

One Year OpenUp! - Discovering Natural History as a Part of Europe's Cultural Heritage

Petra Böttinger, Alexa Michel, Gianna Buchli, Peer Schwirtz, Anton
Güntsch, Walter G. Berendsohn

Botanic Garden & Botanical Museum Berlin-Dahlem, Germany



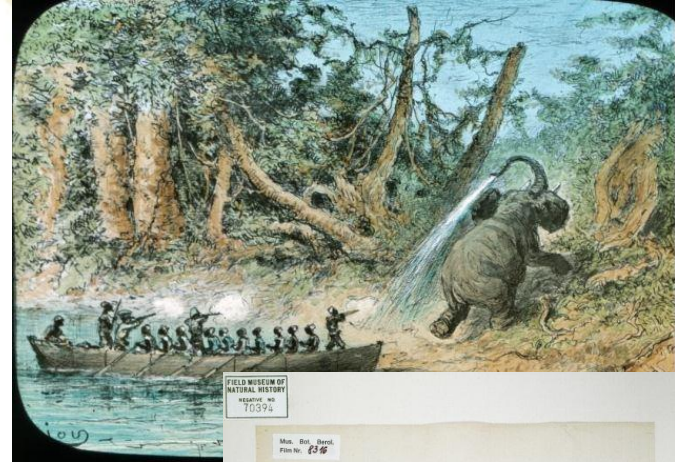
Freie Universität



Berlin

Background

Multimedia objects from the Natural History domain are most certainly part of the European Cultural Heritage but are dramatically underrepresented in public awareness



**Connecting the cultural and the natural history domains is the central idea behind OpenUp!
("Opening up Europe's natural history heritage to EUROPEANA").**



Tapestry called „Krokus“, B. Rendahl, 1976. © Upplandsmuseet, Uppsala, Sweden. www.europeana.eu



Herbarium specimen *Crocus vernus* L., BGBM Collection, Berlin.

Project Details

“OpenUp!” is a 3-year project co-funded by the European Commission and several European Natural History collections. In March 2012 it has successfully passed its first project year.



It is an initiative of CETAF, the Consortium of European Taxonomic Facilities, together with several European GBIF Nodes.

Tasks and objectives of OpenUp!

- Bringing content to Europeana from the Natural Science History domain
- Technical implementation - provide and support technical standards
- development of consistent IPR strategies for future data providers and consortium
- Sustainability plan for the future network maintenance
- Outreach; adding data providers for multimedia content, set up of data provider software



