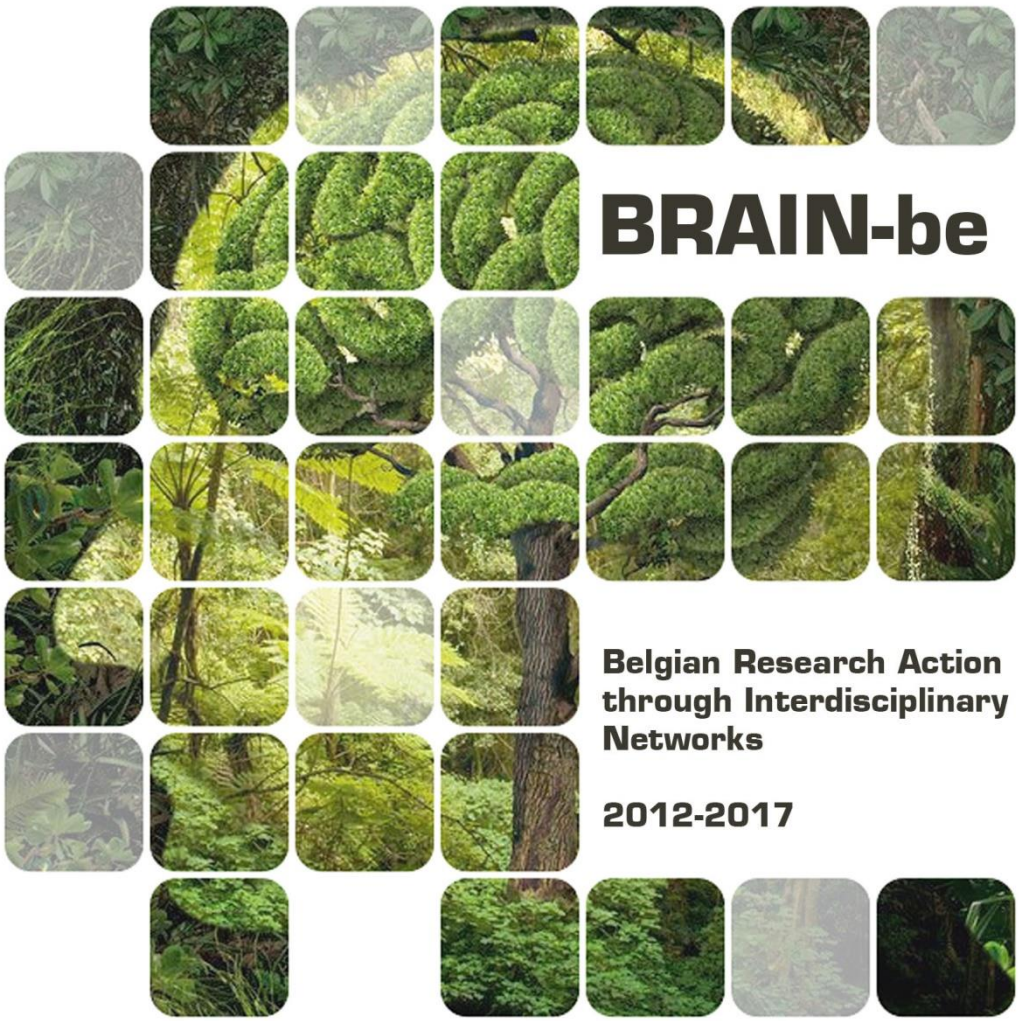


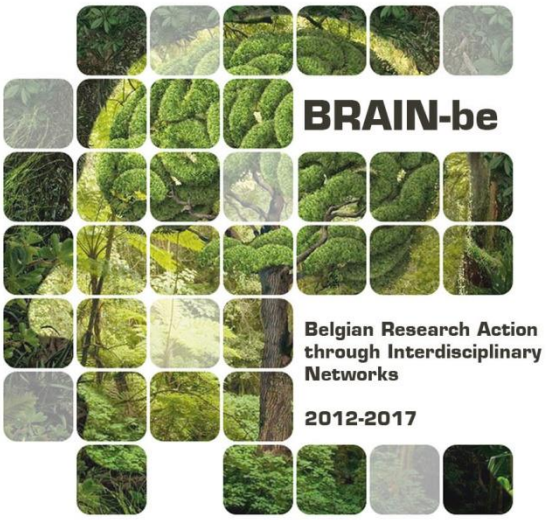
NATURALHERITAGE

NaturalHeritage.be: Modular interoperable database system and Portal for Belgian Natural History Collections

Marielle Adam (RBINS), Ann Bogaerts (APM), Mathias Dillen (APM), Henry Engledow (APM), Jean-Marc Herpers (RMCA/RBINS), Patricia Mergen (APM), Patrick Semal (RBINS), Larissa Smirnova (RMCA/APM), Franck Theeten (RMCA/RBINS), Thomas Vandenberghe (RBINS), Didier Van den Spiegel (RMCA)

Axis 3: Cultural, historical and scientific heritage

belspo .be



NETWORK PROJECT

NATURALHERITAGE

contract BR/175/A3

FINAL REPORT

PROMOTORS: Patrick Semal (RBINS),
Didier Van den Spiegel (RMCA),
Henry Engledow (APM)

AUTHORS: Marielle Adam (RBINS), Ann Bogaerts (APM), Mathias Dillen (APM), Henry Engledow (APM), Jean-Marc Herpers (RMCA/RBINS), Patricia Mergen (APM), Patrick Semal (RBINS), Larissa Smirnova (RMCA/APM), Franck Theeten (RMCA/RBINS),Thomas Vandenberghe (RBINS), Didier Van den Spiegel (RMCA)

museum AFRICA museum Plantentuin Meise belspo .be

Published in 2021 by the Belgian Science Policy Office
WTCIII
Simon Bolivarlaan 30 Boulevard Simon Bolivar
B-1000 Brussels
Belgium
Tel: +32 (0)2 238 34 11
http://www.belspo.be
http://www.belspo.be/BRAIN-be

Contact person: Maaike Vancauwenberghe
Tel: +32 (0)2 238 36 78

Adam M., Bogaerts A., Dillen M., Engledow H., Herpers J.-M., Mergen P., Semal P., Smirnova L., Theeten F., Vandenberghe T. & Van Den Spiegel D. ***NaturalHeritage.be: Modular interoperable database system and Portal for Belgian Natural History Collections***. Final Report. Brussels : Belgian Science Policy Office 2021 – 32 p. (BRAIN-be (Belgian Research Action through Interdisciplinary Networks))

TABLE OF CONTENTS

1. ABSTRACT

European natural science collections contain the largest and most significant part of the world's scientific knowledge of the earth's structure, environment and biosphere. The Belgian collections cover all the fields of Natural History including Zoology, Botany, Geology, Palaeontology and Anthropology. Belgian scientific institutions house around 55 million specimens and form an important collection of Natural Sciences worldwide.

The NaturalHeritage project investigated and evaluated how a modular architecture can address the diversity of data and collection management systems offering a common research tool. This architecture is based on two levels of interoperability between small modules and international authority files. Based on this approach, the project developed a new database ecosystem and a search portal so that there is interoperability amongst the 3 collection partners. The NaturalHeritage.be portal is also interoperable with other reference sources and similar on-line collection systems. This will help scientists and decision makers to address Belgian Natural History Collections issues using cross-linked and big-data approaches, as well as the design of the Belgium data hub for the new DiSSCoResearch Infrastructure (ESFRI Roadmap 2020) .

The project extended the DaRWIN collection management system to new categories of objects and associated data/metadata. It developed also a new common research portal (naturalheritage.be) allowing users to address 3 layers of queries :

- Easy Google-like search,
- WHO, WHEN, WHERE, HOW easy "human" questions
- etailed faceted search.

2. INTRODUCTION

European natural science collections constitute a key international research infrastructure for tackling major socio-economic and scientific challenges.

The 3 collection partners of this project are members of CETAF, the Consortium of European Taxonomic Facilities, which contributes to Europe's knowledge-base by enhancing synergy between the collections and research capabilities of its members. CETAF facilitates the normalization of data and metadata in Natural Sciences databases as well as the implementation of unique identifiers for collection specimens, which is essential for data aggregation.

