

# Sailing to Uncharted Territory - Digitool to Alma Digital Migration

Marina Kalach
Boaz Dotan
University of Haifa

Josh Weisman
Ex Libris



# **Our University:**

- ☐ Located in Haifa, Israel
- **□** 18,000 Students
- ☐ 1 Central Library
- ☐ Over 2M items in our catalog
  - Books
  - Journal
  - E-resources
  - Databases
  - Media
  - Rare books
  - Archives
  - Psychological Tests
  - Children's Lit









#### **Agenda**

- ☐ Library Digital Collections
- ☐ Digitool: The source
- Migrating to Alma
- ☐ Migrating Digitool to Alma
- ☐ The new collections page
- ☐ Special Thanks



## **Library Digital Collections**

- ☐ Cultural heritage photographs
- ☐ Historical Document archives
- ☐ Rare and ancient digitized book



### **Digitool: The source**

- ☐ We have Digitool since 2006
- ☐ All the library collections and digital objects were uploaded and stored in it







$\mathcal{U}$	M Oigital Media Center - The Library		University of Haifa
Search   F	Results   Previous Searches   My Space   User Guides (Hebrew)	About   Conta	ct Us   עברית Login   End Session   Hel Gue
Simple Sea	arch Advanced Search		
Select collec	ction: All	$\overline{v}$	search
A word or ph	hrase:	Exact Starts	With
University of Haifa Dissertations	Haifa University Theses and Dissertations Collection M.A. and Ph.D. dissertations from Haifa University, until 2017.	University of Haifa Institutional repository	Publications of University of Haifa Researchers Publications of University of Haifa researchers, including articles, conference proceedings, research reports, technical reports and assorted internal publications (the database is in the process of being compiled).
	Theatre & the Performing Arts  Texts, photographs, reviews, designs, audio and video recordings, that comprise the Database for Israeli Theater Archives (DITA). The collection includes materials from the archives of prominent Israeli theater institutions, including the Haifa Theater, the Cameri Theater of Tel Aviv, the Acco Fringe Theater Festival, and the Mario Kotliar Archive of the School of Visual Theater, Jerusalem. See also: <a href="mailto:theater">the</a> collection's website		History of art Image collections in the fields of painting, drawing, printing, sculpture, architecture, ceramics, manuscripts and applied arts. These collections are used as accompanying materials for courses in the departments of Art History, Archeology, and General History.
it.m.	Historical Pictures of the Holy Land and the State of Israel Photographs documenting Israel's unique landscapes and people, by professional and amateur photographers, from various archives, institutions and private collections. The images are available to the public, under fair use and copyright limitations.		Archaeology of the Land of Israel Images of archaeological finds (including maritime archaeology) from excavation sites in the Land of Israel, arranged by site name in alphabetical order. The images are available for educational and research purposes for University of Haifa students and faculty.



See also: the collection's website

Images depicting geological phenomena, landscapes, environment and culture, including maps of the Land of Israel and historical maps of the world, aerial photographs and satellite images, divided into material types (photographs, maps) and locations.



#### Crusader Sculpture in the Land of Israel

This collection is the fruit of research conducted by Dr. Zehava Jacoby, of blessed memory. It is arranged by site name in alphabetical order, and divided into building types, such as: churches, monasteries, mosques, tombs, fountains, and museums.

Digital Dariadicale





## Migrating to Alma

- ☐ In August 2016 we migrated to Alma
- ☐ As part of the migration project we told Ex Libris that Digitool needs to be migrated to Alma as well
- ☐ We knew we would need to wait some time until Alma Digital could support roughly the same functions as Digitool



#### **Overview**

- ☐ In Spring 2018 we began the first steps, after new options in Alma Digital module were introduced.
- □ Our collection contains ~250,000 digital objects in various file formats.
- ☐ We needed to plan the migration process from the aspects of both quantity of materials and timing.



