ECP-2008-DILI-518001

BHL-Europe

Progress Report

1 May 2010 – 31 October 2010

Deliverable number/name	*D1.4*
Dissemination level	*Confidential*
Delivery date	*22 November 2010*
Status	*Final*
Author	*Henning Scholz*

*e*Content*plus*

0   Document History

0.0   Contributors

Person	Partner
Henning Scholz	MfN
Melita Birthälmer	MfN
Jana Hoffmann	MfN
Adrian Smales	NHM
Nancy Chillingworth	NHM
Jiri Kvacek	NMP
BHL-Europe members	MfN, NHM, NMP, EDLF, AIT, ATOS, FUB-BGBM, UGOE, NHMW, LANDOE, HNHM, NAT, NBGB, RBINS, MNHN, CSIC, RBGE, SIL, MOBOT, UH-Viikki, UBER

0.1   Revision History

Revision Date	Author	Version	Change Reference & Summary
09 November 2010	Henning Scholz	0.1	First draft
12 November 2010	Henning Scholz	0.2	Second draft including reviewer comments
22 November 2010	Henning Scholz	1.0	Final version including comments and amendments of the BHL-Europe consortium

0.2   Reviewers

This document requires the following reviews and approvals.

Name	Position	Date	Version
Martin Gordon	Project Manager, RBB	11 November 2010	0.1
BHL-Europe consortium		17 November 2010	0.1

0.3   Distribution

This document has been distributed to:

Group	Date of issue	Version
Reviewer and BHL-Europe consortium	09 November 2010	0.1
BHL-Europe consortium	22 November 2010	1.0

1   Table of Contents

2 Summary

The objectives of the reporting period (May to October 2010) of the project were:

(1) Adapt and enhance the prototype of a virtual taxonomic library

(2) Evaluate and review approaches for the establishment and management of multilingual biodiversity digital libraries

(3) Review state-of-the-art technologies used by Europeana and BHL for processing digital content

(4) Initial design of a multilingual access point for the search and retrieval of biodiversity content through at least two portals (Europeana and BHL)

(5) Make BHL-Europe content available for harvesting by Europeana.

(6) Define and agree on requirements of BHL-Europe content providers regarding page images and metadata schema (standards and specifications, best practice, use cases)

(7) Define the user requirements involving the proposed target users of BHL-Europe

(8) Begin implementation of the technical and functional architecture of the BHL-Europe system

(9) Develop operational strategies and processes for long-term preservation and sustainability of the data produced by national biodiversity digitisation programmes

(10) Raise awareness and ensure that the project outputs are known and used by the target users and that the proposed approach directly addresses user needs

(11) Negotiate with Rights Holders and publishers to enable access to in-copyright content

(12) Extend the network of content providers

BHL-Europe provides three access points to the digital biodiversity literature of the partner libraries: the Global References Index to Biodiversity, the multilingual BHL-Europe portal and Europeana. With the end of this reporting period and the end of the first half of the project, we now have prototype solutions for all three access points mentioned above. More than 300,000 library records are available in the GRIB. The German prototype (D3.6) is released at the same time as the delivery of this document, giving access to three collections (BHLUS, NAT, LANDOE). These collections are also available through Europeana.

In addition to the work on the German prototype (D3.6), the work on the final BHL-Europe system is also progressing. The architecture is now defined, the hardware configured and the first content is already copied to the system in London. The identification of the use cases for this system takes into account the results of the first large-scale user evaluation in spring 2010, and is due for completion shortly.

The results of our consensus building process of the first months of the project will be incorporated in the first draft of our best practice guide. We are actively disseminating the project and its result. The public presentation of BHL-Europe has also resulted in potential content providers approaching us. BHL-Europe is now also very well involved in activities and processes of Europeana on various levels (e.g., management, technology, aggregation, business development).

3   Status

3.0   Status and progress of activities

3.0.1   Status and progress of activities – WP1

3.0.1.1 Deliverables and reporting to the EC

Six deliverables were submitted in April/May to the EC. We prepared, revised and finalised D1.2 (Progress report) and D1.3 (Annual report). We also supervised the work on the WP2, 3 and 5 deliverables: D2.3, D2.4, D3.5, and D5.7. As a result of the review meeting, three deliverables have been revised: D1.2, D1.3, and D2.4. The revision was finished on time and all documents were submitted to the EC on 23 July 2010. At the beginning of August, D5.8 was finished and submitted. In November, D1.4 and D3.6 will be submitted.

3.0.1.2 Financial statement, activity and status reports

A major task of the reporting period was the preparation, evaluation and approval of the financial statements and activity reports of all 28 partners for the first year of the project. This was a very long and time-consuming process due to the large number of 28 partners and the variety of individual administrative procedures. We will provide a detailed report on lessons learned from this procedure for the next financial statement.

Based upon the experiences with the financial statement we created new templates for the activity reports, a monthly report of hours of work per person and task, which has been in use since June 2010. Furthermore, we introduced a new written 'status report', due every three months. We have also created an appropriate template including the current status report, the WP5 activity questionnaire and a risk and issue report. The information provided with this report is the basis for the interim and progress reports. This is in order to fulfil the requirement of the EC to report in more detail on the work carried out by every partner.

3.0.1.3 BHL-Europe internal reporting

The activity reports (monthly) and status reports (every 3 months) were collected from the partners and compiled for the internal BHL-Europe Interim report (M15) and the progress reports (D 1.4). In addition, the WP 5 questionnaire information was collected together with the status reports and send to the responsible partner (RBINS) to be put on the BHL Wiki.

3.0.1.4 Project server

Access to the project server (Citrix) was tested for normal and advanced user accounts at MfN, the guidelines for access were revised and issues were discussed with the host of the project server (NHM). Work plans are continuously uploaded to the server as soon as available.

3.0.1.5 Annual Review meeting

The first annual review meeting of BHL-Europe was held in the Natural History Museum in Vienna from 25-30 May 2010[2] and was attended by almost all consortium members. The first

---

[2]   see https://bhl.wikispaces.com/BHL-Europe+Annual+Review+Meeting+Vienna for agenda, minutes and other relevant information

day was spent with small work group meetings (PMG, CWG, user survey). 26 and 28 May were used for content provider and use case working groups. The 27 May was the review meeting. We successfully passed it with a yellow flag. The TMB held its meeting over the weekend to plan the next steps of the technical implementation work.

3.0.1.6 PMG calls and meeting

Regular project management group (PMG) calls were held (Skype conferences) almost every week until end of July 2010 to discuss relevant details on planning, the liaison between BHL-Europe and Europeana, timely production of deliverables and other upcoming issues. More details can be found on the Wiki[3].

Between August and October 2010, PMG calls were only held irregularly due to the absence of various PMG members. However, most PMG members met on various occasions in September and October. The first of these meetings was held in Woods Hole after the Global BHL Technical Meeting. Afterwards, all WP leaders met during the Europeana Plenary (14-15 October 2010) in Amsterdam. These two occasions provided excellent opportunities to discuss the progress of BHL-Europe and plan next steps in more detail.

3.0.1.7 Mailserver

Various tools were used in the past to facilitate communication among partners. We are using the BHL Wiki and the Google Groups in particular. As Google Groups has some disadvantages, we have now set up BHL-Europe mailing lists at the Humboldt University servers (SYMPA). It is now easier to manage the communication within the content providers and also with our external stakeholders. The first mailing list running on SYMPA is the content provider list: bhl-e.cp@lists.hu-berlin.de.

3.0.1.8 WP2 leader of BHL-Europe

MfN successfully finished the recruitment process for the vacant position of the WP2 leader. Melita Birthälmer joined the project team 15 May 2010. After familiarising herself with the project and the team members and the revision of D2.4, she will now focus on project planning and coordination of WP2 work. This include the best practice guide for scanning operations, content analysis, attraction of new content providers, and negotiations with rights holders.

3.0.1.9 Recruitment of new BHL-Europe employees

Several of our partners were able to employ new staff members for BHL-Europe tasks: Michaela Hierschläger (LANDOE), Oriane Boissel (MNHN), Lesa Ng (RBGE). We welcome them to the BHL-Europe team.

The main task of Michaela Hierschläger will be the preparation and compilation of the Best Practice Guidelines and Standards. She also takes care of data refinement for Europeana.

Oriane Boissel will assist in tasks regarding the pre- and post-scanning process. Pre-scanning tasks will include selection of content, description of the bibliographical data and analysis of the intellectual frame. Post-scanning tasks will be quality control of data and metadata.

Lesa Ng's key duties will be the management of RBGE digital content for inclusion in BHL-Europe, creating bibliographic records to be included in the GRIB, promoting BHL-Europe

---

[3] https://bhl.wikispaces.com/BHL-E_PMG_calls

among the Scottish content providers, in order to increase the number of Scottish content providers and attracting funding for future digitisation.

3.0.1.10 Consortium Agreement

The Consortium Agreement was prepared in the first months of the project. Although not required for projects of the *eContentplus* programme, it is strongly recommended. A first draft was available in August 2009 based on the DESCA templates. The document was customised, but the liability section of the document was under discussion for a long time. We had legal consultation to identify the most appropriate wording for our needs. The document is now finalised and sent out for final review. The signed agreements are expected to be available for the next Content Provider and Technical Meeting in London (1-3 December 2010).

3.0.1.11 Sustainability

BHL-Europe is funded until 30 April 2012. During the first year of the project initial ideas for a sustainable BHL-Europe system beyond 30 April 2012 were already discussed and presented in D1.2. During the Europeana Group of Projects Meeting (29 September 2010, The Hague, The Netherlands), a number of projects identified a common need of business development strategies. Europeana will support the collaboration of the Europeana group of projects in developing strategies. We are currently planning a business development workshop for our next Content Provider and Technical Meeting in London (1-3 December 2010).

3.0.1.12 Global BHL Technical Meeting

The first Global BHL Technical Meeting took place from 22-24 September 2010 in Woods Hole/MA, USA. The meeting covered numerous subjects ranging from hardware infrastructure, data synchronisation and harmonisation, content selection, GRIB, OCR, taxonomic intelligence, governance and dissemination. Below, a short report is provided with aspects and action items as far as BHL-Europe is concerned. For the full agenda, notes and action items visit the BHL Wiki[4].

GRIB: The development of a Web database to support analysis of domain content and management of the scanning process is a task within the IT Development of BHL-Europe. As specified in the DoW we developed the Global References Index to Biodiversity (GRIB) based on the ViTaL activity within EDIT (Activity 5.3). The GRIB was presented during the Global BHL Technical Meeting to all other BHL nodes. Until now, no other integral system exists which can deduplicate serials and monographs for metadata management and scanning planning purposes. A joint BHL/BHL-Europe session during the meeting concentrated on the global aspect of the GRIB. A list of useful features of the GRIB in the international context will be provided now and will be crosschecked with the requirements of BHL-Europe.

To date, the GRIB contains BHL-Europe catalogues and metadata and hyperlinks from already-digitised content within BHL-US. It was decided to import more library catalogues from BHL-US and other nodes to the GRIB and to update contributed catalogues regularly in order to avoid duplication of work.

Content Management: During this meeting the global coordination of content management was identified as a further task. It was decided to compile a document that articulates collection issues and workflow tools with information for each regional BHL node. BHL-

---

4   https://bhl.wikispaces.com/GlobalBHLTechMtg_2010

Europe will contribute to this document by describing the activities regarding the Best Practice Guidelines and Standards (D 2.6) and the GRIB.

OCR: Part of the BHL-Europe work plan is the improvement and implementation of OCR techniques. Other projects like IMPACT are working on the improvements of the actual OCR engines and technologies and it would be beyond the scope of BHL-Europe to work on this subject. However, we can support IMPACT and later adapt tools to fit our needs as well. We need to further investigate the current status of IMPACT and will work on a test set of data taken from our corpus of digitised literature. In addition to IMPACT, we are currently looking into the Bibliotheca Alexandrina, which has a well-established OCR procedure which may decrease our OCR error rates. It was decided to write a clear problem statement, so other partners and projects can give us better and more precise information about possible solutions.

We can also involve users in improving the OCR results. This approach has proven to work very well as shown by the example of the Australian Newspaper program[5]. It was decided to set up a Wikisource instantiation and use our OCR testset to test this feature. Once this approach has been proven to be working and useful, we will implement this system for the BHL-Europe Portal as well.

Ratification of GUID for Global Requirements: During the global BHL meeting, BHL-Europe requirements were discussed. The requirements for BHL and BHL-Europe were considered found that similar solutions were required, the development work for the GUID (Globally Unique ID) minting system that BHL-Europe are developing could provide services for the BHL consortium. The GUID system provides a unique identification usually in the form of a Universal Resource Identifier (URI) or Location (URL) for a digital object. Each institution would request a GUID from a central web service and assign the GUID's to each object. As the GUID mint has to scale to encompass the BHL-Europe membership and provide a persistent and resilient service, the additional effort to serve the global requirement was considered minimal. Consideration for the GUID mint extended to commercial services. However, due to the large number of GUID's required by BHL-Europe, commercial minting services were too costly. A number of minting systems systems were evaluated and the current favoured solution is the "Handle" mint[6].

Technical Consultation with Bibliotheca Alexandrina: The wider discussion with global partners has proved very fruitful, and has provided a rich resource for both technical requirements and expertise. The preferred GUID mint solution (Handle) has been implemented and is operational in the Bibliotheca Alexandrina in Egypt. The BHL-Europe technical group in discussion with Bibliotheca Alexandrina consider that a collaborative, technology sharing visit to be a good way of reducing the learning curve and maximising the development opportunity. The Library in Alexandria has developed a URI schema, deployed hardware and maintain a Handle minting service. The requirements that BHL-Europe have identified and need to be developed are very similar to those already deployed. BHL-Europe technical group are hoping that agreement can be granted and a visit to exploit the developments at the Bibliotheca Alexandrina arranged.

Global Replication for BHL-Europe: Global replication of scanned material is a consideration for BHL-Europe. Although, this requirement extends beyond the boundaries of the EU, the preservation and the larger goals of the project require a high degree of

---

[5] http://www.nla.gov.au/ndp/get_involved/

[6] http://www.handle.net

availability, information integrity and continuity. The methods of transfer and synchronisation are currently under review and consideration is being given to various technologies. BHL-Europe has a test replication service in operation between the NHM, London and MBL, Woods Hole. The service is provisioned an "ASYNC" service, the issues that are being worked through are to consider load, fluctuating WAN conditions and long term requirements. ASYNC has issues in that once it "fails" it has to be restarted manually and is not as robust as the service will require in full production, other technologies such as LOCKSS are being evaluated.

Taxonomic intelligence / name finding: One objective of BHL-Europe is the integration of Taxonomic Intelligence web-tools to facilitate the search for taxon-specific biodiversity information. Improving this multilingual search by adding more non-English vernacular names to the Taxonomic Intelligence database is also one of our key tasks. However, a number of international projects are currently dealing with scientific names on various levels. In order to avoid parallel developments and duplication of work, it was decided during the Global BHL Technical Meeting to write a clear problem statement and send it out to all the projects working on related aspects. In addition, we will approach EOL to find out which species pages are the most popular. This will help to estimate the effort required to establish a small thesaurus for testing multilingual taxonomic intelligence across Latin species names and vernacular names.

Business plan: BHL-Europe is the only BHL node with a clear deliverable on business planning in the near future. Therefore, we will continue to compile a plan which includes a solid value proposition statement, the sustainability and business continuity plan for the preservation and archive system as well as an overview of the incorporation of BHL-Europe into the European infrastructure of initiatives. We will have a brainstorming session during the next project meeting in December. We will also include feedback from our global partners and Europeana to facilitate future alignment of the BHL-Europe directions in this international framework.

