Annex 9
The Virtual collections Module

1. INTRODUCTION

The setup of DiSSCO defines new requirements for multimedia displays which need to be covered by the Natural Heritage project.

- DaRWIN can host multimedia files but is not designed for this.
- RBINS uses several Plone 4.3 sites since a decade (mars, collections, virtualcollections). The system is fast and versatile and can link to the DaRWIN illustrated specimens. Nevertheless, none of the Plone image viewers are compatible with IIIF wich is now required by the setup of DiSSCo.
- A website presenting media of the collections of the Royal Museum for Central Africa (RMCA) was developed some years ago (http://digit03.africamuseum.be/) but the process to upload pictures to the website was long and not straightforward.

As RMCA got a new general website, it was decided to create a new website for media, easier to maintain and technically up to date, to be readable on various platforms (pc, tablet, smartphone).
Objectives of this new site called "Virtual Collections" (V.C.) are to present in a first step biological specimens and some anthropological objects. In the future, V.C. should be used to present other collections of the museum and will also be used at Royal Belgian Institute of Natural Sciences (RBINS) as media server and a new tool to present virtual exhibitions.
V.C. will also be used as a media server for the main search portal of the museum' collections, Natural Heritage.

2. METHODOLOGY

In order to save development time, a search has been done to find tools already existing. During the museum' renovation, a tool called "CollectiveAccess" (C.A.) was used to organize the content of the museum. Tests were done to see whether it could be used to create a new public website and it has been chosen to develop V.C.
Based on CollectiveAccess, it consists of a backend allowing management of the MySQL tables, security, lists, various objects and of a frontend which is a public interface.
Several hundreds of pictures have been imported as case study, in Biology and Anthropology.

3. INFRASTRUCTURE

The new multimedia server joins already existing websites.
NaturalHeritage portal will be the main gate to explore data of Royal Museum for Central Africa, Royal Belgian Institute of Natural Sciences and Meise Botanical Garden.
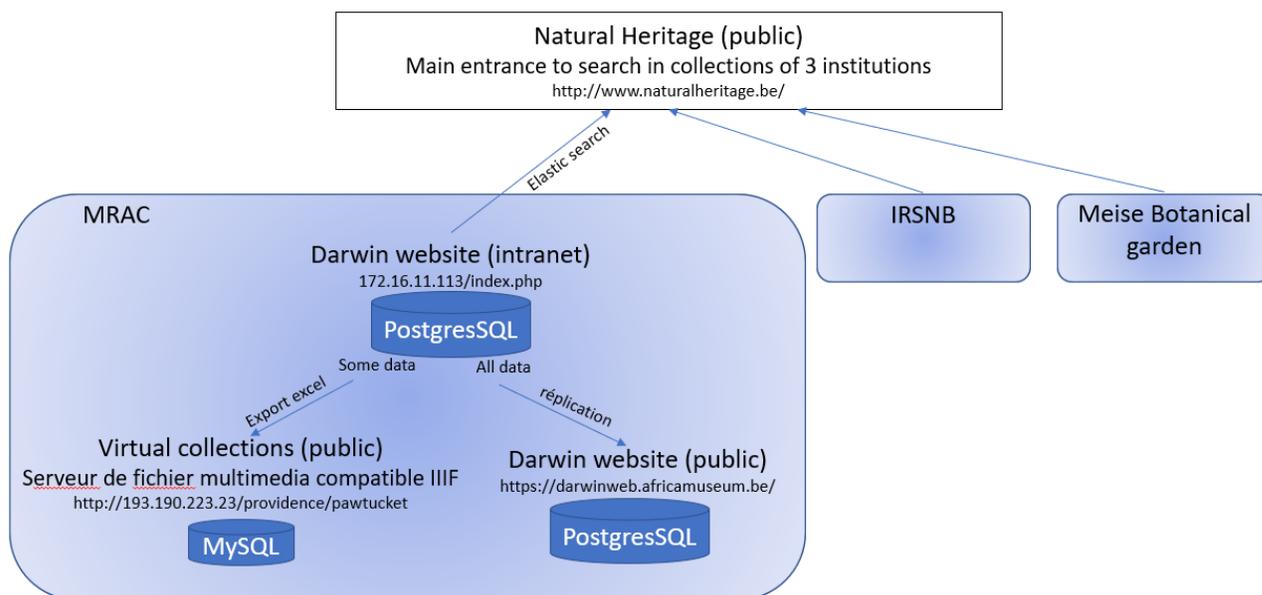
Figure 1. Infrastructure of the different websites under Natural Heritage

4. RESULTS AND RECOMMENDATIONS

Collective Access is a free open-source software for managing and publishing museum and archival collections (https://www.collectiveaccess.org/).
It consists of a MySQL database with 226 tables, of a backend web interface called "Providence" and a frontend called "Pawtucket". The majority of the tables contains basic, system fields and metadata are mainly in tables "attributes"
C.A. has integrated viewers for pictures of various formats, videos, sounds, 3D objects. Other types of files may be imported but can not directly be viewed, as text files, excel, ...
Biological data come from the RMCA database DaRWIN where an UUID is given to every record.
This allows a link between DaRWIN and Virtual Collections.

4.1 Providence
is used as management interface for various points:
- Management of the MySQL tables which are the basis of C.A.:
  creation, deletion, modification of the fields, of their types, and characteristics.
- Indices management
- Interface customization
- Creation of lists, simple or hierarchical.
- Management of the access rights by groups, roles.

- Creation of multiple objects allowing to organize data:
  objects, places, entities, lots, collections, exhibitions, storage locations,…
- Management of the multiple relations occurring between these objects.
- Mass editing with the help of sets allows to finely modify data in every field.
- Management of the various languages.
- Media and data import

The Providence interface of C.A. has not been deeply modified, regarding the original package. Only data and some fields have been added. There are 5 sections in Providence : New, Find, Manage, Import and History.