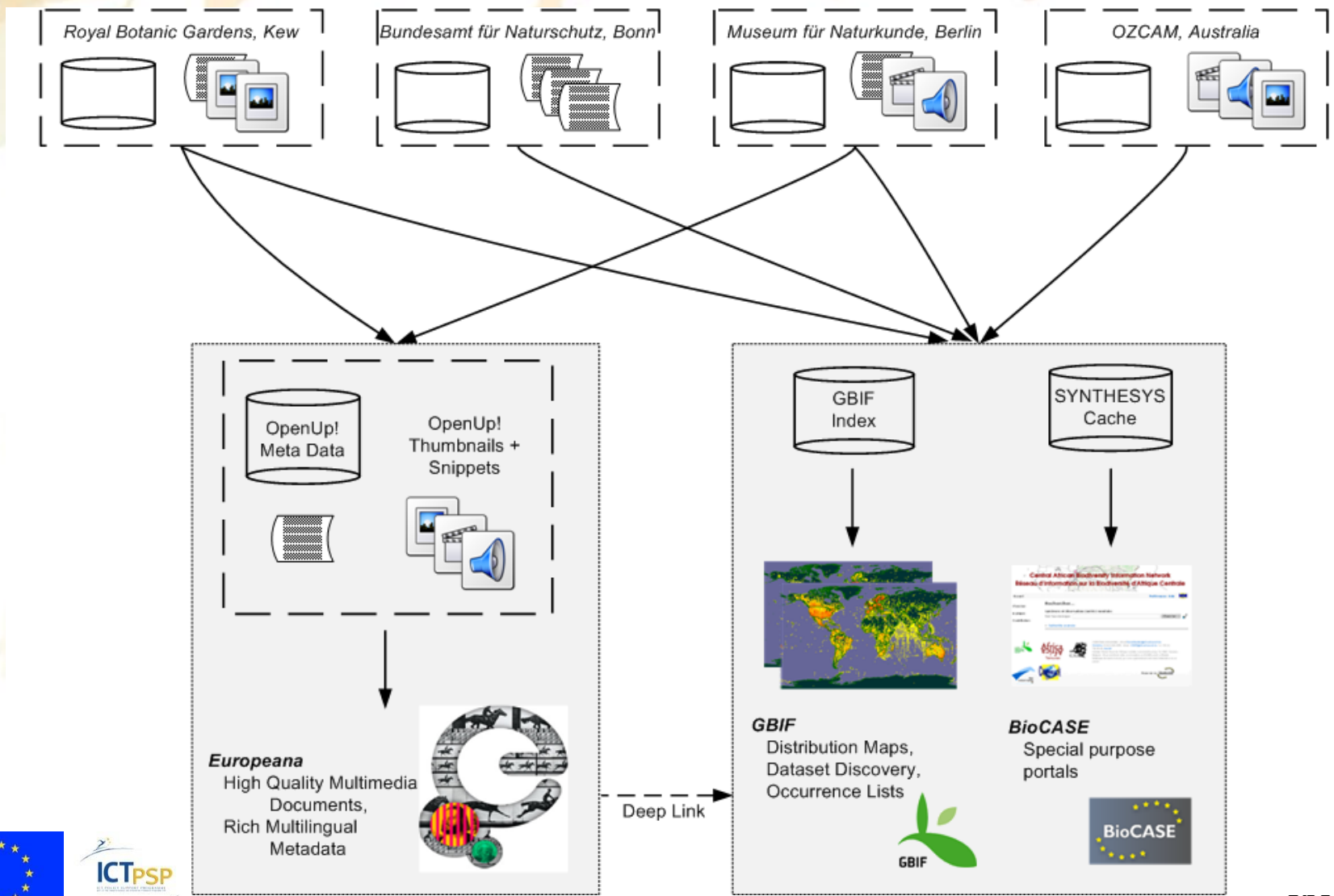
Bronze Cat Coffin.
© The Oriental Museum, University of
Durham, Durham, UK. www.europeana.eu

OpenUp! Task 1: Bringing Content to Europeana

- OpenUp! Is taking advantage of the established infrastructures in the Global Biodiversity Information Facility (**GBIF**), long-term viability
- OpenUp! is making multimedia content in the **BioCASE network** accessible to EUROPEANA
- All content **will automatically be served to GBIF** when mobilised by OpenUp.



Global Context of OpenUp!



Bringing Content to Europeana

OpenUp! committed to serve **at least 1.1 million objects** by the end of the project (Feb. 2014); details and numbers:

High quality object images	881 438
Natural history artwork	119 100
Animal sound files	21 350
Movies	1 200
Other	38 000
Total, at least	1 061 088

223.342 records of multimedia objects were delivered to Europeana via the OpenUp! infrastructure at the end of year 1!

Institution	records	ingested into Europeana
BGBM Herbarium Berlin	100.766 records	Feb2012
Natural History Museum Vienna, HERBARW	26.488 records	Feb2012
University of Vienna HERBARWU	19.678 records	Feb2012
Finnish Museum of Natural History (UH)	3.157 records	Feb2012
LANDOOE Linz	73.253 records	Feb2012

Next to be ingested: from ETI – 27.226 records



Red scheelite crystal on quartz gangue, lecoality Příbram – Březové Hory, central Bohemia, Czech Republic. Collection of National museum Prague.

Example: herbarium specimen search in Europeana

The screenshot displays the Europeana website interface. At the top left is the Europeana logo with the tagline "think culture". The main heading is "Explore Europe's cultural collections". A search bar contains the text "Centaurea cyanus", with a "Search" button and a "Refine your search" link. Below the search bar, it indicates "Matches for: Centaurea cyanus > PROVIDER:OpenUp!".