Dissemination: We discussed the alignment of dissemination activities to maximise the outcome of these activities for BHL-Europe on a global scale. It is obvious that we need to develop a new logo for the product developed by BHL-Europe. This logo will then replace the actual project logo after the end of the project. A requirement for the logo is that it has symbolic aspects and does not include characters, so it can be used and understood by all members of our multilingual community. The core Europeana logo serves as a good example for orientation.

It was decided to set up a BHL calendar of events, meetings and conferences. This will facilitate the organisation of project and work group meetings and maximise the travel efforts of all team members. The use of this calendar will also improve the BHL and BHL-Europe representation at conferences and events that are attended by members of our target users.

Furthermore we will invest more time in the design of a promotion kit including fact sheets and texts for our different target users and project partners. This will be presented on a public Wiki or the project Web site in order to improve the visibility of up-to-date information about the project.

As BHL-Europe has a significant budget for dissemination, we need to intensify our work in disseminating our project to avoid being overshadowed by other BHL nodes.

3.0.1.13 Conference „Deutsches Kulturerbe auf dem Weg in die Europeana" (German Cultural Heritage for Europeana)

From 4-5 October 2010, the German Europeana project partners, members of the German Digital Library, and about 200 members of German archives, museums, and libraries met in the Berlin State Library. The aim of the conference was to gain an overview of existing projects and their tools, services, workflows and best practices in order to identify synergies and discuss the development of digital infrastructures for the German cultural heritage. The presentations covered a wide range of topics from technical solutions over content selection procedures to IPR issues and project management challenges.

During the conference, the variety of content that is available for Europeana in Germany was pointed out by the various projects. After the recent update of Europeana for the Rhine Release, Germany now is contributing 13% of the almost 13 million objects available in Europeana and is second only to France. The conference was organised by ATHENA (Institut für Museumsforschung, Berlin), EuropeanaLocal (Zentral- und Landesbibliothek Berlin), European Film Gateway (Deutsches Filminstitut, Frankfurt a.M.), Biodiversity Heritage Library for Europe (Museum für Naturkunde, Leibniz-Institut für Evolutions- und Biodiversitätsforschung an der Humboldt-Universität zu Berlin) and Stiftung Preußischer Kulturbesitz (Berlin). All information including abstracts is available online[7].

The directors of the organising institutions opened the conference on 4 October. Their opening speeches already underpinned the tasks and challenges for Europeana that were further discussed over the two days of the conference. Jill Cousins gave an excellent overview of the current progress in Europeana towards the upcoming Rhine Release of Europeana v1.0. The participants of this conference were introduced to 17 existing projects. Among these projects were well-established projects like Athena as well as newcomers like Natural Europe. In addition, four institutions not directly involved in a Europeana project presented their experiences with Europeana. The program was topped off by a status report of the German Digital Library and some more technical information on linked data and the Europeana Data Model (EDM). As it was a very successful event with lots of good discussion, it was agreed, to work on establishing this conference as a regular (annual) event.

For BHL-Europe, a number of interesting aspects were presented that will be followed up in the near future:

1) A number of EU projects are also dealing with best practice guidelines for digitisation. This information should be included and considered in our best practice guide to give a complete overview of that topic. We may consult experts from these projects for advice and lessons learnt. Projects working on best practice documents include Athena and EUScreen.

2) The European Film Gateway (EFG) is working on an Authority File Manager (prototype available[8]). This tool allows merging of duplicate records and the creation of master records for databases. It needs to be explored if this tool is an option to be used for the GRIB to merge library records.

3) The deduplication of content in Europeana was discussed with EuropeanaLocal. UGOE is submitting their digital archives to Europeana via EuropeanaLocal Germany as well as via BHL-Europe. After ingesting this content into BHL-Europe, we then deliver it to Europeana, thus creating a duplication of this content in Europeana. It still

---

7 http://www.armubi.de/tagung2010 (in German)

8 http://www.europeana.eu/portal/thoughtlab_improvingmetadata.html

needs to be discussed how this can be solved. Three options are currently discussed: (1) Europeana takes care of duplicates and implements procedures to handle them in a practical way; (2) the cross-domain aggregator (in the above example EuropeanaLocal) filters the content and does not deliver duplicate content to Europeana; (3) BHL-Europe as a domain aggregator filters content from other aggregators which is already present in Europeana.

4) The communication process between a repository and a catalogue in enrichment data is currently implemented by the Dilibri project using the software Visual Library (Semantics). We will check whether this technology and lessons learnt might be useful for the communication of the GRIB with the BHL-Europe system.

5) The Technical University Ilmenau is working on and with OCR technologies and has developed its own tools to improve OCR quality. These tools should to be further investigated during a test run to see if their approach is also valuable for BHL-Europe.

6) The Natural Europe[9] project was launched recently with the Kick-off meeting from 11-13 October in Bayreuth and Eichstätt (Germany). The Natural Europe project suggests a coordinated solution at European level in order to exploit digital content of natural history museums by connecting the digital collections of a number of European NHMs with Europeana, and studying the educational methods and deploy the necessary software tools that will allow museum educators to design innovative online pathways through the digital collections of NHMs. It is funded under the CIP-ICT-PSP programme (follow up of eContentplus). BHL-Europe and Natural Europe have two consortium partners in common: UH-Viikki and HNHM. The connection between two major IT providers for Natural Europe has been made by HS and further areas of collaboration are being investigated.

7) HS made contact with three potential content providers for BHL-Europe. One of them - Dilibri[10] – currently has digital biodiversity content; mobilisation of this content for BHL-Europe will be looked into.

3.0.2 Status and progress of activities – WP2 & WP3

3.0.2.1 GRIB prototype (Task 2.1)

The Global References Index to Biodiversity (GRIB) prototype by VZG at http://grib.gbv.de/ (see also Figure 1) has been enhanced by a Search/Retrieve via URL (SRU)-Interface[11]. The digitisation management tool was changed after D2.3 and has now six different statuses. It can be searched for the digitisation status via [DST] 8300 = Not digitised, 8301 = Should be digitised; 8302 = Will be digitised; 8303 = Digitisation in progress; 8304 = Digitisation completed; 8305 = Document available. A technical documentation is available[12].

A further task during this report time frame was the testing of the GRIB prototype. MfN and UBER worked on a methodology for a user review and technical review of the GRIB, also further requirements for the GRIB where discussed with European partners and international

---

9   http://ec.europa.eu/information_society/apps/projects/factsheet/index.cfm?project_ref=250579

10 http://www.dilibri.de/ (in German)

11 http://grib.gbv.de/sru/DB=1.83

12 http://bhleurope.gbv.de/
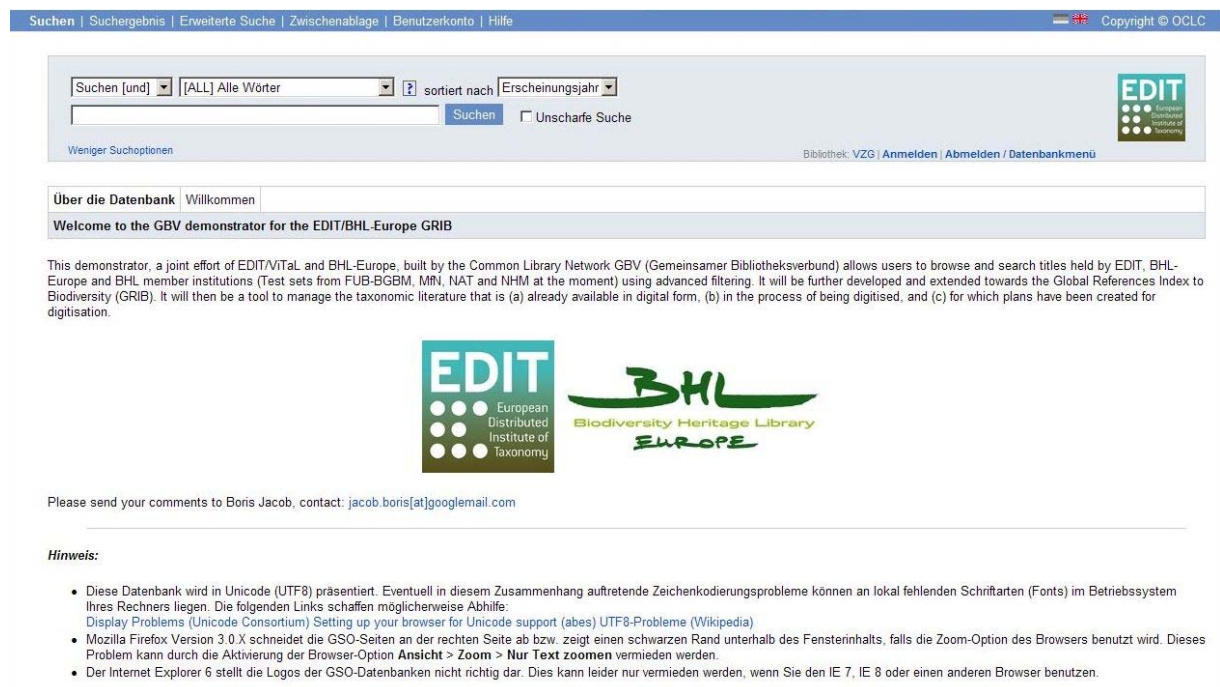
experts. The functionality and documentation of the index web services have been tested by AIT, MfN and BGBM and the results of this analysis has been published in the technical note "Technical Note Prepare Preingest for GRIB prototype API" (TN-SPRINT1-102[13]).

The GRIB will be equipped with a single-sign on and security solution which has been developed by BGBM and MfN within the EDIT (European Distributed Institute of Taxonomy) project. The actual implementation phase will start in September 2010. It still needs to be evaluated if this is a suitable solution for the BHL-Europe Portal as well.

For dissemination purposes an open Twitter-feed has been created[14]. The GRIB is also featured at the Europeana Thought Lab[15].

Figure 1: Screenshot of the current GRIB prototype.

3.0.2.2 Submission of data for the GRIB

All our partners were asked to submit bibliographic records for the initial transfer of our partners library catalogues into the GRIB. In relation to this request RBINS for example discussed the organisation behind their library catalogue and how it can be send to the GRIB. NBGB, RBGE, RMCA, and UH-Viikki submitted sets of their library catalogues for initial GRIB testing.

Metadata from BHL has been harvested and imported into the GRIB and merged with the existing catalogue data, so that the GRIB now links to the already digitised content in BHL. Based on the evaluation of the library questionnaires filled by our partners in 2009 and 2010, BHL-Europe will presumably import 2,500,000 sets of metadata into the GRIB by March 2011.

---

[13] https://bhl.wikispaces.com/WP2_WP3_Sprint_deliverable

[14] http://twitter.com/editvital

[15] http://www.europeana.eu/portal/thoughtlab_digitisation.html

Taxonomists are an important user group for BHL-Europe – biodiversity literature forms the basis of their daily work. Within BHL-Europe, the BGBM aims at integrating BHL and BHL-Europe services with the EDIT Platform for Cybertaxonomy in order to facilitate taxonomists' access to BHL/BHL-Europe resources and taxonomists' feedback to BHL-Europe.

The EDIT Platform for Cybertaxonomy is a collection of tools and services which together cover all aspects of the taxonomic workflow. The workflow is grouped into the following areas: taxonomic editing; publishing of edited data; data storage and exchange; collections and specimens; descriptions; fieldwork; literature; and geography. The key component of the Platform for Cybertaxonomy is the Common Data Model (CDM). This a repository for every conceivable type of data produced by taxonomists in the course of their work and the backend for most EDIT components.

An infrastructure based on OAI-PMH (Open Archives Initiative - Protocol for Metadata Harvesting - v.2.0) which allows BHL-Europe to harvest bibliographic data from CDM Community Stores was established as described in the last status report. The capability of the OAI-PMH (Open Archives Initiative - Protocol for Metadata Harvesting - v.2.0) implemented into the CDM Community Stores has been improved, and is now providing an improved mapping of the CDM data structures to the data exchange formats used by the OAI-PMH.

 The bibliographic data provided by the GRIB was made available to the users of the EDIT Platform for Cybertaxonomy by implementing the option for searching the GRIB directly from within the CDM-based tools of the EDIT Platform. Therefore a web service wrapper has been implemented into the CDM library which allows using the Search and Retrieve via URL (SRU) services of the GRIB directly from within applications like the Taxonomic Editor or the EDIT DataPortal. The SRU service cannot return the result set record in the CDM data format. In order to allow a seamless integration of the data coming from the GRIB into the CDM based components of the EDIT Platform, the records are transformed into according CDM entities on the flow. Web-based applications like the EDIT DataPortals cannot use the CDM library directly. In order to enable these tools to use the SRU service wrapper, a RESTful web service was implemented which exposes this functionality to the web. User interface elements to make this feature available in the web pages published by the EDIT DataPortal are the subject of further implementation tasks.

3.0.2.3 Library questionnaires

In May we conducted a Library Questionnaire. Following on from the questionnaire from 2009, it focused on technical details of the partner's library systems. The results (presented at the Annual Review meeting and in parts described in D2.4) helped us to understand the technical and organisational aspects of the partner libraries better, so the process of collecting library catalogues could eventually start. Most of them will be initially uploaded to the GRIB via manual export, few via XML-/OAI-Server. The most common metadata format is MARC21 and most of the partners have more than one DB for monographs, journals/subscriptions and scans. The questionnaire helped to better understand what partners expect from the GRIB and the Pre-ingest procedure and that there is a need for implementing a scanning workflow software. Those expectations where put into a requirements list and some of them already have been formalised in use cases, to be further discussed and implemented.

3.0.2.4 BHL ScanList (Task 2.1)

Several partners (LANDOE, NAT, RBGE, UH-Viikki) worked on improving the the BHL(-Europe) ScanList. LANDOE enriched the list with items that have been scanned in their institution. RBGE placed two bids for two titles that are scheduled to be included in BHL-Europe.

The maintenance of the ScanList and the provision of enhanced features for our partners, such as a .csv export procedure, is carried out by NHMW. ScanList users can now view and export the current list on their screen of existing bids as a .csv file in order to better coordinate scanning management. These features have been transferred from the test system to the live system.

3.0.2.5 Analysis of domain content (Task 2.2)

During the last six months much work has been done by our partners regarding analysis of domain content. Work has been done on the analysis and the preparation of the monographs and serials that are relevant for the biodiversity community. Our content providers discussed how and which digital publication can be offered to BHL-Europe.

MNHN for example concentrated on identifying how the irregular frame of their old serials can be brought together in one digital structure and continued to describe the bibliographical and physical data of each issue number. Further technical specifications (format, image resolution, indexing, file naming) were written by MNHN in preparation for the outsourced scanning. The two last meetings in Vienna enabled them to have a better understanding of the BHL-Europe technical aspects and to better define their technical requirements and content management. Further, for the outsourced scanning an Excel template for the metadata has been prepared.

MNHN also met several colleagues to discuss the selection of contents and the digitisation policy in the field of higher education and research (Association of University Libraries in Paris, Medical and Dental Academic Library of Paris, and also Persee, a national digital service for journals and works in Humanities and Social Sciences in open access, provided by University of Lyon 2[16]), which was also an occasion to talk about the BHL-Europe project.

RMCA, NBGB and RBINS continued to work in close collaboration on the content selection for digitisation and updating of the local Wiki-based scan list (common for three Belgian institutions). Together they work on the same scanning procedures and IPR handling. They also have a common Wiki[17] where they list the volumes to be scanned and that are already scanned as they share a lot of common literature on Africa in order to avoid duplication of scanning for example.