4.1.1 New

In this section, we can create a lot of different elements : lots, objects, entities, places, collections,… More info in section 2.3.2
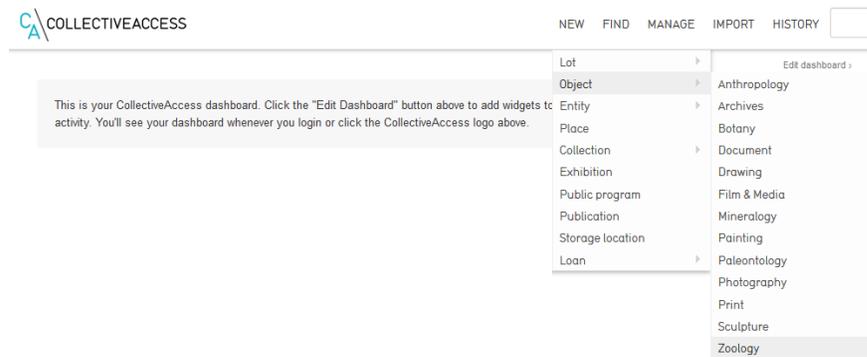
Figure 2 List of the New entries

4.1.2 Find

This section contains the same entries as New, to retrieve and edit elements. The main elements used for V.C. are objects, entities, collections, storage locations.
Each entry can be explored with a basic search, an advanced search or a browse
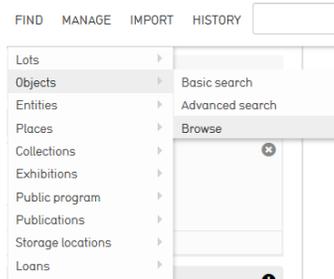
Figure 3 List of the Find entries

The Browse view contains some predefined facets to retrieve easily data.
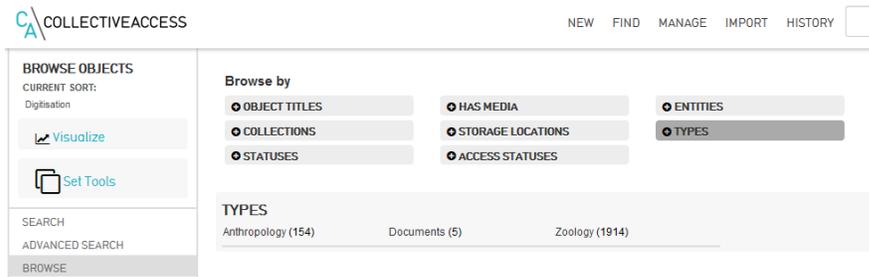
Figure 4 Submenu in Find objects by browsing

Figure 5 Result of a browse objects

A click on the edit icon of a line opens the detail form of the object. Form is divided in sections listed on the left.

Details of each section on the right give the possibility to enter metadata, translations, user access, codes, media, locations,…

Objects may be of different types and each type can get its specific fields. For Virtual Collections, 2 types have up to now been created, for zoology and anthropology.

Figure 6 Details of an object – Basic info

The media section allows users to enter various media :
- pictures,
- videos,
- sounds.

Picture metadata are directly read from the file and access may be given for each media, as well as order of presentation and primary picture.

Figure 7 Details of an object – Media

In Relationships, links are done with other kinds of objects : entities, locations, storage,…

Figure 8 Details of an object – Relationships

A section called Biology has been added to gather all metadata specific to biology and especially taxonomy.

Figure 9 Details of an object – Zoology

4.1.3 Management

The management menu contains a lot of important forms to manage directly from Providence interface various system operations.

It also allows users to manage user lists and preferences.
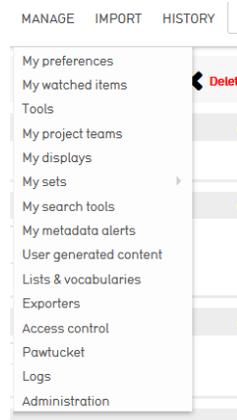
Figure 10 List of Manage entries

The My sets entry allows me to manage lists of objects created from the results of a search. These sets may be used to modify all the records in the set, in 1 operation. All the fields defining the record may be used to update the records in a mass action. Records may also be deleted and the type of document may be changed.

Figure 11 Manage My sets : list of entries

For each field, values can be modified, removed, added.

Figure 12 Set management : list of fields that can be used to modify all records in the set

In Lists & vocabularies, simple or hierarchical lists may be defined. For example, different levels of collections can be created. These levels are then used in the creation of a collection, in the relationships between collection levels.
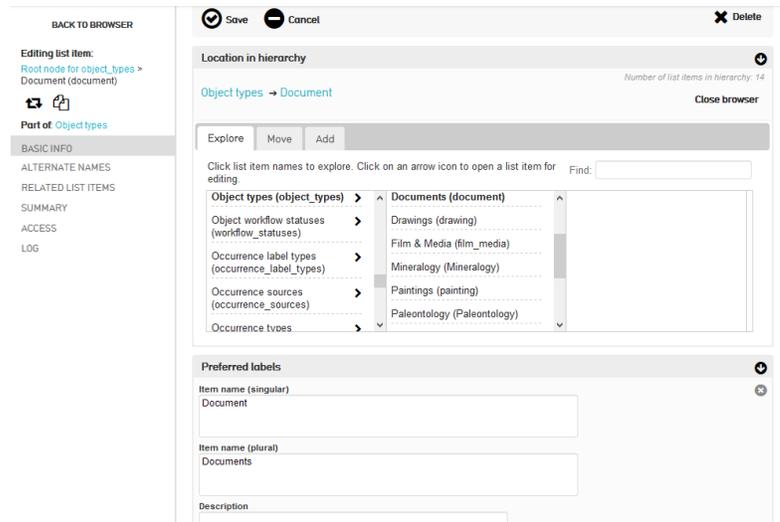
Figure 13 Lists management : hierarchy and parameters

In the Administration entry of Manage, basic elements of the databases and of the interface may be modified..

