Content

Content …… 2
Editorial …… 3
Growing of OpenUp! content on Europeana Portal …… 4
OpenUp! Social Media New Face and Statistics …… 4
News from the Biology Center in Linz, Upper Austria (Land OOE) …… 5
Meetings and Conferences …… 6
Celebrating OpenUp! at the 1st Annual Meeting 2012 in Paris …… 9
BLE as a Part of OpenUp! and Virtual Bridge Between Projects, Natural History Content and Knowledge …… 11
Web-incompatible Image Formats in Europeana …… 14
Survey on Usability of OpenUp! Content in Europeana in the Educational Sector …… 16
Flowing Smoothly into Europeana: the OpenUp! Technical Workflow …… 16

www.open-up.eu
Editorial

Enjoy the third issue of the OpenUp! Newsletter. We have successfully completed the first project year and delivered around 775,000 items to Europeana by the end of November 2012. This great achievement was reached by careful attention to many tasks. I have not enough space to list them all, but as just a few examples, I mention all partners preparing their databases by installing BioCASe, modification of ABCD mapping with EFG extension to incorporate mineralogical and palaeontological data, running the Collection Data Quality Toolkit, daily solving numerous technical issues, disseminating the results of the project at many conferences, installing and running the Helpdesk and last but not least, day to day administrative work on behalf of the project. This issue of the Newsletter is about all these activities.

That we are doing well is not only our own impression, but also the evaluators of the first year stated that:

…OpenUp! is a high-profile project that embodies the very nature of a Best Practice Network, and aims to integrate natural history content within the European digital library Europeana. It not only addresses an existing gap through content provision but also creates an infrastructure and develops specific tools for domain-specific content enrichment. The accomplishments of the project in its first year have been excellent and clearly all critical objectives have been reached…

Our gratitude and best wishes for continuing success during the next year of the project go to all project partners.

Dr. Kamil Zágoršek, National Museum, Prague
Growing OpenUp! Content on the Europeana Portal

J. Frank (NM)

Prepared content from the OpenUp! project is regularly harvested by Europeana. Until the end of November 2012, 765,949 multimedia objects from OpenUp! content provider collections appeared on the Europeana portal. This amount includes 591,000 images and 9,020 sound samples of fourteen content providers from the United Kingdom, Germany, Belgium, the Netherlands, Finland, the Czech Republic, Slovakia, Austria, Denmark, and Estonia. The next harvests are planned for mid-November and January. All our actual OpenUp! content in Europeana can be accessed via this link.

OpenUp! Social Media With New Face

J. Frank & L. Laibl (NM)

In concert with the growing amount of OpenUp! content in Europeana, we are increasing our presence on various social media, like the OpenUp! Blog, Facebook and Twitter. We adapted the design of these social networks and started disseminating the digital natural history collections and OpenUp! project activities. The new series “Content Highlights” was started and from October onward, will be published even more frequently. This series is introducing selected specimens from the natural history col-
lections which are highly interesting or important. Every content highlight blog including the interesting information about the specimen and the species or group which it represents. The images and links to the original source and Europeana are also included. Besides the content highlights articles we are publishing an overview of our meetings, conferences and important events, information about the project progress and project activities. Blog is the ideal platform for longer texts enriched by pictures. These articles are then shared via Facebook, which is only suitable platform for shorter text with pictures or only pictures. We are sharing on Facebook also important information and activities from related projects such as BHL, Europeana and BHL-Europe. In return our activities are shared on their social media. Blogs are shared also on the Twitter platform which is designed for short informative messages. Besides the Blog and Facebook activities OpenUp! publishes Twitter tweets from the OpenUp! Helpdesk including information about new documents and guidelines for new content providers. The increasing traffic on our social media will be also connected to our new virtual exhibition “Poisonous nature” on the BLE platform presenting OpenUp! and BHL content which will be launched in December 2012.

Short statistics:
The OpenUp! Blog has for now published twelve posts since March 2012. The OpenUp! Facebook has 99 total likes and 170 persons weekly total reach. People who like the Open Up! Facebook page are from twenty countries. The OpenUp! Twitter account has 101 tweets, 44 following and 26 followers. Since September, the Twitter account is connected with the OpenUp! Helpdesk and is posting tweets about important documents and news from the project progress.