#### **Collection Tree**

- ☐ Collection tree => How to migrate it without having to recreate it from scratch in the Alma environment?
- ☐ We shared this concern with the Alma-D team in Ex Libris and in April 2018 we got a new migration tool for Digitool
- ☐ This generate\_collection\_EAD\_file script created an EAD XML file format



#### **Collection Tree**

- ☐ First import of the Collection Tree was performed on our Alma SB environment
- ☐ The result was reviewed by our digital collections librarians
- As a result of this test load it was decided that before migrating the collection tree to the Prod environment it should be as clean as possible.



#### **Alma Digital workflows**

- After the load of the collections tree to Alma SB our digital collections librarians started to learn and experiment with the digital import workflows
- □ Also, new "Alma born" records and collections were created via the digital import profile in Alma SB
- ☐ The creation of new records and collections began in the Prod. environment after sufficient testing in Alma SB environment



#### **Our workflows**

- ☐ Historically: catalog the records in Aleph, export them to MARC XML format and ingest them with the relevant files to Digitool
- ☐ In Primo the records were deduped, Aleph was the bib records source and the digital files came from Digitool
- ☐ In Alma, the workflow maintained with the minor change of cataloging and exporting from Alma



# Migrating Digitool 2 Alma Import objects from Digitool

- Now as part of the process Digitool server was planned to be shut down, resulting in the Alma records becoming itemless
- ☐ So the next step in the project was to experiment with import of digital objects from Digitool to Alma without the Digitool metadata, which means adding a digital representation to the Alma bib records



# Migrating Digitool 2 Alma Import objects from Digitool

- ☐ In order to achieve this Ex Libris developed a new process that utilizes Digitool harvesting integration profile using Digitool OAI-PMH
- ☐ On our side we create an import profile based on the Alma-D team guidance
- We did some successful testing on our Alma SB environment until we felt ready for the next challenge



# Migrating Digitool 2 Alma Import objects from Digitool

- With a lot of help from the Alma-D team at Ex Libris we create repository replications in Digitool
- We needed something to do the match between Digitool records and Alma records
- We decided to use Digitool MARC 001 field to do the match; in the replication process using XSL we changed the 001 field from Aleph format to Alma format

Aleph number: 001653267 => Alma MMS ID: 990016532670402791



#### Problem discovered during testing

- ☐ The Digitool version needed to be the latest, so we upgraded our Digitool Prod. environment
- Access rights an object with access rights wasn't imported - we got a JSP file saying you need to login
- ☐ During testing Hebrew labels were displayed as gibberish in Alma viewer, this was fixed during migration period



#### First test process

☐ By the end of July 2018 we ran first sample of 50 records on Alma SB ☐ As a result, with the assistance of the Alma-D team we added both our Alma SB and Alma Prod IPs to all records in Digitool which had access right restrictions ☐ After that, the remote representation job was able to access Digitool and retrieve the digital object ☐ At this point it was decided to stop the daily work in Digitool



#### **Getting ready to start**

- ☐ In early September 2018 the majority of the problems were resolved, a full workflow was tested and a target date for the main migration was set to the end of October 2018
- ☐ This timing chosen due to 2 reasons:
  - The hardest part of the project will be over by the beginning of the academic year
  - The largest part of the migration will be done at the Autumn Holidays



#### Some numbers

- ☐ Our assessment was that about 80% of the digital objects in Digitool had corresponding metadata record in Alma and they should be handled in the main phase of the procedure.
- ☐ The remaining 20% were digital objects that existed only in Digitool, like special collections, METS and complex objects.



- ☐ Repository replication was configured in Digitool. First we created one replication of all Digitool objects but it took too long to import to Alma because of the file size, so we decided to cut the import file to smaller parts
- ☐ We did this according to file types: TIFs, JPEGs, PDFs and JP2s.