On the left side, there is a "Filter your search:" panel. A red arrow points to the "By provider" section, which lists several providers: "OpenUp! (7)", "The European Library (5)", "Slovenské národné... (3)", and "CultureGrid (2)". Below this, there is a checkbox for "Include content contributed by users" which is checked. A "Legend:" section at the bottom of the filter panel shows icons for Image, Video, Text, Sound, and 3D.

The search results are displayed in a grid format. The first row shows four results, each with a thumbnail image of a herbarium specimen and the text "Centaurea cyanus L." followed by a small "OpenUp!" logo. The second row shows three results in the same format. At the bottom of each results section, there is a "Results 1 - 7 of 7 Page: 1" indicator and social media sharing icons (Like, +1, Twitter, +).

For Europeana, OpenUp! provides

- a **single access point** to distributed non-bibliographic multimedia content **in the natural history domain**
- validation mechanisms to ensure **compliance with EUROPEANA standards.**
- **sustained item-level access** by integration with existing networks in the domain (i.e. GBIF, BioCASE and CETAF).
- metadata enrichment by means of **multilingual metadata vocabularies and thesauri** for natural history data (e.g. names) to enhance cross-linking of Europeana content.

For NH Collections, OpenUp! provides

- **Increased relevance** of Natural History domain data by inclusion in EUROPEANA
- Tools for **quality control** of species names and other data
- Help to **extend data provision** for GBIF and BioCASE
- Funding for a multilingual **index of common names**
- **Technical solutions** for shared but distributed information infrastructures

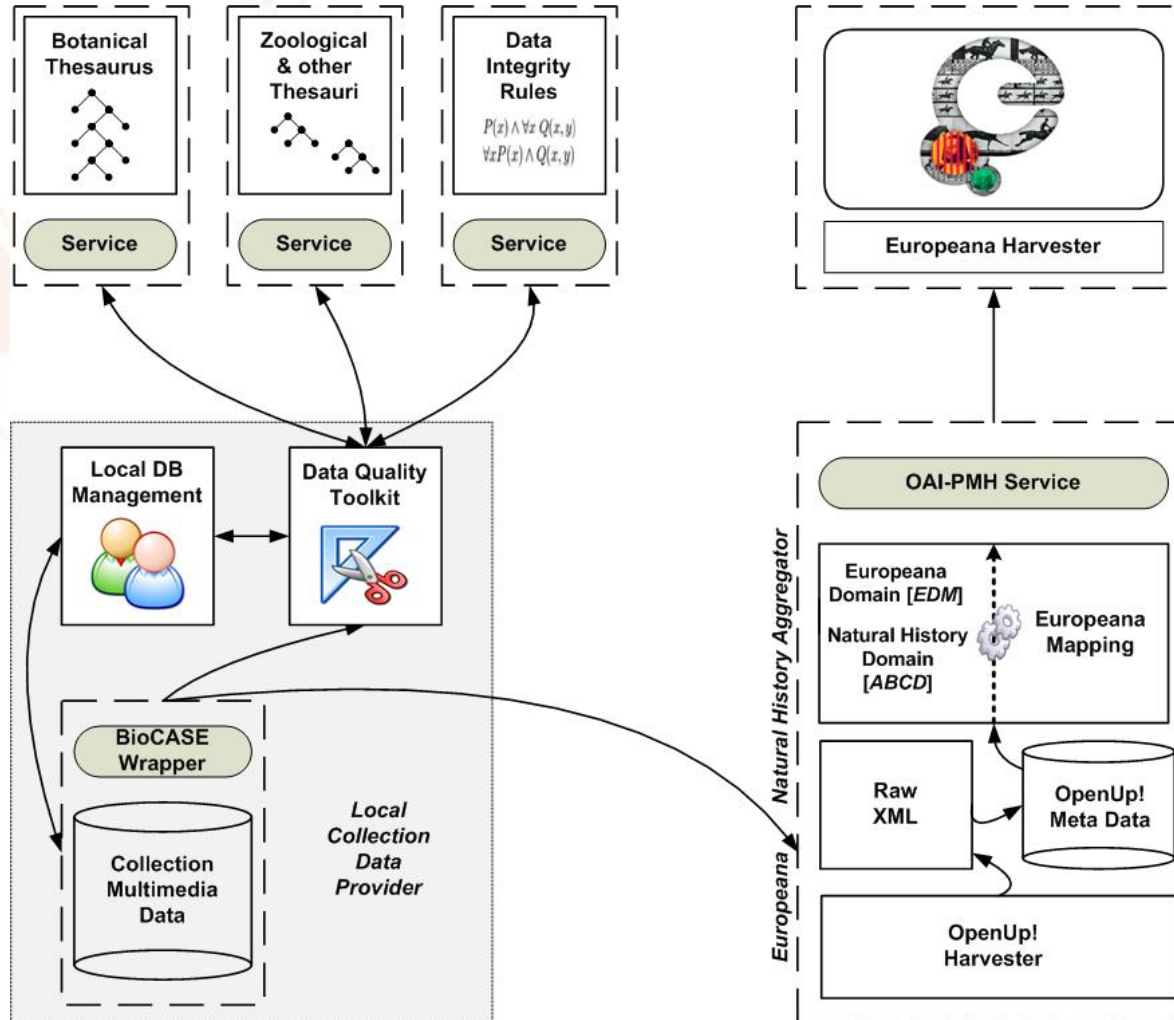
OpenUp! Task 2: technical implementation - provide and support technical standards