News From the Biology Center in Linz, Upper Austria (Oberösterreichische Landesmuseen)

B. Wunder (LANDOOE)

This spring, our institution purchased a new technical device, the stereoscopic microscope Nikon AZ100M, for the digitalisation of zoological collections. Currently, digitalisation and metadata enhancement is in progress, so we will be able to provide this content to OpenUp!. The AZ100M allows us, to produce high quality images with a high depth of focus. The range of magnifications last from 5× to 500×. During the digitalisation process, two observation levels, the lowest and the highest level in which the object is barely focused (e.g. the antenna and the leg of an insect), are defined and within this range the camera takes 20–30 pictures with different focus. These pictures are combined to one single image which has therefore a very high quality with a high depth of focus. For every species, six to eight different images are created, so that the important determination characteristics are visible. Currently we are working on objects from the collection of Klaus Warncke, one of the most productive investigators of bees during the last decades. The collection, which is prepared for OpenUp! consists of type material of bee species (both male and female individuals, if available). Warncke described 57 new taxa (genus group) and 887 new taxa of the species group between 1966 and 1993. Additionally, type material from other collections will be photographed.

Figure 1 Colletes rubellus (Nosk, 1936) – Holotype, LAND OOE collections.
Figure 2 Labels to C. rubellus holotype.
Meetings and Conferences

In 2012, the OpenUp! project was presented during this year at several conferences, meetings and workshops. We list some of them below. The entire overview about where our project has been and will be presented and next is published on our project website. This link and also the events structure will be changed after the project website upgrade on the Scratchpads 2 platform. We will inform you about this upgrade on our social media and in the next newsletter.

OpenUp! project at the Museum für Naturkunde

The OpenUp! project was presented during the official inauguration of the new general director of the Museum für Naturkunde, Prof. Dr. Johannes Vogel, in Berlin 8 May 2012.

The new OpenUp! roll-up screen attracted a number of interested attendees. The ensemble was completed by a show case displaying original objects of the museum’s butterfly collection, already visible in Europeana.

Long Night of Science in Berlin

On Saturday 2 June 2012, the Long Night of Science was held in selected institutions in Berlin. Exhibitions and research facilities were opened to the public. The Museum für Naturkunde participated as well and created the most suitable place for presentation of projects as OpenUp!, BHL-Europe or GBIF. OpenUp! introduced new posters on the GBIF stand during this event. The Biodiversity Library Exhibition was also presented on the BHL-Europe stand.

The Long Night of Science at the MfN was attended by hundreds of people and we are happy to declare this event associated with the presentation of the OpenUp! project as very successful.

BHL-Europe Final Meeting, Europeana Plenary and other events

The Long Night of Science in Berlin was followed by the closely related BHL-Europe final meeting. This meeting was held in the Jerusalem Church in Berlin, 4–6 June 2012, and the first day was a symposium on the topic of the Convention on
Biological Diversity (CBD), which was adopted 20 years ago. A new OpenUp! poster was presented by Jiří Frank. During the BHL-Europe Show and Tell session, on the second day of this event, the OpenUp! project was presented by Petra Böttinger. During the afternoon session, Jiří Frank presented the Biodiversity Library Exhibition (BLE) and its relation to BHL-Europe, OpenUp! and Europeana. You can find more information about this event on the BHL-Europe blog here.

In parallel with the Long Night of Science, OpenUp! was introduced during the 11th Larwood meeting in Brno, Czech Republic by Kamil Zágoršek. From 31 May to 2 June 2012, the Department of Geological Sciences at Masaryk University in Brno (Czech Republic) hosted the 11th Larwood meeting. 29 colleagues from 25 European countries and Bra-
zil took part in the meeting. Kamil Zágoršek introduced the OpenUp! project explaining the ideas behind the project, describing content delivered to Europeana and outlining prospective usage of this content by the scientific community.

From 13 to 15 June the Europeana Plenary event was held in Leuven, Belgium, which was a great opportunity to discuss and share ideas the staff of Europeana. Kamil Zágoršek presented the OpenUp! during the Open Mike session. During the Europeana Plenary, the Hack4Europe 2012 was organised and its preview of prototypes is published [here](#). For more details about this event and programme, please visit the [conference website](#) and [Europeana blog](#).