They are also involved in the setup of DiSSCo (dissco.eu), the new ESFRI related to the Natural History collections aiming to ensure open access to European natural history collections thereby broadening the user community and finding innovative solutions through the use of natural science related data.

More than a decade ago, RBINS developed a bespoke Collection Management System, called DaRWIN. The current version, still used by RBINS and now also by RMCA, is an Open Source system. However, it was built as a system predominantly taxonomically based, to accommodate information for zoological collections. The requirements for other collections (geology, botany, anthropology, living collections) can be different and necessitate adaptations of the DaRWIN system and other specialized CMS.

The NaturalHeritage project aims to establish a new data/metadata ecosystem to (re)create links between the different ressources, improves the standardization, provides a common research portal and develops web services to connect the Belgian Natural History collections to the international infrastructures.

3. STATE OF THE ART AND OBJECTIVES

Digitizing Natural History collections is an enormous task because they typically hold billions of specimens. Belgian institutions are home to more than 55 million specimens representing one of the most important collections in Europe and even worldwide.

The Royal Belgian Institute of Natural Sciences (RBINS, 37 million specimens) decided more than a decade ago, following a review of the Collection Management Systems available at that time, to develop an “in house” solution called ColMaT/DaRWIN. It is used by RBINS and more recently also by Royal Museum for Central Africa (RMCA, 13 million specimens). The database was built to include information concerning biological collections with the focus on zoological taxonomy.

The Earth Sciences collections are notably missing from RMCA and RBINS central databases. RMCA and Agentschap Plantentuin Meise (APM) are also in need of a better system for the digital management of various botanical collections including living specimens and/or lineages and the large wood collections (xylarium and xylotheque). The requirements for these collections are diverse which would necessitate extensive adaptations to the DaRWIN system or the use of another CMS.

The fast evolution in programming and data exchange between different systems and devices (desktops, laptops, tablets, smartphones, etc.), together with the exponential development of web applications providing access to a wide variety of geo-coded or geo-referenced natural sciences information, forces one to adopt an approach with the interconnection of separated modules and data at its core. This requires the use of semi-structured data to cover the scope of natural sciences and related disciplines. The new architecture focuses on interoperability within the NaturalHeritage.be portal (amongst the various sub-modules), as well as data from certified e-resources including international aggregators. The proposed modular organization of NaturalHeritage will better enable scientists, collection managers and others to use the data. By being compliant with recognized European and International standards, protocols and regulations, including INSPIRE, Open Data or ABS, these resources become available, accessible and reusable by other online collection management systems.

While respecting the workflow of each institution, the common platform will facilitate cross-validation of data to avoid multiple encoding, which is time consuming and a source of errors. Interoperable databases using compatible technologies would reinforce synergy between the partners, allowing them to share the resources needed for the further development and maintenance.

The project’s objectives are:

- To achieve a better integrated digital collections management at a Belgian level, based on the interoperability of the different datasets and modules with a common search portal; This concerns the individual data already digitized in the Collection Management Systems, but also the description of collections that have yet to be digitized (about 90% of the collections).

- To provide adequate validation tools, by using web-services compliant with internationally recognized standards for digital curation at source (i.e. cross-checking scientific names, locality names, using standards for import/export of data, using standards for visualisation tools, etc.)
- To extend the data model(s) to new types of collections hosted by partners that are not centred on taxonomic data, in order to reach a broader interoperability and greater application;
- To be compliant with data requirements according to European directives (INSPIRE & OpenData) and international protocols (Nagoya) and regulations (ABS);
- To provide reliable open access to high quality data from the different stakeholders and users through a common research portal allowing an historical track and a global overview on biodiversity and scientific data via a search system sorted geographically and/or thematically;
- To provide curators and scientists with appropriate tools to assist them in their daily research work;
- To promote the new architecture among colleagues and peers using CETAF by enlarging the capacity of the platform both geographically (covering Europe and internationally) and conceptually (to other related disciplines), in the light of existing societal challenges (invasive species, health security, etc.).

4. METHODOLOGY

The project consolidated and broadened the collaboration, integration and interoperability amongst Belgian institutions leading to the development of the technical hub for the Belgian DiSSCo partners.

4.1 IT developing strategy

The IT development was achieved by additional internal developers and biodiversity information data specialists in collaboration with the existing staff and expertise of the IT institutional services. The project used Open Source technologies with a high degree of documentation for further development and smarter use. The project reused all codes available from the existing/previous projects. The working code(s) and the final product(s) are available on GitHub as is the case for the DaRWIN code.

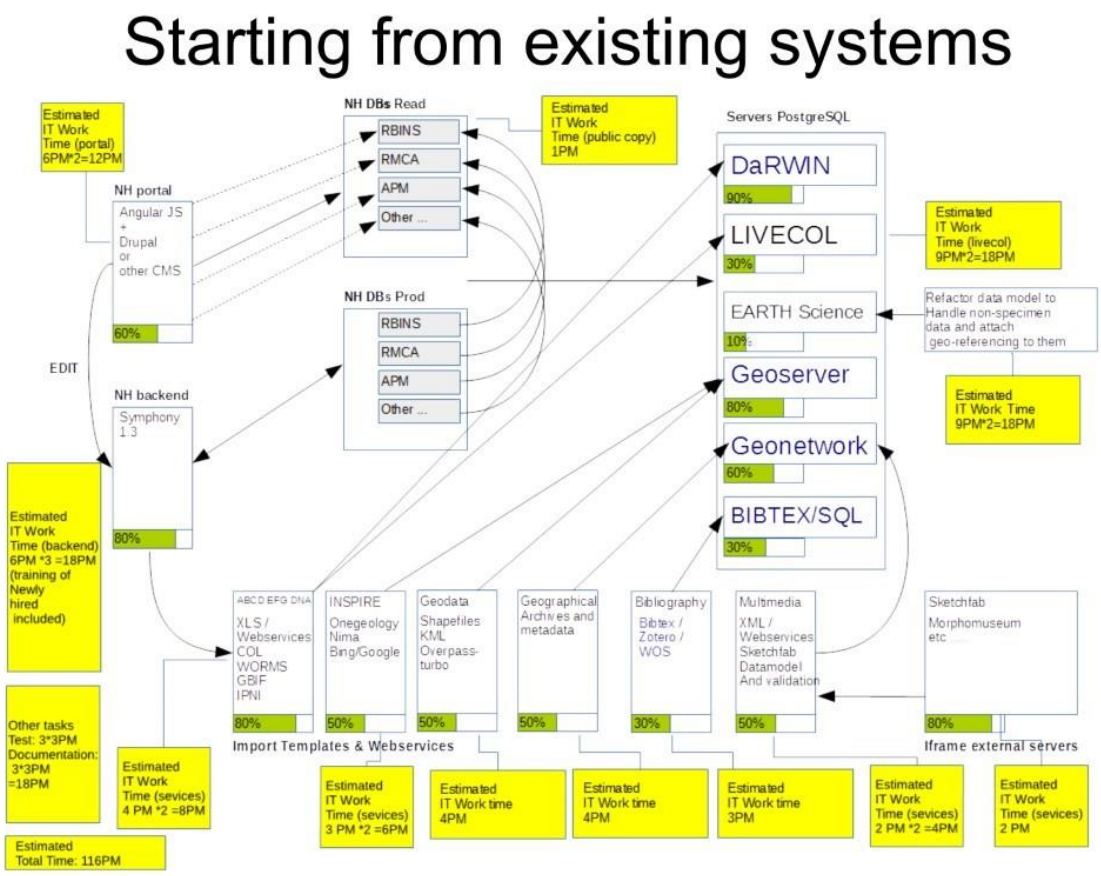


Figure 1: The developing strategy starting from previous developments by the 3 collection partners

4.2 Data Model strategy

The project reused and adapted the existing DaRWIN data model but added several interoperable modules. All modules communicate inside NaturalHeritage at a higher level than the database, through exchange of semi-structured data (JSON or XML) allowing cross linking between information which was not related before due to the use of incompatible systems. All data from the partners are re-indexed using an Elastic Search structure. This approach allows for the addition of new modules or providers without changing the data model or existing collection management systems.