The metadata elements are all the fields used in the different elements of Providence (Objects, collections,…) Each field can be edited to change the type of field, size, access, translations, where it can be used, in which type of element, if it must be unique or not,…

Figure 14 List of metadata elements

Figure 15 Edit form of metadata

In Administration – User interfaces, fields can be added and organized in subsections in the different forms of Providence. For example in Fig.16, subsections Basic info, Administrative info, Media in form Standard object editor.
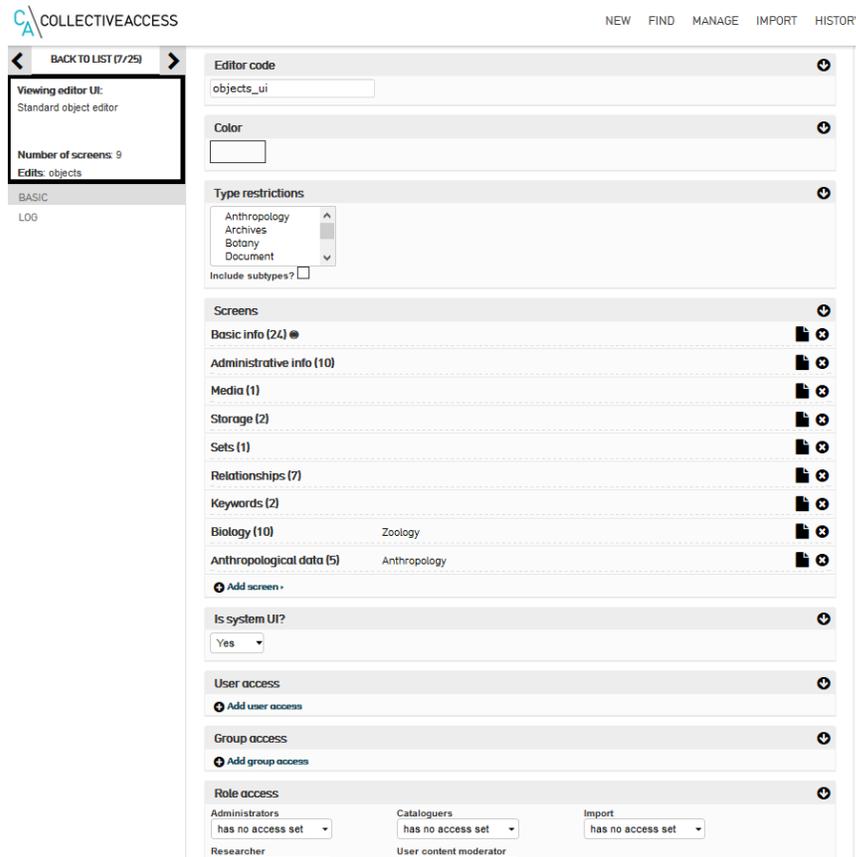
Figure 16 User interface management
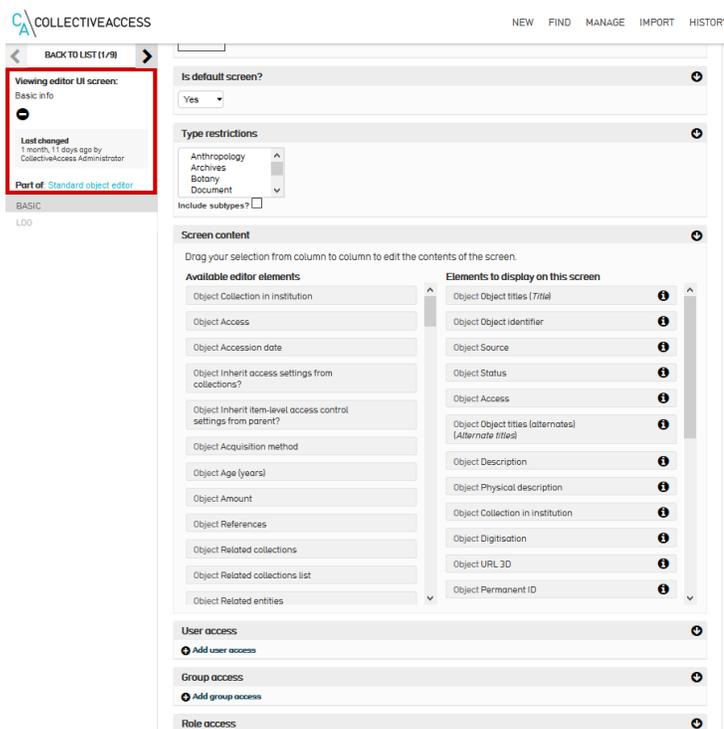
Figure 17 Elements in user interface of object editor

4.1.4 Import

Import of data can be done directly through the forms or via mass import of the Import entry in the menu.

For this project, Import data and Import media have been used.

Import data can be used in nearly all cases. Import media must be used only to import special files as text or excel files.

There are 4 main steps to import data:

- ● Prepare 2 excel files : one for the mapping and the other with data. First one must be done only one time for a type of object. Several mapping files can be created for each corresponding type of object.

Figure 18 Mapping Excel file

Figure 19 Example of excel with data

- Send media files to the server by FTP, in a directory called Import.
- Add the mapping file to the import list

Figure 20 Screen for template choice of data import

- Import data by running import Data

Figure 21 Run import screen

To import media, a mapping file is not necessary : choose Import – Media. A form will open presenting a list of files present in directory Import and you have to finetune the parameters to your needs.

Figure 22 Import medias form

To import data from DaRWIN, more details about the preparation of the excel being imported in CA are given in paragraph 4.5.2.

4.2 Pawtucket

Pawtucket use the media imported via Providence to present them in various ways:

1. Browse facets
2. Advanced search, through customized form
3. Navigation in collections, galleries.
4. Possibility to register to save searches
5. Contact form

4.2.1 Interface Pawtucket

Virtual Collections has been developed to present data easily, clearly and quickly.
CollectiveAccess comes with a default theme. This theme has been copied to a new one and adapted for V.C..
There are 4 main pages and a browse section. Main pages are : Welcome page, Advanced search, Collections and Contact, with a common header and footer.