**Europeana workshop in Prague**

The OpenUp! project was presented by Kamil Zágoršek and Jiří Frank during the Europeana workshop in Prague, which was organised by the National Museum, Prague for new Europeana content providers. This workshop was the first of three workshops, with the aim of presenting project results to different audiences. The second workshop, which was focused on global dissemination of natural history content and knowledge was organised during the TDWG conference in Beijing in October 2012. The third workshop will be organised during the OpenUp! Annual Meeting in Prague in May 2013.

**Digital Fossil Berlin**

Jana Hoffmann presented the OpenUp! project on 24 September during the International conference on fossil digitizing and digital collection data handling which was organised in Museum für Naturkunde Berlin.

**2012 TDWG annual conference**

The OpenUp! project, content and technical workflow were presented at the TDWG (Biodiversity Information Standards) annual conference on October 22–26 in Beijing by Jana Hoffmann and Jiří Frank. Jiří Frank also organised the Global natural history dissemination and Biodiversity Library Exhibition workshop during this conference. We also post updates about these events on our social media, project website and in the next newsletter.
OpenUp! celebrated its first successful project year at the 1st Annual Meeting which was held at the Muséum national d’Histoire naturelle in Paris. It was a great opportunity for almost 60 project participants from almost all 23 partners institutions/organizations and guests to meet and discuss the project face to face.

On the first day, the Outreach and Dissemination Group, the Technology Management Group and the Steering Committee each met and discussed tasks fulfilled during the preceding year, and laid future steps according to the project plan. An update was given on technical aspects of content provision during the OpenUp! Tutorial 1. At the end of the day, there was a conference dinner in the beautiful restaurant „Le Train Blue“ located in the famous Lyon train station.

The second day of the meeting was held in the beautiful wooden Amphithéâtre de Paléontologie. It started with a short Content Providers Meeting followed by the 2nd OpenUp! Project Assembly Meeting. Presentations were structured into thematic groups like technology, content provision, dissemination, and administrative issues. Additional presentations from partners covered virtual exhibitions and cooperation with related projects like BHL-Europe or were given by guest speakers from Natural Europe and Europeana. All presentations can be found at [http://open-up.eu/content/openup-annual-meeting-2012-paris-project-assembly-presentations](http://open-up.eu/content/openup-annual-meeting-2012-paris-project-assembly-presentations).

The Review Meeting with the EC Project Officer and Evaluation Panel was held on the third day. Feedback from the reviewers in the meeting was generally positive and according to the 1st Technical Review Report, OpenUp! made “excellent progress (the project has fully achieved its objectives and goals for the period and has even exceeded expectations).”

The work progress has indeed been completely according to plan. Following is a list with the main results achieved in OpenUp! during its first project year:

- The first 200,000 multimedia objects were delivered to Europeana via the OpenUp! infrastructure at the end of year 1.
- An Outreach and Dissemination Group (ODG) was constituted and templates for promotional materials including the project logo were designed. The OpenUp! communication site was set up using Scratchpad (Drupal content management system). A Dissemination and Publication Plan was submitted and the first 6-monthly OpenUp! Newsletter and two issues of the ODG News were published on the website.
- The mechanisms for Content Provision and
Metadata Enhancement are in place. Preparatory work started at most of the content providers, including identification of additional existing resources, drawing together or harvesting the files and datasets, and defining needs for data transformations, internationalisation, and quality improvements. A Training Workshop for BioCASe providers was organised in Berlin and attended by 17 participants from 14 partner institutions, and a Workshop on Organizing Content Provision for OpenUp! was held on short notice to support collaboration.

- The Technology Management Group coordinated the implementation of the technical infrastructure. All technical developments proceed according to plan:
  - The production versions of the Harvesting and Transformation Component and the OAI-PMH Interface were delivered.
  - The initial OpenUp! metadatabase was set up at the Natural History Museum in London.
  - Mappings of the ABCD data elements to the Europeana Semantic Elements (ESE) were specified and implemented in the Harvesting and Transformation Component. End of February, the new EDM (Europeana Data Model) definition was released and published. This version is now being analysed for its use in the future data provisions of OpenUp! to Europeana. A model for the integration of content from the areas of palaontology and mineralogy was submitted, which uses the ABCD extension for geosciences, EFG, which was also mapped into ESE. The Collection Data Quality Toolkit was specified and implemented. It is an open web-based application for OpenUp! data providers and BioCASe (Biological Collection Access Service) providers in general performing data quality checks on their data. The Data Quality Service for Zoological and Botanical Names and the Data Integrity Service were connected to the quality toolkit.
  - A prototype for metadata validation and enrichment layers was completed. A mock-up for the Common Names Service was designed and implemented. Common names lists will be provided in several different languages including Hebrew, Norwegian, Swedish, Finnish, Icelandic, Czech, Slovak, Maori, English and German.

