

# OpenUp! Opening up the natural history heritage for Europeana

## Background

### (not part of the OpenUp! Memorandum of Understanding)

*OpenUp!* is an international collaboration of some of the most outstanding natural science museums, botanical gardens, and university collections in Europe. Many of these institutions also represent nodes in the network of the *Global Biodiversity Information Facility (GBIF)* and/or are members of the *Consortium of European Taxonomic Facilities (CETAF)*.

The aim of the attached Memorandum of Understanding is to set the principles for upholding the data pipeline from natural history collections to the European digital library, archive and museum *Europeana* and other digital libraries, and, in parallel, to *GBIF* and the *Biological Collection Access Service for Europe (BioCASE)* after the end of the EU-funded *OpenUp!* project period (2011-2014)<sup>1</sup>.

By the end of the project period, *OpenUp!* had made accessible more than 1.6 million multimedia objects from the natural history domain through *Europeana*. Its unique standardised infrastructure remained functional and effective in the post-project phase, so that *OpenUp!* now serves as the European aggregator for digital content from the wider natural history domain for *Europeana* and potentially other digital libraries.

For this purpose, we encourage European institutions and organisations that collect and hold content on natural history and are willing and entitled to make this content available through the *OpenUp!* infrastructure, to join this MoU and become a participant in the *OpenUp!* network.

The software and standards *OpenUp!* uses are based on the established technical infrastructure of the *Global Biodiversity Information Facility (GBIF)*, in particular *BioCASE (Biological Collection Access System for Europe)* and *IPT (Integrated Publishing Toolkit)*.

- *OpenUp!* has set up mappings between the community data standards ABCD (Access to Biological Collection Data)<sup>2</sup> and DwC (Darwin Core)<sup>2</sup> and the *Europeana* metadata scheme EDM (Europeana Data Model).
- In collaboration with other projects, *OpenUp!* has extended the ABCD schema to better accommodate information on paleontological, geological, and mineralogical collection objects. The schema Access to Biological Collection Data Extension For Geosciences (ABCDEFG) is used for the data provision to the Geosciences Collection Access Service (GeoCASE), *GBIF* and *OpenUp!*.
- *OpenUp!* created the *BioCASE Monitor Service (BMS)*<sup>3</sup> to monitor individual content provider's progress in data provision, facilitate checks of data mappings, and verify compliance of provided data for transformation into other target schemas.
- *OpenUp!* set up the Natural History aggregator, a common infrastructure inside the *OpenUp!* architecture to combine ABCD as well as DwC records from content providers. The aggregator engages the *Berlin Harvesting and Indexing Toolkit (B-HIT)*<sup>4</sup> and pushes the records to a common metadata database.

---

<sup>1</sup> OpenUp! was a 3-year CIP-Best Practice Network co-funded by participants and the EU as part of the Information and Communications Technologies Policy Support Programme (ICT PSP) of the European Commission (CIP-ICT PSP Call 4 – 270890).

<sup>2</sup> ABCD and DwC are standards recognised by the organisation Biodiversity Information Standards (TDWG).

<sup>3</sup> Wiczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., De Giovanni, R., Robertson, T. & Vieglais, D. 2012: Darwin Core: An Evolving Community-developed Biodiversity Data Standard. PLoS ONE 7 (1). doi:10.1371/journal.pone.0029715. PMC 3253084. PMID 22238640.

<sup>4</sup> B-HIT is open-source software used for automated harvesting of biodiversity data. It is based on the GBIF HIT. Reference: Kelbert P, Droege G, Barker K, Braak K, Cawsey EM, Coddington J, et al. (2015) B-HIT - A Tool for Harvesting and Indexing Biodiversity Data. PLoS ONE 10(11): e0142240. doi:10.1371/journal.pone.0142240

- *OpenUp!* has complemented the aggregation process with a system for data transformation and semantic enrichment; the latter to the effect that the records provided by the collections can be found in *Europeana* using synonyms and common names of species in various languages as well as alternative geographic area names.
- *OpenUp!* supports new providers in specifying their collection metadata with respect to intellectual property rights<sup>5</sup>.

Several *OpenUp!* participants have committed institutional resources to sustain specific parts of this infrastructure beyond the project period:

- Data-providing institutions maintain their content databases and thus provide a continuously increasing amount of data to the *Europeana*, *GBIF*, and *BioCASE* networks.
- The *OpenUp!* aggregator database is hosted by AIT (Angewandte Informationstechnik Forschungsgesellschaft mbH Graz, Austria) and mirrored at the Botanic Garden and Botanical Museum Berlin (BGBM).
- The open-source and system-independent *BioCASE Provider Software* is maintained and constantly improved by staff at the BGBM Berlin dedicated to that task. In addition, further development will be taking place in ongoing and future projects. The core development and maintenance of the *Integrated Publishing Toolkit (IPT)* is done by the *GBIF* Secretariat, while coding, documentation, and internationalisation of the open source software tool are a community effort.
- The *BioCASE Monitoring Service* is maintained by the Museum für Naturkunde in Berlin (MfN).
- For semantic enrichment, the Royal Botanic Gardens Kew (RBG Kew) provide botanical name services including synonyms. Naturalis (NCBN Leiden) supports the Catalogue of Life, which is used for zoological name enrichment. The Naturhistorisches Museum Wien (NHM Vienna) provides the common name service.
- The *BioCASE* helpdesk at BGBM provides online support for data-holding institutions and organisations.
- For the process of data ingestion and transformation to *Europeana* data standards and semantic enrichment, no free-of-charge solution has been encountered, largely because this is outside the usual area of responsibility of natural history museums. The partner carrying out that step in the *OpenUp!* network is the company AIT (Angewandte Informationstechnik Forschungsgesellschaft mbH Graz, Austria). AIT is engaged in a wide array of *Europeana*-related projects and carries out this function in an excellent and highly efficient manner. AIT offers their work as a contractor by concluding an individual Service Level Agreement (SLA) with the data provider. The calculation of the individual annual fee is based on a scheme that takes into consideration the number of ingestions per year, the total number of records ingested, and the number of static or dynamic data sources to be processed. The SLA can be reviewed annually, so there is an opening for including these costs into future projects and to adapt to changing conditions at the data provider's side.

The attached Memorandum of Understanding provides the framework for the *OpenUp!* network. Together with the preceding introduction it serves to inform potential new data providers about the aims and achievements of the *OpenUp!* network.

---

<sup>5</sup> More information about the application of rights statements can be found in the *Europeana* Publishing Framework, under [www.rightsstatements.org](http://www.rightsstatements.org) and on <http://pro.europeana.eu/share-your-data>.



# Memorandum of Understanding

## PREAMBLE

The signatories to this Memorandum of Understanding are institutions and organisations that collect and hold data on natural history Content and are willing and entitled to make these data available through the *OpenUp!* infrastructure to the European digital library, archive and museum *Europeana*, other European digital libraries, and the networks of the *Global Biodiversity Information Facility (GBIF)* and the *Biological Collection Access Service for Europe (BioCASE)*.

## Paragraph 1: Definitions

**Aggregator:** An aggregator in the context of *Europeana* is an organisation that collects metadata from a group of content providers and transmits it to *Europeana*.

**BioCASE:** Biological Collection Access Service for Europe. A network originated with EU-funded projects, dedicated to providing access to natural history data from European collections. BioCASE functions as the GBIF Participant Node for CETAF.

**CC0 1.0 Universal Public Domain Dedication:** The Creative Commons Universal Public Domain Dedication as published at: <http://creativecommons.org/publicdomain/zero/1.0/>.

**Content:** A physical or digital object that is part of Europe's cultural and/or scientific heritage, typically held by the Content Provider.

**Content Provider:** An institution or organisation sharing its multimedia collection online.

**Europeana:** The *Europeana* website, its data and its machine interfaces, all without necessarily being disclosed under URL [europeana.eu](http://europeana.eu).

**Europeana Licensing Framework:** The *Europeana* Licensing Framework helps govern the relationships of *Europeana*, its data providers and its users. The licences frame what can and cannot be done by human and machine agents with the Metadata and the digital Content it accesses.

**Europeana Data Exchange Agreement (DEA):** The DEA is the central element of the *Europeana* Licensing Framework. It structures the relationship of *Europeana* and its data providers. It establishes rules for updating and deleting metadata stored by *Europeana* and deals with issues such as liability and removal of data at the request of third parties. Participants of the *OpenUp!* network need to sign the DEA with *Europeana*.

**GBIF:** The Global Biodiversity Information Facility is an international open data infrastructure, funded by governments, providing access to biodiversity data over the Internet.

**Metadata:** Textual information (including hyperlinks) that may serve to identify, discover, interpret and/or manage Content.

**Participant:** A member of the *OpenUp!* network, i.e. an institution or organisation that is a signatory to this MoU.

**Party:** The institution signing this document.

**Preview:** A reduced size or length audio and/or visual representation of Content, in the form of one or more images, text files, audio files and/or moving image files.

**Representative of the Participant:** A person authorised by the institution or organisation to sign this MoU.

**Service Level Agreement (SLA):** An agreement between a customer and a service provider formally defining the service to be provided.

## Paragraph 2: Understandings

OpenUp! is a collaboration of European natural history institutions and organisations with the objective of giving access to the world's natural sciences heritage via *Europeana*, other European digital libraries, and the GBIF and BioCASE data networks.

The Participants' involvement in this MoU is subject to the goodwill and effective contribution to OpenUp! activities and operations, either financially or through any of the activities described in Paragraph 4 "Involvements of the Participants", and is subject to the laws and regulations applicable to the individual Participants.

This document is not legally binding and has no effect as a legal precedent.

