

COMPONENT REPORT

Project Acronym: OpenUp!
Grant Agreement No: 270890
Project Title: Opening up the Natural History Heritage for Europeana

C7.3.1 - List of criteria and benefits to become OpenUp! associated partners established

Revision: 4a [Final]

Authors:

Hanna Koivula UH
Boris Jacob MRAC
Anne-Sophie Archambeau MNHN
Patricia Mergen MRAC

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	X
C	Confidential, only for members of the consortium and the Commission Services	

0 REVISION AND DISTRIBUTION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision	Date	Author	Organisation	Description
Draft 1	2012-01-30	H. Koivula, B. Jacob, A. Archambeau	UH, MRAC, MNHN	Draft 1 incorporating ideas from WP7 e-meeting.
Draft 2	2012-02-20	H. Koivula	UH	Draft 2 to receive input from WP7 lead.
Draft 2	2012-02-20	P. Mergen	MRAC	Proof reading and comments.
Draft 3	2012-03-08	H. Koivula	UH	Incorporating feedback from the TMG and from the OpenUp! Annual Meeting.
Draft 4	2012-04-24	B. Jacob, H. Koivula	MRAC, HU	Incorporating ideas from BHL-Europe, Copy editing.
Draft 4	2012-04-24	P. Mergen	MRAC	Proof reading and comments.
Draft 4a	2012-04-25	Coordination Team	BGBM	Minor editing

Statement of Originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Distribution

Recipient	Date	Version	Accepted YES/NO
Work Package Leader (P. Mergen, WP7)	2012-02-20	Draft 2	N.A.
Technology Management Group (TMG)	2012-03-08	Draft 3	N.A.
Work Package Leader (P. Mergen, WP7)	2012-04-24	Draft 4	YES
Project Coordinator (W. Berendsohn, BGBM)	2012-04-25	Draft 4a	YES

TABLE OF CONTENTS

0	REVISION AND DISTRIBUTION HISTORY AND STATEMENT OF ORIGINALITY	1
1	INTRODUCTION	3
2	BENEFITS TO BECOME OPENUP! ASSOCIATED PARTNER	4
2.1	Visibility of your institution, Use of your data, New audiences	4
2.2	Collaboration and community support	4
2.3	Quality control and enhancement of your data	5
2.4	Use of existing technology and protocols, Sustainability, Viability	5
3	CRITERIA TO BECOME OPENUP! ASSOCIATED PARTNER.....	6
3.1	Identification of your natural history data sources	6
3.2	Using the BioCAsE provider software.....	7
3.3	Mapping your data to GBIF and Europeana standards	7
4	OUTLOOK	7
5	LIST OF REFERENCES	7

1 INTRODUCTION

The natural history museums and botanical gardens participating in the OpenUp! project hold a substantial amount of digital objects and related scientific data and metadata that will be provided to the general public via Europe's digital museum, library and archive – Europeana – and to scientific audiences via the Global Biodiversity Information Facility (GBIF).

23 European partners joined together in the OpenUp! project to make this connection between the natural sciences and cultural sector by using existing technologies and standards like the BioCASE provider software, the Access to Biological Collections Data schema (ABCD), an official standard from the Biodiversity Information Standards (TDWG)¹ organisation, and the Europeana schemas 'European Semantic Elements' (ESE), and later 'Europeana Data Model' (EDM).

Currently GBIF provides access to more than 230 million records, more than 60 million of which are concerning objects in natural history collections. Mostly this data is either information from the European territory or has a worldwide scope but is kept in European repositories. Europeana mostly holds objects from the cultural history domain and objects from the natural history domain are dramatically underrepresented, so the OpenUp! project provides a chance to connect these communities.

In 2007, in response to specific questions about their data mobilisation capacity, 33 GBIF Nodes reported that they hold the potential of publishing over 1.325 billion data records. Of these, 204 million (14%) are already in digital format, but are not published through the GBIF network. The remaining 1121 million (77%) records are known to these Participants, but need to be digitised. At this point in time only 136 million records were accessible through the GBIF network.²

However, only a part of the European natural history institutions are partner in the OpenUp! project. There are far more that provide data to GBIF, and even more than that who do neither provide data to GBIF, nor Europeana. It is thus clear, that still a huge amount of our natural history heritage is not moving towards accessibility through one or both of the biggest and most successful digital repositories in the natural sciences and cultural domain, GBIF and Europeana.

One of the aims of the OpenUp! project is to raise awareness of this fact and to help possible new content providers in leveraging their data and making it available via the technologies, standards and workflows that are used and adapted in the OpenUp! project. This will be done in two ways, first via the networks of GBIF, BioCASE, and CETAF (the Consortium of European Taxonomic Facilities). Seven OpenUp! partners are also national GBIF nodes and thus play an important role in information and liaising with possible new partners. CETAF is another important body for outreach. It has 33 members of whom 17 are partners of OpenUp!, and the BioCASE Secretariat at the institution coordinating OpenUp! (BGBM) functions as the GBIF Participant Node for CETAF. The second approach towards increased mobilisation will be via contacting individual institutions as well as projects outside these networks.