3.0.2.6 Analysis and preparation of (Meta)data for Pre-Ingest (Task 2.3.2)

The analysis of metadata files was also one of the main tasks for our content providers. CSIC analysed metadata files regarding EOS - a dead journal with a span of 70 years devoted to Entomology - and processed the Vol. 0 of Fauna Iberica for the Pre-Ingest. Two more volumes of Fauna Iberica have been already digitised. After receiving the feedback from the Pre-Ingest team regarding Vol. 0, CSIC will proceed with the preparation for the Pre-Ingest for the two other volumes.

---

[16] http://www.persee.fr/

[17] http://193.190.223.46/wiki_ext/index.php/BHL

The bulk digitisation of more than 70,000 pages of HNHM journals was carried out during the previous reporting period within a frame of national funded project. Since then HNHM is working on the metadata of their journals (Parasitologica Hungarica, Miscellanea Zoologica Hungarica; Studia Botanica Hungarica; Vertebrata Hungarica). The granularity of metadata is on the journal article level. The metadata (xml files) are imported into a MySQL database where further data cleaning and quality checks were performed. The data are published to the internet. Vertebrata Hungarica was used for experimenting with the metadata creation on the taxonomy. Title and abstract level taxon names are extracted and inserted to the zoosystematic tree. Therefore, various taxon level searches are now available.

RBINS discussed with their library personnel the metadata schema. As a result of working with BHL-Europe, RBINS has employed one person to continue digitisation work on RBINS funds. A priority list of publications to be digitised and contributed to BHL-Europe was prepared in cooperation with the RBINS library personnel.

Processing of metadata (input of new data and enrichment of catalogues with additional information) is an important work for RMCA, because RMCA has worked with the LIBIS catalogue system only from 2005, and this means that most old publications are either still not in the database or that only basic data is available (author, title and year of publication). This summer, students have been hired to enrich the catalogue.

The existence of different catalogue systems in specialised RMCA sub-libraries makes the ingest of metadata difficult. Therefore, some activities have been undertaken to solve this problem: contact with LIBIS IT team in order to complete the library questionnaire, comparison of different sub-catalogues and discussion with IT specialists about different possibilities to transfer these catalogues to BHL-Europe.

NBGB analysed the quality of their bibliographic metadata and completed their bibliographic data. The data, consisting of image files and bibliographic metadata, has been archived to be able to provide it to BHL-Europe as soon as the Pre-Ingest is ready.

MNHN provided test files for the Pre-Ingest Module and also a French written document with technical requirements for outsourcing the scanning.

Quality assurance of scanned content was a further task within the domain of analysis and preparation for the BHL-Europe Pre-Ingest. RBGE began the post-scanning processing of 46 volumes of the Notes of the Royal Botanic Garden Edinburgh (NRBGE). The content of these volumes comprises over 21,000 scanned pages. At the time of writing, 38 volumes of the NRBGE have been quality checked. This involves checking the scans for problems such as skewed text, gutter shadows, blurred images, and missing pages. Lists of corrections were then sent to the third party contractor who had carried out the original scanning for correction.

Structural metadata (page numbers, page types) has also been added to these 38 volumes using an Access database created in-house. This metadata is then 'attached' to each scan, e.g. 0008.tif in volume 1, part 1 of the NRBGE is page type text and has a page number of 5, 0009.tif in volume 1, part 1 of the NRBGE is page type text and has a page number of 6, etc. This metadata for volumes 3 and 4, along with article level metadata extracted from the RBGE library catalogue for these volumes are due to be sent for pre-ingest testing soon. Following completion of this work for the NRBGE, the same quality control and metadata work will begin on the Flora of Bhutan (8 volumes comprising 3,400 scanned pages). Quality issues with the original scans of the colour plates have been identified and these are currently being rescanned by a third party contractor.

Preparation of data for the Pre-Ingest has also been done by UH-Viikki. In Finland, the mapping of subject terms to define a profile of Pre-Ingest criteria was developed on the basis

of the ingest criteria of our colleagues in the US. The result is a table of 34 UDC Class numbers and 2627 subject terms relevant to BHL-Europe. These criteria can also be used for extracting data from their library catalogue for the GRIB.

In August, UH-Viikki started the tendering process for the digitisation on a national level. UH-Viikki received 8 tenders from Finnish companies offering digitisation. The selection criteria were based on both price (60 %) and quality (40 %) estimated from sample pages. The end result was an agreement on the digitisation by a company called Vaasan Micro Copy[18] which has extensive experience in digitisation of historic literature. They use a book scanner manufactured by the German company Zeutschel.

During this reporting period, UH-Viikki started the depositing of their first digitised series for BHL-Europe called "Memoranda Societatis pro Fauna et Flora Fennica". It is placed in the special collection Fauna & Flora[19] of their digital repository Helda.

RMCA digitised 12,000 pages of biodiversity literature from the zoological department in the framework of other projects and in-house activities. The coordinators of these projects are keen on having this content provided to BHL-Europe as well. All files are controlled and renamed following recommendations provided during the last Vienna meeting (page level).

BnF continued the selection of documents to be digitised for BHL-Europe. This activity includes checking of documents to be digitised, making a detailed description of each copy, sending them to the external companies for digitisation and overseeing quality assurance of the outsourced work. Additionally, BnF prepared the first transfer of documents for BHL-Europe: 8 titles of periodicals, 515 digitised documents representing more than 220,000 pages.

3.0.2.7 Content analysis status report

The analysis of the content-specific progress since beginning of this project in May 2009 was a principal task for WP2 in this reporting period. The underlying content of BHL-Europe was analysed based on the signed Memoranda of Understanding and the Description of Work. The analysis was done with regard to quality and quantity of content, and maintained a clear distinction between information from BHL-Europe and that from BHL-US. A further component of the content analysis was to detail the worldwide distribution of European biodiversity literature. The planning of the Europeana ingest was another part of the content analysis status report. The revised D2.4 now contains the complete results of the analysis.

Further work has been done and will continue regarding the estimation of the potential number of BHL-Europe pages available as print in Europe. LANDOE assessed also content existing in related sciences, like nature conservation, agriculture, ecology, palaeontology (and Earth sciences), limnology, medicine (parasites) and molecular biology (concerning taxonomy).

3.0.2.8 Memorandum of Understanding (Task 2.3.1)

As an update to previous reports, three further BHL-Europe content providers have signed the MoU during the reporting period. The data were already included in the revised D2.4 document so it is referred to it in this document.

---

[18] http://www.vaasanmicrocopy.fi/

[19] http://helda.helsinki.fi/handle/10138/14068

3.0.2.9 BHL-Europe content for Europeana (Task 2.3.3, 3.2.1)

EDLF worked on the analysis and promotion of content from BHL-Europe in relation to other material in Europeana and made sure the BHL-Europe content available in Europeana is of good quality and quantity, according to Europeana standards and specifications. A further task during this period was the technical integration with Europeana and development of a distributed data model that enables content to be available through Europeana and BHL-Europe. Securing ongoing access to BHL-Europe material through Europeana and coordinating future developments on Europeana data model was also important.

In the reporting period, 2,568 digital objects and metadata files were prepared for Europeana ingest by LANDOE and in September 2010 this data set was made available in Europeana. The whole data set of LANDOE can be retrieved by using the following search string in the search field on the Europeana portal: *europeana_collectionName:08702**

The existing data sets in Europeana (BHL & NCB Naturalis) were updated to the newest version of the Europeana Semantic Elements (ESEv3.3). This includes the new ESE element *europeana:dataProvider,* which allows more visibility of the BHL-Europe content providers in the Europeana service. A part of the updated BHL-Europe data also includes the new ESE element *europeana:rights*, which is an element for the rights statement applied to the objects. BHL-Europe was the first aggregator to publish data in Europeana with the new *europeana:rights* field.

After having ingested more than 13 million objects, Europeana now also put more effort into data quality management. We are discussing with the ingestion team approaches to improve the data and the representation of data in Europeana. In doing this, we are monitoring the data quality of our content providers in order to deliver a constant quality across all our partners.

3.0.2.10   Content acquisition (Task 2.3.2)

Attracting new partners during the next year is a major task within WP2. A new approach for content acquisition was identified in which our consortium members will take part. More precisely, the identified method will take advantage of our consortiums connections in order to attract new content providers within their own countries. For this purpose a "BHL-Europe in a nutshell" document is currently in preparation.

This document aims to be a very brief but informative document including the most important facts about BHL-Europe, benefits and IPR issues. This document is created in collaboration with WP4. The document will be distributed first to one of our partners to test the method and after evaluation of the new content acquisition method, more partners will be asked to use their contacts and connections within their own country for attracting new content providers to join BHL-Europe. This new method is designed to be a sustainable method in order to ensure the continued growth of BHL-Europe.

RMCA compiled a table of Belgian biological journals and societies. Furthermore, RMCA discussed with other Belgian institutions, societies and other projects to provide additional content to BHL-Europe (for example with the Royal Belgium Zoological Society). Further work has been done on Charles Lemaire expeditions to the Congo (1901-1903) on selecting and digitising of related material.

CSIC approached Dr. Fernando Palacios concerning the availability of the Proceedings of the "Simposio Internacional y Primer Congreso Mundial sobre Preservación y Conservacion de Colecciones de Historia Natural" (International Symposium and First World Congress on the Preservation and Conservation of Natural History Collections). The three volumes of the congress are now available for digitisation and ingestion to BHL-Europe.

3.0.2.11   Assist partners in scanning operations (Task 2.3.2)

Assisting partners in implementation and evaluation of scanning operations is a task within WP2. For this purpose the scanning workflow management tool GOOBI[20] was taken into account. During evaluation of the system it was found out that two of our partners (CSIC, UBER) already considered GOOBI as a scanning workflow management tool for their institutions. Therefore, the evaluation of GOOBI was passed to our partner UBER for further evaluation and test implementation within their own institution. The evaluation and test implementation is planned to be finished by beginning of 2011.

3.0.2.12   Use Case working groups (Tasks 2.3.1 and 3.4)

Work on use cases was another part of the work carried out. An use case discussion lead to the establishment of a use case working group subdivided into content users (Scientists, Citizen Scientists) and technology users (Librarians, IT). A standardised template for recording use cases was developed. It was agreed that the use case group leaders will fill in the final use case template based upon the discussion and input from other use case group members. Group leaders will then submit the use case descriptions to the developers. A Wiki page[21] was created to compile all information on use cases and list all BHL-Europe use cases described so far. Google groups were established to facilitate the communication between use case group members. During the meeting in Vienna (May 2010) the technology user use case group met to discuss those use cases dealing with scanning management, content ingest into the Pre-Ingest and library catalogue export to the GRIB.

In order to build a powerful portal for search and retrieval of biodiversity literature, taxonomic use cases have to be described to provide a guideline for the portal development team. In this context it is essential to integrate user needs and requirements. Therefore, a catalogue of requirement was established representing the maximum requirements for a biodiversity search portal from the user's perspective. The catalogue of requirements is a compilation of existing information from the BHL Wiki, BHL-Europe meeting presentations and minutes, as well as the results of the BHL user surveys. The catalogue was discussed and improved during intensive discussions (Skype, BHL-Europe Technical Meeting London) between the portal developers (AIT) and JH. Furthermore content user needs were explained in more detail to the developers and taxonomic background information was provided. In a next step, four core features were identified according to the DoW and a wish list of possible portal functionalities was established. There is an ongoing search for other tools and projects useful to be integrated into the BHL-Europe Portal functionalities, e.g. EOL. In addition, examples for portal front ends were collected and discussed with the developers.

In order to discuss relevant information with members of the BHL-Europe consortium the Google Content User Use Case Group (UCUG) was established. This is a forum to further discuss content user related issues and collect feedback from the consortium members with taxonomic expertise.

3.0.2.13   Best Practice guidelines and standards (Task 2.3.2)

One task within WP2 is assisting partners in implementation and evaluation of scanning operations and to deliver a first draft of the Best Practice Guidelines and Standards. This delivery aims to give a first standardised guideline of the workflow from the book/journal (either print or digital version) to the downloadable item within the BHL-Europe portal.

---

[20] http://gdz.goobi.org

[21] https://bhl.wikispaces.com/Use+Case+Workgroup

A further objective is to compile an easy-to-understand document for our consortium. Therefore, it was decided to involve experienced partners to contribute to the document. The Best Practice Guidelines and Standards organisational tasks took place during the second half of this reporting period. Firstly, our experienced partner and content provider LANDOE (Land Oberösterreich) was chosen as the overall coordinator of the Best Practice Guidelines. In cooperation with LANDOE several content providers were identified to contribute to the document and were asked for their contribution. As not all identified partners responded to our requests, the Best Practice Guide was also announced over the Content Provider mailing list (bhl-e.cp@lists.hu-berlin.de). All other content providers with experience of scanning and digitisation management were asked to contribute to the document. Furthermore, in order to facilitate communication between all contributors, a BHL-Wiki page was created and can be found under the WP2 Wikipage[22].

A first meeting for the Best Practice Guide will be held on 8-9 November 2010, hosted by LANDOE in Linz, Austria with the aim to bring together all participating parties and to discuss existing workflows. The outcome of the workshop will be reported at the next meeting in London in the beginning of December.

Furthermore a first draft of table of contents was established and contributors were asked to indicate chapters to which they will contribute. The table of contents and an up-to-date working document can be found on the above mentioned Best Practice Guideline Wikipage.

3.0.2.14   The Europeana Council of Content Providers and Aggregators

The Council of Content Providers and Aggregators (CCPA) is a pan-European forum which reflects the views of content holders, ranging from big aggregators to individual institutions. It represents museums, libraries, archives and the audio-visual sector across a wide range of European Member States. The structure of the CCPA is embodied in the Europeana Foundation statutes. The CCPA contributes to the decision-making of the Foundation and advises its members on strategy and policy from the content providers' point of view. Any content provider or aggregator is welcome to become a member of the CCPA. All institutions of an aggregator, as well as all individual content providers, can participate in the CCPA. Registration is available online[23].

In spring 2010, six officers were elected from the members of the CCPA to represent the four domains, libraries, museums, archives and audio-visual archives, as well as one aggregator and one commercial content provider. Three of the officers manage the CCPA. The elections and first meeting took place in June 2010 in Frankfurt/Main:

- Nick Poole, Collections Trust, UK – Chair of the CCPA
- Anne Bergman-Tahon, Federation of European Publishers – Vice-Chair of the CCPA
- Henning Scholz, Museum für Naturkunde, Leibniz Institute for Research on Evolution and Biodiversity – Secretary of the CCPA
- Francisco Barbedo, Portuguese Archives
- Kjell Nilsson, National Library of Sweden
- Hans van der Linden, Flemish Government. Art and Heritage agency

---

[22] https://bhl.wikispaces.com/BHLE_WP2_BPG

[23] http://version1.europeana.eu/web/europeana-foundation/content-council

The vision of the CCPA is to encourage Europeana and to celebrate European cultural identity by delivering rich and meaningful online experiences to the widest possible audience. This vision will be achieved by:

- Building a thriving, dynamic community of content providers and aggregators across Europe
- Connecting this community with the Board and Executive committee of Europeana
- Providing resources which identify and help overcome barriers to participation
- Reaching out to non-participants and proactively making the case for participation
- Championing Europeana to cultural institutions and their users throughout Europe

The CCPA has formed a number of working groups in order to achieve the goals set in the action plan. These working groups will mainly deliver advocacy papers, guidance papers and white papers to promote participation in Europeana and remove the barriers for contribution to Europeana focussing on different aspects of participation. The working groups are:

- WG1 – User engagement group
- WG2 – Value proposition group
- WG3 – Ethical group
- WG4 – Technical group
- WG5 – Legal group

The WG setup was agreed during the first general meeting of the CCPA in The Hague, 13 October 2010. Registration for the WG will open soon. Each WG will have a chair and a core group to manage the group and create the proposed documents.