- The OpenUp! Helpdesk was set up at http://openup.helpdesk.africamuseum.be. It features a first implementation of a question and answer dispatching system, provides the OpenUp! Guidelines v.1 and other documents from related projects and networks like GBIF, CETAF, BioCASe, Europeana and BHL-Europe, as well as information on upcoming events in the digital library, collections management and standardization domain.
- The Outreach and Dissemination Group coordinated the collaborative efforts in these areas; achievements include the setup of the OpenUp! website (http://www.open-up.eu), the dissemination and publication plan, the newsletter and promotional materials.

We are very pleased with the results of the first project year and to have received such a positive acknowledgment of our collective efforts from the Evaluation Panel. We will continue working to achieve the same level of success next year. In closing, we wish to thank everyone involved for this excellent work. We are looking forward to the next year of joint activity.
BLE as a Part of OpenUp! and Virtual Bridge Between Projects, Natural History Content and Knowledge

J. Frank (NM)

We are glad to inform you that the Biodiversity Library Exhibition (BLE) is now connected to the Europeana portal via all images on informative cards. We already introduced the BLE platform in the last newsletter in spring. Just as a reminder, BLE is a virtual exhibition focused on disseminating natural history content (literature, illustrations and images) and knowledge to the broader public via curated col-

Figure 1. BLE is now connected to the Europeana portal

Figure 2. Schema of content connection between BLE, BHL US/UK and Europeana
lections centred around a variety of themes. BLE also creates interconnections between portals/archives which provide content under the public domain or variable CC licenses.

BLE was officially launched in February as a virtual exhibition for Biodiversity Heritage Library Europe and BHL US/UK. Since then, BLE has disseminated natural history literature via two thematic topics: Spices and Expeditions, using impressive illustrations, interesting information and high quality images. These exhibitions also link to the BHL-Europe and BHL-US/UK portals. The BLE platform was developed by a team from the National Museum in Prague and developers from the IT4Care Company under the BHL-Europe project. Since end of the BHL-Europe project in April 2012, the development and refinement of BLE continue under the OpenUp! project in which the Prague team is also involved. A new BLE topic, ‘Poisonous Nature’, is in preparation and will use content from BHL (illustrations) and OpenUp! (images) that are displayed on the Europeana portal. ‘Poisonous Nature’ offers a great opportunity for OpenUp! and Europeana to use this platform and reach larger audiences, including those of BHL-Europe and BHL US/UK.

Development work on BLE is still continuing and additional platform improvements are being implemented. The new version of BLE was launched on 27 September 2012 under the same URL. At the global BHL meeting in Berlin in June 2012, project participants decided to share BLE with all possible global BHL nodes, including Australia, China, Egypt, Brazil, US and Europe (with possible participation in sub-Saharan Africa in the future), allowing these nodes to create new, interesting thematic topics, link them to variable portals including Europeana, BHL-Europe, BHL US/UK, EOL etc., and use it as a dissemination tool. Once the platform is populated with a wider variety of thematic exhibitions, the potential to reach broader audiences will rapidly grow, particularly if disseminated via global partners. BLE will serve as a central repository, allowing content providers to reach new audiences, while simultaneously being promoted by a variety of BHL nodes.