- We wrote a script that used OAI protocol in order to harvest the records' metadata from Digitool into MARC XML format
- ☐ The script creates file for every result page and contains approximately 50 records per file
- ☐ We used the **resumptionToken** field to advance to the next page until the end of the set



# Migrating Digitool 2 Alma The main phase

```
<subfield code="a">HTA</subfield>
         </datafield>
        -<datafield ind1=" " ind2=" " tag="992">
           <subfield code="a">DATI</subfield>
         </datafield>
        -<datafield ind1=" " ind2=" " tag="999">
           <subfield code="a">DIMAGE</subfield>
         </datafield>
        -<datafield ind1=" " ind2=" " tag="999">
           <subfield code="a">DOC</subfield>
         </datafield>
        -<datafield ind1=" " ind2=" " tag="999">
           <subfield code="a">DTEXT</subfield>
         </datafield>
        -<datafield ind1=" " ind2=" " tag="999">
           <subfield code="a">DMC</subfield>
         </datafield>
        -<datafield ind1=" " ind2=" " tag="999">
           <subfield code="a">NOULI</subfield>
         </datafield>
        -<datafield ind2="0" ind1="4" tag="856">
          -<subfield code="u">
             http://digitool.haifa.ac.il.1801/webclient/DeliveryManager?pid=163468&custom att 2=simple viewer
           </subfield>
         </datafield>
       </record>
     </metadata>
    </record>
   </ListRecords>
</OAI-PMH>
```



- ☐ After experiencing import of 50 records files we realized that it would be an extremely long process using a file for each 50 records
- ☐ So we added another part to the script that enable creating 5,000 records in each file, and used those files for the import

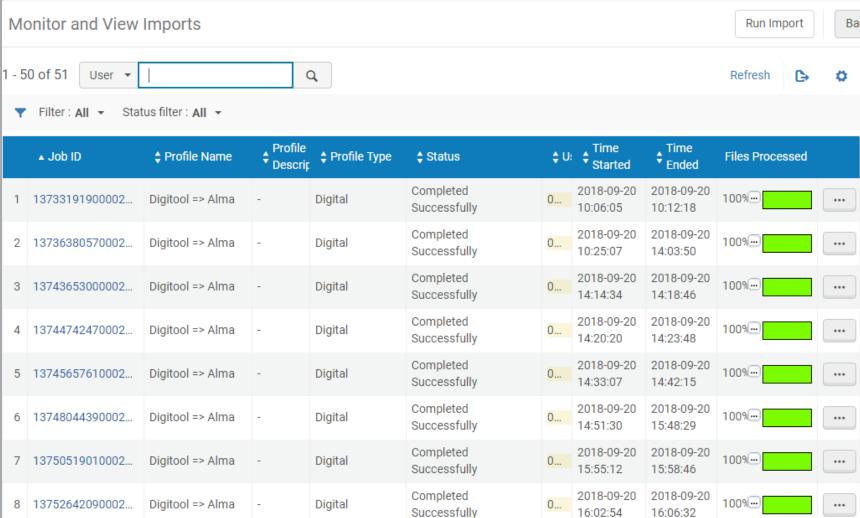


- ☐ In Alma the import profile was created using the following:
  - The match was based on 001 field
  - If there was no match => no records were imported
  - If there was a match => the Alma record metadata was kept, and a remote representation was added to the record



- □ During Succoth Holiday the actual migration was done; we imported ~110,000 records
- ☐ Importing the digital object as a remote representation (a representation linked to Digitool) using the import profile
- ☐ Copying the files to Alma using the Remote Representation Migration job
- ☐ Withdrawing the remote representation







#### **Problems during the import**

- Only one import job could run at a given time
- ☐ The remote representation migration job created some unusual load on the whole environment => This was fixed by the Alma-D team and only then we could resume the process
- ☐ Digital files larger than 40Mb couldn't be copied, creating numerous empty representations
- ☐ Labels that were too long (over 256 characters) caused the migration to fail.



#### What next?

- ☐ Given the situation we stopped all additional attempts until a fix was implemented by Ex Libris
- After few days, the migration job resumed and the jobs runs were completed with errors, the job reports were unclear to us, so we had many calls with the Alma-D team to figure out how to handle those errors
- ☐ It was agreed that the Alma-D team will help us import the failed records files.