At the moment, the CCPA has 140 members including six members of BHL-Europe. As Europeana is including more and more science content, it would be beneficial for us, to be represented in the CCPA with more people to represent the science domain adequately. Therefore, all current and potential content providers of BHL-Europe are encouraged to register for the CCPA. The CCPA will provide relevant information on Europeana and provide the chance to get involved in the decision-making process of the Europeana Foundation. With the setup of the CCPA it is not longer Europeana speaking to the contributors, but the community speaking to one another. We are all part of this community and our input, knowledge and expertise as well as our questions and problems are very welcome and highly appreciated.

3.0.2.15   *Management of the technical development (Task 3.1.1)*

Overall, the development work is divided into 12 Sprints with specific topics (SCRUM). We are currently in Sprint 4. As outlined in D3.5 - Technical architecture status and progress report with particular focus on the development of the German prototype - many of the functional requirements are defined incrementally and hence the reason why the Agile Software Development Methodology (SCRUM) is used for the distributed development. The tasks defined in the 'Scrum backlog' are translated into the Prince2 Methodology in line with the DoW for the BHL-Europe project.

The SCRUM planning work is project-managed through the backlog in order to maintain & track the progress of WP3/WP2 development tasks, which has proved to be working well. However in order to ensure that we have a consistent plan in alignment with the DoW, it was agreed that the WP3 Project Manager will be responsible for translating the work into MS Project to be published through Project Server (see also Figure 2 for a simplified roadmap).

Furthermore, the ongoing development work carried out in each of the Sprints has lead to the output of various Sprint Deliverables in the form of 'technical notes' which explains the results and analysis of outputs circulated amongst the group to give better understanding and information - this has proved to be very useful. This includes:-

- **TN_Sprint 102** – Prepare Pre-Ingest for GRIB prototype API
- **TN_Sprint 106** – Proof of Concept for LOCKSS – this note describes administrative and technical needs to implement LOCKSS for BHL-Europe.
- **TN_Addendum_DataFlow_v04**. - The note explains the dataflow and mapping implemented in the BHL-Europe German Prototype and points out architectural changes needed for the development of the Pre-Ingest Tool and the BHL-Europe community portal.
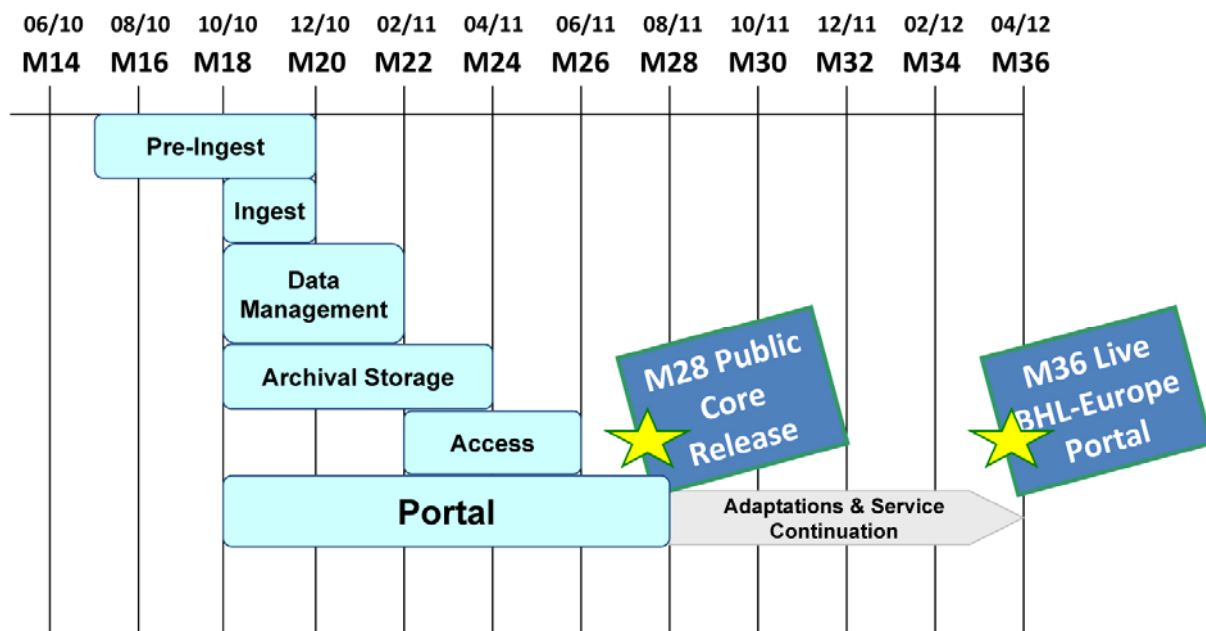
Figure 2: BHL-Europe development roadmap for the next 18 months.

3.0.2.16   Technical Board Meeting London

The BHL-Europe technical board meeting was held at the Natural History Museum (NHM) in London from 25-27 August, 2010 and was attended by all members of the technical group: AIT, Atos, NHM, NHMW, MfN, UBER and MOBOT. The main purpose of this 3 day meeting was to review the progress of the project and to plan for the next phase of the Technical Implementation work. Topics discussed included:

- Latest updates on the GRIB development and the development of the German Prototype with possible interaction with the Portal Use Case elaboration.
- Further discussions on the GUID and deciding which ID-system to use.
- Data mapping and the Schema Mapping Tool.
- Data Synchronisation and Communication and Planning were also discussed.

It was also a good forum to integrate new key players involved in the development of the system including Bernd Sproger from AIT who is currently working on the Pre-Ingest

activities and finalisation of the German Prototype in preparation for the next deliverable D3.6 Release of the German Prototype.

3.0.2.17 Technical Board Meeting Amsterdam

Following the Europeana Plenary, a TMB meeting was held in Amsterdam from 15-16 October 2010. In that meeting we discussed the outcomes of the BHL Global Meeting in Woods Hole and the implications for BHL-Europe. We agreed on the configuration of the NHM servers for the upcoming ingest process of BHL-Europe. We agreed on the next steps to finalise the metadata discussions. A GUID session was held as well, and further details on GUID are provided below this section. Finally we revised and approved the German Prototype of BHL-Europe and agreed on the final steps to prepare this deliverable for submission.

3.0.2.18 Metadata walkthrough and guidelines (Task 3.1)

AIT contributed further to the metadata discussion and provided texts and input. A first draft of a metadata walkthrough and guidelines for the content providers has been developed. This documentation will be further enhanced and will be basic helpdesk literature for content providers wanting to ingest their metadata and scans to the BHL-Europe system.

3.0.2.19 Development and adaptation of specific tools, in particular for Pre-Ingest

For the creation of a Pre-Ingest Tool Alpha 1: Analysis was carried out reviewing software libraries, tools and specifications used by Archivematica[24] (e.g. easy-extract, UUID, detox, ClamAV, NLNZ Metadata Extractor) and by California Digital Library[25] (e.g. BagIt, JHOVE, dflat, storage web service).

Achievement: Technical feasibility and suitability was evaluated from numerous sources for BHL-Europe Pre-Ingest.

Provision of an FTPS upload facility for Pre-Ingest Editor: Apache FTPS was installed and tested on server *prototype.bhle.eu*. This provided the facility to manage FTPS accounts for the content providers, enabling uploading of content. The server was provisioned with an NFS share, allowing access to 45Tb of BHL-Europe storage.

Achievement: A secure mechanism for the uploading of metadata and content was provided.

Final version of use cases for the Pre-Ingest: Research has been carried out to analyse Archivematica and Merritt Micro-Services to derive Pre-Ingest use cases based on standards (OAIS) and best-practice examples.

Achievement: Producer Use Cases, Archivist Use Cases, OAIS activities and Archivematica Micro-Service Categories have been developed.

During the iPres 2010 convention in Vienna, contact with producers of OAIS-compliant archival and preservation systems provided various opportunities for collaboration. Discussions with Archivematica provided possible co-operation regarding the introduction of

---

[24] http://archivematica.org/

[25] http://www.cdlib.org/

workflows. The California Digital Library (UC3/Merritt) provided possible co-operation within the pre-ingest development using the CDL-micro services. Tessella/Siemens provided a demonstration of the system installed at the National Archives in Austria.

Coordination with partners and data analysis: The coordination of the content providers' sample data sent for the Pre-Ingest was provided by the updating of relevant coordination overview sheets at BHLWiki and Google Groups.

To undertake the preliminary mapping each partners' metadata records were first converted to XML (if they are not in XML format). Then custom xsl stylesheets map the metadata to the METS format, which includes MODS. Logical construction of items/collections is done according to the 'Logical Construction' row. Afterwards ESE records are automatically generated from the METS records. It is intended to implement automatically enrichment processes in the mapping procedure.

Achievement: The analysis of sample data provided by CSIC, MNHN, Naturalis, RMCA, UBBI (University Bielefeld), UBER and UH-Viikki was preliminarily data mapped using the MODS/METS standard.

The implementation of high level curation and preservation workflow: The Pre-Ingest Tool Workflow was presented and discussed at the BHL-Europe Technical Group Meeting in London. The workflow serves as a basis for further development of the Pre-Ingest Tool and has been approved at the BHL Tech Group Meeting.

Achievement: Technical feasibility tests were conducted using Archivematica and Merritt Micro-Services.

Architecture: Further development were made on the technical and functional architecture of the BHL-Europe system and on integration of the GRIB with the architecture of the BHL-Europe system.

3.0.2.20 Optical Character Recognition (Task 3.1.5)

The project coordinator met with the IMPACT coordinator during the Europeana Plenary in Amsterdam on 14 October 2010. It was agreed to intensify the collaboration between BHL-Europe and IMPACT. IMPACT had recently released their evaluation tools and is looking for test cases. With our test set in hand we discussed how to adapt and rekey it for application in the IMPACT evaluation process.

3.0.2.21 Development work on the VMs hosted on NHM servers

The System Infrastructure team at NHM continue to support the configuration requirements for the ongoing development work carried out by Atos/AIT in conjunction with the setup of the development, test and integration environment which is hosted on NHM servers. This includes the installation of packaged software components (Apache for Access Part (Drupal); Apache for Data Management component; the creation of the basic UID Generator and the basic GUID Resolver. Handle.net has been configured on the servers and will be used as the GUID resolver. For the entire developer information it is referred to the BHLWiki[26].

---

[26] https://bhl.wikispaces.com/Developer_Information

3.0.2.22 Pre-Ingest test and metadata mapping

Creation of small data set (AIP) from Ingest to AS: MODS and METS AIP examples were generated based on sample data received from users. We used a preliminary METS profile and MODS schema for development. Fedora Commons was installed to test ingest of AIP example.

File submission guidelines have been checked and further refined. Revised guidelines will be published on *prototype.bhle.eu* enabling content providers to begin uploading content and metadata. The revised guidelines now cater for specific content provider needs (such as material that cannot be categorised as title or item).

A feasibility analysis was conducted for the Metadata editor used during Pre-Ingest and Orbeon Forms has been identified as a suitable solution, since there are already best-practice examples (such as MODS editors) based on Orbeon Forms.

BHL-E Dataflow & Mapping documentation and revision: During the BHL-Europe meeting in London it was suggested that content and metadata data-flows should be documented and explained for the whole project group. Documenting the flows for the whole group will make sure that dataflow and in due course metadata mappings are transparent. To facilitate the transparent flows, a technical note (BHL-E_TN_Addendum_DataFlow_v04.pdf) was created and published on the BHLWiki[27], and additionally sent to the BHL-Europe Tech Group. The technical note explains the dataflow and mapping implemented in the BHL-Europe German Prototype and points out architectural changes needed for the development of the Pre-Ingest Tool and BHL-Europe community portal.

3.0.2.23 Data management at RMCA (Task 3.2)

A RMCA ICT committee meeting was held in Tervuren on 3 June. The follwong items were discussed: hosting of digital content, including the books and monographs for BHL-Europe; global infrastructure of BHL and benefits for RMCA; overall catalogue and IT related management of RMCA library. The conclusion was that the BHL-Global network can be used as mirrors and back-ups for RMCA digital information. A local infrastructure either at RMCA or at the Belgian government, where a full copy of the digital information is kept, was discussed.

3.0.2.24 Globally Unique IDentifiers (GUID) (Tasks 3.2 & 3.3)

The requirements for data management and access & storage require each object within the data repository to be uniquely identifiable. From the original requirements, page metadata needs to be uniquely addressable for example when a search is undertaken for taxonomic key words and we return the associated OCR text from the page.

Noid was investigated for use by BHL-Europe as an identify service. A web service based on WSDL has been implemented to test creation of identifiers ('minting'). The handle.net API has been investigated and it has been confirmed that Noid is fully compatible with the handle.net's Prefix system. Handle.net is going to be used to create GUIDs out of Noids.

It is still under discussion what the preferred solution for BHL-Europe GUID's will be. The technology group is considering various options and systems. Currently, the best fit identified for BHL-Europe is utilising the open source GUID mint called "Handle"[28]. This open source software is used by other global ID systems, such as the Digital Object Identifier (DOI)[29].

---

[27] https://bhl.wikispaces.com/Minutes+-+Thursday+26th+August

[28] http://www.handle.net/

On Oct 13, Chris Freeland had a meeting in Oxford with Ed Pentz and Geoffrey Bilder from CrossRef to discuss assigning DOIs to our content. The outcomes of this meeting helped out clarifying some doubts about DOIs and sparked a fruitful discussion through the BHL-Europe list and call to help decide later on how to move forward assigning GUIDs for the project elements.

3.0.2.25   Addressing distributed access and storage – long-term sustainability (Task 3.3)

In May this year the hardware infrastructure for BHL-Europe was built and configured, and is hosted at NHM London. Discussions are currently underway regarding the design of a larger data centre (Slough) so that long-term sustainability of the data is secured.

3.0.2.26   BHL-Europe German Prototype (Task 3.4.2)

Following the Europeana Plenary in October, BHL-Europe held a technical review meeting to discuss developments and to finalise the German Prototype.

D3.5 was prepared by ATOS in April 2010 which describes the key components needed within the system as well as the technical architecture status and the development of the German Prototype. The planning for the system in terms of identifying the steps to develop and implement the prototype system was also confirmed in this document. AIT enhanced the German Prototype (integration of Solr-Indexer) as documented for D3.5.

A first version of the prototype was recently deployed on the new BHL-Europe infrastructure at NHM[30]. Development work for the German language prototype is currently progressing well. The prototype will be delivered together with this progress report. The work is being done by AIT, ATOS, NHM and MfN.

The distributed development of the system is being managed using the SCRUM Agile software development approach. The development work has been split into 12 sprints each lasting 2 months. Conference calls of the development team every two weeks were held to discuss progress and next steps. The development team is progressing well in delivering the required outputs.

3.0.3 Status and progress of activities – WP4

3.0.3.1 Overview

Activity in WP4 during this period has focussed upon metadata and data sharing, and drawing distinction between copyright-protected material and data/metadata (which, although unlikely to be protected by copyright, may be protected by sui generis EU Database Right).

WP4 leader has had involvement with Europeana and the continued development of their Data Provider and Data Aggregator agreements through to the realisation of the future of Linked Open Data. This work impacts D4.1 agreements with BHL-US and Europeana for reciprocal access and Rights metadata.