Many of these potential content providers also have digital content that is not yet published. Independent from GBIF, the OpenUp! partners have their own local data infrastructure, long term existence and well established networks both nationally and internationally. By providing assistance, participating in the

¹ See: <http://www.tdwg.org/activities/>

² See Chavan et. al 2010: State-of-the-Network 2010: Discovery and Publishing of the Primary Biodiversity Data through the GBIF Network (http://www.gbif.org/orc/?doc_id=2960)

helpdesk facilities and promotional communication, all partners will play a major role in extending the network and mobilizing more content in the appropriate formats for Europeana. This list of criteria and benefits to become an OpenUp! Associated Partner serves as a mean to disseminate this idea and to attract possible new content providers.

2 BENEFITS TO BECOME OPENUP! ASSOCIATED PARTNER

2.1 Visibility of your institution, Use of your data, New audiences

- OpenUp! increases the visibility of your content, and thus increases the visibility of your institution, both to the general public and to scientific audiences.
- OpenUp! makes content available to the general public via Europeana – Europe’s digital museum, archive and library at www.europeana.org.
- Natural history content is not adequately presented in Europeana. OpenUp! will increase the visibility of this content, and it will also increase the visibility of the content providers.
- OpenUp! facilitates content provision to the Global Biodiversity Information Facility at www.gbif.org, the unique world-wide infrastructure for our data, reaching out to scientists all over the world.
- OpenUp! is repatriating biodiversity knowledge by sharing data that have been produced from collections taken from other countries.
- OpenUp! and Europeana dissemination activities will reach a large group of new target users to increase the use of all data and content. These activities will lead to increased awareness about the content, which is in the direct interest of the content providers, who themselves pursue constant promotional activities to attract the public to their institution’s collections and exhibitions.
- Access to multimedia data of voucher material held in natural history museums will also support monitoring programs for nature conservation. Significant parts of monitoring programs are currently carried out by crowd sourcing. Direct access to the voucher material can help to increase the skills in determining biological objects during field work or the content can be used teaching these skills.

2.2 Collaboration and community support

- Increased visibility and interoperability through joining the consortium will enhance organizations own dissemination of data, in-house services and use of standards. Documentation and adaptation of these procedures offers partners functional tools and materials that can be used in funding applications. For all of these aspects (technical, workflow, IPR), partners of OpenUp! can rely on a number of existing networks and helpdesks, with the OpenUp! Helpdesk acting as an aggregator and editor of relevant information.
- The OpenUp! Helpdesk facility offers data providers and users access to documentation, online tutorials and trainings. It will provide access to existing and newly produced documentation and tutorials for the services and assist institutions in becoming an associated partner in the OpenUp! Consortium.

- The OpenUp! Helpdesk collaborates with the GBIF Helpdesk in supporting content providers in mobilizing and publishing information about the occurrence of organisms to GBIF (<http://www.gbif.org>).
- The BioCAsE Helpdesk and the OpenUp! Helpdesk are closely working together in supporting partners to set up and configure the BioCAsE provider software (<http://www.biocase.org/>) and map the data to ABCD. The BioCAsE provider software is used as a technical solution to simultaneously provide data to GBIF and enriched metadata and multimedia content to Europeana.
- OpenUp! partners are strongly involved in the Biodiversity Information Standards (TDWG) organization, disseminating standards such as ABCD and enhancing them on an international level. OpenUp! relies on commonly accepted standards and aims at wider institutionalising of these data exchange standards for better interoperability and data quality. (<http://www.tdwg.org/>).
- OpenUp! is also well represented in CETAF, the Consortium of European Taxonomic Facilities, which promotes research in biology and paleobiology as well as cooperation among the European natural history institutions.
- OpenUp! also has close relations with the Europeana Network (formerly the CCPA) which is an open forum for experts across Europe, for content providers and aggregators and for providers of technical, legal and strategic knowledge (<http://pro.europeana.eu/web/guest/about/network>).

2.3 *Quality control and enhancement of your data*

- OpenUp! uses and develops tools to check your data via data quality and integrity services.
- OpenUp! sets up tools with which you can enrich your metadata so that every user, but also every content provider can retrieve and reuse the enhanced data.
- OpenUp! provides you with access to ‘Taxonomic Intelligence’ tools, and other means for data enhancement to facilitate the search of natural history content.
- OpenUp! provides multilingual access to your content, including multilingual common names.
- A multilingual vocabulary of organism names, linked together by a backbone of scientific names can be used for cross-linking objects from a variety of domains.

2.4 *Use of existing technology and protocols, Sustainability, Viability*

- OpenUp! leverages the content to be shown in the existing portals of GBIF (www.gbif.org), Europeana (www.europeana.org) and BioCAsE (<http://search.biocase.org/europe/>) and thus provides tools and technologies to cost-effectively present the digital content.
- OpenUp! builds upon the existing workflows and infrastructure from GBIF, BioCAsE and Europeana, ensuring sustainability and interoperability via international commitment also after the end of the project.