Provision of a framework for distributed content from local databases by a constantly updated aggregator meta-database, accessible through the OAI-PMH Interface (Open Archives Initiative Protocol for Metadata Harvesting)

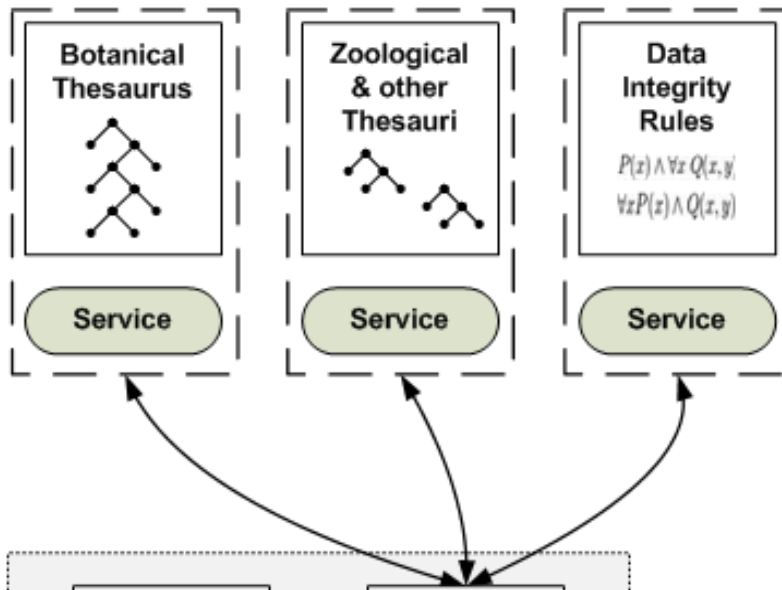


A drawer with tropical butterflies collected by Alfred Russel Wallace (Natural History Museum, London)