4.2.1.1 Header

Header contains the menu, a multi search field, an icon to register/login and 3 flags to change language (English, French, Dutch)
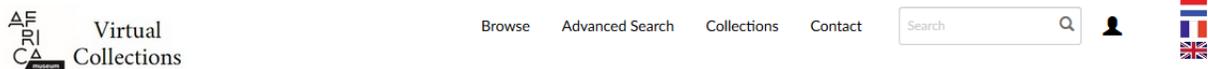
Figure 23 Header of Virtual Collections pages

4.2.1.2 Footer

Footer contains a copyright notice and links to social networks.

Figure 24 Footer of Virtual Collections pages

4.2.2 Welcome page

Welcome page contains 3 sections : a welcome word, news with links for more details and a banner presenting some specimens at random. A click on the specimens opens the detail form.

Welcome to Virtual collections!

You can browse through our digital collections and find detailed pictures, 3D models, sounds and videos of our collection items.

LATEST NEWS!

23/07/2020

How to digitize natural science collections?

Researchers from the Royal Belgian Institute of Natural Sciences and the Royal Museum for Central Africa have created a manual for the digitization of natural science collections. And it can be done literally in fifty different ways, with fifty more nuances. A guide.

By digitizing museum pieces in 2D and / or 3D and publishing them online, they are made accessible to researchers around the world. Thanks to the high-resolution images, specimens no longer have to leave the museum's deposit with its ideal conservation conditions. In addition, they are not likely to be damaged by handling and transport.

More...

Figure 25. Welcome screen

4.2.3 Advanced search

Advanced search form allows to find objects according to their main metadata. Displayed fields depend on the chosen collection. If no collection is chosen, all fields are displayed.
An explanation to use the form is given on the right.
Types of media can also be chosen and a check in the Boxes field limits search to boxes.

Figure 26 Advanced search screen

Results of a search are presented in various displays:
- thumbnails,
- lists,
- tables,
- maps,
- timeline.

The choice of the display is done by clicking on the icons on the right, above the facets.

Content of each element can be adapted, as well as content of the tables.

On the right, results can be filtered by facets.

Figure 27 Results of a search in view "Thumbnails"

Figure 28 Results of a search in view "List"

85 RESULTS ⚙

SEARCH: Type: Zoology; Keywords: arachn*

ID	Title	Collection	Phylum	Class
RMCA_ARA_54259	Titanatemnus ugandanus - RMCA_ARA_54259	Arthropoda	Arthropoda	Arachn
RMCA_ARA_209260	Thoriosa spadicea - RMCA_ARA_209260	Arthropoda	Arthropoda	Arachn
RMCA_ARA_239916	Thomisus unidentatus - RMCA_ARA_239916	Arthropoda	Arthropoda	Arachn
RMCA_ARA_239914	Thomisus unidentatus - RMCA_ARA_239914	Arthropoda	Arthropoda	Arachn
RMCA_ARA_215886	Thaumastochilus termitomimus - RMCA_ARA_215886	Arthropoda	Arthropoda	Arachn
RMCA_ARA_174690	Thaumastochilus termitomimus - RMCA_ARA_174690	Arthropoda	Arthropoda	Arachn
RMCA_ARA_171953	Thaumastochilus martini - RMCA_ARA_171953	Arthropoda	Arthropoda	Arachn
RMCA_ARA_143319	Tetragnatha infuscata - RMCA_ARA_143319	Arthropoda	Arthropoda	Arachn
RMCA_ARA_245726M	Tarne dives - RMCA_ARA_245726M	Arthropoda	Arthropoda	Arachn
RMCA_ARA_245726F	Tarne dives - RMCA_ARA_245726F	Arthropoda	Arthropoda	Arachn
RMCA_ARA_245858	Synema viridisternum - RMCA_ARA_245858	Arthropoda	Arthropoda	Arachn
RMCA_ARA_114787	Solpuga upembana - RMCA_ARA_114787	Arthropoda	Arthropoda	Arachn
RMCA_ARA_245365	Selenops radiatus - RMCA_ARA_245365	Arthropoda	Arthropoda	Arachn
RMCA_ARA_233843	- RMCA_ARA_233843	Arthropoda	Arthropoda	Arachn
RMCA_ARA_134817	Ricinoides karschi - RMCA_ARA_ 134817	Arthropoda	Arthropoda	Arachn
RMCA_ARA_133729M	Psammoduon deserticola - RMCA_ARA_133729M	Arthropoda	Arthropoda	Arachn
RMCA_ARA_133729F	Psammoduon deserticola - RMCA_ARA_133729F	Arthropoda	Arthropoda	Arachn
BE_RMCA_INV_ARA.Ara.228612	Plexippus_minor_BE_RMCA_INV_ARA.Ara.228612	Arthropoda	Arthropoda	Arachn
BE_RMCA_INV_ARA.Ara.228611	Plexippus_minor_BE_RMCA_INV_ARA.Ara.228611	Arthropoda	Arthropoda	Arachn
BE_RMCA_INV_ARA.Ara.228610	Plexippus_minor_BE_RMCA_INV_ARA.Ara.228610	Arthropoda	Arthropoda	Arachn

FILTER BY

THEME
  Zoology (85)

PERSON
  Jonathan Brecko (4)
  J. Fourmy (1)
  Arnaud Henrard (80)
  Aurore Mathys (1)
  A. Van Harten (3)
  Lourenço W.R. (1)
  Wesolowska & Van Harten (3)

MEDIA
  2D (85)

COLLECTION
  Arachnids (85)

TYPE
  Holotype (15)
  Paratype (16)

PHYLUM
  Arthropoda (85)

CLASS
  Arachnida (85)

ORDER
  Araneae (69)
  Ixodida (1)
  Opiliones (1)
  Prostigmata (3)