The addition of new topics and information will also expand the potential for use in the education
sector, introducing the content and knowledge in attractive ways for future generations. In order to share BLE via other partners, such as Global BHL nodes, an international workshop as part of the TDWG conference in Beijing in October 2012 was organized and coordinated. If you would like to know more about this workshop, BLE potential usage and news, please contact us via the BLE contact form. We will inform about the workshop results via our social media, project website and in the next newsletter. You can check the new BLE version and new virtual exhibitions on www.biodiversityexhibition.com. If you would like to participate on existing or new exhibitions, please contact us via the contact form on the website.
Web-incompatible Image Formats in Europeana

P. Daneš (NM)

In the National Museum, we are currently providing data and images to OpenUp (and consequently Europeana) from six different departments, with several more under consideration. As the sole technical support for this activity, I am responsible for the storage and delivery of all this data. Trying to support all this manually would be nearly impossible, so, like most DBAs, I do as much as possible with automated scripts. One of the problems we have to address is that Europeana, as a web portal, supports a limited number of image formats. Since our internal databases and storage areas must deal with all possibilities that the various departments choose to utilize, it was necessary to invent some transformation mechanism.

A simple possibility would have been to upload only images stored in supported formats, but this would mean reducing the number of images our institution provides – not a desirable outcome. Another path would be to insist that our data suppliers provide images in web-compatible formats, but this would mean additional work and storage (and the inevitable attendant confusion) necessary to maintain a high-quality original and a lower-quality public version – more work for users with no discernible (to them) benefit. For me to manually create web-compatible versions would mean the same problems, except for me instead of our suppliers.

To resolve this, I have written a VBScript routine that scans a folder and looks for all images in TIF, BMP or PSD formats. Those are the three formats which our data suppliers currently use; others could be added to the routine with minimal effort. When such an image file is found, the routine looks for a JPG image file of the same name, in the same folder. If none such is found, or if one is found, but has an older ‘LastModified’ date than the non-JPG image, the script calls IrFanView to create a new copy of the image in JPG format, in the same folder. If such a JPG image already exists, with a newer ‘LastModified’ date, the non-JPG image is skipped and the script goes on to the next non-JPG image file. The subsequent transfer routines, which copy image files from our internal server to the public server then copy only JPG files. Non-JPG images stay on our internal server, for use by the department owning them, but are not available to the public.

This script runs automatically every night, but in most cases does nothing. Only when data suppliers have placed new, non-JPG images on the server does the script create new, JPG versions. Otherwise it scans the folders, sees that everything is in order and exits. The transfer routines also copy only newer files – those that have not changed on either server remain untouched, so on most nights, the actual work performed is just a compare. The method is not foolproof, but it deals well with our situation.

I have identified two possible problems:

1. The user could supply two completely different images with the exact same name, one in JPG format and one some other format. This routine would overwrite the JPG image with a JPG version of the non-JPG image, if the supplied JPG image had an older ‘LastModified’ date than the non-JPG image.

2. The user could ‘update’ the non-JPG image with a newer one, but with an older ‘LastModified’ date than the JPG copy created by my routine. The routine would not create an updated version of the JPG copy, since it would not see it as more recent.

The first possibility does not exist in our datasets. I wrote a script to scan for such instances and there are none. Although this is no guarantee for the future, it does not seem likely that user would create such a situation. The second does not seem likely either – it would mean that user had unearthed an old image and decided to use it to replace an existing image. Much more probable is that the second image would be added, rather than be a replacement. There is no certainty that either of these will always be the case, so occasional manual monitoring will still be prudent, but this script takes care of the vast majority of routine updates. The NeDiakritik function is used to remove Czech accent marks from filenames, since they can cause problems with some file system managers.
UB Script Document

Option Explicit
Dim fs, ScanDir
Const JGP = "jpg"
ScanDir = WScript.Arguments(0) ' Pass in (top level) folder to be scanned
Set fs = CreateObject("Scripting.FileSystemObject")
RenameFiles(ScanDir)

Function RenameFiles(FolderName)
Dim ObjFolder, ObjSubFolders, ObjSubFolder, ObjFiles, ObjFile, x
Set ObjFolder = fs.GetFolder(FolderName)
Set ObjFiles = ObjFolder.Files
' Write all files to output files
For Each ObjFile In ObjFiles
    x = ObjFile.Name
    If LCase(Right(x, 3)) = JGP Then
        x = MeDiakritik(x)
        If x <> ObjFile.Name Then
            ObjFile.Name = x ' Diak in JPG filename, rename
        End If
    End If
Next
' Getting all subfolders
Set ObjSubFolders = ObjFolder.SubFolders
For Each ObjFolder In ObjSubFolders
    ' Getting all files from subfolder
    RenameFiles(ObjFolder.Path)
Next
End Function