## Paragraph 3: Objectives

The purpose of this MoU is to agree on principles for sustaining a continuous flow of multimedia data from European natural history institutions and organisations to *Europeana*, the European digital library and museum ([www.europeana.eu](http://www.europeana.eu)), to other European digital libraries, and to the *GBIF* and *BioCASE* networks. Digital libraries such as *Europeana*, with their main orientation towards culture, have a much broader audience than the average natural history museum site, and their significance, therefore, lies in the connection of natural history content and institutional websites with people who would otherwise not be reached. *GBIF* and *BioCASE* in turn are mainly directed to the research community, which represents the main original audience of natural history content.

## Paragraph 4: Involvement of the Participants

*OpenUp!* is based on digital Content provision of the Participants using the established technical infrastructure of the *Global Biodiversity Information Facility (GBIF)* including the *BioCASE* network (*Biological Collection Access System for Europe*). Sustaining the *OpenUp!* data pipeline rests with the Participants. To become a Participant, the Party agrees to:

- Communicate its participation and the availability of its content to other natural science museums, universities and the general public and inform them about this unique opportunity to explore the world's natural heritage online.
- Mobilise and prepare multimedia data from its natural history collections for online publication via *Europeana* and other digital libraries, and the *GBIF* and *BioCASE* networks.
- Install the *BioCASE* provider software which is an open-source, system-independent software module developed to enable collection database holders to link their collections to international data networks without having to modify the implementation of their systems. Once set up, the *BioCASE* provider software automatically picks up new records in institutional databases and makes them

accessible to the networks. As an alternative *GBIF's Integrated Publishing Toolkit (IPT)* can be used.

- Map its data to one of the two accepted biodiversity data standards (ABCD or DwC). ABCD contains a comprehensive metadata section for individual objects as well as entire datasets and covers the metadata elements presently required by *Europeana*.

## **Paragraph 5: Intellectual Property**

All participants allow access to their multimedia content through the *OpenUp!* technical infrastructure. They will thus provide that content via the *BioCASE* provider software or the *GBIF IPT* to *Europeana* and other European digital libraries, and to the *GBIF* and *BioCASE* networks. Accessing Content and sharing information has intellectual property (IP) implications, which have been considered by the Party's policy advisors.

The Party hereby certifies that the textual metadata provided with the *OpenUp!* Aggregator to *Europeana* will be under the terms of the Creative Commons Universal Public Domain Dedication (CC0 Universal), i.e. free from any restrictions on re-use. A minimum set of textual metadata has to be provided including the organism name, the URL that links to the multimedia content to which the metadata refers, the kind of the underlying object (usually a specimen), a rights statement regarding the linked multimedia object, and the set of institutional and technical metadata.

All other Content, such as previews, soundbits, and the Content referenced by URI (web addresses) in the Metadata remain completely under the control of the Content Provider. This is not to exclude the possibility that the Content Provider makes this Content available under the CC0 Universal Public Domain Dedication.

## **Paragraph 6: Sustainability**

Several institutions have agreed to support and provide specific functionalities needed for the *OpenUp!* data pipeline. *Europeana* itself supports the harvesting of the data from the *OpenUp!* aggregator database. Before harvesting, however, the data are ingested and transformed into the *Europeana* data standard. This also includes the process of semantic enrichment. In contrast to the other steps in the *OpenUp!* workflow, no sustainable institutionalised free-of-charge solution has been encountered. For the time being these costs have to be covered by the Data Providers or an organisation representing them. For this purpose, the Data Provider will enter into a Service Level Agreement with a service provider capable of carrying out data harvesting and transformation.

## Paragraph 7: Other matters and signature page

### **OpenUp! Secretariat**

The institution designated by *CETAF* as host for *CETAF*'s participant node for *GBIF* annually designates the host of the *OpenUp!* Secretariat among the *OpenUp!* Participants. The *OpenUp!* Secretariat calls upon Participants with respect to matters of liaison, of providing guidance and direction in daily operational matters, and of general coordination. The *OpenUp!* Secretariat will name a contact person and provide an email address for contact.

### **List of Participants**

A list of all current Participants will be maintained by the *OpenUp!* Secretariat and published on a website.

### **Association and Disassociation of Participants**

Association with this MoU is open to all institutions or organisations that may assist *OpenUp!* in achieving its goals as outlined in this MoU. Such association becomes effective upon signature of this MoU, submission of the signed MoU to the *OpenUp!* Secretariat and listing of the new Participant in the List of Participants.

Any Participant may remove itself from this MoU by advising the *OpenUp!* Secretariat in writing of its intention to do so and providing an effective termination date.

### **Not a Legal Partnership**

The Participants do not intend to create a legal partnership, principal/agent, or joint venture relationship, and nothing in this MoU shall be construed as creating such a relationship between the Participants. Neither Participant may incur any obligation on behalf of the others.

### **Non-exclusive Agreement**

This MoU is non-exclusive, and in no way restricts Participants from participating in similar activities or arrangements with other public or private initiatives, organisations, or individuals. Each Participant is wholly responsible for its Content and data provision to the *Europeana*, *GBIF*, and *BioCASE* networks.

For

Place, Date

Signature