Figure 29 Results of a search in view "Table"

85 RESULTS ⚙

SEARCH: Type: Zoology; Keywords: arachn*
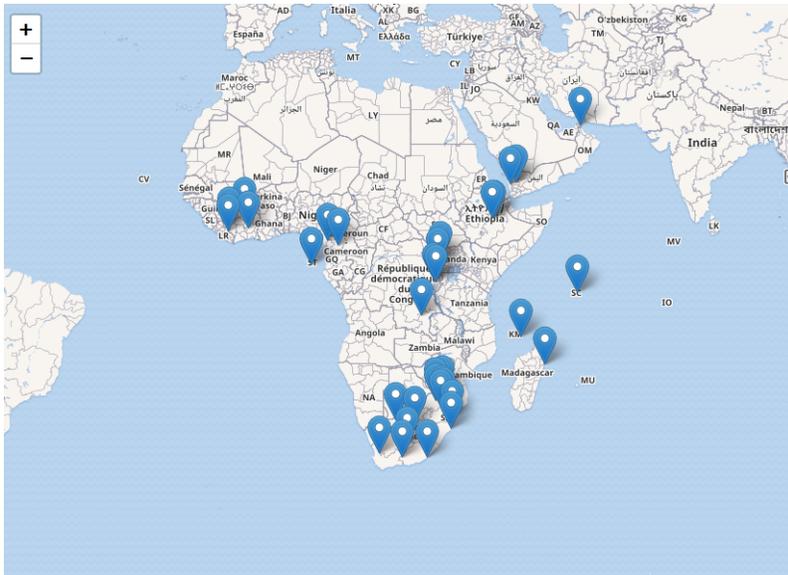
FILTER BY

THEME
  Zoology (85)

PERSON
  Jonathan Brecko (4)
  J. Fourmy (1)
  Arnaud Henrard (80)
  Aurore Mathys (1)
  A. Van Harten (3)
  Lourenço W.R. (1)
  Wesolowska & Van Harten (3)

MEDIA
  2D (85)

COLLECTION
  Arachnids (85)

TYPE
  Holotype (15)
  Paratype (16)

PHYLUM
  Arthropoda (85)

CLASS
  Arachnida (85)

ORDER

Figure 30 Results of a search in view "Map"

Figure 31 Results of a search in view "Timeline"

Presentation of the results can be adapted by the reader by clicking on the cogwheel on the right of the number of results.

In the menu that appears, you will see that results can be saved in a "lightbox" (if user is registered), sorted by title or identifier, or exported to a pdf or an excel.
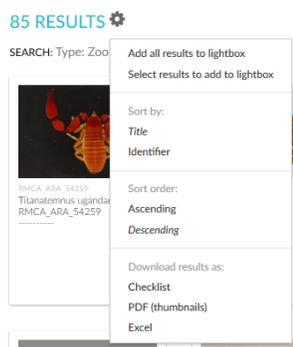
Figure 32 Menu in search results

Figure 33 Excel file generated from search results

4.2.4 Collections

A second way to find media after advanced search is "Collections".
A first screen presents the main themes according to which media are organized in V.C.
A small presentation of the collection is given when the mouse passes over the photos.
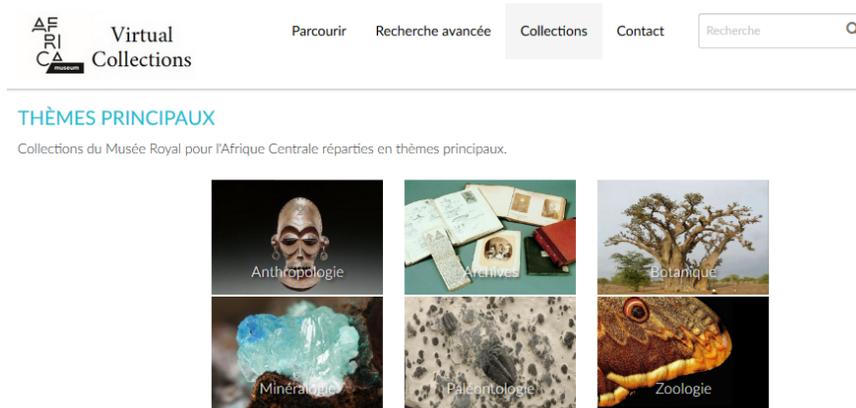
Figure 34 Screen Collections

When we click on a theme, a hierarchical view of subcollections is given and a click on a subcollection goes further in the tree or opens the collection.
A pdf can be generated with a listing of all specimens of that level of collection.

Figure 35 Hierarchical view of the collections

When a collection is opened, you see the results in thumbnails. Path with upper collection levels is given and the pdf download is still present.

Figure 36 Content of a collection

Figure 37 Export listing of a collection

4.2.5 Contact

Contact page presents the institution (e.g. Royal Museum for Central Africa), a notice to know how to get high resolution pictures, contact information and maps.
A form allows also to send an email to the contact person defined in the parameters.

CONTACT

ABOUT THE ROYAL MUSEUM FOR CENTRAL AFRICA

The Royal Museum for Central Africa must aspire to be a world centre of research and knowledge dissemination on past and present societies and natural environments of Africa, and in particular Central Africa, to foster – among the public at large and the scientific community – understanding of and interest in this area and, through partnerships, to contribute substantially to its sustainable development. Thus the core endeavours of this Africa-oriented institution consist of acquiring and managing collections, conducting scientific research, implementing the results of this research, disseminating knowledge, and mounting selected exhibitions of its collections.

The AfricaMuseum is a centre for knowledge and resources on Africa, in particular Central Africa, in an historical, contemporary, and global context. The museum exhibits unique collections. It is a place of memory on the colonial past and strives to be a dynamic platform for exchanges and dialogues between cultures and generations.

REQUEST FOR MEDIAS

For any request for a high resolution photo or other media, please use the contact form below or contact the curator concerned directly (Collections).