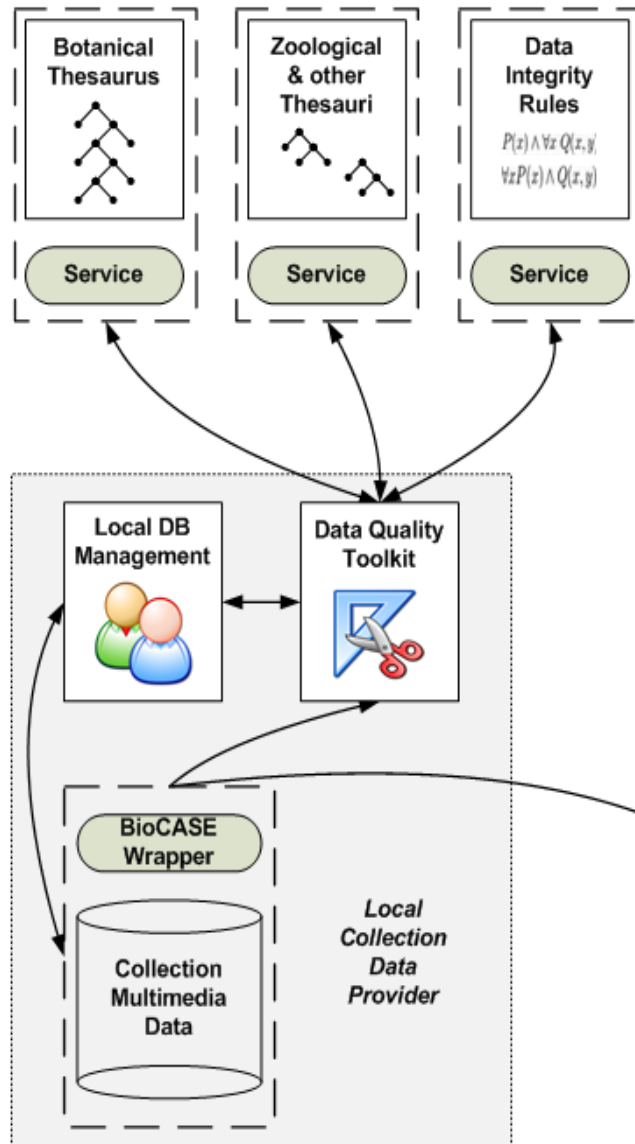
Data processing steps in the mobilisation of multimedia content from natural history institutions for Europeana



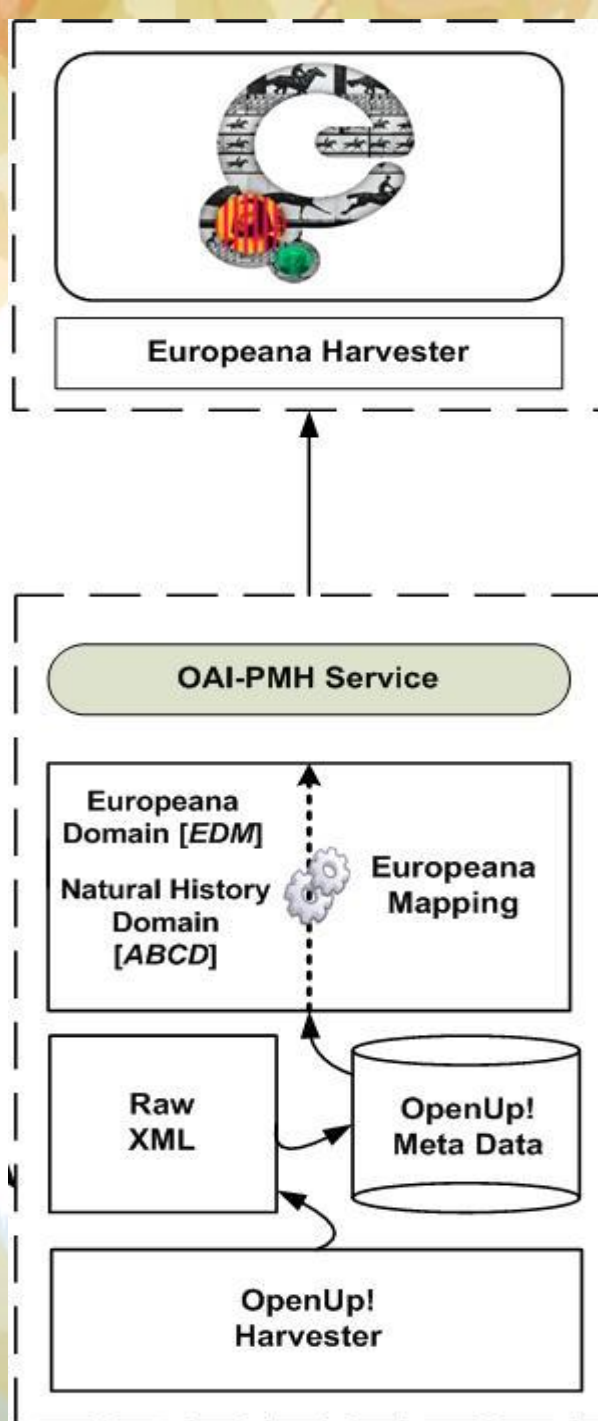
Data processing steps in the mobilisation of multimedia content from natural history institutions for Europeana



- **Local collection databases** holding records with references to multimedia objects
- managed with **local tools**



- The local database is accessed by the **BioCASE Wrapper**, that performs the **transformation to ABCD standard XML** (allowing access by standard tools)
- The **Data Quality Toolkit** software interacts with the BioCASE Wrapper to conduct **integrity checks** with respect to **names** and other **data integrity rules**.
- **Results are fed back** to the local DB management, including an interface, explaining XML errors that are discovered in the course of the harvesting process.



