ideas and profitable collaborations. Looking to the future, all the above emphasizes the importance of constant updating and development on the part of the library both in the technological field and in the content aspects to respond to the emerging needs in the world of research and even anticipate them if possible. To this end, an active acquaintance with

innovative fields in the academic world and in the commercial world, examination analysis and adoption of innovative systems and tools, attention and a deep connection with the research staff are required. The adoption of these steps will improve the support of the academic library to the users of the library - the researchers and students and will allow them to fulfill their academic goals and thereby expand human knowledge.

Thanks

Thanks to Naomi Gridinger, director of the Younes and Soraya Nazarian Library, University of Haifa, to Dr. Ricky Greenberg, head of the TADAM branch, to Prof. J. H. Chajes

and Dr. Eliezer Baumgarten, Ilanot Project, University of Haifa to Dr. Moshe Lavie, head of the Eliyahu Laboratory, scientific The digital spirit, the University of Haifa and a big thank you to the MMS team, my librarian colleagues, and the wonderful scanners for their generations for their crucial contribution to the digital collections.

Bibliography

Alma for digital resources. ExLibris. Retrieved from

<https://exlibrisgroup.com/products/alma-library-services-platform/digital-resources/>

Barner, K. (2011). [The Library is a Growing Organism: Ranganathan's Fifth Law of Library Science and the Academic Library in the Digital Era](#), *Library Philosophy and Practice* ,9p.

Retrieved from

<https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1579&context=libphilprac>

Barner, K (2021) Webinar Alma-D workprocess

<https://meli.org.il/wp-content/uploads/2021/04/ALMA-D.pdf>

Cannelli, B. et al. (2021) [Next Generation Careers in and from Digital Humanities: Third Discussion Paper of the UK-Ireland DH Network](#). Discussion Paper. Digital Humanities Research Network.

Chanjong I., Yongho K. ,Thomas M. (2022), Deep learning for historical books: classification of printing technology for digitized image. *Multimedia Tools and Applications* 81:5867–5888
retrieved from: <https://link.springer.com/content/pdf/10.1007/s11042-021-11754-7.pdf>

Chajes, J.S (2015) Reading from Maps of God

<https://isf.org.il/layperson-reports/%D7%9C%D7%A7%D7%A8%D7%95%D7%90-%D7%90%D7%AA-%D7%9E%D7%A4%D7%95%D7%AA-%D7%94%D7%90%D7%9C%D7%95%D7%94%D7%95%D7%AA-%D7%A9%D7%9C-20190101153357.html>

Costis, D. et al. (2017) European Survey on Scholarly Practices and Digital Needs in the Arts and Humanities. Project Report. DARIAH; DIMPO. (doi: [10.5281/zenodo.260101](https://doi.org/10.5281/zenodo.260101)). Retrieved from <https://zenodo.org/record/260101#.Yx657LTP2Uk>

Digital collections – Younes & Soraya Nazarian Library, University of Haifa. https://haifa-primo.hosted.exlibrisgroup.com/primo-explore/collectionDiscovery?vid=HAU&inst=972HAI_MAIN&lang=en_US

Ilanot – maps of god. https://ilanot.haifa.ac.il/site/?page_id=37

Malaperdas, G. (2021). [Standards for Digitization in Cases of Maps, Documents, and other Relics in the Service of Cultural Heritage](#). *Open Journal of Educational Research*, 1(1), 24–31. Retrieved from <https://www.scipublications.com/journal/index.php/ojer/article/view/99>

Metis-EDS-Gamma www.metis-digital.com retrieved from <https://iguana-idm.com/wp-content/uploads/iGuana-Metis-EDS-Gamma-Planetary-Book-Scanner-Brochure-Specifications.pdf>

Pandey, R. & Kumar V. (2020). Exploring the Impediments to Digitization and Digital Preservation of Cultural Heritage Resources: A Selective Review. [Preservation, Digital Technology & Culture](#), 49(1): 26–37. Retrieved from <https://www.degruyter.com/document/doi/10.1515/pdte-2020-0006/html?lang=de>

Prescott, A. & Hughes, L. (2018) Why Do We Digitize? The Case for Slow Digitization – case study. *Archive Journal*. retrieved from <https://eprints.gla.ac.uk/165726/7/165726.pdf>

[Ranganathan](#) S. R. (1931) The Five Laws of Library Science. Madras Library Association