Address:
Leuvensesteenweg 13
3080 Tervuren
Belgium

Phone:
+32 2 769 52 11
Email:
info@africamuseum.be

Belgian company registration number:
BE 0306.562.560

- A - Welcome Pavilion
- B - Museum Building
- C - Director's Pavilion
- D - Stanley Pavilion
- E - CAPA Building
- F - Finances Building
- G - Africa Palace (former Colonial Palace)
- H - Villa

NAME

EMAIL ADDRESS

SECURITY QUESTION

4 + 4 =

MESSAGE

SEND

Figure 38 Screen Contact

4.2.6 Browse

A third way to search media is Browse. Browse appears as an overlay above the opened screen. There are currently 3 tabs : all themes, zoology and anthropology. All themes show all the facets while the others show only the facets linked to the theme. With this interface, you can quickly reach the media you want without doing a search. You can also browse all media by clicking on the "Browse all".

Figure 39 Browse overlay screen

4.2.7 Details form

The details form presents all media linked to the specimen/object, with some info about it.

Media are shown below the left column with a thumbnail or an icon (for sound). When you click on one of these, it is displayed in a bigger format on top of the left column.

You can again click on that bigger picture to open it as full screen and zoom into it. Close that full screen by clicking on the cross at bottom right.

If there are other files than pictures, sounds, videos, they are displayed under "Related objects". A click on the link opens a window where you can download the file.

On the right are displayed information about the specimen/object : code, link to more details in the original database (click on 🔗), related people and other info such as taxonomy, storage, amount, sampling.

Comments can be added if the user is logged in.

Figure 40 Form with details of an object

Figure 41 Zoom on a picture

4.3 Security

Several levels of security can be defined in Providence, with an impact on data displayed in Pawtucket. This protects data by creating groups and roles that are allowed to work on and see specific data.

Users, groups, roles are created and managed in Providence, under the menu Manage – Access control.

4.3.1 Users

First entry in the left menu of Access control is User logins. You have there a list of users, with info about email, last login and if it is active or not

Figure 42 User management : List of users

When you create or edit a user, you get this form where you can enter all info about the user and his permissions

Figure 43 Management form for a user

4.3.2 Groups

Second entry is User groups. Groups are only there to gather users and give them common roles.

Figure 44 Management form of user groups

4.3.3 Roles

Access roles form contains tens of entries allowing to define very accurately the rights of users on objects and actions. These rights can be defined at different levels: actions, metadata, types of objects, sources, and users in Pawtucket can get different rights than in Providence.

Figure 45 Management form of access roles

These roles can be added to each element of Providence, objects, collections, fields,...

Figure 46 Rights management on an object

4.3.4 Registration

In Pawtucket, visitors can register to create a permanent account. By default a role "public user" gives him only the right to write comments and save searches in the lightbox but the administrator can change that role to give him for example a role which will allow him to see hidden objects or files.
To register, visitors have to click on icon of a person at left of the flags in the main menu.
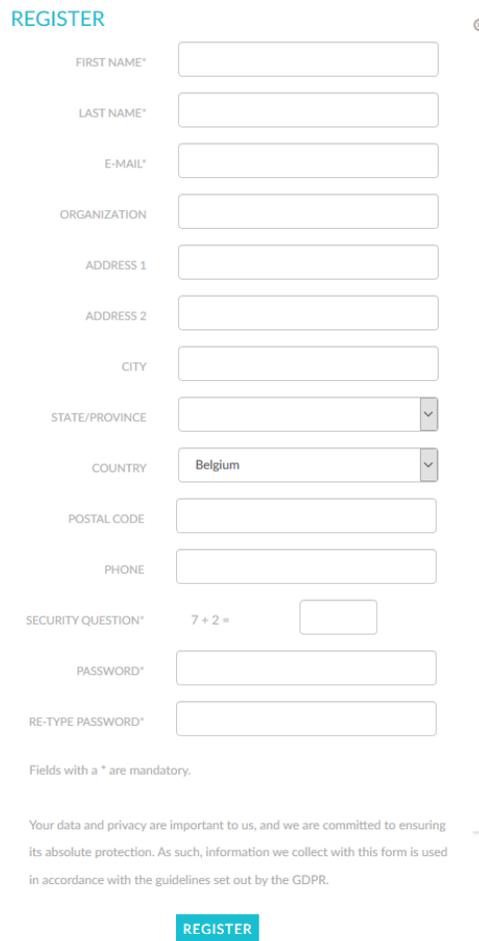
Figure 47 Popup window for registration

4.4 The IIIF Image viewer

Pictures can be saved on CA in different formats and resolution high resolution pictures can be displayed C.A.. This can be done because CA is IIIF compatible. IIIF is an international standard to share high resolution pictures (https://iiif.io/). Pictures are tiled to avoid time consuming download and also to protect them from unwanted download at high resolution. CA uses 2 of the main important viewers to display them : UniversalViewer and Mirador. UniversalViewer has been developed for the British Library while Mirador has been developed by Stanford University and is maintained by many contributors, with among them Harvard University.

Photos stored on the IIIF server can be accessed through a manifest request, a json document containing metadata about the characteristics of the image and other information. This manifest was partially present in CollectiveAccess (integrated in an html page) and a development had to be done to complete it. This development consists of a PHP / cURL script correcting the syntax problems of the original manifest, and functioning as a proxy. The script generates a correct manifest, having the MIME type Json. This development made it possible to use Virtual Collections as a fully functional IIIF server, allowing in particular to display photos in DaRWIN. On the DaRWIN server, the Mirador viewer was chosen because the DaRWIN interface can be seen through an iframe and UniversalViewer already working with iframe, this caused viewing problems. Mirador did not present these problems. Mirador has also a more important community of developers. A test was also carried out to create an annotation server for the photos.

4.5 Establishing the links between CMS DaRWIN / Plone and CA

4.5.1 Collective Access as a stand alone picture server

Development has been done in C.A. to complete the compatibility by adding a IIIF service to retrieve pictures of an object with its UUID from a foreign website.