#### **Ending main phase**

- ☐ At the end of October 2018 we had ~200,000 digital representation with files
- ☐ At this point we exported the collection tree from Digitool Prod and imported it to Alma Prod environment
- ☐ Testing and assessment of the migration to Alma took several weeks, including checking the objects manifestation in Primo

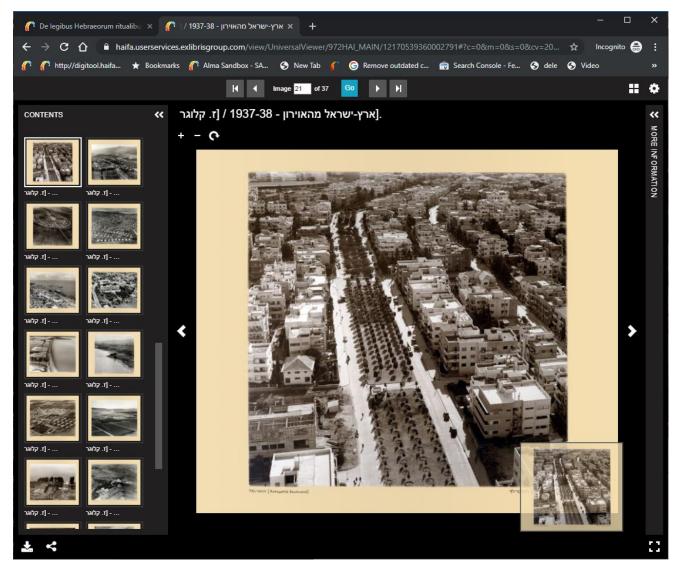


#### **Assessment and testing**

- ☐ Analyzing the amount of exported data from Digitool vs. the amount of data actually imported in Alma
- ☐ The following conclusions were drawn:
  - Numerous and recurrent issues with publishing and display of digital records and collections in Primo, especially using Universal Viewer (IIIF) => Fixed after June 2019 release
  - Changes made in the records or collections in Alma were not published regularly to Primo => Fixed after November 2018 release

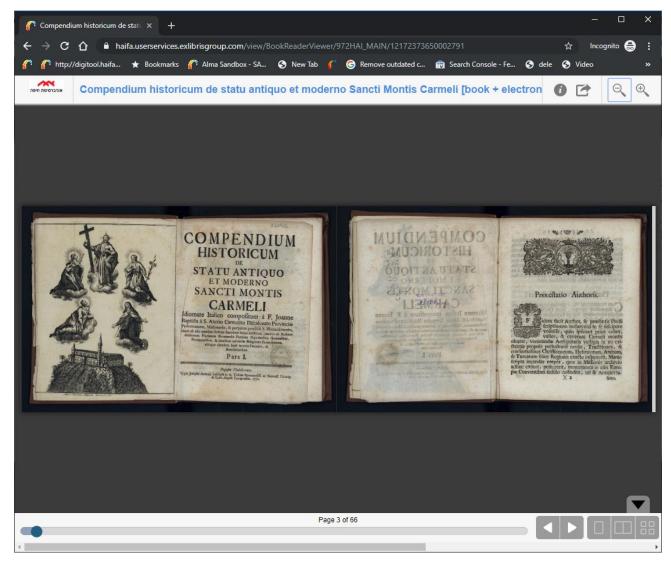


### **Universal Viewer (IIIF)**





#### **Internet Archive Book Reader**





#### **Assessment and testing**

- ☐ Records that were not migrated properly to Alma due to various reasons:
  - Missing 001 field
  - Records in Digitool that still had access rights
  - Too large files
  - Too long labels
- ☐ We had a lot of help from the Alma-D team dealing with all those issues



#### **Second phase**

- By December 2018 we were ready to face the second and more complicated phase of the project
- ☐ Importing METS and Complex objects
- ☐ Along with Ex Libris Alma-D team we ran an assessment check of the complex object inventory in Digitool
- ☐ The Alma-D team prepared a script that creates 2 directories with the complex records to upload, one for MARC format and the other for DC format