In summary, Europeana has recognised a need to move away from a non-commercial licence model for data (as currently exists in the Data Provider Agreements), and acknowledges the need for all data to be made available as Linked Open Data, to maximise re-use, innovation

---

[29] http://www.doi.org/

[30] http://prototype.bhl-europe.eu/

and potential. To this end WP4 leader of BHL-Europe attended Europeana's 'Your Metadata on the web' workshop, led by Jill Cousins, with representation from Patrick Peiffer and Paul Keller and Wikipedia representatives among others, to ascertain how this will impact BHL-Europe as an aggregator to Europeana. This workshop prompted the need for a review of BHL-Europe documentation including the MoU for new partners to ensure compatibility with the Linked Open Data model.

WP4 leader has also been researching data licensing models, including monitoring development of new UK Government Open Data Licence for Crown and other public body data. This licence is compatible with Creative Commons Attribution licences so may potentially be adapted for use by BHL-Europe.

WP4 leader has reviewed Europeana's draft User Generated Content Policy as part of activities for the Users Working Group.

3.0.3.2 Rights clearance

We currently discuss how we might sensibly try to plan rights clearance in such a way that there is as little duplication of effort as possible when contacting publishers and other rights owners. We currently consider using GRIB and deduplication of library catalogues to inform this. The detailed procedures need to be worked out in the near future.

3.0.3.3 IPR survey and preparation work (RMCA)

BHL-Europe content will be available under Open Access terms. It means that content is not only free to access but also free to re-use. There are three groups of risk material: low, medium and high. RMCA has very few literature from the first group and a little from the second group. For many of the publications the copyright status should be cleared at first (identification of rights owner, finding the rights owner if possible and obtaining permission from the rights owner). Thereupon the list of authors and third party rights holders (photographers, artists etc.) has been created and is constantly updated. Rights owners are traced via different sources (local databases of ethnographic departments: TMS and DE; internet), all available information was collected. If the date of birth and death of author is determined, it is added as MARC tag in the catalogue. A template permission letter to the rights owner was compiled by Patricia Mergen, Lutgard Kenis and Isabelle Gerard. Some permissions have already been obtained.

3.0.4 Status and progress of activities – WP5

For more details on certain activities and an overview of awareness and dissemination activities including conferences attended by BHL-Europe members see section 4 below.

3.0.4.1 CWG meeting in Vienna

During our annual meeting in Vienna on May 25 we had a dissemination/communication workshop. We prepared strategies to reach public users and aligned our activities with Europeana. We discussed the preparation of press releases about cooperation and content provision between BHL-Europe and Europeana in the context of the Rhine Release for this autumn. Furthermore, we discussed possibilities to promote BHL-Europe on Facebook or Wikipedia. We also started to prepare e-exhibitions and short stories about extinct or endangered species. We started the discussion for the product branding of BHL-Europe. The Europeana communication tools might serve as an example on how to proceed.

3.0.4.2 User requirement survey

BHL-Europe carried out an extensive user requirement survey at the end of the last reporting period (15 March to 3 May 2010). Announcements for the user survey were sent to various institutional lists to reach a wide range of users[31]. The results of this survey were evaluated in the first half of this reporting period and published as D5.8 in the first days of August 2010[32]. The outcomes of this survey influenced to a large extent the use cases discussion in WP3, thus being a valuable source of information for the development of the BHL-Europe Portal.

3.0.4.3 Project publications in the current reporting period

Gilissen, T. 2010. NCB Naturalis literatuur in Europa verspreid.- Internal newsletter.

Gilissen, T. 2010. Zoektips voor gebruik in BHL.- Internal newsletter.

Götze, M. & Scholz, H. 2010. Tagung der Wegbereiter des Digitalen Kulturerbes.- press release, 30 September 2010.

Hoffmann, J. & Scholz, H. 2010. BHL-Europe: Biodiversity Heritage Library for Europe.- In: Nimis, P.L. & Vignes Lebbe, R. (eds.). Tools for identifying biodiversity: Progress and problems.- pp. 43-48.

Koskinen, K.; Kärki, S. & Oker-Blom, T. 2010. BHL-Europe – Biologisen lajitiedon ja kulttuuriperinnön keidas (BHL-Europe – an oasis of biological knowledge on species and cultural heritage).- Electronic newsletter Verkkari.

Mallaerts, T. 2010. Biodiversity Heritage Library for Europe towards a global digital library of life/Un patrimoine scientifique mondial à la portée de tous!.- Bulletin de la Société Royal Belge D'Entomologie/ van de Koninklijke Belgische Vereniging voor Entomology, 146(I-IV):1-74.

Scholz, H. 2010. Eine globale Bibliothek des Lebens – A Global Library of Life.- MfN Annual Report for the year 2009: 16-17.

Valdecasas, A.G. 2010. La Biblioteca del Patrimonio Natural (BHLE).- FECYT, Fundación Española de la Ciencia y la Tecnología, En: Unidad Didáctica sobre Biodiversidad.

3.0.5 BHL – Extract from the Program Director's Report (Tom Garnett)

Chris Freeland and Martin Kalfatovic met in Vienna with Tod Stussy, the Director of the International Association of Plant Taxonomy (IAPT) to discuss the planned digitisation of the *Taxonomic Literature, 2nd edition* (TL-2). TL-2 is the standard reference work for plant taxonomic literature from Linnean times to 1940. Martin is preparing a permissions document. SIL has submitted a proposal to the Atherton Seidell Endowment for scanning and correcting of the OCR texts as well as parsing the results for further processing.

From June 1 - 4 May, Chris Freeland, Martin Kalfatovic, and Anthony Goddard attended a series of meetings in Australia for the establishing of a BHL-Australia. The project includes cooperation of the Atlas of Living Australia and Museum Victoria. A MoU was signed by all parties. The Australia partners have already implanted a prototype site. The agenda is on the BHLWiki[33].

---

[31] https://bhl.wikispaces.com/Survey+Targets

[32] https://bhl.wikispaces.com/file/view/BHL-E_5pt8_100731.pdf

[33] https://bhl.wikispaces.com/BHLAu_June2010

On June 27, the Association for Library Collections and Technical Services awarded the BHL their award for Innovation in Outstanding Collaboration at their annual meeting during the American Library Association in Washington D.C. On June 28, the Smithsonian Institution Libraries hosted a well-attended meeting for ALA attendees with presentations by Grace Duke, Connie Rinaldo, Bianca Lipscomb, Jane Smith, Suzanne Pilsk, Matthew Person, Chris Freeland, and Tom Garnett. Many very positive responses were provided by the attendees.

One of the most prevalent themes throughout the range of responses to the Survey was that our users want to be able to submit requests for scanning. Thanks to the contributions of multiple staff, especially Mike Lichtenberg at MOBOT, we now have on the BHL Portal a new scanning request form[34].

---

[34] http://biodiversitylibrary.blogspot.com/2010/06/new-feature-user-submitted-requests-for.html or http://bit.ly/anJW6b

3.1 Resources employed

Resources employed for the reporting period (person-months)[35]						
Beneficiary short name	WP 01	WP 02	WP 03	WP 04	WP 05	TOTAL
MfN	4.74	7.79	1.22	0.00	1.47	15.22
NHM[1]	0.01	0	3.44	0.98	0.05	4.48
NMP	0	0	0	0	8.47	8.47
EDLF[2]	0.16	0.32	3.58	0	0.16	4.22
AIT	0.13	0	9.28	0	0	9.41
ATOS	0	0	2.01	0	0	2.01
FUB-BGBM	0.06	0	1.67	0	0	1.73
UGOE	0.09	1.01	0.22	0.01	0.68	2.01
NHMW	0	0.30	3.91	0	0	4.21
LANDOE	0.14	3.59	0.13	0	0	3.86
HNHM[3]	0	5.31	0	0	0.11	5.42
MIZPAS[4]	0.07	0.21	0	0	0	0.28
UCPH	0	0	0	0	0	0
NAT	0.02	0.60	0.02	0	0.05	0.69
NBGB	0.06	0.37	0.21	0.01	0.19	0.84
RMCA	0	3.41	1.24	1.13	0.27	6.05
RBINS	0	1.52	0	0	1.44	2.96
BnF	0.03	2.47	0.08	0	0.03	2.61
MNHN	0	2.83	1.93	0.03	0.05	4.84
CSIC	0.09	1.20	0.05	0	0.18	1.51
MSN	0	0	0	0	0	0
RBGE	0.06	3.10	0	0	0.04	3.20
Sp2000	0	0	0	0	0	0
Wiley	0	0	0	0	0	0
SIL	0	0	0	0	0	0
MOBOT	0	0	0	0	0	0
UH-Viikki	0.07	1.35	0.01	0.08	0.18	1.69
UBER	0	1.58	1.75	0	0.12	3.45
TOTAL	5.73	36.96	30.75	2.24	13.49	89.16

[1] NHM erroneously included subcontractors in the activity reports, which results in higher numbers for the hours of work/ person months. The hours contributed by subcontractors to the project are now excluded.

[2] EDLF reported some of the task to a wrong work package. Only the distribution of the hours changed, hence the total of the person months remains the same.

[3] HNHM erroneously did not include the institutional work contribution in the activity reports, which has been corrected here.

[4] MIZPAS did not report correctly for month 13 (May 2010) of the project with the activity report, which results in a slight change in person month in WP1.

[35] The calculation of PM is based on productive hours mentioned in CPD form of partners.

3.2 Work Package Overview

<table>
<tr><td colspan="6" align="center">Work package description</td></tr>
<tr><td>Work package number :</td><td>1</td><td>Start date:</td><td>M1</td><td>End date:</td><td>M36</td></tr>
<tr><td>Work package title:</td><td colspan="5">Project Coordination and management</td></tr>
</table>

Objectives for the period

Administrative objective: Ensure adherence of the consortium to the rules, regulations, and financial guidelines of the *eContentplus* programme; establish the project in the European biodiversity community.

Technical objective: Implement the project as set out in the work plan; ensure exchange of information and communication between partners; ensure progress of the project; guarantee timely deliverables.

QA objective: Ensure verifiable progress of the project and be committed to high quality output that has tangible impact on *eContentplus* programme objectives.

Description of work carried out and achievements

Task 1.1 – Administrative coordination

1.1.1 – Resource planning (financial, personnel, material), monitoring and controlling.

Achievements: For the technical development of the project (WP3, Task 2.1), task briefs and SCRUM are used. The backlog is in Google docs so all team members can access it. This process is established and well integrated in the work planning.

1.1.2.1 – Liaison between the Commission, consortium members, EUROPEANA, BHL, and external experts; effective communication with the consortium members, Work Package leaders, the Commission, and interested external parties; coordination of meetings and progress reviews.

Achievements: Email, phone, Skype, BHLWiki, and regular personal meetings ensure effective communication between all parties involved in BHL-Europe and related to the work carried out in the project. Every Tuesday, the Project Management Group of BHL-Europe has a conference call to update on progress and to discuss recent developments and issues. The IT team has a conference call every second Wednesday to review work progress and adjust the SCRUM planning. For special purposes and working groups, Google Groups is used to ensure effective communication. Furthermore, mailing lists were established using facilities available at UBER to facilitate the communication of the content providers. Internal reports are produced quarterly to summarise the project progress for all consortium members. PCO of BHL-Europe is communicating with Europeana on a very regular basis on various levels (e.g. project management, content ingestion, CCPA). Since autumn 2010, the

collaboration of the Europeana Group of Projects was intensified as documented by the Group of Projects meeting in The Hague and the German Europeana conference in Berlin.

1.1.2.2 – Coordination of meetings and progress reviews.

Achievements: One contractual project meeting was held in this phase, in Vienna from 25-30 May 2010. Another large meeting was held in Vienna from 23-27 March to discuss content provider requirements and use cases. Three TMB and technical meetings were held in addition to further discuss WP2 and WP3 developments, review work progress and plan next steps.

1.1.3 – Production and consolidation of periodic external reports, including cost-statements; internal quality assurance; set-up of the Web-based project management portal.

Achievements: There are no specific achievements in this period as this is an ongoing task.

1.1.4 – Representing the project.

Achievements: The project was presented at various conferences and meetings (see below). The Project Coordinator is now secretary of the newly established Council of Content Providers and Aggregators of Europeana.

Task 1.2 – Technical coordination

1.2.1 – Work-package and task coordination.

Achievements: This is an ongoing task using the tools and services implemented before (Project Server, BHLWiki, Google groups).

1.2.2 – Project plan maintenance; monitoring of project progress and milestones; identification and trouble shooting of technical and organisational problems.

Achievements: This is an ongoing task using the tools and services implemented before (Project Server, MS Project). To include the SCRUM planning work, a project manager of the WP3 leader at NHM is translating this into MS Project to have a consistent master plan. This procedure is now fully established and the WP2 and WP3 development process is managed with SCRUM.

1.2.3 – Timely production of deliverables.

Achievements: The Project Office of BHL-Europe is the final instance for ensuring timely production of deliverables. In cooperation with the WP leaders, the members of the consortium were asked to fulfil their commitments according to the Description of Work.

1.2.4 – Quality control against the technical and contractual aspects; coordination with EDL Foundation Office.

Achievements: We now have a more rigorous internal reporting procedure in place. Partners have to report all their activities on a monthly basis and provide a narrative report of their activities on a quarterly basis. This way we ensure partners work as planned and we are able to identify issues as early as possible.

Task 1.3 – Quality assurance: Definition and communication of quality assurance procedures via project management portal, coordination of quality assurance process (internal as well as external review procedures for the various project results and deliverables).

Achievements: External reviewers are identified for most deliverables and will be consulted to review the deliverables of the project.

Task 1.4 – Business plan

1.4.1 – Investigate services that help the sustainability of project results

Achievements: BHL-Europe is currently working on the "First overview of business plan for long-term sustainability" (D1.5) for M24 of the project. A closer collaboration with Europeana and the Group of Projects was established as this is a very important task of all larger projects and aggregators. The Europeana CCPA is also investigating options for aggregators. This task is an agenda item for the next BHL-Europe Content Provider and Technical Meeting in early December. All input will be used to further elaborate the ideas presented in D1.2.

1.4.2 – Identify related networks or organisations interested in the implementation of project results

Achievements: BHL-Europe is currently in discussion with two new EU projects of the biodiversity domain to investigate areas for collaboration or options for implementing project results, i.e. NaturalEurope and OpenUp!. The closer alignment of the Europeana Group of Projects since autumn 2010 will potentially result in new opportunities for the implementation of project results.

1.4.3 – Develop a business plan for long term sustainability with WP2 and WP3

Achievements: See above under Task 1.4.1.

Task 1.5 – Networking and clustering activities

1.5.1 – Coordinate with relevant partner networks; Identification of new important networks as potential disseminators of project results; Identification and analysis of stakeholders

Achievements: The most important partner networks currently work are Europeana (data ingest), EDIT (ViTaL implementation), BHL (alignment of standards and technologies, content ingestion procedures), and Vifabio. In addition, the collaboration of BHL-Europe with the numerous projects of the Europeana group is now much better established after a number of meetings and conferences in September and October (Group of Projects Meeting, Europeana conference in Berlin, Europeana Plenary with side events).

1.5.2 – Contribute to the clustering activities of the EC.

Achievements: No cluster meetings of the EC were held.

Deviation from work plan & remedial action
No deviations from work plan.

Work package description

Work package number :	2	Start date:	M1	End date:	M36
Work package title:	Analysis of domain content and management of the content acquisition process				

Objectives for the period

IT objectives: Establish bibliographic database system, metadata repositories and Web-based content management system.

Management objectives: Ensure that all content providers agree on the technical architecture of the project; ensure effective scanning in all content providing institutions; ensure linkage to Europeana; ensure extension of the content providing network.