In RMCA,Collective Access is used with the front end part, Pawtucket but Providence can be installed alone to be a picture server for other applications.

RBINS will use CA only to store and display images with IIIF viewer but it will not be directly accessible by the public with an independant search interface.

4.5.2 DaRWIN

In CA, the link with DaRWIN is established using the CETAF stable identifier used by DaRWIN for each specimen or part of specimen. The UUID must be entered as metadata to enable the link with DaRWIN. If it's there, a link can be found on the detail page in the form of a chain icon. By clicking on it, public DaRWIN opens:
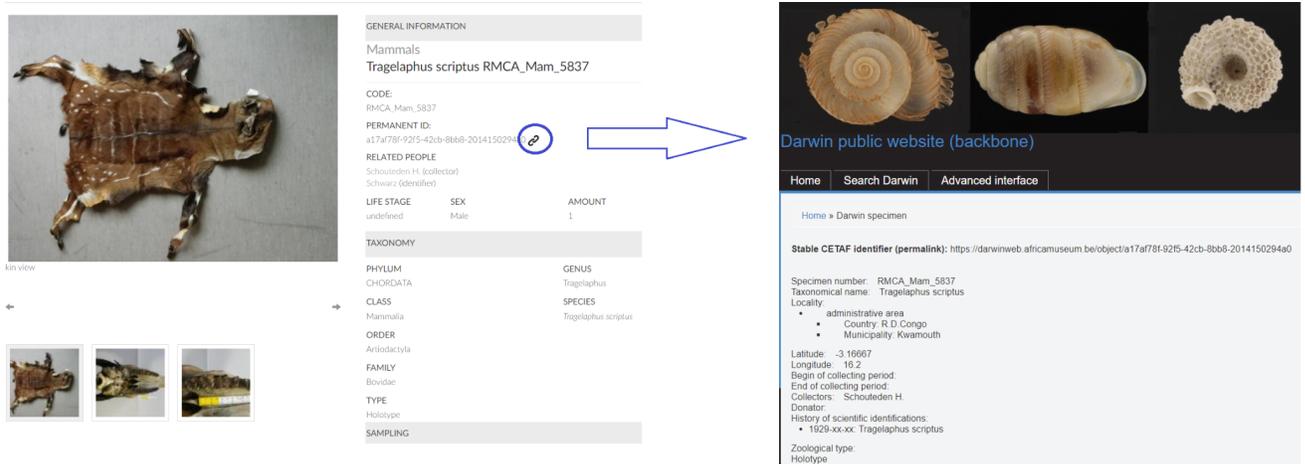
Figure 48 DaRWIN External link edition widget

In the opposite direction, DaRWIN also has links toward CA.

It is now possible to define external links in DaRWIN redirecting to specific multimedia servers. The Picture (IIIF) option can link to the Collective Access image url and displays it directly in DaRWIN. All zooming options are activated in the DaRWIN widget. A link allows you to open the image in a separate window.
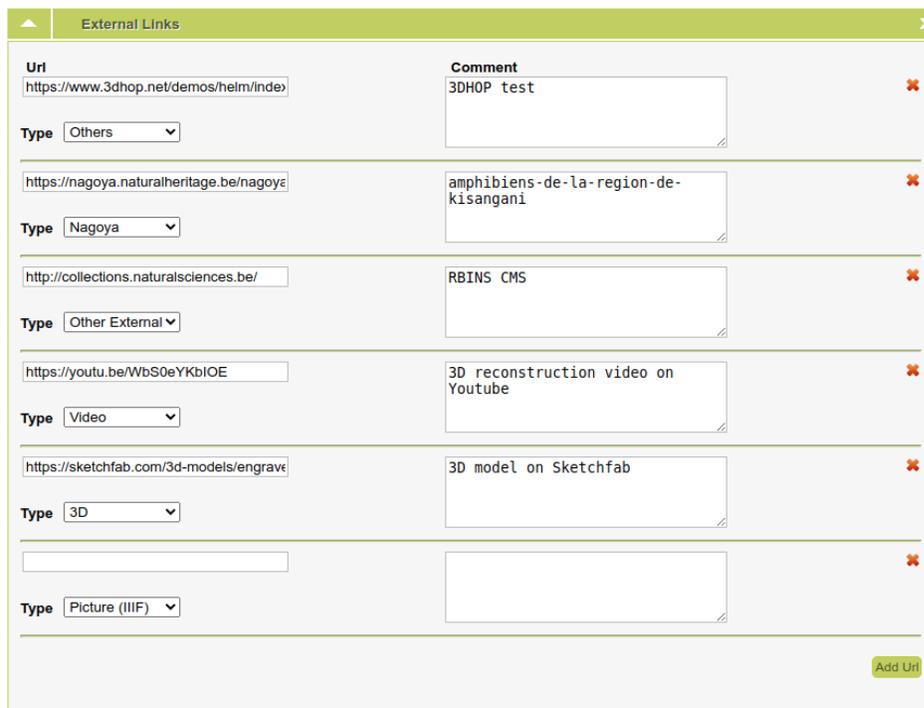
Figure 49 DaRWIN External link edition widget

Figure 50 Embedded view of the CA image in DaRWIN widget

From DaRWIN, data can also be exported, to populate CA or to update it.