The data model produces data and metadata compatible with other international efforts in the environmental domain (GBIF, BIOCASE, GEOGASE, OAI-PMH) and fulfil the obligations of the European Union Member States (derived basically from INSPIRE, Open Data and Nagoya Protocol).

4.3 Traceability Strategy

The different modules of NaturalHeritage not only manage the value(s) describing objects but also the argumentation of the attribution(s) and the complete history of them. The use of unique indentifiers, DOIs or a permalinks help to merge data and metadata from different sources. NaturalHeritage addresses this issue, following the recommendations of the CETAF working group. The use of UUID allows one to link to the original data. This is important for identifying original data ownership for the correct citation of the source.

4.4 Multiplatform Open source technologies

The first version of DaRWIN was developed with the proprietary SQL database Oracle and Oracle forms.

The current (and rebuilt) version of DaRWIN is based on open source technologies (PostgreSQL & the Symfony framework) and the PHP programming language.

RBINS is also using the Plone open source Content Management System (CMS) for bibliographic references, and for some collections. RMCA developed in the framework of the NaturalHeritage Project a Multimedia Database built on the Collective Access framework.

The NaturalHeritage Portal is based on the reindexing of data using the ElasticSearch technology (www.elastic.co/) which allows fast and agile indexation of different resources. The visual interface comprises a simple view of the search and results forms, which is included in existing websites using Iframes. This approach iseasier to maintain whilst being independent of the technical evolutions by the partners.

NaturalHeritage uses open source technologies that are already used by its ICT partners. This will not only increase interoperability between the Belgian partners, but also with international systems like the GBIF portal and DiSSCo Elvis system. The Open Sources technologies used in NaturalHeritage are:

- DaRWIN Collection Management System
  - Php, Symfony 1.4 (using a recent fork compliant with PHP 7)
  - PostgreSQL
- VirtualCollections
  - RMCA Collective Access (MySQL, with Mirador IIIF viewer)
  - RBINS Plone CMS 4.3 but will migrate to CollectiveAccess and IIIF viewer
- Nagoya/ABS Management system
  - RBINS Plone CMS 4.3
- Bibliography database and Open Access depot
  - RBINS Plone CMS 4.3
- NaturalHeritage Portal
  - RBINS Plone CMS 4.3
  - Symfony 3
  - ElasticSearch
  - Python scripts with Cron jobs (to refresh and update the data)
  - R Shiny
- CETAF Collections Registry
  - RBINS Plone CMS 4.3
  - ElasticSearch
  - Python scripts with Cron jobs (to refresh and update the data)
  - R

Figure 2: The open source technical architecture of Natural Heritage (RBINS & RMCA)

Figure 3: The technical architecture of Natural Heritage (APM)

4.5 Hardware architecture

The portal server is hosted by the Coordinator (RBINS) on the same hardware as the other RBINS collection applications. This allows for the easier maintenance of the NH Portal. Each server is virtualized, allowing a fast migration of each component to new hardware. Each "collection" partner hosts their own collections server and collection management system.

RBINS and RMCA tested and used different approaches to access to the back-end interface of DaRWIN. It is accessible through:

- the internet for RBINS
- the RMCA intranet with VPN for external access

Both architectures were tested and are fully functional for all developments including the embedding of the new DaRWIN interface in the Natural Heritage portal.

Figure 4: The open source technical architecture of Natural Heritage

5. SCIENTIFIC RESULTS AND RECOMMENDATIONS

The NaturalHeritage project produced several developments related to the different goals of the project.

5.1 The DaRWIN Open source Collections Management System

1. It created new functionalities in the DaRWIN Collection Management System: the complete description of the new features is available in annex 1 - BBR/175/A3 NaturalHeritage: The DaRWIN Collection Management System

2. It developed a new module for the labels (direct printing with specific printer drivers or indirect printing using CSV file): the complete description of the new features is available in annexe 2 - BBR/175/A3 NaturalHeritage: The Label module of the DaRWIN Collection Management System

3. It developed a new module for the loans: the complete description of the new features is available in annex 3 - BBR/175/A3 NaturalHeritage: The Loans module of the DaRWIN Collection Management System

4. It developed a new import module based on XLS templates and a 3 steps import: the complete description of the new module is available in annexe 4 - BBR/175/A3 NaturalHeritage: The Import module of the DaRWIN Collection Management System

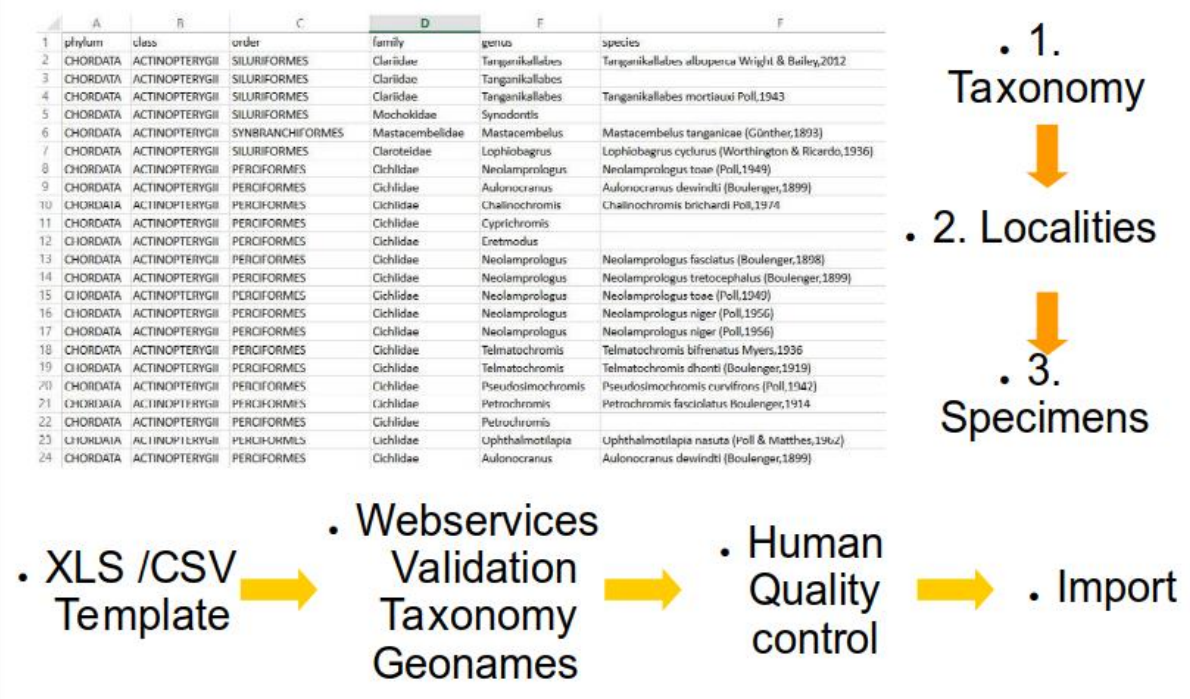


Figure 5: The new import procedure in 3 steps using XLS templates

5. It designed an alternated user interface for the DaRWIN database which can be embedded in the Institutional website using simple Iframes: the complete description of the new interface is available in annex 6 - BBR/175/A3 NaturalHeritage: The new interface of the DaRWIN Collection Management System