- At the **aggregator level** the **OpenUp! Harvester** will access the local wrappers
- **ABCD data records with linked multimedia content** are transformed to **Europeana standards (EDM)**
- and offered to Europeana (and others) for **harvesting via OAI-PMH**

Data check with the OpenUp! OAI-PMH platform

The screenshot displays the OpenUp! OAI-PMH platform interface. At the top, there is a banner with the 'OPEN UP!' logo and the text 'Opening Up the Natural History Heritage for Europeana'. Below the banner is a navigation bar with links for 'Search', 'Advanced Search', 'Browse', 'Collections', 'Login', 'OAI', and 'Help Wiki'. The main content area is divided into two columns. The left column contains a 'Main Menu' with options like 'OAI Provider', 'User Menu', 'Login', 'Select Language' (set to 'English'), and 'Select Template' (set to 'openup'). The right column shows a 'Detail' view for a record with the identifier 'ESE'. The record information includes: Title: Accipiter gentilis; Europeana Type: Sound; Type: Observation; Identifier: TSA - MFN - Anima; Europeana Is Shown By: http://www.tierstimmenarchiv.de; Rights: CC BY-NC-ND (Attribution for non commercial use); Europeana Rights: http://creativecommons.org/licenses/by-nc-nd/3.0/; Europeana Object: http://www.tierstimmenarchiv.de; Europeana Data Provider Name: Museum fuer Nat. The right column also displays the original ABCD2.06 Unit metadata in XML format, including elements like <abod:Identifications>, <abod:RecordBasis>, <abod:MultiMediaObjects>, <abod:FileURI>, <abod:Format>, <abod:IPR>, <abod:Licenses>, <abod:GatheringAgent>, and <abod:Person>.

The **OAI-PMH Protocol for Metadata Harvesting** is the preferred method for Europeana of capturing metadata. The OAI-PMH Interface displays both: the result of the ESE transformation (data in ESE format) and the view on the original ABCD2.06 Unit metadata.

Content Checker Tool

The screenshot illustrates the Content Checker Tool workflow. On the left, a browser window shows the source image at www.globis.insects-online.de/images/Berlin/Berlin. The browser's address bar and tabs indicate the user is on the 'Content Checker' page. A blue arrow points from the source image to a central orange box containing a magnifying glass icon, representing the content check process. Below this box, a blue arrow points to the resulting Europeana record for *Parnassius actius* (Eversmann, 1843). The record includes metadata such as 'Type: specimen', 'Data provider: Museum für Naturkunde Berlin', and 'Provider: OpenUp!'. It also features a 'User login' section with fields for 'Username' and 'Password', and a 'Log in' button. The record is displayed within a Europeana interface, which includes a search bar, a 'Translate details' section, and a 'Species Tree' showing the classification of the specimen. The record is also linked to a 'Similar content' section and a 'User login' section. The record is displayed within a Europeana interface, which includes a search bar, a 'Translate details' section, and a 'Species Tree' showing the classification of the specimen. The record is also linked to a 'Similar content' section and a 'User login' section.

Images: GloBIS, MfN

When the final checks are successful and the content provider agrees to the display, the data is released for subsequent publication in the Europeana portal.

Enhancing the Metadata

- The innovation lies in the way OpenUp! implements **data quality and enrichment measures** to achieve a high level of completeness and correctness of metadata associated with the multimedia objects.
- A dedicated **Data Quality Toolkit** is developed which integrates a number of quality services including comprehensive scientific name catalogues for botany and zoology. It also validates information on taxonomy, geography, language, and type of multimedia.