Function MeDiakritik(ByVal diak)
' Nahradit diakritikké známinka prostým písmen
diak = Replace(diak, "á", "a")
diak = Replace(diak, "č", "c")
diak = Replace(diak, "ď", "d")
diak = Replace(diak, "ě", "e")
diak = Replace(diak, "í", "i")
diak = Replace(diak, "ň", "n")
diak = Replace(diak, "ú", "u")
diak = Replace(diak, "ů", "u")
diak = Replace(diak, "ž", "z")
diak = Replace(diak, "ž", "z")
diak = Replace(diak, "ć", "c")
diak = Replace(diak, "ć", "c")
diak = Replace(diak, "ę", "e")
diak = Replace(diak, "ł", "l")
diak = Replace(diak, "ń", "n")
diak = Replace(diak, "ń", "n")
End Function
Survey on Usability of OpenUp! Content in Europeana in the Educational Sector

J. Frank (NM)

The primary task of the OpenUp! project is to provide natural history content to Europeana. This content has excellent potential to be used by scientists and academics. We also see potential for this content to be used in the educational sector, since it offers metadata curated by specialists, and clear copyright/licensing terms. Especially the unclear copyright licensing of the multimedia content provided for example by Google, makes it often problematic to use these data. This is not the case of OpenUp! content.

How can we reach the educational sector? One possibility could be a massive PR campaign in European countries using media like TV and radio, but this would require far more resources than are available in this project. Our chosen strategy is to use options like social media (Facebook, Twitter, Blog) and sharing these media via project and Europeana partners. We are also providing dissemination materials like leaflets, flyers, posters and newsletters, running virtual exhibitions and the project website, and present the project at international conferences and meetings. To better understand how teachers and students might use OpenUp! content, we prepared a survey in the form of an online questionnaire that was distributed to schools and universities.

The survey is addressed at natural history teachers from primary and secondary schools throughout Europe, and to university students and teachers from faculties of education. Its main purpose is to advance and explore the use of Natural History multimedia objects provided by the OpenUp! project to Europeana. A questionnaire is usually the most effective tool to get feedback from a target audience. We chose an online questionnaire via the professional web service SurveyMonkey and intend to continue this survey and get more feedback in the future as the amount of OpenUp! content on the Europeana portal grows. Currently, three language versions of the questionnaire have been prepared:

English: www.surveymonkey.com/s/OpenUp_questionnaire
German: www.surveymonkey.com/s/OpenUp_survey_German
Czech: www.surveymonkey.com/s/OpenUp_questionnaire_Czech

We expect that the time for completion to be about 15 minutes and but it is necessary for the respondent to have at least some experience with the Europeana portal. The deadline for this questionnaire is 2 December 2012. Your participation in this survey is more than welcome. If you have any questions, please do not hesitate to contact us or if you are interested in this survey and its results. The contact person is Jiri Frank – open_up@nm.cz.

Flowing Smoothly into Europeana: the OpenUp! Technical Workflow

J. Frank (NM), J. Hoffmann (MfN), G. Koch, W. Koch (AIT) & F. Theeten (MRAC)

One of the main tasks in the OpenUp! project is to harvest the standardised metadata of multimedia objects of natural history data providers and to transform this data into the Europeana schema. The transformed data is aggregated in the OpenUp! Metadata Database of the Europeana Natural History Aggregator established by the OpenUp! project and subsequently handed over to Europeana (Berendsohn & Günstsch 2012).

Data or metadata?

We need to explain our view on the term, ‘data’ and ‘metadata’ in the OpenUp! project. For example: Natural History domain data is included in the metadata for multimedia objects (physical object information). Metadata usually refers to the technical data of a multimedia object, e.g. aperture, camera type, etc. However, Europeana calls domain data (= records)
related to the physical object by the term ‘metadata’, and features this associated metadata along with the digital object. This metadata is distributed under CC zero licence in Europeana – under the full control of the provider. Only a minimum set of mandatory concepts is required for Europeana (fig 1). The OpenUp! architecture is divided in two integral parts. The first part addresses the data provision, including the set-up of the BioCASe Provider software and the mapping to the domain standard ABCD (Access to Biological Collections Data) and its extension EFG (Extension for Geosciences). The second part is the Europeana Natural History Aggregator which builds the OpenUp! Metadata Database, assures the transformation of the domain standard ABCD (EFG) to the Europeana standard ESE (Europeana Semantic Elements) and enables the harvest by Europeana.