A special report has been created to easily export the specimens that have to be imported in CA. This report is similar to the other reports and generates a CSV file. Choose "Tab-delimited (Virtual Collections)" to get that report:
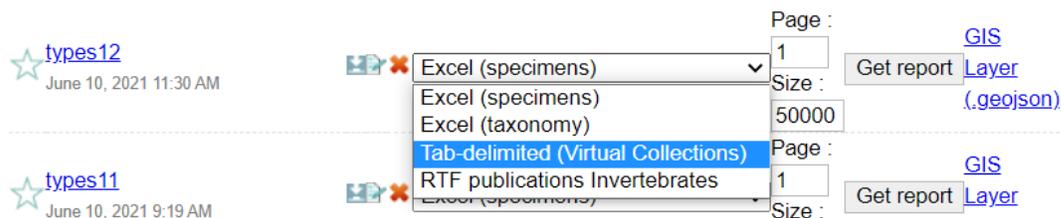
Figure 51 Report choice in DaRWIN

This CSV file can then be imported in an excel, to add media. This excel is simple and automatizes the filling of media filenames and infos for each specimen. It contains 3 tabs: a first one ("ImportDarwin") with only buttons, a second one ("temp") empty and a third one("Generate photo filename") to help to compose complex filenames:
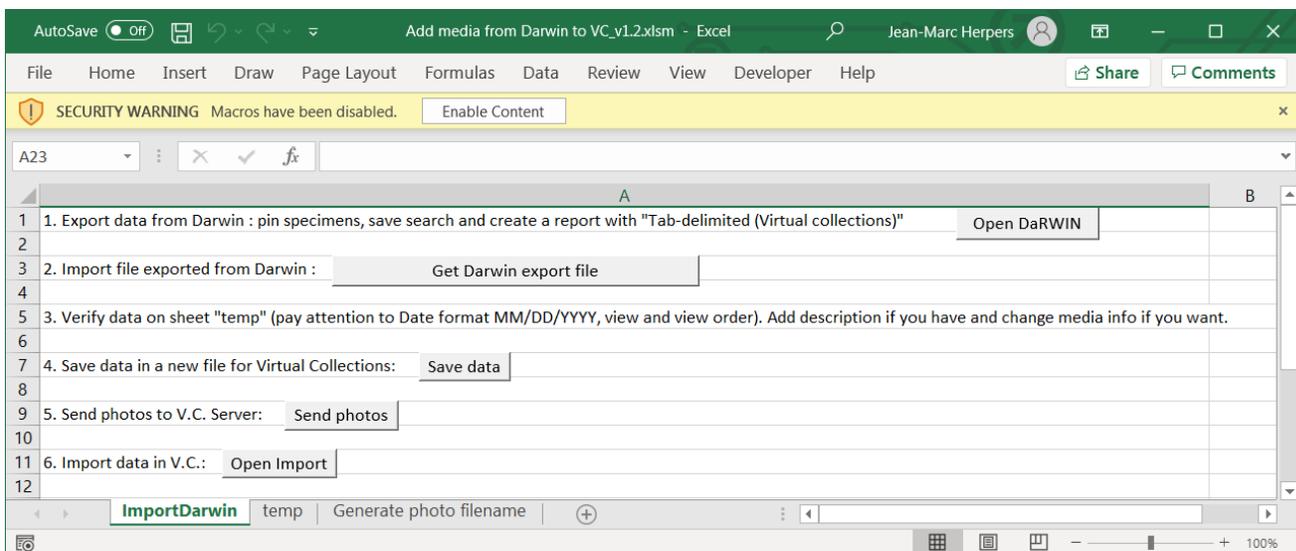
Figure 52 Excel template to import data from DaRWIN to Collective Access

The first tab contains actions to try to minimize the work of the user:

- 1. Open DaRWIN : to avoid searching for the link to DaRWIN, this button opens DaRWIN where the user can pin the specimens he wants, create the list and export the report as a CSV.

- 2. Get Darwin export file : Opens a window to get the CSV file and then opens a second window to enter the root path to the photos and the name of the collection in CA :

Figure 53 Screen to enter path to photos and collection name.

The photos files must be in directories of which name must be exactly the code in DaRWIN. This code is indeed used by CA as the main code.

When OK is clicked, data from CSV are inserted in the excel, on tab "temp", and photo filenames are automatically also filled, with one filename by row. So, we get the same format as the template described in 4.1.4.

This is done without intervention of the user to fill any data, and therefore speeds the process and also avoids errors that can occur by typing again filenames in the excel.

No need also to copy manually data from DaRWIN: all is inserted automatically.

- 3. This step is there to verify data of tab "temp"
- 4. Save data : Tab "temp" is saved as a new excel file, ready to be imported in CA.
- 5. Send photos : opens filezilla to send the photos to the CA server. To avoid losing time, connection is automatically done. User has to drag and drop files to the server.
- 6. Open Import. This step is described in 4.1.4 to import the excel.

This excel template greatly speeds up the process of creation of the import file to CA, by minimizing the risk of errors.

The last tab is there to help to create the filenames of the photos, which can be complex. It's used independently of the rest of the excel template. It contains formulas to automatically create the name. If tab "temp" is already filled in, this tool can be used to generate the photo description by copying automatically the data with button "copy data from sheet 'temp'":

Figure 54 Tool to generate filename and description of the photos.

Since the installation of the new media server at RMCA, 4102 specimens/objects (3946 in zoology and 156 in anthropology) and 9561 photos have been imported. In the last months of the project were imported:

- 144 mammals,
- 26 amphibians and reptiles,
- 25 birds,
- 1895 boxes of Lepidoptera.

4.5.3 Plone

Plone can display the CA images in an iframe with zooming and annotations.

Testing demonstrated that the default IIIF viewer Universal Viewer of CA doesn't support the Iframe because it uses it already. The Mirador IIIF viewer is working fine both in Iframe and direct link. Mirador was therefore chosen for next development.

5. CONCLUSION

During the test of Collective Access, more than 4000 objects and 9000 pictures from the RMCA zoology and anthropology collections were imported with success.

The evaluation of the CA solution is very positive:

- The new public website Virtual collections presents media from Africa Museum in a clear and easy design.
- Various gates exist to explore the data, save them or export them.
- Registration allows us to protect some more sensible data from a general public view by restricting access to some registered people.

- The backend interface is an easy-to-use solution to manage the MySQL database, import data and media, edit metadata and interfaces.

The use of Collective Access will be extended in 2021-2022 to the RBINS Virtualcollections allowing a better common policy and a common management of the infrastructures.

Authors : Jean-Marc Herpers, Franck Theeten & Patrick Semal