Digital herbarium**Barcode:** B 10 0294605**ImageId:** 285393**Family:** OPILIACEAE**Storage name:**Melientha suavis subsp. macrocarpa Hiepko
det. P. Hiepko 02.1979**Type status:** holotype**Herb. sub-label:** Royal Society North Borneo Expedition
1961**Collection:**Malaysia: Mt. Kinabalu; Ulu Liwagu and Ulu Mesilau. Leg.:
W. L. Chew, E. J. H. Corner & J. D. A. Stainton s.n.**Notes:** Ex Herb. Hort. Bot. Reg. Kew

Contact

Imprint

Disclaimer

Overview

Index

internet: konstruktiv

Melientha suavis subsp. macrocarpa Hiepko

Save as JPG...



**Quality check – Data cleaning:
275 Person-months reserved for
local data cleaning**

See OpenUp Newsletter #2

Image should be cited as: Röpert, D. (Ed.) 2000- (continuously updated): Digital specimen images at the Herbarium Berolinense. - Published on the Internet <http://ww2.bgbm.org/herbarium/> (Barcode: B 10 0294605 ImageId: 285393) [accessed 18-Aug-10].

[Direct link for this Image](#)


© Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin

Collection Data Quality Toolkit (main form)

Web-based integrated interface for a useful set of automatic quality checks:

- simple syntax rules for selected ABCD elements (e.g. country, date, latitude, longitude)
- comparison of scientific names with international checklists.

Additional rules are continuously added



The screenshot shows a web-based interface for the Collection Data Quality Toolkit. It is divided into three main sections: Provider Settings, Quality checks, and Filter. The Provider Settings section includes a text input field for 'Provider URL' and a dropdown menu for 'Sync' with 'yes' selected. The Quality checks section lists several criteria with checkboxes and a '2' next to each: 'Collection date fields', 'Site coordinates', 'Syntax of email elements', 'ISO country element', 'Scientific name (zoology)', 'Scientific name (botany)', 'Mime type for multimedia objects', and 'Multimedia object availability'. The Filter section has text input fields for 'Scient. name' and 'Ids', and a checkbox for 'Multim. only' which is checked. At the bottom of the form are 'Reset' and 'Go' buttons. The footer of the interface contains the text 'Opening Up the Natural History Heritage for Europeana' and the BGDm logo.

<http://services.bgbm.org/DataQualityToolkit>

Enhancing the Metadata

- ABCD specimen data records contain many elements that are useful for semantic linking
- Multicultural context
- Enhance scientific names by multilingual common names
- Enhance names with synonyms (using same services as data quality toolkit)

OpenUp! Welcomes New Data Providers

OPEN UP! OpenUp! Helpdesk - test

File browser Forums

Search this site:

User login

Username or e-mail: *

Password: *

- Log in using OpenID
- Create new account
- Request new password

Sign in with Windows Live ID

Opening up the Natural History Heritage for Europeana - Helpdesk

This is a beta version of the OpenUp! helpdesk facility. You can find documentation on project related matters in the File browser and a Question and Answer section in the Forums.

"Opening up the Natural History Heritage for Europeana" is a EC project under the CIP-ICT-PSP call. It aims to make available natural history collection objects and data to the Global Biodiversity Information Facility (GBIF) and to Europeana.

The OpenUp! Heldesk is managed by the Cybertaxonomy Services of the Royal Museum for Central Africa and will provide access to documentation, online tutorials, trainings and advisory helpline for partners, providers and users within and beyond the Consortium.

If you have any questions please post them in the Helpdesk General Forum or send them to

[Read more](#)

co-funded by the Community programme #Contentplus

GBIF

BioCASE

For further information, please visit the OpenUp! Helpdesk:
<http://open-up.cybertaxonomy.africamuseum.be/>
or contact them at: ***openup.helpdesk@africamuseum.be/***



The **OpenUp!** communication website set-up, using a Scratchpad (Drupal content management system) hosted by the NHM. <http://www.open-up.eu>