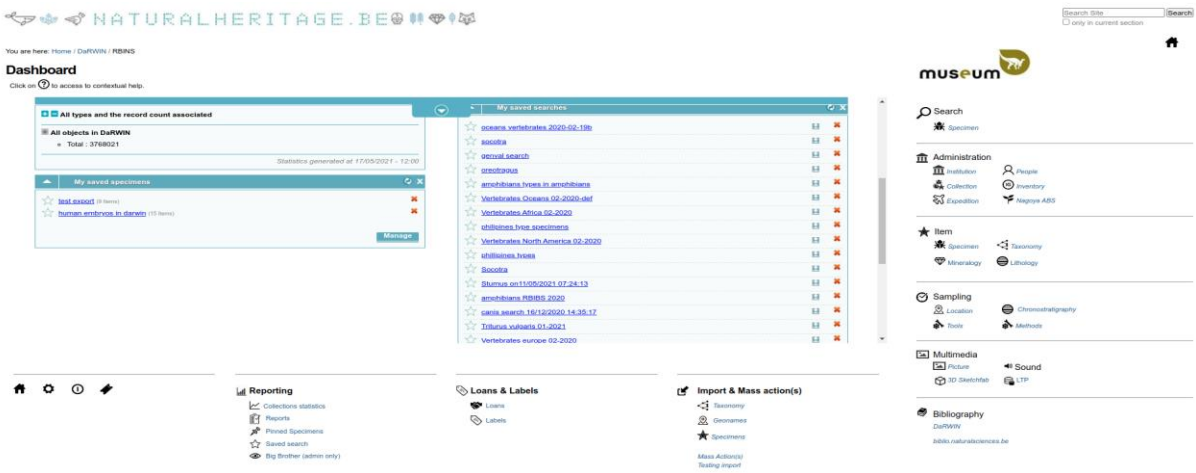


Figure x: The new interface of DarWIN CMS integrated in the NH Portal

5.2 The GeoDaRWIN Open source Collections Management System

6. It developed a new module to manage the information related to geological data and collections: the complete description of the new module is available in annex 7 - BBR/175/A3 NaturalHeritage: The GeoDaRWIN Collection Management System

5.3 The commercial CMS BG-BASE (living collections)

7. It evaluated several solutions and finally opted for BG-BASE, a specialized CMS for the botanical preserved and living collections: the complete description of the new module is available in annex 8 - BBR/175/A3 NaturalHeritage: The Botanical collections Management System

5.4 The ABS / Nagoya Open source Archives Management System

8. RBINS and RMCA developed a module for the ABS Nagoya and Cites Documents: the complete description of the new module is available in annex 5 - BBR/175/A3 NaturalHeritage: The ABS and CITES module of the DaRWIN Collection Management System

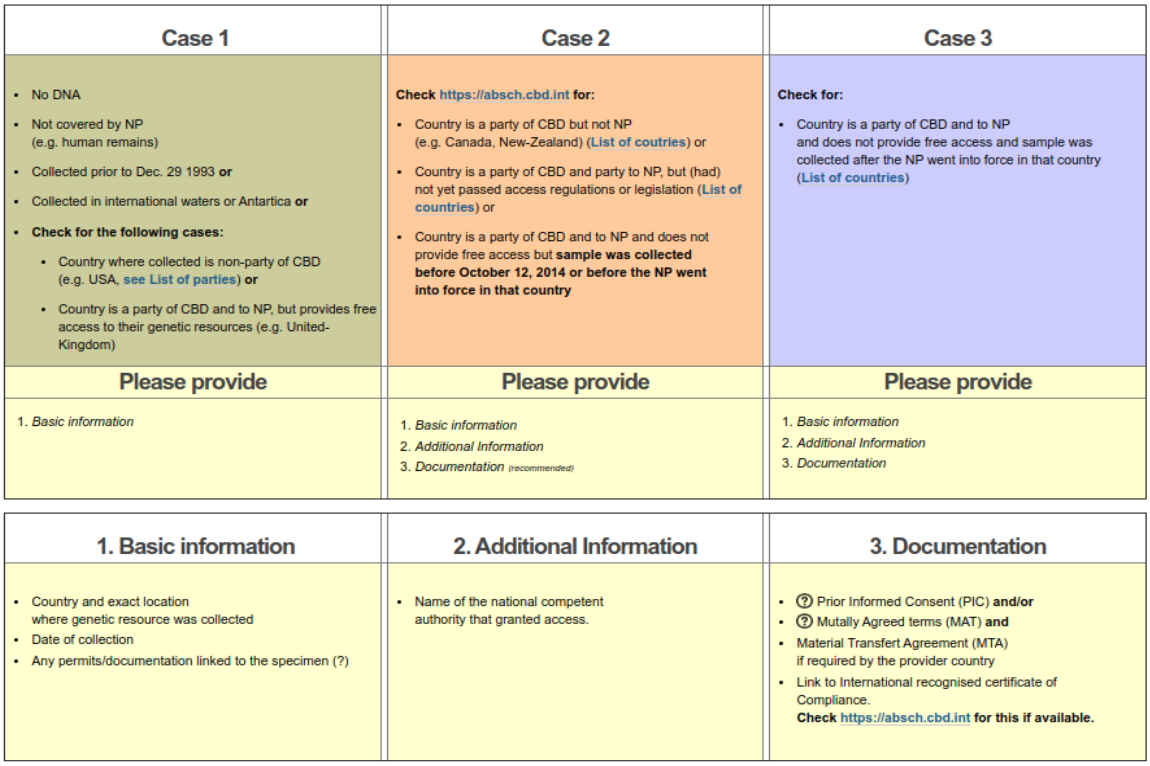


Figure 6: The new Homepage of the Nagoya module of NaturalHeritage

APM stores the scanned documents on a cloud drive (Google), following a specific naming protocol, which are linked to their relevant metadata and specimens in the Collections Management System (see Annex 5).

5.5 The Multimedia Server(s)

9. It developed a new Multimedia module using IIIF compliant viewers: the complete description of the new features is available in annex 9 - BBR/175/A3 NaturalHeritage: The Multimedia module of the DaRWIN Collection Management System

Figure 7: The IIIF image viewer from DaRWIN CMS

5.6 The Centralized Indexing System and Research portal

10. It created a new centralised indexing system with a multi-level search interface using ElasticSearch and a new public search portal: the complete description of the new features is available in annex 10 - BBR/175/A3 NaturalHeritage: The (re)Indexing System and Research portal

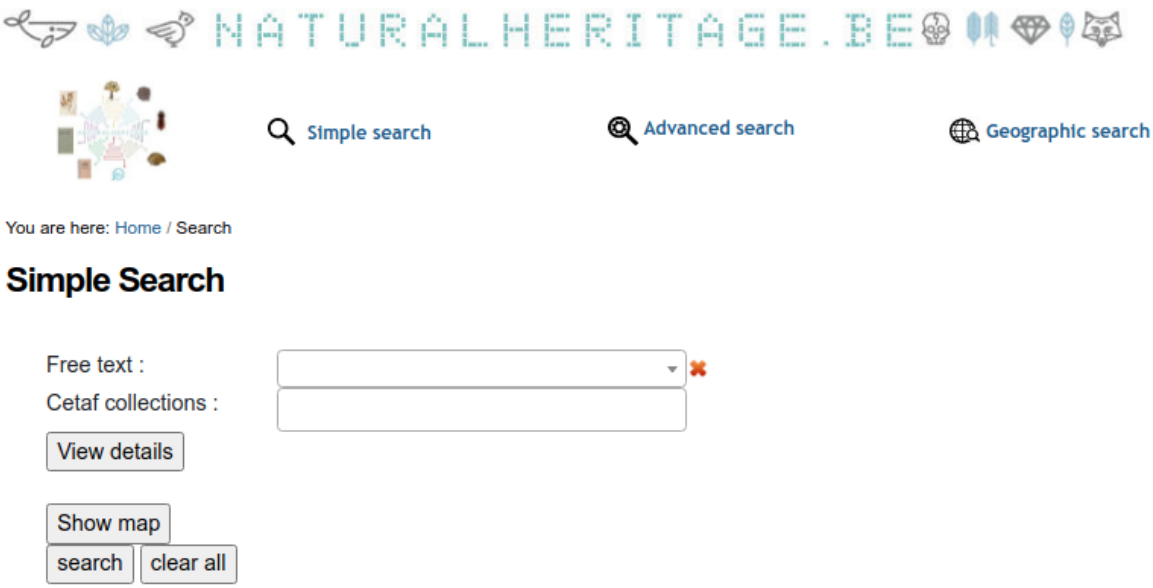


Figure 8: The simple search interface of Natural Heritage portal

5.7 The integration of the NH portal in the international landscape

5.7.1 DiSSCo ESFRI (https://www.dissco.eu/)

DiSSCo represents the largest ever formal agreement between natural history museums, botanic gardens and collection-holding universities in the world. This Distributed System of Scientific Collections is a new Research Infrastructure (RI) for the natural science collections. The DiSSCo RI works for the digital unification of all European natural science assets under common curation, access, policies and practices, and aims to ensure that the data is easily Findable, Accessible, Interoperable and Reusable (FAIR). As such, DiSSCo will transform a fragmented landscape of crucial natural science collections into an integrated knowledge base that provides interconnected hard evidence of the natural world.