Description of work carried out and achievements

Task 2.1- IT Development

2.1.3 – Develop a Web database to support analysis of domain content and management of the scanning process based on 2.1.1 and 2.1.2

Achievements: The GRIB prototype has a new URL[36] and has been enhanced with several functionalities in cooperation with EDIT. The GRIB provides a SRU-Interface for data export, an enhanced digitisation management tool with six different digitisation statuses. Further work has been started to equip the GRIB with the EDIT SSO-System (Single sign-on). Metadata from BHL has been imported allowing to link to BHL content. A newsfeed for information on the GRIB is available[37] and the index is also featured at the Europeana ThoughtLab[38].

Tasks 2.2 – Analysis of domain content

2.2.2 – Use of the Web-database to identify the distribution of this relevant literature in the libraries of the content providers.

Achievements: Several partners are running individual processes to establish a list of relevant literature. RMCA, NBGB and RBINS collaborate on a local Wiki based scan list. Further partners (LANDOE, NAT, RBGE, UH-Viikki) updated the BHL ScanList. To date, the GRIB is in prototype stage and some BHL-Europe content providers have already provided their bibliographic catalogue for import to the GRIB. Once the GRIB reaches the productive stage, the BHL ScanList will be merged into the GRIB.

Task 2.3 – Management of the content acquisition process

---

[36] http://grib.gbv.de/

[37] http://twitter.com/editvital

[38] http://www.europeana.eu/portal/thoughtlab.html

2.3.1 – Identify content holder requirements

Achievements: A use case working group has been established and a standardised template for recording use cases was developed. For this purpose a BHLWiki page[39] was created and a Google group has been set up.

2.3.2 – Assist partners in implementation and evaluation of scanning operations; control duplicate scanning of literature (using results of task 2.2); discussion and distribution of data standards and specifications; work with individual donors and governments to facilitate the funding of the scanning; attracting new content providers.

Achievements: To assist partners in implementation and evaluation of scanning operations, we started to evaluate the scanning workflow management tool GOOBI. This task started during the second half of the reporting period and is one main task of our partner UBER during the next reporting period.

In order to give a standardised guideline of the workflow from the book/journal either as print or digital version to the downloadable item within the BHL-Europe portal, we started with the work for the delivery D2.6. Up to now we created a Wiki page with all necessary information regarding the Best Practice Guide. A first draft of the table of contents already exists and the first face-to-face meeting will be hold on 8-9 November 2010, hosted by LANDOE in Linz, Austria with the aim to bring all participating parties together and to discuss already existing workflows, already written chapters and the assignment of each chapter to the participating partners.

A new approach for content acquisition was identified in which our consortium will be part of. For this purpose a "BHL-Europe in a nutshell" document is currently in process and is expected to be finished soon. This document aims to be a very brief but informative document including the most important facts about BHL-Europe, benefits and IPR issues.

2.3.3 – Coordinate with EUROPEANA, BHL and national scanning projects to ensure that material scanned by BHL-Europe is available through these portals.

Achievements: To date a corpus of totally 89,860 objects is available in Europeana from BHL-Europe and can be retrieved by using the following search string: *europeana_collectionName:087**.

2,568 digital objects and metadata files have been prepared for Europeana ingest by LANDOE and in September 2010 this data set was made available in Europeana. Already-existing data sets in Europeana (BHL & NCB Naturalis) were updated to the latest version of the Europeana Semantic Elements (ESEv3.3). BHL-Europe was the first aggregator to publish data in Europeana with the new europeana:rights field.

2.3.4 – Addressing IPR issues in cooperation with WP4.

Achievements: The "BHL-Europe in a nutshell" document addresses IPR issues for prospective BHL-Europe content providers. This document is created in collaboration with WP4.

Deviation from work plan & remedial action

---

[39] https://bhl.wikispaces.com/Use+Case+Workgroup

To date there is no deviation from work plan.

Work package description

Work package number :	3	Start date:	M1	End date:	M36
Work package title:	Technological implementation				

Objectives for the period

Ensure German prototype objectives are met. Package technologies to provide functional demonstration of German Prototype. Consider prototype refactorisation requirements for actual portal.
Ensure the development of BHL-Europe data model is consistent with ESE (Europeana Semantic Elements).

Description of work carried out and achievements

Task 3.1 – Technological implementation (Overall Coordination)

3.1.1 – Management of the technical development team.

Achievements: In the reporting period, three meetings of the technical team were held (Vienna, London, Amsterdam) to review the progress and to decide on the next steps. Over the past months, the technical group has been holding regular conference calls on a fortnightly basis to discuss and exchange ideas, report on task updates and on the overall development progress. The SCRUM planning work is project managed through the backlog in order to maintain and track the progress of WP3/WP2 development tasks, which has proved to be working well.

In order to ensure that we have a consistent plan in alignment with the DoW this is now translated into MS Project to be available via the Project Server and the Wiki for our partners. The ongoing development work carried out in each of the Sprints has lead to the output of various Sprint Deliverables in the form of 'technical notes' which explains the results and analysis of outputs circulated amongst the group to give better understanding and information.

3.1.2 – Adaptation of EUROPEANA and BHL data model, workflow, harvesting procedure, standards, specifications.

Achievements: Europeana is validated since June 2010 and published now. BHL-Europe was represented in this process by AIT. Currently consideration is being given on how to provide extensions for EDM to incorporate BHL-Europe requirements. AIT is also our partner with the best connection to Europeana for questions related to content ingestion based on the experiences with the Dismarc and EuropeanaLocal projects.

AIT is aligned very closely with the Europeana content ingestion team to make BHL-Europe content accessible via Europeana. We successfully established a harvesting procedure and have a first set of data submitted to Europeana.

3.1.3 – Definition of own standards for images, metadata, harvesting according to partner requirements (consensus-building).

Achievements: We are currently in progress to finalise the BHL-Europe schema.

3.1.4 – Liaise with scanning centres of the national initiatives for post-processing of content.

Achievements: Based on the decision during the content provider meeting in Vienna (March 2010), we continuing to sample page image data from our partners to test harvesting procedures and define the effort for post-processing of content. 14 of our content providers have so far provided data for this test[40].

3.1.5 – Development and adaptation of specific tools; implementation and adaptation of taxon finder and name recognition tools; improvement and implementation of OCR techniques.

Achievements: Additional features for the Schema-Mapping-Tool developed by NHMW, like the XSOM parser were finalised. AIT contributed further to the metadata discussion and provided texts and input. A first draft of a metadata walkthrough and guideline for the content providers has been developed. We are currently investigating the options for improving the name recognition tools. We now have the OCR test-set to be used for testing OCR precision using various tools. We are also beginning to investigate two tools enabling OCR correction by users.

Task 3.2 – Technical integration with EUROPEANA, BHL and national platforms

3.2.1 – Ensure that the standards for data management and image formats are consistent with international requirements and specifically the EUROPEANA and BHL, adoption of EUROPEANA Semantic Elements (ESE).

Achievements: BHL-Europe data was mapped successfully to ESE and was harvested by Europeana for the Europeana Rhine. Publishing date for the data update including an update according to ESE v3.3 (BHL-US, Naturalis, LANDOE data: Σ 89.860 records > Object type: text; item/book level) in Europeana was the last week of September 2010.

3.2.2 – Develop a distributed data model that will allow countries to retain control of their data, while enabling the material to be available through EUROPEANA, the BHL Portal (with Taxonomic Intelligence tools), and any national portals.

Achievements: A LOCKSS test installation has been examined and a technical note (BHL-E_TN_sprint01_106_v01.pdf) was developed which describes administrative and technical needs to implement LOCKSS for BHL-Europe (proof of concept). This was further discussed during follow up meetings in London, Woods Hole and Amsterdam.

3.2.3 – Build a prototype distributed data system for the German language material, and integrate with the BHL Portal.

Achievements: The prototype has been built. The functionality is consistent with being a prototype and refactorisation will be required for the final portal.

Task 3.3 – Addressing distributed access and storage – long-term sustainability:
Develop a distributed access and storage system to enable national and international storage of the scanned materials; to develop the storage system in such a way that

---

[40] https://bhl.wikispaces.com/BHL-E_WP3_PREINGEST

long-term sustainability of the data is secured.

Achievements: The transfer of data from Woods Hole (BHL-US) to London (BHL-Europe) using portable hard drives proved to be both economical and viable; to date 7 TB of data has been transferred and completed successfully (zero failure rate).

Task 3.4 – Enabling BHL Portal access in European languages – interfaces, usability and mobility

3.4.1 – Work with the EUROPEANA to create multiple language access to the BHL Portal.

Achievements: We use the experiences of projects related to Europeana that already have multilingual interfaces. Multilingual components from Dismarc were refactored to provide functional German Prototype platform technology. The key IT development partner of Dismarc (AIT) is also a key development partner in BHL-Europe. The German Prototype provides a functional framework to demonstrate the various technologies working together. The final portal may look very different, as the technologies in the final portal will need to be refactored from the prototype.

3.4.2 – Create a prototype portal in German to allow access to the BHL Portal.

Achievements: The documentation for the German Prototype has been further developed and is finalised for deliverable D3.6 Release of the German Prototype. Use cases fully document the German prototype system (name, actors, use case identifier, description, and screenshots). The German Prototype user guide is finalised, which describes the implemented technologies and German Prototype user interface interaction. The prototype is published on the NHM servers[41].

3.4.3 – Apply the model to the key European languages (English, French, German, Italian, Portuguese, Spanish, Danish, Dutch, Czech, Hungarian, Polish, Swedish), enabling access to the BHL Portal throughout Europe.

Achievements: Preparation of Use Cases for BHL-E Community Portal: The creation of use cases was started based on the results of the User Survey (see D5.8) and input from the use case working group. The BHL-Europe prototype use cases have been merged with the list of all use cases. Wording has been changed from 'use case' to feature, since the list of use cases was in reality a list of features. Features have been categorised as core features or been put on the wish list. Core features are requirements for BHL-Europe, which are defined by and within the DoW (BHL-Europe contract).

Deviation from work plan & remedial action

WP3 is well on track with no deviations from the work plan.

---

[41] http://prototype.bhl-europe.eu/

Work package description

Work package number :	4	Start date:	M1	End date:	M36
Work package title:	**Intellectual Property Rights**				

Objectives for the period

Ensure deliverables for WP4 meet the needs of the project. Ensure that BHL-Europe, EUROPEANA, and the BHL are using common approaches and common agreements, such that data can be exchanged between these partners without further Rights activity.

Keep abreast of legal developments that may help or hinder the ability of project partners to supply legal content to the project.

Description of work carried out and achievements

Task 4.1 – IPR Framework: Establish IPR working documents – including best practice guide, due diligence guide, pro-forma agreements, and process for formally agreeing rights management with rights holders; align the approach with EUROPEANA and BHL and maximise level of interoperability including Rights metadata exchange.

Achievements:

- WP4 leader attended M12 meeting reinforcing the principles of the IPR working documents delivered in M9 to partners, particularly in relation to Risk Management and licensing solutions.
- WP4 leader has been working with Europeana colleagues on the development of the next phase of data licence to Data Providers and Aggregators. This has involved participation in 'Your Metadata on the Web' workshop in The Hague and investigation into Linked Open Data as a suitable model for Europeana to adopt with its data providers and aggregators in order to maximise the potential of the data in Europeana and the functionality of the portal. WP4 leader also participated in the Europeana Users workgroup with a view to understanding and representing the needs of users and how different data models impact the potential user experience of the site.
- Study of the suitability of the new UK Government Open Data Licence as a model for licensing data between BHL-Europe, Europeana and BHL projects. This licence is compatible with Creative Commons copyright licences so a similar licence is potentially suitable for use by BHL-Europe.
- WP4 leader has been studying the suitability of the new Creative Commons Public

Domain Mark, in the light of its promotion by Europeana at the 2010 Plenary. Advice has been given to some partners about the restrictions on use of this mark for digital content created by UK or other common law countries which have a low threshold of originality for a work to be protected by copyright.

Task 4.2 – IPR Agreements with Data Providers: Complete formal IPR agreements with data providers identified in WP2; keep records of all agreements secure.

Achievements:

- WP4 leader has working with WP2 leader to clarify data responsibility, submission of summary responsibilities for new Data Providers for inclusion in WP2 Guide for new Data Providers.
- Project partners have been progressing in identifying public domain material and material in need of rights clearance within their collections. NBGB, MNHN, RBGE, RMCA, UGOE and UH-Viikki in particular have begun negotiations with rights holders (journal publishers and individual authors) with positive responses to date.
- CSIC managed to get the permission to digitise the publications of the Fundación Otonga (a series of monographs on Birds, Mammals, Insects, etc from Ecuador).

Task 4.3 – IPR Agreements on projects results and outcomes: Develop IPR framework for long-term sustainability and long-term access to the digitised content of BHL-Europe, and ensure that EUROPEANA and BHL have long-term access to the material and associated Rights.

Achievements:

All EC, UK and International consultations and debates contribute towards sustainability and long-term access. Included in this are the following that WP4 is abreast of/has contributed to:

- Lobbying Jose Manuel Barroso, President of the European Commission, regarding the proposed amendment to the EU Copyright Directive to facilitate the legal use of 'orphan works', with the objective of broadening the scope of the amendment to include artistic and audio-visual works as well as the literary works already under consideration. News suggests this has been successful and that the proposed amendments will now include such works. A draft amendment to the Directive is to be ready by end 2010. Progress is watched with interest.
- DACS UK proposal to provide collective licensing solution to Orphan Works.
- Continued research into protection of databases (original databases as copyright works, the sui generis database right in following cases such as Football Dataco Ltd v Brittens Pools)
- Attendance at ORGCON (Open Rights Group conference) in July 2010 where the UK Digital Economy Act (DEA) was discussed in regard to its impact on public sector intermediaries and the provisions for blocking access to websites that may contain infringing content.
- Representing the views of museums, libraries, archives and galleries in responding to the UK OfCOM consultation of the practical implementation of the DEA.
- Consideration of the long term EU legislative developments and to inform sustainability of the project by attendance at the lecture "The Future of

European Copyright and the potential development of a European Copyright Code" by Professor Lionel Bently (Professor of IP at University of Cambridge)

Deviation from work plan & remedial action
D4.1 Data-sharing agreements with BHL and Europeana are yet to be completed. The Europeana Data Aggregator Agreement will fulfil the purpose and function of an agreement with Europeana and it is unlikely that there will be a need for a separate agreement with Europeana in addition. The Aggregator Agreement developed by Europeana will be superseded by an agreement allow for Linked Open Data (LOD) as is desired and being led by Europeana, consequently delivery schedule is not within BHL-Europe's control. WP4 leader has contributed to the debate about LOD with Europeana and continues to do so.
A data sharing agreement with BHL colleagues in the US will require legal input from Farrer & Co due to international dimension. This will be co-ordinated by WP4 leader. The amicable and continued co-operative working with BHL colleagues in the US means a minimal risk of delay to the completion of the document.

Work package description

Work package number :	5	Start date:	M1	End date:	M36
Work package title:	Dissemination, Exploitation and Evaluation				

Objectives for the period

Dissemination objective: Raise awareness, understanding and action of the project among the community and stakeholders in EU member states; ensure effective dissemination of project goals and results to the target users; ensure good communication within the European scientific community

Presentation and demonstration objective: Ensure dissemination of project results at conferences, public events and among the networks of the consortium members

Evaluation objective: Monitor the level of use of BHL; survey the users of the BHL Web Portal

Description of work carried out and achievements

Task 5.2 – Develop and establish Web environment

5.2.1 – Plan, design, publish, maintain, and update BHL-Europe Web site for internal and external communication with publication of all relevant results and links to partner networks.