- OpenUp! builds upon the existing data standards from TDWG (ABCD³) and Europeana (ESE/EDM). Both standards are supported by large communities with good support mechanisms.
- The use of these standards ensures that the data of all OpenUp! partners will become interoperable, not only with Europeana but also among the different partner institutions.
- Using the OAI Standard and the and a common metadata scheme that maps to Dublin Core (OAI basic requirement) as well as the BioCASE protocol that is firmly established in the natural history community ensures long-term sustainability of the technical infrastructure.
- The OpenUp! Helpdesk is and will stay a sustainable source of information and documentation after the end of the project.
- The project's data quality tools will be available on the OpenUp! site for at least 5 years after the project ends. Equally, the central metadata access point for OpenUp! has been guaranteed maintenance and a full caching of the multimedia objects will also be offered at their storage facility for institutions that prefer that option.

3 CRITERIA TO BECOME OPENUP! ASSOCIATED PARTNER

3.1 Identification of your natural history data sources

Associated partners have natural history data connected to multimedia content available. There is a broad selection of types and formats of content, that OpenUp! will provide to GBIF and Europeana. Here are a few examples of the more than 1,000,000 natural history objects:

- The Botanical Garden and Botanical Museum in Berlin-Dahlem has already put 100,766 scans of botanical specimen⁴ online to Europeana.
- You can also see 3,768 JPEG photographs of zoological specimen⁵ from the University of Helsinki in Europeana and it is planned to also provide scans of historical accession books from their digital repository⁶.
- More than 10,000 natural history artwork images from zoology and botany will be made available by the Natural History Museum in London.
- The Natural History Museum in Berlin will provide several thousand animal sound files from their Tierstimmenarchiv to OpenUp!⁷
- ETI will provide zoological movies in Quicktime format and digitized species drawings.

³ See: <http://www.tdwg.org/standards/115/>

⁴ See:

http://www.europeana.eu/portal/search.html?query=europeana_dataProvider%3A%22Botanic+Garden+%5Cand+Botanical+Museum+Berlin-Dahlem%22

⁵ See:

http://www.europeana.eu/portal/search.html?start=61&query=europeana_dataProvider%3A%22Finnish+Museum+of+Natural+History%22

⁶ See: <http://digit.luomus.fi/>

⁷ See: <http://www.tierstimmenarchiv.de/>

3.2 Using the BioCASE provider software

Associated partners have the capability to install, use and maintain a BioCASE provider software and its configurations (<http://www.biocase.org/>). The only prerequisite is an institutional website where the provider software can be installed. The technical requirement for joining as an associated partner is to install and use the BioCASE provider software. That connects your existing database to make it available on the internet as a web service. The partner can use BioCASE not only for sharing data with Europeana, but simultaneously also with GBIF, if desired.

The BioCASE provider 3.0 version is used as a web service to directly access your data. It can also be used to create an XML representation of your data, storing all information published by your BioCASE web service in a single file. If required, this can be transformed into a DarwinCore Archive in a subsequent step. The DarwinCore-Archive format facilitates the harvesting and indexing of the data by data aggregators such as GBIF.

3.3 Mapping your data to GBIF and Europeana standards

Multimedia content has to have basic information attached that needs to be mapped to the ABCD and ESE data exchange standards. Data and content provided via the OpenUp! project has to be compliant with TDWG's ABCD standard, and Europeana's ESE/EDM standard. For example, the licensing terms under which the multimedia content is made available have to be specified; several models for this are in place. The textual data provided to Europeana will be made available under the CC0 Creative Commons Public Domain Dedication (<http://creativecommons.org/publicdomain/zero/1.0/>). Documentation and help in this regard is provided by the OpenUp! Helpdesk.

4 OUTLOOK

One major milestone for the third project year, which we are already start preparing now, is an outreach campaign to attract additional associated partners. The findings of this component directly support the tasks of work package 7 in attracting new partners. The OpenUp! project aims to extend the participation in content provision to 10 providers beyond the original consortium. But we also like to strengthen the natural history and biodiversity informatics domain by providing support and help via the OpenUp! Helpdesk and our partners at GBIF, BioCASE, TDWG, CETAF, and Europeana. In that sense we welcome content contribution but also contribution in expertise.

5 LIST OF REFERENCES

BHL-Europe in a nutshell URL: <http://www.bhl-europe.eu/en/information-for-partners>

BioCASE network URL: <http://www.biocase.org/>

Biodiversity Information Standards (TDWG) URL: <http://www.tdwg.org/> and <http://www.tdwg.org/standards/>

Consortium of European Taxonomic Facilities (CETAF) URL: <http://www.cetaf.org>

Europeana network URL: <http://pro.europeana.eu/web/guest/about/network>

GBIF. 2010. State-of-the-Network 2010: Discovery and Publishing of the Primary Biodiversity Data through the GBIF Network. Authored by Chavan, V. S., Gaiji, S., Hahn, A., Sood, R. K., Raymond, M., and N. King. 2010. Copenhagen: Global Biodiversity Information Facility, 36 pp. ISBN: 87-92020-13-5. Accessible online at http://www.gbif.org/orc/?doc_id=2960

Global Biodiversity Information Facility (GBIF) URL: <http://www.gbif.org>

OpenUp! Description of Work and Annex