The scope of DiSSCo encompasses Earth sciences and Life sciences collections. Earth sciences focuses on studying the evolution and composition of planet Earth, and are thus generally understood to comprise subjects that are related to geology. Life sciences focus on the study of living organisms (animals, plants, fungi and microorganisms), their biology, their ecology and their evolutionary histories.

Several projects have already started in order to prepare the setup of DiSSCo. The partners of NaturalHeritage (APM, RBINS, RMCA and CETAF) are involved in several of these projects.

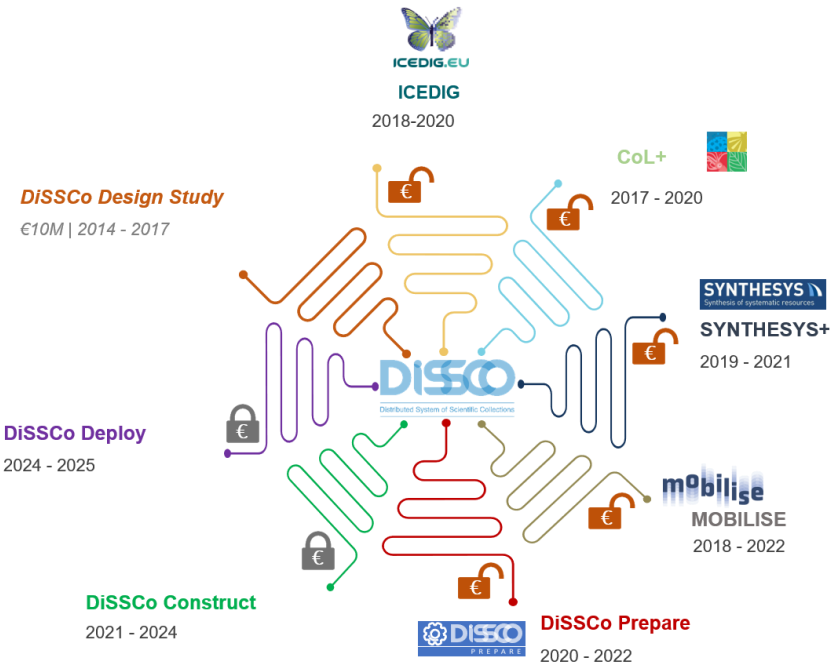


Figure 9. Schematic overview of DiSSCo-linked projects

Several Belgian scientific institutions already signed the European MOU. They collaborate in the framework of several infrastructure initiatives at the Federal or Flemish regional level.

Figure 10: DiSSCo Belgian partners having signed the MOU or ongoing

Figure 11: Timeline of DiSSCo setup

Since the beginning of the NaturalHeritage project, all developments were made in alignment with existing tools for the requirements of DiSSCo. The results of the NaturalHeritage project can be used by the DiSSCo Belgian partners community.

The NaturalHeritage portal is the common search interface for collections housed by Belgian institutions. This could also be used by the Belgian node in the DiSSCo infrastructure. The new version of DaRWIN and its embedding in the NaturalHeritage portal as well as the setup of Multimedia server with IIIF viewer and a dedicated CMS for the Nagoya / ABS documents prepare common tools which can be reused/shared by the DiSSCo community. This is important especially for smaller collections which can not invest significant time or money in the development of a specific Collection Management System aligned with the DiSSCo requirements.

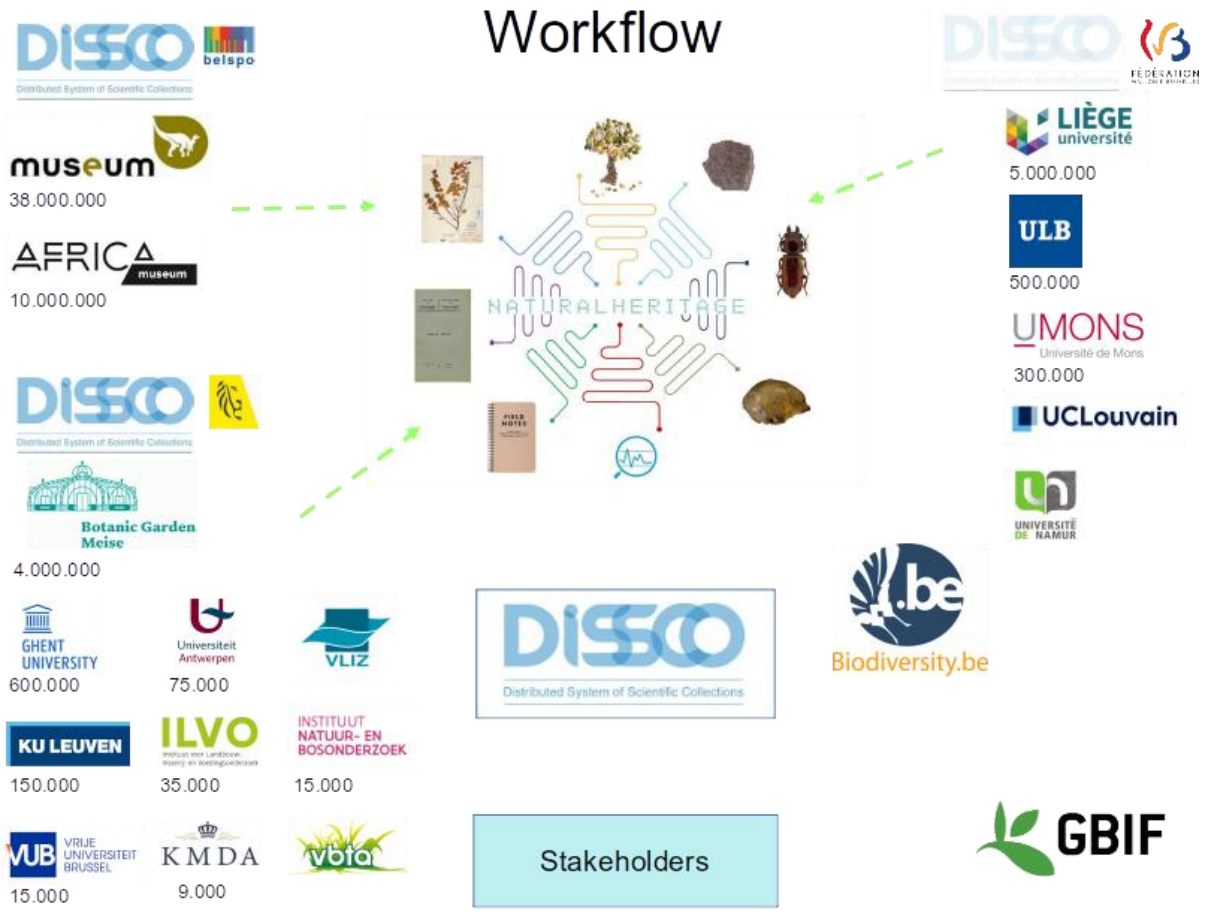


Figure 12: Interaction of DiSSCo Belgian partners with the NaturalHeritage portal

The DaRWIN, PLONE CMS and Collective Access systems developed by RBINS and RMCA could be used by the other collections of Zoology, paleontology, geology and anthropology.

The University of Liège which is the third Belgian collection in size (> 5.000.000 specimens) is interested in benefiting from the expertise and infrastructure of the FSI's, while most of the preserved botanical collections and many of the living botanical collections have been digitized and managed by the APM.

The Flemish universities and research institutions are part of the the DiSSCo Flanders project financed from 2021-2024 by the Flemish Science Foundation FWO lead by APM. As their collections become available they will also be connected to the Natural Heritage portal, provided to GBIF and GeoCase and become part of the DiSSCo services.

5.7.2 CETAF Community (https://www.cetaf.org/)

71 European institutionsare represented by 40 CETAF members: Europe's leading natural science museums, natural history museums, botanic gardens and research centres. All CETAF members curate, host and manage biological collections and engage in taxonomic research. Collectively, CETAF institutions manage half of the world's biological collections. The range of specimens in CETAF taxonomic facilities holds a wealth of stories about the past, present and future of the natural world around us. This unique infrastructure represents a precious resource for research on global biodiversity.