Achievements: The BHL-Europe Web site was moved into a CMS (Joomla). This will also help to manage the available multilingual versions of the site. Many texts are now translated into German, Spanish and Czech. Translations to French and Italian are in progress. The transfer to the CMS also improved the search functionality and facilitates the management of authors of Web site contributions.

Task 5.3 – Develop and prepare promotional materials and newsletters, press relations

5.3.1 – Develop, prepare, and update target group specific promotion kits (fact sheets, flyers, posters, presentations) and related printed matters (e.g. illustrated book on best practice guidelines and standards for the public).

Achievements: For conferences, we prepared a number of target specific posters throughout the reporting period. In cooperation with Europeana, we prepared a concept for a virtual exhibition. Furthermore, draft promotion material was prepared for dissemination of BHL-Europe. Among these drafts are comics about the extinct Dodo.

5.3.2 – Prepare a quarterly newsletter to be published online.

Achievements: Newsletter No. 5 was finalised in August and is now combined with No. 6 to be published in November 2010. The summer was not very busy so this combination was considered to be more effective for the dissemination of news of BHL-Europe and related projects.

5.3.3 – Ensure proper press relations; maintain the press review; distribution of articles for various types of print and online media (journals, newspapers, blogs).

Achievements: We are currently preparing material to be presented on various online platforms. The Facebook site went live recently with updates on BHL-Europe dissemination activities[42]. During the Europeana Plenary we were able to discuss with Liam Wyatt (Wikipedia) the presentation of BHL-Europe in Wikipedia. A Weblog was set up, and will shortly be formalised and launched.

Task 5.4 – Demonstration and awareness raising

5.4.2 – Maintain and update the database of conferences and public events relevant for BHL-Europe; identify consortium members to present the project at selected events.

Achievements: Currently the events are managed on the BHL Wiki[43]. We are currently implementing a feature in our CMS for the Web site to be used for the management of conferences and public events relevant for BHL-Europe.

5.4.3 – Alignment of public awareness activities of related projects (e.g. EDIT).

Achievements: We aligned our public awareness activities with related projects. The Europeana Group of projects are collaborating more closely since this autumn. One example is the conference of the German Europeana partners co-organised by BHL-Europe. As another example, BHL-Europe prepared a press release about Europeana activities in the Czech Republic and a workshop on digitisation in Brno (both in cooperation with Athena). The workshop will show the contribution of Czech

---

[42] http://www.facebook.com/pages/BHL-Europe/151086001600041?v=wall

[43] https://bhl.wikispaces.com/dissemination+of+project+by+partners

museum and library workers to BHL-Europe and Athena. Europeana is communicating BHL-Europe within the Europeana network (also in public events like the Europeana Plenary) and Europeana within BHL-Europe network. Europeana is maintaining the information page for BHL-Europe on the Europeana Group page[44].

5.4.4 – Organise public events and talks.

Achievements: No BHL-Europe specific public events and talks were organised so far - this activity will take place mainly in the last year of the project. However, BHL-Europe co-organised a conference of all Europeana partners in Germany. This was a public conference with about 200 participants. With press releases and radio interviews, this conference generated good media feedback.

5.4.5 – Represent BHL-Europe at conferences and special events (responsibilities depend on conference location and theme).

Achievements: We presented BHL-Europe in numerous conferences to scientific public (see table of attended conferences below) using presentations, posters and flyers (see our slide library[45] and posters[46] on the Wiki). We also used these events for networking and promoting BHL-Europe in individual talks and discussions.

Task 5.5 – Plan and organise user evaluation activities

5.5.1 – Put in place Web-based evaluation tools to survey users (to determine target group of the users, country of the users, most interesting content, page views, etc.).

Achievements: As BHL-Europe has no portal to date, we monitor the usage of the BHL-Europe content through Europeana using the usage statistics of the BHL Portal as a proxy (see also section 3.4 below).

5.5.3 – Establishing user testing groups; evaluation of this group using interaction logging.

Achievements: As BHL-Europe has no portal to date, we are not using this method at the moment to work with our users.

5.5.4 – Plan and prepare the two special evaluation periods during the project.

Achievements: The user evaluation was planned and prepared in February and March 2010 (previous reporting period). The evaluation of the user survey took place in this reporting period. The results are published in D5.8.

Deviation from work plan & remedial action

The update of the project Web site is still slower then expected. The CMS is now in place, but we need a better tutorial for the project members. This will also help to complete the translations, increase the number of languages and deliver more content and information to the Web site. With the CMS an event management tool will be implemented to better manage the conferences and meetings we attended or are relevant for BHL-Europe.

We also need to improve the representation of BHL-Europe on other online platforms like Wikipedia. The work plan has been adapted to implement this by the end of 2010.

We have made no progress with establishing external user testing groups. However, BHL-Europe team members are continuously working with our products and providing ongoing feedback. The user survey was another source of very valuable input and directly influences

---

[44] https://version1.europeana.eu/web/guest/details-bhl/

[45] https://bhl.wikispaces.com/BHL-Europe-slide-library

[46] https://bhl.wikispaces.com/BHL-Europe-posters

our development process.

3.3 Deliverables Status

Deliverables List			
Deliverable No[1]	**Deliverable title**	**Delivery due date[2]**	**Actual date of delivery**
D5.8	First user evaluation report	M 15	M 16
D1.4	Progress Report 3	M 18	M 19
D3.6	Release German prototype	M 18	M 19

---

[1] Deliverable numbers in order of delivery dates: D1 – Dn. Deliverable numbers must indicate which workpackage they relate to, e.g. D2.1 for the first deliverable from workpackage 2).

[2] Month in which the deliverables will be available according to the DoW. Month 0 marking the start of the project, and all delivery dates being relative to this start date.

3.4 Performance Indicators

One important success indicator needs to be mentioned, which is the "Percentage of literature available through Europeana".

After year one, content of BHL-Europe was not accessible through Europeana. In the meantime, 89,860 objects have been ingested. This exceeds the number of objects we planned in the beginning of the project and includes about 78% of all content currently available over all BHL-Europe content providers. Thus, we have already reached the target expected for year 2 of the project, and will continue to increase the quantity of content accessible via Europeana.

As of today, it is estimated that we will not reach 1 million page views through the Europeana Portal at the end of year 2 of the project. After the BHL-Europe content went online on 10 June 2010, only 15,050 page views were counted through Europeana (10 October 2010). It is foreseeable that we will fail with this success indicator. Six reasons for that are discussed:

1) The content ingestion of BHL-Europe was delayed so there were missed opportunities scheduling-wise.

2) Europeana is not yet used to the extent forecast. If BHL-Europe would have had 1 million page views through Europeana, this would be about 15% of the annual page views of Europeana, which is not realistic - BHL-Europe presently provides the relatively small amount of 0.5% of the objects forming the Europeana corpus of 13 million items. With increased usage of Europeana, we can also expect enhanced page views for BHL-Europe. The indexing of the Europeana database by search engines is expected to improve the general situation. Europeana will also try to increase inbound links from other web sites to full object views in the Europeana database. Finally, it is expected that with the new data model (EDM), linked data will significantly increase usage of BHL-Europe data via Europeana.

3) The availability of BHL-Europe content via Europeana has not yet been seriously promoted to date - Europeana has been promoted to a wider public since the Rhine release programme in autumn 2010. We are planning a virtual exhibition to showcase the usability and significance of our content for the general public, which should increase the number of users. Two other virtual exhibitions ('Reading Europe' and 'Art Nouveau') have shown the potential of virtual exhibitions in attracting visitors. In addition, Europeana started their end-user marketing campaign after the Rhine release (incl. road show events) and will continue promotion in the future.

4) The level of access to our content may also have an effect on the general usage of BHL-Europe content through Europeana. In the planning phase of the project we assumed Europeana would present BHL-Europe content on a page level, while it turned out to be more appropriate (consistent with other library aggregators of Europeana) to present BHL content such as books as objects in Europeana.

5) At the moment it is not possible to identify the usage of BHL-Europe content on the Europeana side. Europeana Connect is currently working on log files to be produced for the individual aggregators. Meanwhile we are recording the usage of the BHL-Europe content by recording the usage of the BHL Portal by Europeana as a proxy. We are not recording the usage of LANDOE and NAT content by Europeana. We are also not recording the usage of the BHL-Europe specific content of BnF and UGOE that is currently available in Europeana via other aggregators than BHL-Europe. Thus we are missing users of content of BHL-Europe partners through Europeana. For the

next report we will include statistics for LANDOE and NAT, but for BnF and UGOE the statistics are difficult to generate because their usage statistics in Europeana cover not only the biodiversity content but the entire content set that they deliver.

6) The general user is expected to search for vernacular names rather than Latin names when looking for information on animals and plants. To find this information either needs an improved Taxonomic Intelligence or a full text search. Both items are identified as important features for the final BHL-Europe system, but it was not planned in the beginning at what stage of the project these features could be implemented. We are now working on this and plan for the functionality to be available for the public core release in M28 of the project.

Indicators	Expected vs Actual						
	Year 1 (exp.)	Year 1 (act.)	Year 1.5 (act.)	Year 2 (exp.)	Year 2 (act.)	Year 3 (exp.)	Year 3 (act.)
Number of total accessible pages of biodiversity literature (in million)	17	17.4	35.3	21		25	
Number of pages of biodiversity literature from European content providers to BHL-Europe (in million)	3.4	3.4	3.5	4.1		4.8	
Percentage of literature available through Europeana	20%	0%	78%	50%		100%	
Amount of metadata sets to be imported into the GRIB (in million)	0.15	0.15	0.3	2.5		4.0	
Number of interconnected repositories	7	14	14	20		30	
Number of content providers	20	21	19[1]	25		30	
Number of portal languages	1[2]	1	2	7[3]		12[4]	
Page views through BHL Portal (in million)	1.0	3.7	4.1[5]	2.0		2.5	
Page views through BHL-Europe Portal (in million)	0.0	0.0	0.0	0.5		2.5	
Page views through Europeana Portal (in million)	0.25	0	0.015	1.0		3.0	
Case studies of successful usage of the material by non-scientists	5	5	5	10		15	
Agreements with Rights Holders / Publishers	2	7	7	4		6	

[1] Two potential content providers have signed the MoU but have no funding to digitise content yet. Therefore, they have been removed from the list. A large number of potential new content providers have been identified over the last weeks and negotiations with them are in progress.

[2] English; [3] + French, German, Italian, Portuguese, Spanish; [4] + Danish, Dutch, Czech, Hungarian, Polish, Swedish

[5] This is for the period of one year, i.e. 1.11.2009-31.10.2010.

4 Awareness and Dissemination

4.0 Overview of awareness and dissemination activities

The BHL-Europe project Web site has been online since July 2009. BHL-Europe is now using a CMS to manage the Web site, its multilingual content and regular content updates (Figure 3). The BHL-Europe Newsletter is published on a quarterly basis - issue 5 is published together with issue 6 in a double issue as an exception.
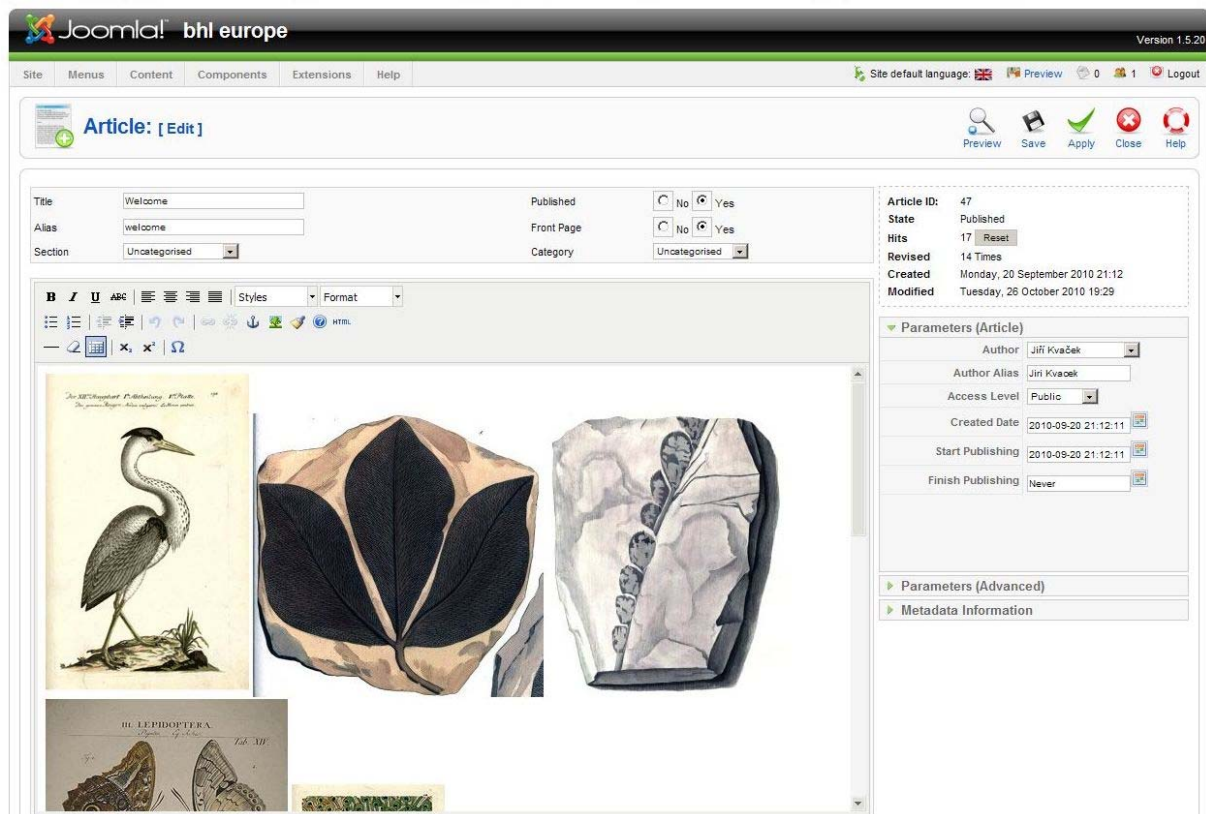
Figure 3: The BHL-Europe CMS.

In recent months, BHL-Europe has been very active in approaching various target users through official conferences and publications. Depending on the audience, presentations or posters were adapted accordingly. We also sent experts from the consortium depending upon the topic of the conference, in order to adequately represent the project.

Two dissemination activities are mentioned particularly in this section. The first is the temporary anniversary exhibition of the Museum für Naturkunde Berlin, where BHL-Europe is present in a multimedia exhibit. The museum is now 200 years old and is presenting this exhibition between 14 September 2010 and 27 February 2011. The exhibition covers the development of science in that time period and how this influenced the collections and scientific programme of the museum. In six steps the visitor is taken on a journey through time: the museum until 1810, the establishment of the new building in 1889, the time of large expeditions until 1918, the museum during World War 2 until 1945, the museum in the former GDR, the Fall of the Wall and the establishment of new structures. In this last section,

a large multimedia station has been designed with texts, objects and videos to illustrate the objectives of large scientific initiatives, projects and networks. Among these projects is BHL-Europe (Figure 4). A banner presents the key objectives of the project, a historic book is on display[47] describing the objects of the museum collection and a video is shown with an interview in which the BHL-Europe project coordinator explains the objectives and benefits of BHL-Europe.

Figure 4: The multimedia station in the anniversary exhibition of the Museum für Naturkunde Berlin: the video (left), the banner (top right), the book (bottom right).