#### **Second phase**

☐ In January 2019 we run the new processes and jobs and after some more tuning we managed to successfully import the MARC complex objects



#### **Alma Viewer**



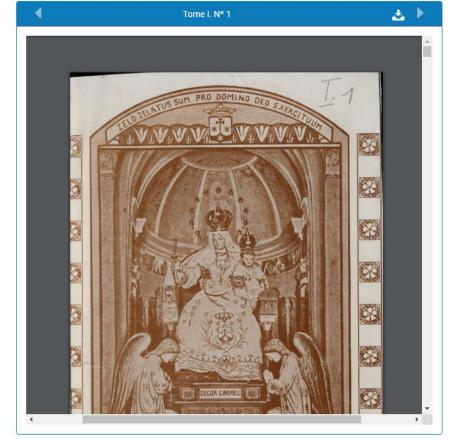
C Language -

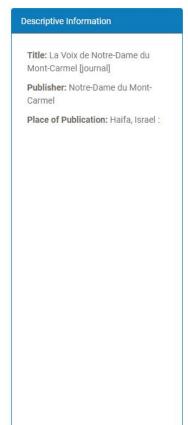


La Voix de Notre-Dame du Mont-Carmel [journal]

Share







Ex Libris, a ProQuest Company, 2019



#### **Second phase**

- ☐ After some debates, the Library decided to abandon the idea of importing DC records, which were Digitool original records with no match in Alma
- ☐ So instead of importing DC records, we converted them to MARC records to maintain the unified records format in our catalog



### **Second phase**

- □ To accomplish the conversion, we changed our approach and created the replication according to the collections
- ☐ We converted them to MARC format using XSL mapping which was adjusted to every collection individually, by adding to each record the URL to the digital objects
- ☐ Then, using the same script as before creating files with 5,000 records in it and ran the import process



#### The End

- ☐ The migration project was completed in May 2019
- We did a thorough analysis and comparison of Digitool repository vs. Alma with the help of Alma-D team
- ☐ The digital collections librarian approved that the amount of the migrated records and objects was as expected.



# The new Collection page



**NEW SEARCH** 

INTERLIBRARY LOAN

REMOTE ACCESS

BROWSE

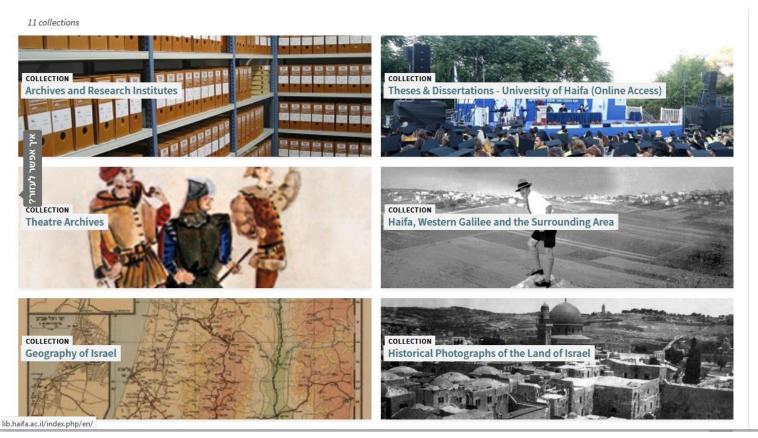


Sign in



#### **Digital Collections**

HAVE A COMMENT? PLEASE CONTACT US





### Special thanks to the Alma-D team

- ☐ Many thanks to the Alma-D team in Ex Libris for there help and collaboration with this huge and complicated project.
- ☐ The Ex Libris Alma-D team:
  - Nir Kashi
  - Opher Kutner
  - Yifat Tanami
  - Moti Perl



### **Questions?**



Source: https://www.inc.com/jim-haudan/20-questions-to-make-meaningful-connections.html



Younes & Soraya Nazarian Libran