The digitization of Natural History collections is one of the biggest challenges for this community. Some specific collections like herbariumsheets allow massive digitization programs but the global ratio of digitizing most of the collections at specimen level does not exceed 10%. The complete digitisation of the specimens is thus a very long term objective and therefore it is very important to provide information at collection and sub-collection levels.

Providing data/metadata at collection level giving information about the size, the diversity and the taxonomic and/or geographic coverage of these collections becomes very important. The NaturalHeritage consortium developed such a tool reusing the approach and the results of the NH project.

11. Development of an “Institution Passport” and a “Collections Registry” for CETAF and DiSSCo partners using the NaturalHeritage technologies (Plone & ElasticSearch): the complete description of the new features is available in annex 11 - BBR/175/A3 NaturalHeritage: The CETAF Collection Registry. CETAF Collection.

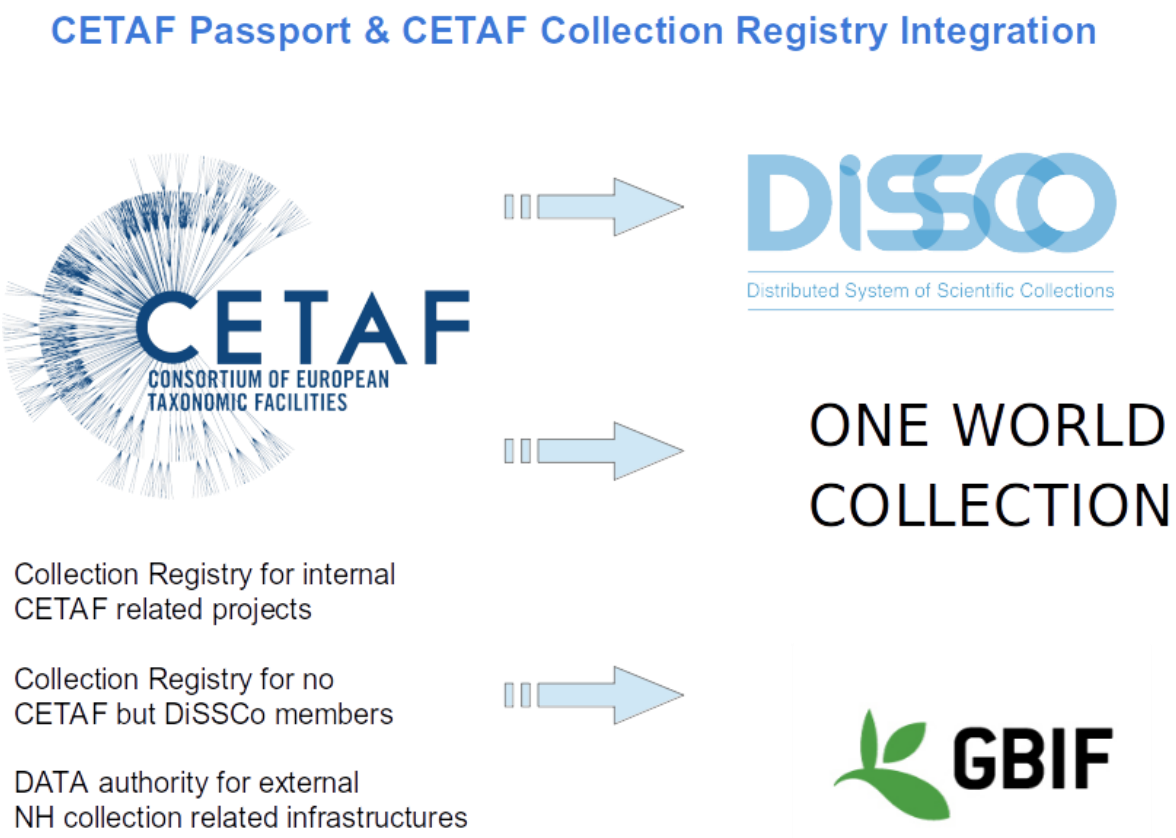


Figure 13: CETAF collections registry as authority about collection descriptors

This registry allows to encode this information and to distribute it to an international aggregator as GBIF or an infrastructure as DiSSCo.

The information is organised in 11 main collections and 69 subcollections. The user can navigate through countries, institutions or collection categories. The encoding

system was developed using the open source Plone CMS and the dexterity tool for the definition of the data model of the objects. The REST API of Plone was used for collecting all public information which is re-indexed by an Elasticsearch index as developed for the Natural Heritage portal. The Elasticsearch index can be reused by any external “client”/aggregator which will use the CETAF collections registry as an authority.

Figure 14: CETAF collections registry and Institutions menu

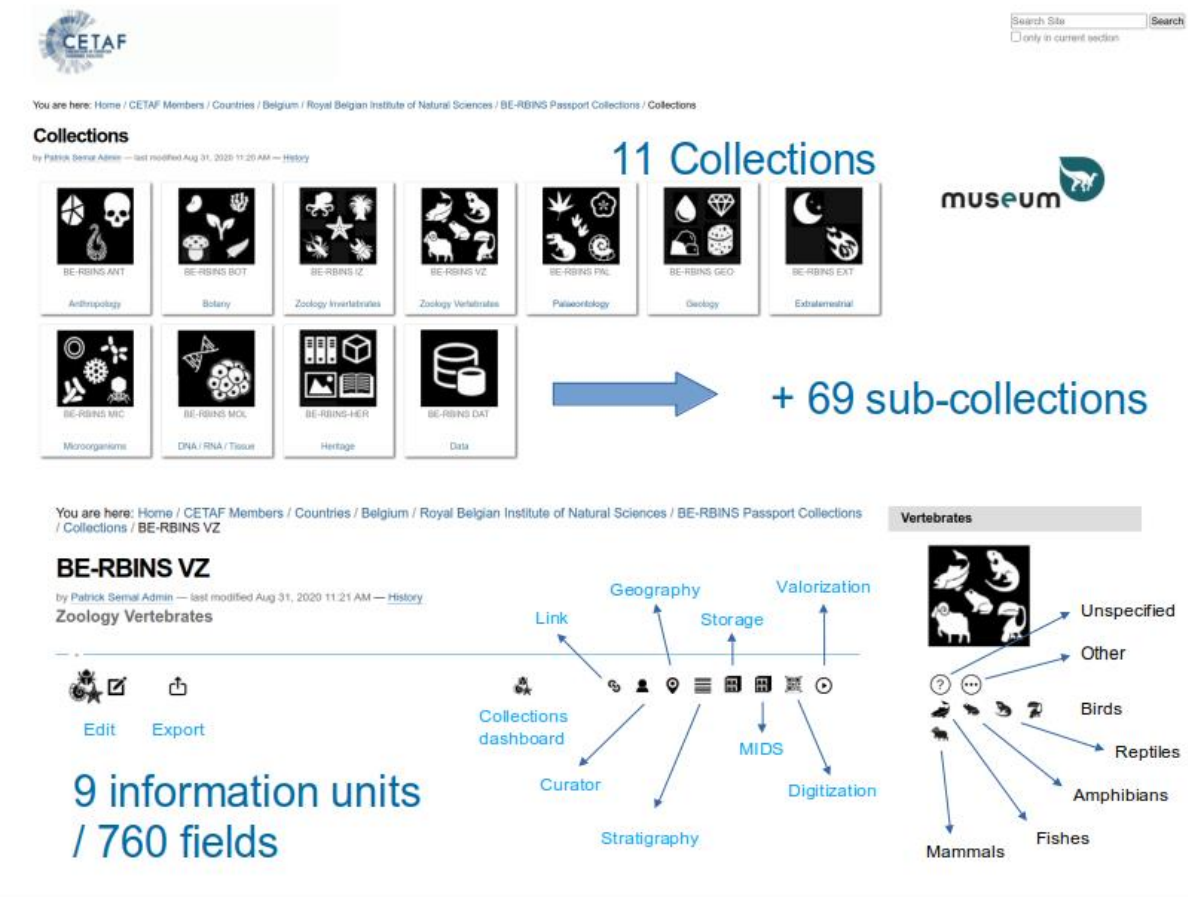


Figure 15: Organisation of the information and collections in the CETAF collections registry

5.8 Recommendations

The BRAIN project allowed partners to accomplish important developments in the framework of the preparation of the Belgian collections to the DiSSCo infrastructure:

- the DaRWIN Collection Management Systems are now able to export automatically the public data to GBIF using IPT;
- the connections to the CMS are now using https:// protocols;
- the CMS are compliant with international standards (Darwin Core, Dublin Core, OAI-PMH, IIIF, CETAF stable identifiers);
- the indexation with Elasticsearch combines the information from the different CMS respecting the original ownership of the data;
- the CETAF collection registry provides tools for the description of the collections which are not yet digitized at specimen level.