The second activity was the Conference „Deutsches Kulturerbe auf dem Weg in die Europeana" (German Cultural Heritage for Europeana). From 4-5 October 2010, the German Europeana project partners, members of the German Digital Library, and about 200 members of German archives, museums, and libraries met in the Berlin State Library. BHL-Europe (MfN) co-organised the conference, chaired one session, gave a talk and presented flyers and other printed material with information about the project. All information including abstracts are available online[48] (in German). Further information is also provided above in section 3.

The Botanic Garden and Botanical Museum Berlin-Dahlem (Freie Universität Berlin) and the Museum für Naturkunde Berlin will host the BioSystematics[49] conference in February 2011. The scope of the congress is to bring together evolutionary biologists and systematists working on plant, animal, and microscopic organisms to discuss and debate topics of common

---

[47] http://www.biodiversitylibrary.org/item/29960

[48] http://www.armubi.de/tagung2010

[49] http://www.biosyst-berlin-2011.de/

interest. The focus will be on innovative and forward-looking ideas, concepts, and methods in systematic and evolutionary biology. BHL-Europe is organising a symposium on 22 February 2011 entitled "Worldwide Access to Biodiversity Literature: Digital and Virtual Libraries for Taxonomists". This symposium will provide an overview of BHL, BHL-Europe and related activities as well as other digital and virtual library initiatives relevant for taxonomists. BHL-Europe is also organising a workshop entitled "Biodiversity Heritage Library: Hands-on training on search and retrieval of digital literature".

BHL-Europe activities were communicated throughout the Europeana consortium. Europeana provided input to the BHL-Europe newsletter and website and continuously updated the information page for BHL-Europe on the Europeana Group page[50]. There were two short items published about BHL-Europe via RSS feed: Darwin Among Scientific Treasures On Europeana (June 14, 2010 11:56 AM) and Biodiversity Heritage Library Survey (April 20, 2010 11:33 AM). Furthermore, Europeana used Twitter to distribute the call for participation in the Biodiversity Library Survey[51]. The first host images representing BHL-Europe can be seen on the Europeana Imagery Gallery[52]. In addition there is an item published in the Europeana newsletter about the availability of BHL-Europe content[53]. The Europeana Communication group meeting in Edinburgh on 29 April 2010 was attended by representatives of the BHL-Europe consortium. BHL-Europe is now also represented in the Europeana ThoughtLab with the GRIB and the ScanList[54]. This space was established to feature the work of Europeana's partners and to showcase technologies for the improvement of access to Europe's cultural and scientific heritage. The BHL Scan List and the GRIB are presented here among eight other demonstrators of three other projects.

On a national level, MNHN attended training by the National High School of the Information Sciences and Libraries. MNHN also submitted an activity report to the Ministry of Higher Education and Research, regarding the "CADIST" mark, i.e. Centre for Acquisition and Circulation of Scientific and Technical Information. This report also mentions the partnership with BHL-Europe. Naturalis provided training on use of the BHL portal for internal users.

RMCA participated in local and international meetings, events and discussion groups to elucidate and promote BHL-Europe. Future activities include an exhibition (open mid first semester 2011) about Charles Lemaire's expeditions in Congo (1901-1903). The material about Lemaire and his publications is collected for a web site of the exhibition and some of this will also be displayed. Collected and scanned literature (also 'grey' literature) in the framework of the preparation of this exhibition will be made available to BHL-Europe and Europeana. It is also planned to show the specimens collected during the expedition and their original descriptions.

During the Europeana meeting in Edinburgh, the focus was mostly on attracting the general interest user to Europeana. We evaluated the outcomes of the Edinburgh meeting during our review meeting in Vienna in May 2010. We discussed the implications of the Edinburgh meeting for BHL-Europe and agreed to further align our dissemination strategy with Europeana. The concept of virtual exhibitions was introduced in Edinburgh and we are

---

[50] https://version1.europeana.eu/web/guest/details-bhl/

[51] http://twitter.com/europeanaEU

[52] http://version1.europeana.eu/web/guest/imagery/

[53] http://app.e2ma.net/app2/campaigns/archived/1403149/daf0b589c262f18dc941829e2b8ba219/

[54] http://www.europeana.eu/portal/thoughtlab.html

building a virtual exhibition in cooperation with Athena. Based on the discussion in Edinburgh, we are preparing other promotion material purely designed to attract the general public. Some first draft cartoons on the most famous, extinct bird, the Dodo, are already available. Another intention is to enhance information on convenience goods with information on their biodiversity context. In addition to this promotional material, we are also working on the BHL-Europe representation on various online platforms. The Facebook site went live recently with updates on BHL-Europe dissemination activities (Figure 5)[55]. During the Europeana Plenary we were able to discuss with Liam Wyatt (Wikipedia) the presentation of BHL-Europe in Wikipedia. A Weblog was set up, but needs to be formalised and launched (Figure 6). A netvibes account was created to collect all project related websites and social networks (Figure 7)[56].
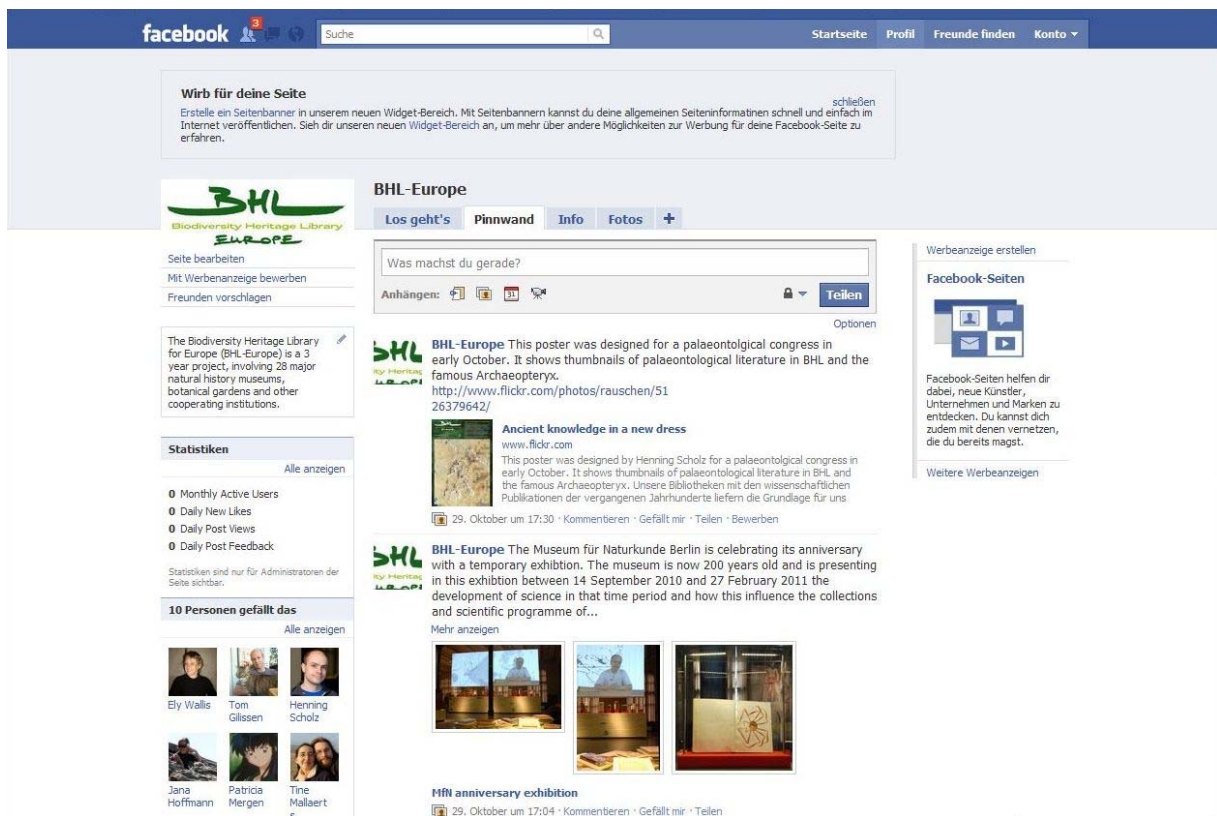
Figure 5: The BHL-Europe Facebook site.

---

[55] http://www.facebook.com/pages/BHL-Europe/151086001600041?v=wall

[56] http://www.netvibes.com/bhl-europe#Projects
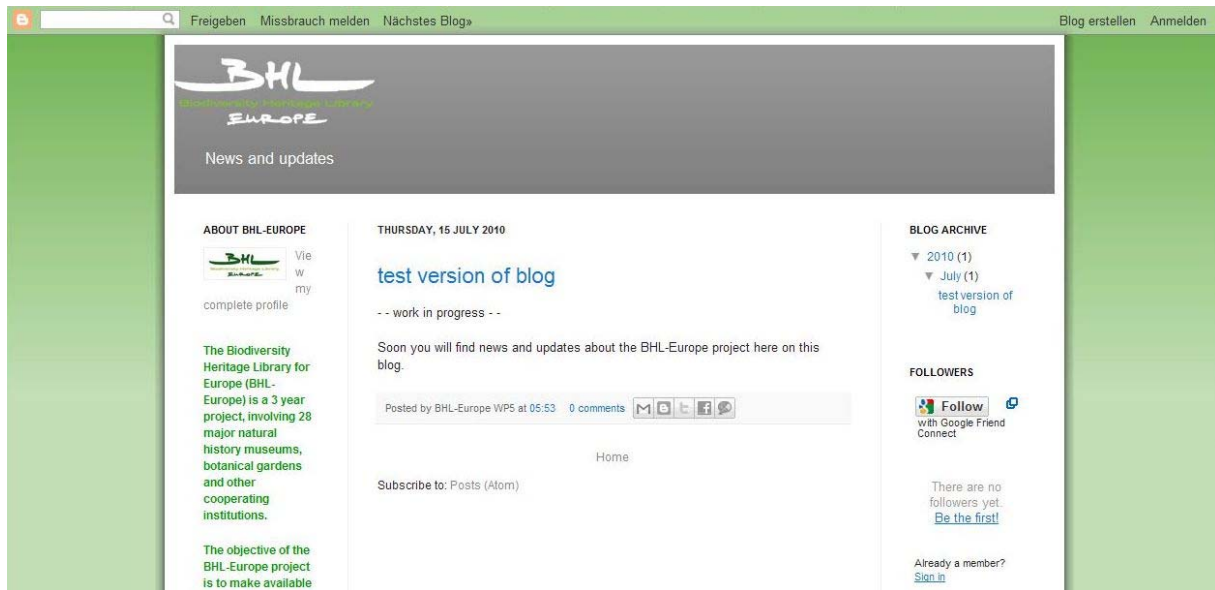
Figure 6: The test version of the BHL-Europe weblog.

Figure 7: The EDIT/Vital & BHL-Europe netvibes site.

4.1 Events and meetings

List of Events & Meetings

Below is a list of meetings and workshops of members of the BHL-Europe team to discuss and work on various aspects of the project.

Name	Location	Start Date	End Date
BHL-Europe Annual Review	Vienna, AT	25/05/2010	28/05/2010
TMB and technical meeting	Vienna, AT	29/05/2010	30/05/2010
TMB and technical meeting	London, UK	25/08/2010	27/08/2010
TMB and technical meeting	Amsterdam, NL	15/10/2010	16/10/2010

Below is a list of conferences and events attended by members of the BHL-Europe team in order to disseminate the project to the various audiences or to collect important information for the progress of the project.

Partner	Title	Start date	End date	Location	Level of dissemination
NBGB	World congress of the International Association of Agricultural Information Specialists (IAALD) joined with the 17th annual meeting European Botanical and Horticultural Libraries group (EBHL)	26/04/2010	29/04/2010	Montpellier, France	Poster, flyers
LANDOE, RMCA	CETAF meeting	27/04/2010	28/04/2010	Linz, Austria	Talk
MfN	Berliner Wirtschaftsgespräche – Forschung vor Ort	05/05/2010	05/05/2010	Berlin, Germany	Talk
RMCA	STERNA follow-up meeting	19/05/2010	19/05/2010	Salzburg, Austria	Talk
NHM	ALA Annual Conference 2010	24/06/2010	30/06/2010	Washington, USA	Talk
NMP	8th Palaeobotany Palynology Conference	06/07/2010	11/07/2010	Budapest, Hungary	Poster, flyers
CSIC	EDIT Summer course	07/2010		Funchal, Madeira	Talk
RBINS, HNHM	IXth European Congress of Entomology	22/08/2010	27/08/2010	Budapest, Hungary	Talk, poster, flyer
NMP	Prague Summer Conference on IP & Innovation	13/09/2010	14/09/2010	Prague, Czech Republic	Flyers
MfN	Klasse, Ordnung, Art – 200 years Museum für Naturkunde (temporary exhibition)	14/09/2010	27/02/2011	Berlin, Germany	Multimedia exhibit
NMP	11th Palaeontological Conference	14/09/2010	16/07/2010	Prague, Czech Republic	Poster, flyers
RBINS	Final conference of Digital Extra: online exhibitions, low cost, big results	16/09/2010	17/09/2010	Brussels, Belgium	Talk
NBGB	International Symposium Botanical Diversity: exploration, understanding	16/09/2010	18/09/2010	Meise, Belgium	Poster, flyers

	and use				
AIT	iPres2010	19/09/2010	24/09/2010	Vienna, Austria	Bilateral business meetings
UBER	Digitale Wissenschaft 2010	20/09/2010	21/09/2010	Köln, Germany	Flyer
MfN	Bioidentify - Tools for identifying biodiversity: progress and problems	20/09/2010	22/09/2010	Paris, France	Talk, flyers
MfN, NHM, RMCA	Global BHL Technical Meeting	22/09/2010	24/09/2010	Woods Hole, USA	Talks
MfN, NHM, RMCA, NHMW, FUB-BGBM	TDWG 2010	26/09/2010	01/10/2010	Woods Hole, USA	Posters
MfN	Europeana Group of Projects Meeting	29/09/2010	29/09/2010	The Hague, The Netherlands	Bilateral business meetings
UBER	Open Access Tage	04/10/2010	05/10/2010	Göttingen, Germany	Flyer
MfN	Deutsches Kulturerbe auf dem Weg in die Europeana	04/10/2010	05/10/2010	Berlin, Germany	Talk, chair, co-organiser
MfN	Jahrestagung der Paläontologischen Gesellschaft	05/10/2010	08/10/2010	Munich, Germany	Poster, flyers
MfN, NBGB, RMCA	3rd meeting Scientific Publishing in Natural History Institutions	07/10/2010	08/10/2010	Copenhagen Denmark	Talks
UH-Viikki	Autum event of the Finnish Biological Society Vanamo	09/10/2010	09/10/2010	Helsinki, Finland	Poster
MfN, NHM, NMP, RMCA	Europeana Plenary and side meetings	13/10/2010	15/10/2010	The Hague/ Amsterdam, The Netherlands	Bilateral business meetings
MfN	11. Oracle Bibliotheken Summit	27/10/2010	28/10/2010	Weimar, Germany	Talk
MfN, UBER	Internal meeting of workgroup electronic publishing at Humboldt-Universität zu Berlin	29/10/2010	29/10/2010	Berlin, Germany	Talk

5 Conclusions

We are now past the halfway point of the project. A number of important decisions have been made and we have achieved the objectives of the project so far. We now have to build the products based on the established foundations. The next six months of the project is a very critical period for the success of BHL-Europe. At the end of this period, much more content will be available in Europeana, the Global References Index to Biodiversity will be released as a tool for librarians and scientists, and we will have made significant progress towards implementing the live BHL-Europe system.