The overall OpenUp! to Europeana (technical aggregation) workflow consists of seven major steps that are visualized in the following graphic (p. 19) and described below.

Workflow Description:
Content provider and coordination (Steps 1–3): The technical set-up for data provision in OpenUp! can be used/is used to provide data to the GBIF network.

Step 1: Domain standard ABCD and its extension EFG
As the first step the multimedia object associated
metadata of the provider (collection data) is mapped to the ABCD domain standard (zoology and botany) and its extension EFG (paleontology, mineralogy and anthropology). The mapping to the ABCD standard is carried out using the BioCASe Provider Software. Finally the BioCASe Provider Software serves as a web-interface for providing the data for harvesting.

Step 2 (optional): Data quality check
Before harvest (Step 4), providers can check their data with the Data Quality Toolkit, which provides a service for automated testing of their data quality, e.g. conformity of the data or check of scientific names against reference services. After testing the data, providers can apply necessary changes in their source data or in mapping between the database and the BioCase Software tool.

Step 3: Compliance check and monitoring of data provision
Providers can check their mapping and the correctness of the used concepts in the BioCASe Monitor Service by attaching their data source access point URL to this URL. Sample values for each concept are displayed and concept values are counted on demand, which helps detecting inconsistencies or incorrect use of concepts according to the ABCD documentation.

Furthermore, the BioCASe Monitor provides a compliance check for Europeana and displays error messages if mandatory concepts for the ABCD to ESE transformation are missing. The providers should assure they have a functional data source and correct mapping before requesting a test-harvest. The OpenUp! Helpdesk provides documentation and technical assistance for the setup of the BPS and the ABCD (EFG) mapping, and assists the providers in troubleshooting, in close collaboration with the BioCASe Helpdesk and the GBIF team.

The progress in content provision is monitored in the BioCASe Monitor Service by the coordination teams of the content providing Work Packages 4 & 5 in OpenUp!.

Step 4: (Test) Harvest
Once the mapping is quality checked by the coordination teams of the content providing Work packages and the OpenUp! Helpdesk, a test harvest with the GBIF Harvester, the HIT (Harvesting and Indexed Toolkit), is initiated. Test results and valid content is communicated back to the provider in order to allow for further adjustments. Technical problems encountered during test-harvest are fixed in collaboration with AIT, the OpenUp! Helpdesk and the BioCASe Helpdesk team. A harvest of the entire data source is initiated after successful completion of the test-harvest and confirmation by the provider.

The data provider can check the visualization of their content in Europeana by the Europeana Content Checker tool. This tool is also used by the WP coordination for a final quality check and to detect issues in the display of data/content in Europeana. Encountered problems in display of the data in the Europeana portal not related to the data provided are communicated back to the Europeana.

Step 5: HIT Harvest
The HIT Harvester stores bulks of ABCD (EFG) records into the central aggregator OpenUp! metadata database. This database stores only the metadata, including the URLs of the multimedia data.

Step 6: ABCD (EFG) transformation to ESE
The metadata from the ABCD (EFG) standard used by the natural history domain are transformed into ESE, which is used as a cross-domain metadata standard in Europeana. The transformation is carried out using Pentaho Data Integration (Pentaho Kettle). The mapping tool picks up the metadata, transforms them and stores them in a metadata database.

Step 7: OAI-PMH and Europeana harvest
The metadata are periodically harvested by Europeana via a single OAI-PMH (The Open Archives Initiative Protocol for Metadata Harvesting) access point at the metadata database. Previews of multimedia objects for presentation and queries in the Europeana portal are generated by Europeana from full object URLs given in metadata. This is the final step in the workflow when providing data in the flat ESE standard.

This is the actual implemented technical workflow (November 2012). We will publish updated workflow including the semantic enrichment and EDM transformation in next newsletters and on our Blog.

References:
Figure 2: OpenUp! technical architecture with main steps indicated.
OpenUp! Newsletter


This communication tool is issued by OpenUp! at the National Museum, Cirkusová 1740, 193 00 Praha 9, Czech Republic.

All natural objects figured in the newsletter come from collections of the OpenUp! content providers and will be displayed on the Europeana portal.

www.open-up.eu