All these developments are based on Open source technologies and available on the Github folders of the project partners for reuse by other institutions/collections in Belgium or abroad.

Nevertheless, the division of competences and funding resources make it (very) difficult to transfer technologies to interested partners which are not in the same institutional pillar (regional or federal).

- APM has a huge expertise related to botanic preserved and living collections;
- RBINS and RMCA have a huge expertise related to zoological, geological, paleontological and anthropological collections;
- Universities and regional institutions have collections with a high level of diversity which correspond to APM, RBINS and RMCA collections and could benefit from their expertise;
- The BRAIN program allows collaboration between FSI's and regional institutions in the research aspects but there is no funding opportunity for the transfer of technologies/expertise. The infrastructure program of the FWO does not allow the funding of federal institutions and the federal program supporting the infrastructures is only funding the FSI's.

A specific funding, independent of the political levels, using a direct financing of BELSPO or a co-financing of the different political levels could be the solution. The Belgian Platform for Biodiversity is a good example of what we need. The scope of the Platform could be extended and/or RBINS/RMCA can play this role among the Belgian DiSSCo community but this needs "extra money" as these costs cannot be covered by the FSI's dotations due to several cuts during the last decade.

6. DISSEMINATION AND VALORISATION

Alignment workshop of the DiSSCo ESFRI proposal (P. Semal, Leiden, 19/01/2018)

2017

Internal APM Natural Heritage meeting 25/01/2017 (Meise)
First Consortium meeting 27/01/2017 (RBINS)
Consortium meeting: 17/02/2017 (Meise)
Consortium meeting: 17/03/2017 (Tervuren)
CETAF ISTC/DWG: 27-28/03/2017 (Stuttgart)
Belspo PO and Consortium meeting: 18/04/2017 (Belspo)
Participation to CETAF Earth Sciences group: 24/04/2017 (Bonn)
Internal APM Natural Heritage meeting: 4/05/2017 (Meise)
Consortium meeting: 10/05/2017 (RBINS)
ABCDEFG and GeoCASE workshop 30/05/2017 (Berlin)
Participation to DiSSCo Connect Meeting: 15-16/05/2017 (Leiden)
Pluto F evaluation visit: 29-31/05/2017 (Meise)
Nagoya-ABS expert meeting: 5-6/07/2017 (EU, Brussels)
Internal APM Natural Heritage meeting: 15/06/2017 (Meise)
Internal APM Natural Heritage meeting: 27/07/2017 (Meise)
Demo multiple Taxonomies 4/08/2017 (Tervuren)
Internal APM Meeting on Botalista 21/08/2017 (Meise)
TDWG annual meeting 2-6/10 2018 (Ottawa)
Internal APM Meeting 19/10/2017 (Meise)
Internal APM Meeting on Botalista analysis 21/11/2017 (Meise)
Working visit of Botalista staff 27-30/11/2017 (Meise)

2018

Internal APM Meeting 4/01/2018 (Meise)
Consortium meeting 10/01/2018 (RBINS)
Internal APM meeting 18/01/2018 (Meise)
DiSSCo Align meeting 19/01/2018 (Leiden)
European Open Science Cloud meeting 22-25/01/2018 (Porto)
CETAF ICT/DWG meeting 21-22/02/2018 (Copenhagen)
ICEDIG Opening Conference 6/03/2018 (Helsinki)
Internal APM meeting 28/03/2018 (Meise)
General assembly of CETAF (P. Mergen, London, 24-25/04/2018)
CETAF meeting on stable identifiers (Copenhagen 19-20/06/2018)
General assembly of CETAF (P. Mergen, London, 24-25/04/2018)
TDWG/SPNCH annual meeting 26/08-01/09/2018, Dunedin

2019

CETAF ISTC/DWG 13-14/02/2019 (Vienna)
SYNTHESYS Plus Kick off 18-19/02/2019 (London)
COST Mobilise General Meeting (11-15/03/2019 (Sofia)
TimeMachine kick off 18-19/03/2919 (Brussels)

Natural Heritage consortium meeting 5/04/2019 (RBINS)
DiSSCo National Nodes meeting 15-16/04/2019 (Den Haag)
Digital Sequencing Information (DSI) 9/05/2019 (Brusssls)
Time Machine workshop 9-10/05/2019 (Brussels)
CETAF Spring General Meeting 13-16/05/2019 (Tartu)
SPNCH 26-31/06/2019 (Chicago)
Belgian ESFRI day 5/06/2019 (Brussels)
Open Science Cloud in Flanders 30/08/2019 (Meise)
DiSSCo meeting with UAntwerpen 02/09/2019 (Meise)
Bishop (DARIAH) launch 05/09/2019 (Brussels)
CIS- Open Science 06/09/2019 (Brussels)
ELIXIR infoday 16/09/2019 (Brussels)
Natural Heritage meeting 19/09/2019 (RMCA)
Data Management Plan for Belgium 10/09/2019 (Gent)
ESFRI Roadmap 2021 infoday 25/09/2019 (Brussels)
Museum.be 11/10/2019 (Brussels)
DiSSCo Synchronizing group 16/10/2019 (virtual)
DiSSCo Gent museum 16/10/2019
Biodiversity Next , TDWG, CETAF 20/10/2019-25/10/2019
GBIF BID and EU DEFCO 13-14/11/2019 (Brussels)
EOSC symposium 25-28/11/2019 (Budapest)
ICEDIG All hands meeting 2-5/12/2019 (Meise)

2020

LTER Belgium 9/01/2020 (Brussels)
Bishop meeting 14/01/2020 (Brussels)
FWO IRI call infoday 15/01/2020 (Brussels)
EOSC sustainability working group 22/01/2020 (Meise)
COST Mobilise Annual meeting and training 10-14/02/2020 (Warshaw)
DARIAH FWO call 17/02/2020 (Brussels)
Time Machine CSA final meeting 17/02/2020 (Brussels)
ICEDIG/DiSSCo/DiSSCo-Prepare 25-28/02/2020 (Helsinki)

Arvanitidis C., Hollingsworth P., Mergen P., Semal P., Keklikoglou K., Chatzinikolaou E., Addnik W., Smith V. & Koureas D. Combined High-Throughput Imaging and Sequencing: Addressing the collections on demand requirement in SYNTHESYS+ project. Poster at the Biodiversity next Conference (Leiden, 22-25 October 2019)

Theeten F., Adam M., Vandenberghe T., Dillen M., Semal P., Scory S., Herpers J., Van den Spiegel D., Mergen P., Smirnova L., Engledow H., Casino A. & Gödderz K. NaturalHeritage: Bridging Belgian natural history collections. Poster at the Biodiversity next Conference (Leiden, 22-25 October 2019)

Adam M., Theeten F., Herpers J., Vandenberghe T., Semal P., Van den Spiegel D. & Duchesne P., 2019. DaRWIN: An open source natural history collections data management system. Poster at the Biodiversity next Conference (Leiden, 22-25 October 2019)

Semal P., Adam M., Van den Spiegel D., Theeten F., Engledow H., Mergen P., Gödderz K. & Casino Rubio A. CETAF Collection Dashboard: Mapping natural history collections diversity. Poster at the Biodiversity next Conference (Leiden, 22-25 October 2019)

Lahogue P., Herpers JM., Theeten F., Van den Spiegel D. GeoDarwin, an Open-Source Geological Data Management Tool. Poster at the Biodiversity next Conference (Leiden, 22-25 October 2019)