The screenshot shows the OpenUp! website interface. At the top, there is a navigation bar with links: Home, About us, News, Events, Partners, Related projects, Forum, Contact us, FAQ, and Helpdesk. The main header features the 'OPEN UP!' logo and the text 'Opening Up the Natural History Heritage for Europeana'. Below the header, there is a search bar on the left with the text 'Search this site:' and a 'Search' button. To the right of the search bar is a 'Content Highlights' section. This section has a green banner with the text 'botanical collection' and 'PLANTS'. Below the banner is a photograph of a botanical specimen, *Abrus precatorius*, with a scale bar. To the right of the photograph is a text box with the title '*Abrus precatorius*' and the text: 'This specimen of *Abrus precatorius* subsp. *africanus* is a duplicate (isotype) to its holotype stored at the herbarium in Kew.' Below this text box is a green box with the title 'Did you know that...?' and two bullet points: '• *Abrus precatorius* is also called *Jequirity* or *Crab's eye*?' and '• *Abrus* leaves are consumed as *vegetable* in Central and'. To the right of the content highlights is a 'Calendar' section showing a calendar for June. Below the calendar is an 'Upcoming events' section with a list of events: 'PIT Summer School_12 Sun, 2012-07-01 (All day) - Sun, 2012-07-08 (All day)', '12th DIVERSE conference Tue, 2012-07-03 (All day) - Fri, 2012-07-06 (All day)', and '12th International Coral Reef Symposium'.

Thank you for your attention!



Lateral view of *Epimetopus mendeli* from Peru, which is currently under description as new for science in a collaborative paper by the scientists of National Museum in Prague and the Museum of Natural History in London.

„Open access to relevant metadata is the base for semantic linking of scientific content with all other domains“

Data Access Rights

In OpenUp! access to metadata is based on CC by license (according to EU Grant Agreement 270890)

IPR issues related to the project outcome including IPR clearing methods for content and tools (output)

All specialised software tools that will be required to operate the access system are or will be available under open source licenses and on public websites:

GBIF makes its software available under the Apache License, Version 2.0 (www.gbif.org)

BioCASE uses a Mozilla Public License. (www.biocase.org)

The ScratchPad website is based on the Drupal CMS, which operates under the the GNU General Public License. (www.scratchpads.eu) If AIT's contribution to the project is based on own developed proprietary components

that are not yet in the open domain, these will be re-factored and delivered under an open source license.

The GeoNames server considered as an initial tool for geographic name checking (needed mainly for the locality where the natural history objects comes from, not were it is housed today) operates under a Creative Commons by license for both, data and web services (REST, JSON). (www.geonames.org)

The textual metadata as well as thumbnail images or soundbits provided for EUROPEANA will be under open access without restrictions (Creative Commons cc by license).



On January 1st, 2012, the new **Europeana Data Exchange Agreement (DEA)** came into effect and replaced all existing agreements between Europeana and its content providers and aggregators. Europeana now established CC0 as standard.



Europeana
foundation

www.europeana.eu

Koninklijke Bibliotheek
Prins Willem-Alexanderhof 5
Postbus 90407
2509 LK Den Haag
Nederland

Europeana Data Exchange Agreement

Way of proceeding for OpenUp! partners

- (i) **All OpenUp! content providers will continue to provide their full GBIF/BioCASE metadata to the OpenUp! meta-database hosted by the NHM in London.** So the full metadata can still be made available to other portals that fulfil the provider's licensing terms - e.g. digital libraries on the national level, GBIF and BioCASE.
- (ii) All records of those **providers who want to restrict data provision will be flagged** in the OpenUp! meta-database. Data provision to Europeana will then be restricted
- (iii) a formal statement will be provided that metadata provision to Europeana will be restricted for data from those institutions that wish to do so.



Attribution

CC BY

This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.



Attribution-NonCommercial

CC BY-NC

This license lets others remix, tweak, and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they don't have to license their derivative works on the same terms.

About CC0 — “No Rights Reserved”

CC0 enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law.

In contrast to CC’s licenses that allow copyright holders to choose from a range of permissions while retaining their copyright, CC0 empowers yet another choice altogether – the choice to opt out of copyright and database protection, and the exclusive rights automatically granted to creators – the “no rights reserved” alternative to our licenses.

„Eine gute Rede hat einen guten Anfang und ein gutes Ende – und beide sollten möglichst dicht beieinander liegen.“

(Mark Twain)