2020

DiSSCo Coordination team, Presentation of the CETAF Collections registry (P. Semal, Online, 14-04-2020)
Meeting CETAF GBIF, Presentation of the CETAF Collections registry (P. Semal, Online, 31-08-2020)
CETAF 47, Presentation of the CETAF Collections registry (P. Semal, MNHN Paris online, 31-08-2020)
CCG CETAF meeting, Presentation of the CETAF Collections registry (P. Semal, Online, 16-12-2020)

7. PUBLICATIONS

Vissers, J, Bosch, FVD, Bogaerts, A, Cocquyt, C, Degreef, J, Diagre, D, Haan, MD, Smedt, SD, Engledow, H, Ertz, D, Fabri, R, Godefroid, S, Hanquart, N, Mergen, P, Ronse, A, Sosef, M, Stévart, T, Stoffelen, P, Vanderhoeven, S & Groom, Q 2017, 'Scientific user requirements for a herbarium data portal' PhytoKeys, vol 78, pp. 37-57. DOI: 10.3897/phytokeys.78.10936

Hidvegi, F, Es, KCR, Groom, QJ, De Smedt, S & Bogaerts, A [PRESS RELEASE] De toekomst van digitalisatie van museumcollecties, van “DOE!” over “DoeDat”... tot “WeDigBio”.

Ellwood, ER, Kimberly, P, Guralnick, R, Flemons, P, Love, K, Ellis, S, Allen, JM, Best, JH, Carter, R, Chagnoux, S & Groom, QJ 2018, 'Worldwide Engagement for Digitizing Biocollections (WeDigBio): The Biocollections Community's Citizen-Science Space on the Calendar' BioScience, vol 68, no. 2, pp. 112–124. DOI: 10.1093/biosci/bix143

De Smedt, S, Stoffelen, P, Bogaerts, A, Engledow, H, Groom, QJ, Sosef, MSM, Van Wambeke, P & Dessein, S 2017, '[poster] Doe! Mass Digitization of the Herbarium Collection BR at Botanic Garden Meise', Gent, Belgium, 6/09/17 - 9/09/17.

Groom, QJ, Hyam, R & Güntsch, A 2017, 'Stable identifiers for collection specimens' Nature, vol 546, pp. 33.

Mergen P., Saarenmaa H., Lahti K., Dugénie P., Addink W., Koureas D., Groom Q., De Smedt S., Engledow H., Bogaerts A., Stoffelen P., Dessein S., Bloothoofd J., Casino A., Schulman L. DiSSCo: Towards a pan-European Infrastructure of Scientific Collections 2018). EUDAT Conference "Putting the EOSC vision into practice" 22-25/01/2018, Porto, Portugal.

Arvanitidis C., Hollingsworth P., Mergen P., Semal P., Keklikoglou K., Chatzinikolaou E., Addnik W., Smith V. & Koureas D., 2019. Combined High-Throughput Imaging and Sequencing: Addressing the collections on demand requirement in SYNTHESYS+ project. Biodiversity Information Science and Standards, 3:e37290.

Theeten F., Adam M., Vandenberghe T., Dillen M., Semal P., Scory S., Herpers J., Van den Spiegel D., Mergen P., Smirnova L., Engledow H., Casino A. & Gödderz K., 2019. NaturalHeritage: Bridging Belgian natural history collections. Biodiversity Information Science and Standards, 3:e37854.

Adam M., Theeten F., Herpers J., Vandenberghe T., Semal P., Van den Spiegel Spiegel D. & Duchesne P., 2019. DaRWIN: An open source natural history collections data management system. Biodiversity Information Science and Standards, 3:e39054.

Semal P., Adam M., Van den Spiegel Spiegel D., Theeten F., Engledow H., Mergen P., Gödderz K. & Casino Rubio A., 2019. CETAF Collection Dashboard: Mapping natural history collections diversity.Biodiversity Information Science and Standards, 3:e39667.

Lahogue P., Herpers JM., Theeten F., Van den Spiegel D., 2019. GeoDarwin, an Open-Source Geological Data Management Tool,3: e35946.

Hobern D., Paul D., Robertson T., Groom Q., Thiers B., Asase A., Luo M., Semal P., Woodburn M. & Zschuschen E., 2020. Advancing the Catalogue of the World's Natural History Collections Biodiversity Information Science and Standards, 4:e59324.

8. ACKNOWLEDGEMENTS

The NaturalHeritage project is a 24-month project funded by the federal science policy office (BELPSO). We would like to thank the project evaluators of the BRAIN programme and the members of BELSPO who monitored the project.

NaturalHeritage has developed a strategy, technologies and tools that optimise the encoding and valorisation of data/metadata related to the natural history collections and their accessibility by the different stakeholders.

We would like to thank all the members of the partners (RBINS, RMCA and APM) as well as the CETAF secretariat who have contributed to feed this project on a daily basis through their experience and through regular exchanges in the framework of the digitization efforts of the collections and the preparation of the ESFRI DiSSCo.

We would like to thank the IT departments of the partners for their support in installing the physical servers and setting up the https access as well as the Makina Corpus company for the maintenance and help in the development of the Plone sites.

DiSSCo is an important infrastructure for the valorization of natural science collections and data. The consortium thanks in advance any initiative that will allow the further development of this infrastructure in the near future and to promote collaboration between the participating Belgian institutions, whatever their level of political authority.

9. ANNEXES

Title	Authors	pages
annex 1 - BBR/175/A3 NaturalHeritage: The DaRWIN Collection Management System	Franck Theeten (RBINS/RMCA), Jean-Marc Herpers (RMCA), Didier VandenSpiegel (RMCA) & Patrick Semal (RBINS)	33
annex 2 - BBR/175/A3 NaturalHeritage: The Label module of the DaRWIN Collection Management System	Jean-Marc Herpers (RMCA), Franck Theeten (RBINS/RMCA), Marielle Adam (RBINS) & Patrick Semal (RBINS)	7
annex 3 - BBR/175/A3 NaturalHeritage: The Loans module of the DaRWIN Collection Management System	Jean-Marc Herpers (RMCA) & Patrick Semal (RBINS)	15
annex 4 - BBR/175/A3 NaturalHeritage: The Import module of the DaRWIN Collection Management System	Jean-Marc Herpers (RMCA), Franck Theeten (RBINS/RMCA), Marielle Adam (RBINS) & Patrick Semal (RBINS)	31
annex 5 - BBR/175/A3 NaturalHeritage: The ABS and CITES module of the DaRWIN Collection Management System	Jean-Marc Herpers (RMCA), APM, Marielle Adam (RBINS) & Patrick Semal (RBINS)	15
annex 6 - BBR/175/A3 NaturalHeritage: The new interface of the DaRWIN Collection Management System	Patrick Semal (RBINS) and Franck Theeten (RMCA)	54
annex 7 - BBR/175/A3 NaturalHeritage: The GeoDaRWIN Collection Management System	Jean-Marc Herpers (RMCA) & Pascale Lahogue (RMCA), Franck Theeten (RBINS/RMCA) & Didier Van den Spiegel (RMCA)	18
annex 8 - BBR/175/A3 NaturalHeritage: The Botanical collections Management System	Henry Engedow (APM) Mathias Dillen (APM), Ann Bogaerts (APM), Thierry Vanderborght (APM), Piet Stoffelen (APM), Elke Bellefroid (APM), Marc Reynders (APM), Larissa Smirnova (APM), Quentin Groom (APM), Frederik Leliaert (APM), Niko de Meeter (APM) & Patricia Mergen (APM)	5
annex 9 - BBR/175/A3 NaturalHeritage: The Multimedia module of the DaRWIN Collection Management System	Jean-Marc Herpers (RMCA), Franck Theeten (RBINS/RMCA) & Patrick Semal (RBINS)	36
annex 10 - BBR/175/A3 NaturalHeritage: The (re)Indexing System and Research portal	Franck Theeten (RBINS/RMCA) & Patrick Semal (RBINS)	26

annex 11 - BBR/175/A3 NaturalHeritage: The CETAF Collection Registry	Patrick Semal (RBINS), Franck Theeten (RBINS/RMCA), Laura Tilley (CETAF) & Ana casino (CETAF)	56
Total		296